TABLE 1. Summary of molecular analyses

Patients				*POR* mutations	
Case	Karyotype	Age (yr)	Inheritance	Nucleotide changes[a]	Aminoacid changes
Group A: homozygotes for R457H					
1	46,XY	5.0	Sporadic	1370G>A/1370G>A	R457H/R457H
2	46,XY	23.8	Familial-A	1370G>A/1370G>A	R457H/R457H
3	46,XY	22.6	Familial-A	1370G>A/1370G>A	R457H/R457H
4	46,XY	6.7	Sporadic	1370G>A/1370G>A	R457H/R457H
5	46,XY	0.4	Sporadic	1370G>A/1370G>A	R457H/R457H
6	46,XX	0.4	Sporadic	1370G>A/1370G>A	R457H/R457H
7	46,XX	0.4	Sporadic	1370G>A/1370G>A	R457H/R457H
8	46,XX	2.0	Sporadic	1370G>A/1370G>A	R457H/R457H
9	46,XX	14.1	Sporadic	1370G>A/1370G>A	R457H/R457H
10	46,XX	15.0	Familial-A (P)	1370G>A/1370G>A	R457H/R457H
11	46,XX	3.0	Sporadic	1370G>A/1370G>A	R457H/R457H
12	46,XX	0.2	Sporadic	1370G>A/1370G>A	R457H/R457H
13	46,XX	0.1	Sporadic	1370G>A/1370G>A	R457H/R457H
14	46,XX	18.0	Sporadic	1370G>A/1370G>A	R457H/R457H
Group B: compound heterozygotes for R457H and an apparently null mutation					
15	46,XY	16.8	Familial-B (P)	1370G>A/601C>T	R457H/Q201X
16	46,XY	15.7	Familial-B	1370G>A/601C>T	R457H/Q201X
17	46,XY	14.8	Sporadic	1370G>A/1329-1330insC	R457H/I444fsX449
18	46,XY	17.5	Sporadic	1370G>A/(15A>G)	R457H/Non-transcribed (G5G)[b]
19	46,XY	2.1	Sporadic	1370G>A/143delG	R457H/R48fsX63
20	46,XY	0.2	Sporadic	1370G>A/1665delG	R457H/Q555fsX612
21	46,XY	13.1	Sporadic	1370G>A/(−)[c]	R457H/DeltaExons 2–13[d]
22	46,XX	9.0	Sporadic	1370G>A/IVS7+1G>A	R457H/IVS7+1G>A
23	46,XX	14.8	Sporadic	1370G>A/1698-1699insC	R457H/Y567fsX574
24	46,XX	13.2	Sporadic	1370G>A/1329-1330insC	R457H/I444fsX449
25	46,XX	12.9	Familial-B	1370G>A/601C>T	R457H/Q201X
26	46,XX	6.6	Sporadic	1370G>A/(−)[c]	R457H/Non-transcribed[b]
27	46,XX	4.2	Sporadic	1370G>A/(−)[c]	R457H/Non-transcribed[b]
28	46,XX	17.0	Sporadic	1370G>A/1329-1330insC	R457H/I444fsX449
Group C: other compound heterozygotes					
29	46,XY	0.4	Sporadic	1370G>A/1386-1387insATCGCC	R457H/A462-S463insIA
30	46,XY	23.5	Familial-C (P)	1370G>A/1835-1858del[e]	R457H/L612-W620delinsR
31	46,XY	18.0	Familial-C	1370G>A/1835-1858del[e]	R457H/L612-W620delinsR
32	46,XY	17.9	Familial-D	1733A>G/1329-1330insC	Y578C/I444fsX449
33	46,XX	0.8	Sporadic	1370G>A/1738G>C	R457H/E580Q
34	46,XX	0.7	Sporadic	1370G>A/1042-1044delGTC	R457H/348delV
35	46,XX	0.5	Familial-D (P)	1733A>G/1329-1330insC	Y578C/I444fsX449

The genomic position corresponding to each mutation based on NC_000007.12 sequence at the National Center for Biotechnology Information database (Bethesda, MD) is as follows: R457H, 75452433G>A; Q201X, 75448386C>T; I444fsX449, 75452391-2insC; G5G, 75421261A>G; R48fsX63, 75421389delG; Q555fsX612, 75453099delG; IVS7 + 1G>A, 75448861G>A; Y567fsX574, 75453205-6insC; A462-S463insIA, 75452349-50insATCGCC; L612-W620delinsR, 75453432-55delTAAAGCAAGACCGAGAGCACCTGT; Y578C, 75453237A>G; E580Q, 75453245G>C; and 348delV, 75451086-88delGTC. Cases 1–3, 6–10, 15–18, 22–26, 29–33, and 35 have been reported previously (6, 8, 9), and the remaining 12 cases were first examined in this study. P, Proband.

[a] The A of the ATG encoding the initiator methionine residue of the predicted translation product is denoted position +1.

[b] The allele with G5G and the apparently normal alleles are not trasncribed into mRNA.

[c] The (−) symbol indicates the absence of a recognizable mutation on the exonic sequences.

[d] An intragenic microdeletion involving exons 2–13.

[e] 1835-1858delTAAAGCAAGACCGAGAGCACCTGT.

subjects of both sexes (854 males and 909 females), the M data of the patients were expressed as the SD score to allow for the comparison among patients of different sexes and ages.

Statistical analysis

Statistical significance of the frequency of clinical features was analyzed by the Fisher's exact probability test, and that of the median of nonpaired and paired variables was examined by the Mann-Whitney's *U* test and the Wilcoxon signed-rank test, respectively. $P < 0.05$ was considered significant.

Results

POR mutations

The results are summarized in Table 1. Direct sequencing revealed 12 types of mutations and one silent substitution (G5G) (Fig. 1A), with R457H being identified in 40 of the 58 alleles (~70%) in 25 sporadic cases and four probands of families A–D. Of the 12 mutations, R48fsX63, Q555fsX612, and 348delV were first identified in this study. These mutations were absent in 100 control subjects.

balt1/zeg-jcem/zeg-jcem/zeg00509/zeg6496-09a | angnes | S=6 | 4/7/09 | 21:45 | 4/Color Figure(s): F1,F2,F4 | Art: 08-2816 | Input-ss

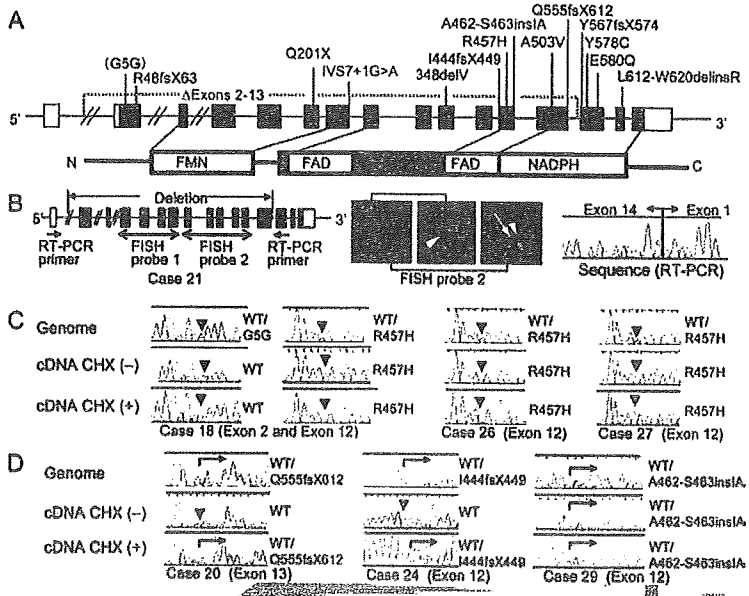


FIG. 1. Mutation analysis of *POR*. A, Schematic representation of the *POR* gene and the positions of identified mutations. The Japanese founder mutation R457H is shown in *red*, other disease-causing mutations in *black*, and the common A503V variant in *blue*. *Upper diagram*, The genomic structure comprising 16 exons. The *black* and *white boxes* denote the coding and the untranslated regions, respectively. *Lower diagram*, The protein structure consisting of the cofactor binding domains (FMN: flavin mononucleotide; FAD: flavin-adenine dinucleotide; and NADPH: nicotinamide-adenine dinucleotide phosphate; reduced) and the connecting domain (*stippled area*). B, FISH and RT-PCR sequencing analyses in case 21. *Left diagram*, The positions of the two FISH probes and those of the primers for RT-PCR. *Middle diagram*, FISH findings showing two signals for *D7Z1* (*arrowheads*) and a single signal for *POR* (*arrow*) delineated by the FISH probe 2. *Right diagram*, RT-PCR sequencing indicating the fusion between exons 1 and 14 (the deletion of exons 2–13). C, Transcription failure in cases 18, 26, and 27. Although heterozygosity for R457H is delineated for the genomic DNA, RT-PCR sequencing indicates absent expression of the wild-type (WT) alleles in the three cases. Similarly, although heterozygosity for G5G is shown for the genomic DNA of case 18, RT-PCR sequencing reveals no expression of the G5G allele. Such lack of transcripts is not recovered by CHX. D, Nonsense-mediated mRNA decay in cases 20 and 24 but not case 29. Although heterozygosity for the mutations is shown for the genomic DNA, RT-PCR sequencing delineates the WT alleles only before CHX treatment and the heterozygosity after CHX treatment in cases 20 and 24. The NMD is not observed in case 29.

Fifteen cases were apparently homozygous for R457H, and hemizygosity was excluded in 14 of the 15 cases by parental analysis indicating heterozygosity for R457H in both parents (cases 1–3, 6–11, and 13) and by FISH analysis with two FISH probes (cases 4, 5, 12, and 14). Notably, however, FISH analysis delineated a heterozygous microdeletion in case 21, and RT-PCR sequencing analysis revealed loss of exons 2–13 in this case (Fig. 1B). The mother was heterozygous for R457H, and the father was heterozygous for the intragenic microdeletion.

Three cases were apparently heterozygous for R457H (cases 18, 26, and 27), although case 18 also had G5G. However, RT-PCR sequencing analysis using lymphoblastoid cell lines showed nearly complete absence of mRNA derived from the apparently normal alleles in the three cases (Fig. 1C). The mRNA remained undetected after CHX treatment, indicating transcription failure.

Of the 11 other types of mutations, the nonsense and four frameshift mutations (Q201X, R48fsX63, I444fsX449, Q555fsX612, and Y567fsX574) leading to premature termination and the conserved splice donor site mutation (IVS7+1G>A) appeared to be null mutations, whereas the remaining five mutations (Y578C, E580Q, 348delV, A462-S463insIA, and L612-W620delinsR) were unknown for residual activities. Indeed, RT-PCR sequencing analysis performed before and after CHX treatment in three cases with available lymphoblastoid cell lines demonstrated that the alleles carrying Q555fsX612 and I444fsX449 underwent NMD, whereas the allele harboring A462-S463insIA escaped NMD (Fig. 1D).

The common A503V variant was absent from cases of group A and was identified in four cases of group B (cases 22, 23, 26, and 27) and four cases of group C (cases 29–31, and 34). The eight cases with A503V were all compound heterozygotes with R457H and another mutation, and direct sequencing for subcloned PCR products encompassing both 457th and 503rd codons revealed lack of coexistence of R457H and A503V. Thus, it was indicated that the A503V variant was absent from all of the 47 alleles carrying R457H and was present on alleles carrying IVS7+1G>A, Y567fsX574, A462-S463insIA, L612-W620delinsR, and 348delV and on the two nontranscribed alleles.

Classification of the patients

On the basis of the above results, the 35 cases were classified into three groups: group A, homozygotes for R457H (cases 1–14); group B, compound heterozygotes for R457H and one apparently null mutation (cases 15–28); and group C, other types of compound heterozygotes (cases 29–35) (Table 1). The residual POR activity was predicted to be higher in group A than group B, although it was unknown for group C. In addition, group B was subclassified into A503V-positive cases (cases 22, 23, 26, and 27) and negative cases (cases 15–21, 24, 25, and 28).

Clinical features

The prevalence of each clinical feature in groups A–C is summarized in Table 2, together with its comparison between groups A and B. The sex ratio was similar between groups A and B, as was the median age. T2,AQ:D

ABS-compatible skeletal features were definitely more prevalent in group B than group A (Table 2 and supplementary Fig. 1, published as supplemental data on The Endocrine Society's Journals Online Web site at http://jcem.endojournals.org). In particular, severe brachycephaly, elbow joint synostosis, and choanal stenosis were exclusively identified in group B.

Adrenal steroidogenic dysfunction was biochemically identified in all cases, with some difference between groups A and B. Blood ACTH was normal or elevated at the baseline, 17-OHP was normal or elevated at the baseline and above the normal range after ACTH stimulation, and cortisol was normal at the baseline but barely responded to ACTH stimulation (Fig. 2A). F2 Significant difference between groups A and B was identified for basal 17-OHP value ($P = 0.044$) and basal and ACTH-stimulated cortisol values ($P = 0.018$ and $P = 0.022$). Urine Ms of progesterone and 17-OHP were elevated, whereas those of an-

TABLE 2. The prevalence of each clinical feature in groups A–C and its comparison between groups A and B

	Group A (n = 14)	Group B (n = 14)	Group C (n = 7)	Groups A *vs.* B (*P* value)
Sex (male:female)	5:9	7:7	4:3	0.35
Age (median, range, yr)	4.0 (0.1–23.8)	13.1 (0.2–17.5)	0.8 (0.4–23.5)	0.19
Skeletal features				
Any skeletal feature	7/14	14/14	7/7	0.0029
Brachycephaly (overt)	0/14	14/14	6/7[a]	0.000000025
Elbow joint synostosis[b]	0/14	7/14	4/7	0.0029
Arachnodactyly (overt)	5/14	14/14	7/7	0.048
Choanal stenosis	0/14	5/14	1/7	0.020
Joint contracture	7/14	14/14	7/7	0.0029
Adrenal dysfunction				
Adrenal crisis	0/14	4/14	1/7[c]	0.049
Detection by mass screening[d]	5/8	3/8	2/4	0.31
46,XY DSD				
Any genital feature at birth	1/5[e]	3/7[f]	3/4	0.42
Hypospadias	0/5	2/7	1/4	0.32
Cryptorchidism	0/5	3/7	2/4	0.16
Micropenis	1/5	2/7	3/4	0.64
46,XX DSD				
Any genital feature at birth	9/9[e]	7/7[f]	3/3	1.0
Clitoromegaly	8/9	5/7	3/3	0.40
Labial fusion	8/9	5/7	2/3	0.40
Common urogenital sinus	2/9	2/7	0/3	0.61
Maternal virilization	8/14	5/14	4/7	0.22
Pubertal failure, 46,XY				
Delayed (>2 SD) or no pubertal sign	0/2[g]	3/4[h]	2/3	0.20
Small testis (<2 SD)	0/2	2/4	1/3	0.40
Primary hypogonadism[i]	0/2	2/2	3/3	0.17
Pubertal failure, 46,XX				
Delayed (>2 SD) or no pubertal sign	3/3[g]	4/4[h]		1.0
Delayed (>2 SD) or no menses	0/2[j]	2/2		0.17
Primary hypogonadism[i]	3/3	3/3		1.0
Polycystic ovary	4/9	3/6	1/3	0.62

The denominators indicate the number of patients examined for the presence or absence of each feature, and the numerators represent the number of patients assessed to be positive for that feature; thus, the differences between the denominators and numerators denote the number of patients evaluated to be negative for that feature.

[a] Severe craniosynostosis is absent in case 33 with two missense mutations.

[b] Humeroradial, humeroulnar, or radioulnar synostosis.

[c] Adrenal crisis has been manifested by case 35 with Y578C and I444fsX449.

[d] The measurement of 17-OHP in the mass screening for 21-hydroxylase deficiency has been performed since 1988 in Japan.

[e,f] DSD is more frequent in 46,XX cases than 46,XY cases in groups A ($P = 0.0050$) and B ($P = 0.035$).

[g,h] The *P* values between 46,XY and 46,XX cases are 0.19 for group A and 0.50 for group B.

[i] Elevated gonadotropins (LH and/or FSH) and/or decreased T or E_2, as compared with age- and sex-matched reference data.

[j] Only a few vaginal spottings.

drostenedione, 11-deoxycortisol, cortisol, and aldosterone grossly remained within the normal range (Fig. 2B). The M ratio indicating 17α-hydroxylase activity remained almost normal, consistent with the elevation of both substrates and products, whereas the M ratios indicating 17,20 lyase and 21-hydroxylase activities were grossly decreased. Significant difference between groups A and B was identified for MS of progesterone ($P = 0.044$), those of 17-OHP ($P = 0.022$), those of aldosterone ($P = 0.0084$), and M ratio indicating 17,20 lyase activity ($P = 0.011$). Adrenal crisis was observed only in group B with a significant difference between groups A and B, whereas the detection frequency of elevated 17-OHP in mass screening was similar between groups A and B (Table 2).

DSD was more prevalent in 46,XX cases than 46,XY cases in both groups A and B (Table 2, footnote, and supplementary Fig. 2). 46,XY DSD in group A was micropenis in one case, and that in group B included more severe phenotypes. By contrast, 46,XX DSD was invariably identified in both groups A and B. Maternal virilization during pregnancy was often found in groups A and B with a similar prevalence. Serum T of case 20, aged 0.2 yr in group B, was 6.5 and 7.6 nmol/liter (1.9 and 2.2 ng/ml) before and after hCG stimulation, respectively.

Pubertal development was apparently normal in two 46,XY cases of group A and one of four 46,XY cases in group B and was invariably affected in 46,XX cases in both groups A and B (Table 2). In family A of group A, cases 2 and 3 exhibited full pubertal development with testis volume of 20 ml, whereas case 10 had obvious pubertal failure with Tanner B2 stage. T value of case 18, aged 17.5 yr in group B, was low at the baseline (0.7 nmol/liter,

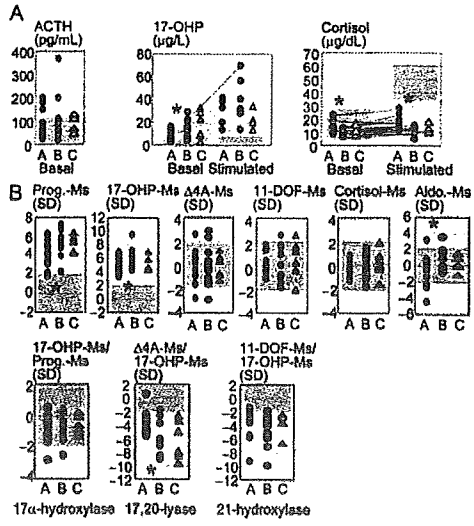


FIG. 2. Adrenal steroidogenic dysfunctions in groups A–C. *Light blue areas* represent the normal ranges. *Red asterisks* indicate the presence of significant differences between groups A and B, A: Basal and ACTH-stimulated blood hormone values. B, Basal urine steroid M values. Prog, Progesterone; Δ4A, androstenedione; 11DOF, 11-deoxycortisol; Aldo, aldosterone.

0.2 ng/ml) and poorly responded to hCG stimulation (1.0 nmol/liter, 0.3 ng/ml). PCO was observed in infantile or pubertal cases with a similar frequency between groups A and B, and cases 22 and 24 had ovarian torsion. Notably, bilateral ovarian cysts of case 10 markedly reduced in size after treatment with estradiol (E_2) (supplementary Fig. 3).

Long-term growth patterns were obtained in eight cases (Fig. 3). Whereas childhood heights tended to be high in both groups A and B, pubertal growth was different between the two groups. Cases in group A lacked obvious pubertal growth spurt but continued to grow for a long term, attaining tall adult heights, whereas those in group B showed rather compromised pubertal growth with worsening of scoliosis (supplementary Fig. 1).

There was no phenotypic difference between A503V-positive and -negative cases of group B (supplementary Table 2). In addition, the phenotypes in group C were grossly similar to those in group B (Table 2). In particular, craniosynostosis was identified in all cases except for case 33 with R457H and E580Q, and adrenal crisis was manifested by case 35 with Y578C and I444fsX449.

Discussion

Molecular studies

Detailed molecular studies were performed in this study, providing two notable findings. First, all 35 cases were found to be homozygotes or compound heterozygotes for *POR* mutations including intragenic microdeletion and transcription failure. Because the microdeletion was found in case 21 with apparent R457H homozygosity, such a microdeletion might be hidden in the previously reported patients with apparent homozygosity (1, 5). Similarly, because transcription failure was invariably identified in cases 18, 26, and 27 with apparent heterozygosity, it may also underlie in the previously reported patients with apparent heterozygosity (4, 5, 10). In this regard, it is likely that the three cases carry a mutation in a hitherto unidentified *cis*-regulatory sequence(s) for the transcription of *POR*, as has been reported for several genes (24).

Second, RT-PCR sequence analysis indicated the occurrence of NMD in the two frameshift mutations (I444fsX449 and Q555fsX612). In this context, all the premature termination codons caused by the nonsense and the four frameshift mutations satisfy the positional conditions for the occurrence of NMD that functions as an mRNA surveillance mechanism to prevent the formation of aberrant proteins (13, 14). Thus, it is likely that the remaining three mutations (Q201X, R48fsX63, and Y567fsX574) are also null mutations subject to NMD *in vivo*.

Genotype-phenotype correlations

Genotype-phenotype correlations also provide several informative findings. Skeletal features were clearly different between groups A and B. Because cholesterol production in skeletal tissues is carried out in a simple one way manner (Fig. 4), this would explain why the skeletal phenotype is obviously dependent on the R457H dosage, reflecting the residual activity. It is likely that the threshold level for the development of severe skeletal phenotypes resides between a single copy and two copies of the R457H residual activity.

Adrenal steroidogenic dysfunction was grossly similar between groups A and B, although it was somewhat milder in group A than group B. Such a relatively minor role of R457H dosage in adrenal steroidogenesis

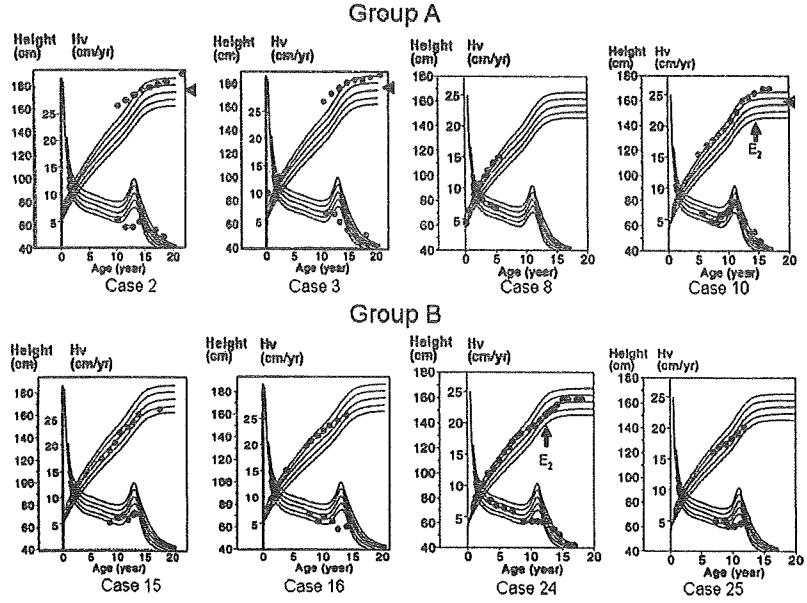


FIG. 3. Growth charts of eight cases plotted on the sex-matched longitudinal growth curves for the normal Japanese children (+2 SD, +1 SD, the mean, −1 SD, and −2 SD). The *triangles* in cases 2, 3, and 10 represent the target heights. Cases 10 and 24 are placed on E_2 replacement therapy. Hv, Height velocity.

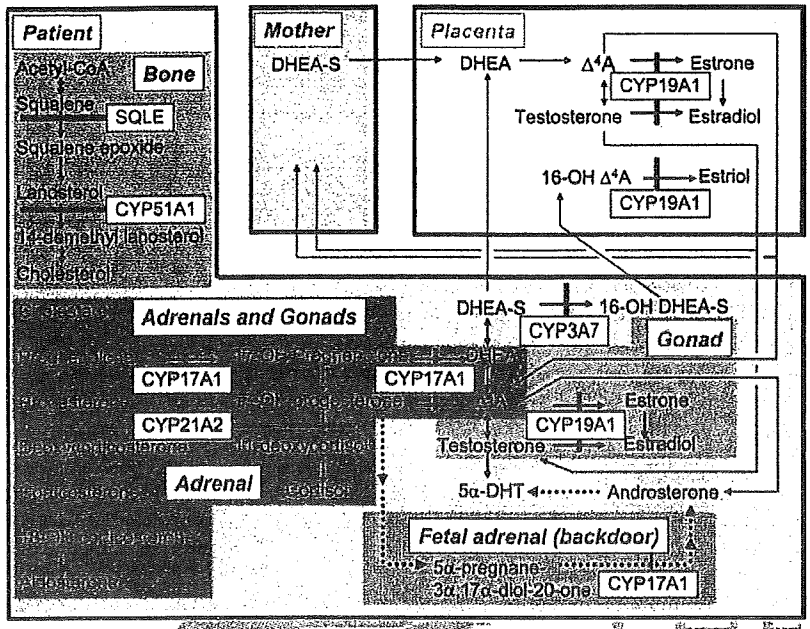


FIG. 4. Simplified schematic representation indicating impaired cholesterologenesis and steroidogenesis in PORD. DHEA, Dehydroepiandrosterone; DHEA-S, DHEA sulfate; Δ^4A, androstenedione; DHT, dihydrotestosterone; SQLE, CYP51A1, CYP17A1, CYP21A2, CYP19A1, and CYP3A7 are POR-dependent enzymes. The important Ms only are shown, and the reaction steps in which some Ms are omitted are indicated by two tandem arrows. Note that the amount of estriol synthesized in the placenta far exceeds the total amount of estrone and E_2 (~10 times) (34).

may primarily be due to the complexity of steroidogenesis in PORD (Fig. 4). For example, both production and degradation of 17-OHP are carried out by POR-dependent enzymes, and such enzymatic reactions would depend on the R457H dosage and the differential supporting activity of the R457H protein for target enzymes as well as the amount of substrates and products. Furthermore, the basal cortisol values imply that the baseline steroidogenic capacity can grossly be sustained, even in group B. Indeed, whereas basal blood 17-OHP values were significantly higher in group B than group A, some of them remained within the normal range, and several cases of both groups were not detected in neonatal mass screening. Nevertheless, the R457H dosage would have important clinical relevance, because the ACTH-stimulated blood cortisol was drastically reduced especially in group B, and adrenal crisis was observed only in group B. Furthermore, because 17,20 lyase activity alone was significantly different between groups A and B (Fig. 2B), this would provide further support for the previous finding that 17,20 lyase activity is the most sensitive index of defective POR activity (5, 15).

46,XY DSD was not so remarkable, whereas 46,XX DSD was invariably identified. This suggests a mildly reduced androgen production in genetic males and a definitely excessive androgen production in genetic females. In this context, there are three androgen sources during the fetal life in PORD, *i.e.* the fetal testis, backdoor pathway, and placenta (3, 4, 9, 25, 26) (Fig. 4). For fetal testicular T production specific to 46,XY cases, placental hCG-stimulated T production around the critical period for sex development would be more compromised in group B than group A because testicular T production is performed in a simple one-way manner, as in cholesterologenesis. Furthermore, because T responses to hCG stimulation were reduced, at least in the two examined cases of group B, this implies the compromised maximum T production capacity. By contrast, the backdoor- and placenta-derived androgen productions common to both 46,XY and 46,XX cases may be similar between groups A and B: 1) whereas 17-OHP as the source metabolite for the backdoor pathway is higher in group B than group A, the supporting activity for fetal adrenal CYP17A1 involved in the backdoor pathway would be lower in group B than group A; and 2) whereas fetal adrenal derived dehydroepiandrosterone as the source metabolite for placental androgens would be lower in group B than group A (4, 9, 25), the residual supporting activity for placental CYP19A1 would be lower in group B than group A. Thus, the total amount of androgens would be relatively well preserved in 46,XY cases with a mild difference in the fetal testis-derived T between groups A and B and invariably and similarly increased in 46,XX cases of both groups A and B. Furthermore, this notion explains why maternal virilization during pregnancy was similar between groups A and B because it is primarily due to androgens of the placental origin rather than the fetal gonadal or the backdoor origin (3, 4, 25).

Assessment of pubertal development was possible in a limited number of patients. However, pubertal development appeared to differ between groups A and B and between 46,XY and 46,XX cases. In this regard, T and E_2 biosynthesis during puberty is also performed in a simple one-way manner, and T production is mediated by CYP17A1 and E_2 production is mediated by both CYP17A1 and CYP19A1 (Fig. 4). Thus, gonadal steroid production would depend on the R457H dosage, with T production being less compromised than E_2 production. In addition, our observation suggests the frequent occurrence of PCO in infancy and puberty when gonadotropins are physiologically elevated (27) and the beneficial effect of estrogen replacement therapy in the amelioration of PCO.

Evaluation of growth pattern also remained fragmentary. However, two implications are possible. First, the intrinsic skeletal abnormalities may be relevant to the growth pattern. Indeed, relative tall stature in childhood may be compatible with the elongation of long bones as indicated by arachnodactyly and dolichostenomelia, and worsening of scoliosis during puberty in group B would also be consistent with the low POR activity (supplementary Fig. 1). Second, the spontaneous pubertal growth pattern of cases 2 and 3 without scoliosis is considered to represent a mild form of that of male patients with aromatase deficiency (28, 29). Such a qualitatively similar but quantitatively different pubertal growth pattern would be explained by assuming a drastically attenuated but not abolished *in vivo* supporting function of the R457H protein for aromatase.

Lastly, clinical features were similar between A503V-positive and -negative cases in group B. However, this would not argue

against a possible phenotypic effect of mildly hypomorphic A503V, because A503V of the four cases in group B was present on the alleles carrying apparently null mutations. Thus, it remains unknown whether A503V can modify phenotypic features in PORD, although the previous study argues against a modifying effect of A503V on clinical phenotypes in 21-hydroxylase deficiency (30). Furthermore, because A503V was absent from all of 47 alleles carrying R457H, this would provide further support for the previous notion that R457H is a founder mutation accompanied by a specific haplotype (6, 7). Thus, whereas A503V was identified in only eight of the 70 alleles (11.4%) in this study, this frequency is obviously biased by the high prevalence of R457H in Japanese patients. Rather, the frequency of A503V in R457H-negative alleles suggests that the prevalence of A503V is considerably high in the Japanese population, as reported in other populations (from 19.1% in African American to 36.7% in Chinese American) (15).

Remarks and conclusion

It should be pointed out that the results are totally based on the studies of Japanese patients. In this regard, A287P is common in Caucasian patients (4, 5), and clinical studies in 10 A287P-positive patients including three homozygotes (five with 46,XY and five with 46,XX) have suggested phenotypic similarities and differences between R457H-positive patients and A287P-positive patients: 1) skeletal phenotype is usually obvious and appears to be grossly dependent on the A287P dosage; 2) 46,XY DSD is variable and is apparently independent of the A287P dosage; 3) 46,XX DSD is also variable and absent in one A287P homozygote and one of four compound heterozygotes with A287P; and 4) maternal virilization during pregnancy is not described (1, 2, 5, 31, 32). Thus, skeletal phenotype would be explained by assuming that both R457H and A287P have drastically lost supporting activities for CYP51A1 and/or SQLE involved in cholesterologenesis, although functional studies have not been performed. Furthermore, clinical features relevant to steroidogenic dysfunction would be grossly consistent with the previous *in vitro* functional data. It has been reported that R457H yields only 1–3% supporting activities for 17α-hydroxylase and aromatase, and virtually no activity for 17,20 lyase, whereas A287P provides supporting activities of about 40% for 17α-hydroxylase, about 20% for 17,20 lyase, about 70% for 21-hydroxylase, and about 100% for aromatase (1, 5, 11, 33). Thus, the relative activities of frontdoor and backdoor pathways would be different largely between R457H-positive and A287P-positive patients, and placental T production would remain minor, if any, in A287P-positive patients. Collectively, the Japanese data would not apply simply to other populations.

In conclusion, the present study in Japanese patients argues against the heterozygote manifestation and suggests that the residual POR activity reflected by the R457H dosage constitutes the underlying factor for the clinical variability in some features but not other features, probably because of the simplicity and the complexity of the POR-dependent metabolic pathways relevant to each phenotype. Further studies including genotype-phenotype analyses in various ethnic groups will permit a better clarification of the molecular and clinical characteristics of PORD.

Acknowledgments

We thank Drs. T. Ohashi, R. Takeda, and I. Fujiwara for providing us with clinical data and blood samples of the patients and their family members.

Address all correspondence and requests for reprints to: Dr. M. Fukami, Department of Endocrinology and Metabolism, National Research Institute for Child Health and Development, 2-10-1 Ohkura, Setagaya, Tokyo 157-8535, Japan. E-mail: mfukami@nch.go.jp.

This work was supported by Grants 20C-2, H18-005, and H20-004 from the Ministry of Health, Labor, and Welfare and Grants 16086215, 19390290, and 20390265 from the Ministry of Education, Culture, Sports, Science, and Technology.

Disclosure Summary: The authors have nothing to declare.

References

1. Flück CE, Tajima T, Pandey AV, Arlt W, Okuhara K, Verge CF, Jabs EW, Mendonca BB, Fujieda K, Miller WL 2004 Mutant P450 oxidoreductase causes disordered steroidogenesis with and without Antley-Bixler syndrome. Nat Genet 36:228–230
2. Arlt W, Walker EA, Draper N, Ivison HE, Ride JP, Hammer F, Chalder SM, Borucka-Mankiewicz M, Hauffa BP, Malunowicz EM, Stewart PM, Shackleton CH 2004 Congenital adrenal hyperplasia caused by mutant P450 oxidoreductase and human androgen synthesis: analytical study. Lancet 363:2128–2135
3. Miller WL 2004 P450 oxidoreductase deficiency: a new disorder of steroidogenesis with multiple clinical manifestations. Trends Endocrinol Metab 15:311–315
4. Scott RR, Miller WL 2008 Genetic and clinical features of P450 oxidoreductase deficiency. Horm Res 69:266–275
5. Huang N, Pandey AV, Agrawal V, Reardon W, Lapunzina PD, Mowat D, Jabs EW, Van Vliet G, Sack J, Fluck CE, Miller WL 2005 Diversity and function of mutations in P450 oxidoreductase in patients with Antley-Bixler syndrome and disordered steroidogenesis. Am J Hum Genet 76:729–749
6. Fukami M, Horikawa R, Nagai T, Tanaka T, Naiki Y, Sato N, Okuyama T, Nakai H, Soneda S, Tachibana K, Matsuo N, Sato S, Homma K, Nishimura G, Hasegawa T, Ogata T 2005 Cytochrome P450 oxidoreductase gene mutations and Antley-Bixler syndrome with abnormal genitalia and/or impaired steroidogenesis: molecular and clinical studies in 10 patients. J Clin Endocrinol Metab 90:414–426
7. Adachi M, Asakura Y, Matsuo M, Yamamoto T, Hanaki K, Arlt W 2006 POR R457H is a global founder mutation causing Antley-Bixler syndrome with autosomal recessive trait. Am J Med Genet A 140:633–635
8. Fukami M, Hasegawa T, Horikawa R, Ohashi T, Nishimura G, Homma K, Ogata T 2006 Cytochrome P450 oxidoreductase deficiency in three patients initially regarded as having 21-hydroxylase deficiency and/or aromatase deficiency: diagnostic value of urine steroid hormone analysis. Pediatr Res 59:276–280
9. Homma K, Hasegawa T, Nagai T, Adachi M, Horikawa R, Fujiwara I, Tajima T, Takeda R, Fukami M, Ogata T 2006 Urine steroid hormone profile analysis in cytochrome P450 oxidoreductase deficiency: implication for the backdoor pathway to dihydrotestosterone. J Clin Endocrinol Metab 91:2643–2649
10. Scott RR, Gomes LG, Huang N, Van Vliet G, Miller WL 2007 Apparent manifesting heterozygosity in P450 oxidoreductase deficiency and its effect on coexisting 21-hydroxylase deficiency. J Clin Endocrinol Metab 92:2318–2322
11. Pandey AV, Kempna P, Hofer G, Mullis PE, Flück CE 2007 Modulation of human CYP19A1 activity by mutant NADPH P450 oxidoreductase. Mol Endocrinol 21:2579–2595
12. Agrawal V, Huang N, Miller WL 2008 Pharmacogenetics of P450 oxidoreductase: effect of sequence variants on activities of CYP1A2 and CYP2C19. Pharmacogenet Genomics 18:569–576
13. Carter MS, Doskow J, Morris P, Li S, Nhim RP, Sandstedt S, Wilkinson MF 1995 A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts *in vivo* is reversed by protein synthesis inhibitors *in vitro*. J Biol Chem 270:28995–29003
14. Kuzmiak HA, Maquat LE 2006 Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges. Trends Mol Med 12:306–316
15. Huang N, Agrawal V, Giacomini KM, Miller WL 2008 Genetics of P450 oxidoreductase: sequence variation in 842 individuals of four ethnicities and activities of 15 missense mutations. Proc Natl Acad Sci USA 105:1733–1738
16. Homma K, Hasegawa T, Masumoto M, Takeshita E, Watanabe K, Chiba H,

Kurosawa T, Takahashi T, Matsuo N 2003 Reference values for urinary steroids in Japanese newborn infants: gas chromatography/mass spectrometry in selected ion monitoring. Endocr J 50:783–792
17. Matsuo N 1993 Skeletal and sexual maturation in Japanese children. Clin Pediatr Endocrinol 2(Suppl):1–4
18. Fujieda K, Matsuura N 1987 Growth and maturation in the male genitalia from birth to adolescence. I: change of testicular volume; II: change of penile length. Acta Paediatr Jpn 29:214–223
19. Yokoya S, Kato K, Suwa S 1983 Penile and clitoral sizes in premature newborns, full term neonates, infants, and children. Horumon to Rinsho 31:1215–1220 (in Japanese)
20. Suwa S, Tachibana K, Maesaka H, Tanaka T, Yokoya S 1992 Longitudinal standards for height and height velocity for Japanese children from birth to maturity. Clin Pediatr Endocrinol 1:5–13
21. Lashansky G, Saenger P, Fishman K, Gautier T, Mayes D, Berg G, Di Martino-Nardi J, Reiter E 1991 Normative data for adrenal steroidogenesis in a healthy pediatric population: age- and sex-related changes after adrenocorticotropin stimulation. J Clin Endocrinol Metab 73:674–686
22. Lashansky G, Saenger P, Dimartino-Nardi J, Gautier T, Mayes D, Berg G, Reiter E 1992 Normative data for the steroidogenic response of mineralocorticoids and their precursors to adrenocorticotropin in a healthy pediatric population. J Clin Endocrinol Metab 75:1491–1496
23. Hasegawa Y 2003 Normal range of GnRH and hCG stimulation tests. In: Let's enjoy pediatric endocrinology. 3rd ed. Tokyo: Sindan-to-tiryou-sha; 260–262 (in Japanese)
AQ: E
24. Kleinjan DJ, van Heyningen V 1998 Position effect in human genetic disease. Hum Mol Genet 7:1611–1618
25. Shackleton C, Marcos J, Arlt W, Hauffa BP 2004 Prenatal diagnosis of P450 oxidoreductase deficiency (ORD): a disorder causing low pregnancy estriol, maternal and fetal virilization, and the Antley-Bixler syndrome phenotype. Am J Med Genet A 129:105–112
26. Auchus RJ 2004 The backdoor pathway to dihydrotestosterone. Trends Endocrinol Metab 15:432–438
27. Achermann JC, Hughes, IA 2008 Disorders of sex development. In: Kronenberg HM, Melmed M, Polonsky KS, Larsen PR, eds. Williams textbook of endocrinology. 11th ed. Philadelphia: Saunders; 783–848
28. Morishima A, Grumbach MM, Simpson ER, Fisher C, Qin K 1995 Aromatase deficiency in male and female siblings caused by a novel mutation and the physiological role of estrogens. J Clin Endocrinol Metab 80:3689–3698
29. Carani C, Qin K, Simoni M, Faustini-Fustini M, Serpente S, Boyd J, Korach KS, Simpson ER 1997 Effect of testosterone and estradiol in a man with aromatase deficiency. N Engl J Med 337:91–95
30. Gomes LG, Huang N, Agrawal V, Mendonça BB, Bachega TA, Miller WL 2008 The common P450 oxidoreductase variant A503V is not a modifier gene for 21-hydroxylase deficiency. J Clin Endocrinol Metab 93:2913–2916
31. Wudy SA, Hartmann MF, Draper N, Stewart PM, Arlt W 2004 A male twin infant with skull deformity and elevated neonatal 17-hydroxyprogesterone: a prismatic case of P450 oxidoreductase deficiency. Endocr Res 30:957–964
32. Williamson L, Arlt W, Shackleton C, Kelley RI, Braddock SR 2006 Linking Antley-Bixler syndrome and congenital adrenal hyperplasia: a novel case of P450 oxidoreductase deficiency. Am J Med Genet A 140A:1797–1803
33. Dhir V, Ivison HE, Krone N, Shackleton CH, Doherty AJ, Stewart PM, Arlt W 2007 Differential inhibition of CYP17A1 and CYP21A2 activities by the P450 oxidoreductase mutant A287P. Mol Endocrinol 21:1958–1968
34. Sarda IR, Gorwill RH 1976 Hormonal studies in pregnancy. I. Total unconjugated estrogens in maternal peripheral vein, cord vein, and cord artery serum at delivery. Am J Obstet Gynecol 124:234–238

RESEARCH LETTER

AMERICAN JOURNAL OF medical genetics PART A

Mutation Analysis of *SOX9* and Single Copy Number Variant Analysis of the Upstream Region in Eight Patients With Campomelic Dysplasia and Acampomelic Campomelic Dysplasia

Yuka Wada,[1]* Gen Nishimura,[2] Toshiro Nagai,[3] Hideaki Sawai,[4] Mayumi Yoshikata,[5] Shinichirou Miyagawa,[6] Takushi Hanita,[7] Seiji Sato,[8] Tomonobu Hasegawa,[9] Shumpei Ishikawa,[10] and Tsutomu Ogata[1]

[1]Department of Endocrinology and Metabolism, National Research Institute for Child Health and Development, Tokyo, Japan

[2]Department of Radiology, Tokyo Metropolitan Kiyose Children's Hospital, Kiyose, Japan

[3]Department of Pediatrics, Dokkyo Medical University, Koshigaya, Japan

[4]Genetic Counseling and Clinical Research Unit, Kyoto University, Kyoto, Japan

[5]Department of Neonatology, Hyogo Children's Hospital, Kobe, Japan

[6]Department of Pediatrics, National Hospital Organization Kure Medical Center, Kure, Japan

[7]Department of Pediatrics, Tohoku University School of Medicine, Sendai, Japan

[8]Department of Pediatrics, Saitama City Hospital, Saitama, Japan

[9]Department of Pediatrics, School of Medicine, Keio University, Tokyo, Japan

[10]Genome Science Division, Department of Pathology, Research Center for Advanced Science and Technology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

Received 21 August 2009; Accepted 16 July 2009

TO THE EDITOR:

Campomelic dysplasia (CD; OMIM 114290) is a rare skeletal disorder characterized by hypoplastic scapulae, 11 pairs of ribs, pelvic abnormalities, and bowing of the lower limb bones [Maroteaux et al., 1971]. Affected patients often die shortly after birth due to respiratory distress, and roughly two-thirds of affected genetic males have disorders of sex development (DSD) due to dysgenetic testes [Mansour et al., 1995]. Acampomelic campomelic dysplasia (ACD) is associated with similar but milder skeletal features and lacks long bone curvature [MacPherson et al., 1989].

SOX9 on chromosome 17q24 is a member of SRY-related gene family [Harley et al., 2003]. It encodes a 509-amino acid protein that harbors a high mobility group (HMG) domain with a DNA-binding capacity and a proline/glutamine/serine-rich domain with a transactivation function [Harley et al., 2003]. Furthermore, putative *cis*-control elements have been mapped within the 1 Mb region upstream of *SOX9* [Hill-Harfe et al., 2005].

To date, it has been shown that both CD and ACD can be caused by heterozygous intragenic *SOX9* mutations or chromosomal aberrations (translocations, inversions, or deletions) affecting *SOX9* or the putative enhancer region [Pfeifer et al., 1999; Thong et al., 2000; Moog et al., 2001; Harley et al., 2003; Pop et al., 2004; Leipoldt et al., 2007]. However, the frequency and the type of mutations and chromosomal aberrations are quite different

How to Cite this Article:
Wada Y, Nishimura G, Nagai T, Sawai H, Yoshikata M, Miyagawa S, Hanita T, Sato S, Hasegawa T, Ishikawa S, Ogata T. 2009. Mutation analysis of *SOX9* and single copy number variant analysis of the upstream region in eight patients with campomelic dysplasia and acampomelic campomelic dysplasia.
Am J Med Genet Part A 149A:2882–2885.

*Correspondence to:
Yuka Wada, M.D., Department of Endocrinology and Metabolism, National Research Institute for Child Health and Development, 2-10-1 Ohkura, Setagaya, Tokyo 157-8535, Japan. E-mail: ywada@nch.go.jp
Published online 16 November 2009 in Wiley InterScience (www.interscience.wiley.com)
DOI 10.1002/ajmg.a.33107

TABLE I. Clinical and Molecular Findings in Patients Examined in This Study

	Campomelic dysplasia				Acampomelic campomelic dysplasia			
	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7	Patient 8
Patient								
Gastational age (weeks)	25	42	38	38	39	40	42	38
Birth weight (g)	625	2490	2670	2060	3400	2700	2680	2306
Present age (y:m)	Stillbirth	0:11	(0:5)[a]	(1:5)[a]	11:6	19:8	3:2	3:9
Karyotype	46,XY	46,XX	46,XX	46,XX	46,XY	46,XY	46,XX	46,XX
Phenotype								
Cleft palate	−	−	+	−	+	+	+	+
Micrognathia	+	+	+	+	+	−	−	+
Scapular hypoplasia	+	+	+	+	+	+	+	+
Tibial bowing	+	+	+	+	−	−	−	−
Femoral bowing	+	+	+	+	−	−	−	−
11 pairs of ribs	−	+	+	+	−	+	+	+
Small thoracic cage	+	+	+	+	+	+	+	−
NM thoracic pedicles	+	+	+	+	−	−	+	+
Scoliosis	−	−	−	−	+	+	+	−
Narrow iliac wings	±	+	+	+	±	±	±	+
Clubfeet	+	+	+	+	−	−	−	+
46,XY DSD	+	−	−	−	−	−	−	−
Mutation								
cDNA	771 772insGGCGC	1330 1333delGACC	T338C	G442T	C509T	−	−	−
Amino acids	G257fsX296	T443fsX468	M113T	E148X	P170L	−	−	−

NM: non-mineralized; DSD: disorders of sex development.
[a]Deceased at 5 months and 1 year and 5 months, respectively.

between CD and ACD. CD is predominantly caused by nonsense or frameshift mutations or by chromosomal aberrations disrupting *SOX9*, although missense mutations and chromosomal aberrations impairing the enhancer region are also occasionally identified. By contrast, ACD is almost exclusively caused by missense mutations or by chromosomal aberrations affecting the enhancer region. Thus, while missense mutations are exclusively identified within the HMG box in both CD and ACD [Kwok et al., 1995; Cameron and Sinclair, 1997; Meyer et al., 1997; Hageman et al., 1998; Moog et al., 2001; Thong et al., 2000], these findings imply that severe mutations usually result in CD whereas mild mutations usually lead to ACD.

However, the underlying causes remain to be determined in several patients, especially those with ACD, and such patients may have hidden perturbation in the putative enhancer region. Thus, we performed mutation analysis of *SOX9* in eight patients with CD or ACD and single copy number variant (CNV) analysis [Redon et al., 2006] of the upstream region in *SOX9* mutation negative patients.

Clinical features of the eight patients are summarized in Table I, and representative roentgenograms are shown in Figure 1. Patients 1–4 showed CD-compatible severe clinical features, whereas patients 5–8 exhibited relatively mild ACD-compatible clinical features. In addition, patient 1 ended in a stillbirth, and patients 3 and 4 died of respiratory insufficiency during infancy, although patient 2 aged 11 months was alive. By contrast, patients 5–8 have survived a relatively long period. Among genetic males, patient 1 exhibited DSD with nearly complete female external genitalia, while patients 5 and 6 showed male external genitalia.

We first performed mutation analysis of *SOX9*. This study was approved by the Institutional Review Board Committees at National Center for Child Health and Development, and performed after obtaining written informed consent. Genomic DNA samples

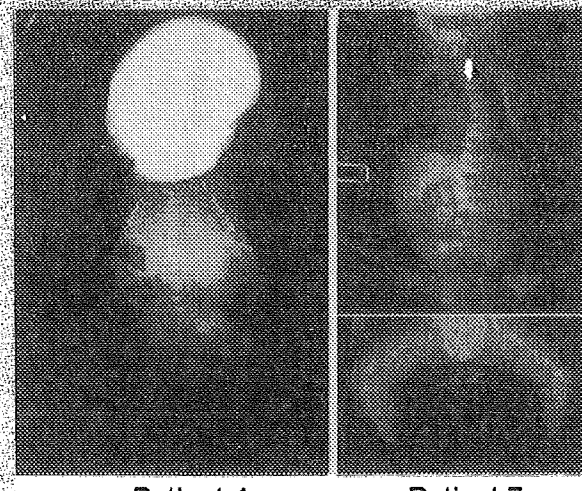


FIG. 1. Representative roentgenograms indicating CD in patient 1 at birth and ACD in patient 7 at 3 months of age.

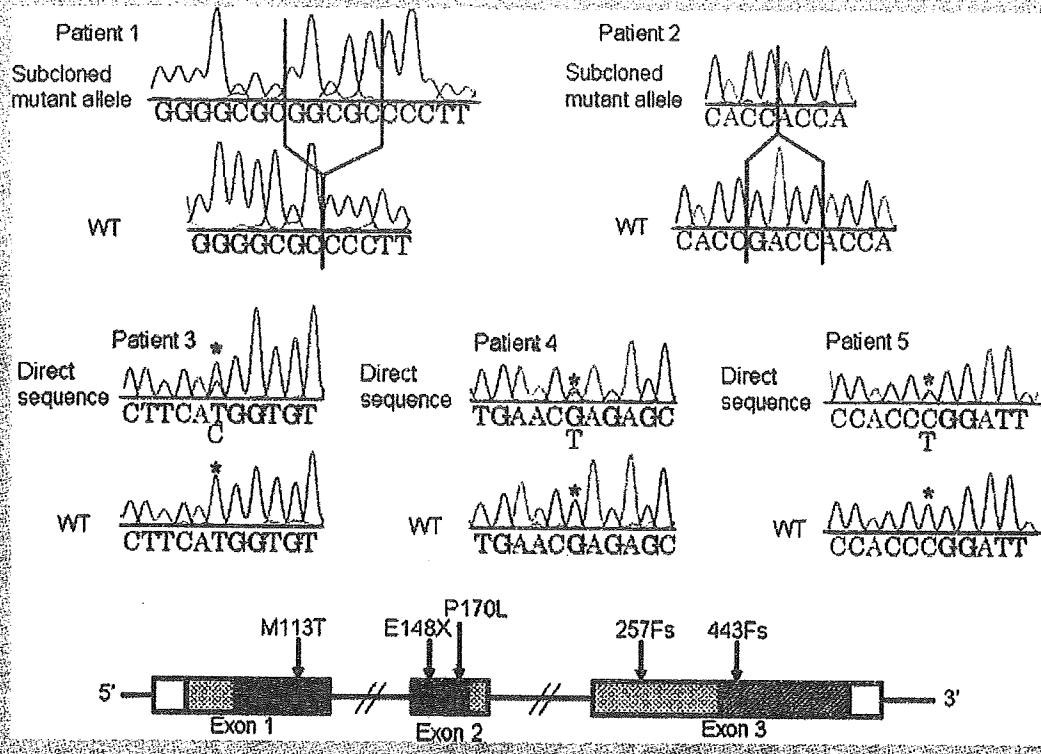


FIG. 2. Molecular findings in patients 1–5 with *SOX9* mutations. Upper part: Electrochromatograms showing the mutations in patients 1–5. In patients 1 and 2, the subcloned mutant alleles and the corresponding wildtype (WT) alleles are shown. In patients 3–5, the direct sequences are shown, together with the corresponding wildtype sequences; the asterisks indicate the mutant and the corresponding wildtype nucleotides. Lower part: The position of the mutations on the genomic sequences. Exons 1–3 are depicted with boxes; the black, the striped, the stippled, and the white areas indicate the HMG domain, the transactivation domain, other translated regions, and the untranslated regions, respectively.

extracted from cord blood cells (patient 1) or peripheral blood cells (patients 2–8) were amplified by PCR for all the three coding exons and were subjected to direct sequencing on a CEQ 8000 autosequencer (Beckman Coulter, Fullerton, CA) (the primer sequences are available on request). To confirm frameshift mutations, the corresponding PCR products were subcloned with TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) and normal and mutant alleles were sequenced separately.

Consequently, we identified a novel heterozygous 5-bp insertion mutation at exon 3 that is predicted to cause a frameshift at the 257th glycine codon and resultant termination at the 296th codon (G257fsX296) in patient 1, a novel heterozygous 4-bp deletion mutation at exon 3 that is predicted to cause a frameshift at the 443rd threonine codon and resultant termination at the 468th codon (T443fsX468) in patient 2, a novel heterozygous missense mutation at exon 1 (M113T) in patient 3, a recurrent heterozygous nonsense mutation at exon 2 (E148X) in patient 4, and a novel heterozygous missense mutation at exon 2 (P170L) in patient 5 (Fig. 2). The two missense mutations resided within the HMG. The mutations of patients 1–4 were absent in their parents. In addition, while mutation analysis was refused by the parents of patient 5, the P170L missense mutation was absent in 200 control subjects. No mutations were identified in patients 6–8.

Then, to examine for a small deletion, we carried out the whole genome CNV analysis in patients 6–8 and their parents, using custom high density oligonucleotide microarray based on Affymetrix platform [Redon et al., 2006]. In brief, 25 bp oligonucleotide probes are designed on 1,330,354 *Nsp* I restriction fragments with average and median spacing of 2,271 and 776 bp. The experimental protocol is the same as the Affymetrix 500K arrays. Ninety microgram of target was hybridized overnight to the arrays [Fujii et al., 2007]. The signal intensity ratio of the sample to reference was calculated by Genome Imbalance Map Algorithm [Ishikawa et al., 2005], using NA10851 HapMap DNA samples from Coriell Cell Repositories (Camden, NJ) as the reference samples. Consequently, no deletion was indicated in the whole genome including the 5′ region of *SOX9* in patients 6–8.

The results are primarily consistent with the previous data. Three of four patients with CD died during fetal life or infancy, whereas patients 5–8 with ACD survived into childhood or puberty. 46,XY with DSD was observed in patient 1 with CD but not in patients 5 and 6 with ACD. Similarly, truncating mutations of *SOX9* were identified in patients 1–3 with CD, together with a missense mutation in patient 4 with CD, whereas only one missense mutation was found in patients with ACD.

We could not detect a microdeletion in patients 6–8 with ACD in whom no intragenic mutations were identified. Although the underlying causes remain to be clarified in patients 6–8, there are several possible explanations for the development of ACD in patients 6–8. First, a mutation(s) may exist in the unexamined intronic or the downstream region. Second, a tiny deletion may remain undetected. Third, there may be a mutation in some gene(s) other than *SOX9*. Further studies will identify underlying mechanisms involved in the development of ACD in *SOX9* mutation negative patients.

REFERENCES

Cameron FJ, Sinclair AH. 1997. Mutation in SRY and SO X9: Testis-determining genes. Hum Mutat 9:388–395.

Fujii K, Ishikawa S, Uchikawa H, Komura D, Shapero MH, Shen F, Hung J, Arai H, Tanaka Y, Sasaki K, Kohno Y, Yamada M, Jones KW, Aburatani H, Miyashita T. 2007. High-density oligonucleotide array with sub-kilobase resolution reveals breakpoint information of submicroscopic deletions in nevoid basal cell carcinoma syndrome. Hum Genet 122:459–466.

Hageman RM, Cameron FJ, Sinclair AH. 1998. Mutation analysis of the SOX9 gene in a patient with campomelic dysplasia. Hum Mutat Suppl 1:S112–S113.

Harley VR, Clarkson MJ, Argentaro A. 2003. The molecular action and regulation of the testis-determining factors, SRY (sex-determining region on the Y chromosome) and SOX9 [SRY-related high-mobility group (HMG) box 9]. Endocr Rev 24:466–487.

Hill-Harfe KL, Kaplan L, Stalker HJ, Zori RT, Pop R, Scherer G, Wallace MR. 2005. Fine mapping of chromosome 17 translocation breakpoints >900 kb upstream of *SOX9* in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. Am J Hum Genet 76:663–671.

Ishikawa S, Komura D, Tsuji S, Nishimura K, Yamamoto S, Panda B, Huang J, Fukayama M, Jones KW, Aburatani H. 2005. Allelic dosage analysis with genotyping microarrays. Biochem Biophys Res Commun 333:1309–1314.

Kwok C, Weller PA, Guioli S, Foster JW, Mansour S, Zuffardi O, Punnett HH, Dominguez-Steglich MA, Brook JD, Young ID, Goodfellow PN, Schafer AJ. 1995. Mutations in SOX9, the gene responsible for Campomelic dysplasia and autosomal sex reversal. Am J Hum Genet 57: 1028–1036.

Leipoldt M, Erdel M, Bien-Willner GA, Smyk M, Theurl M, Yatsenko SA, Lupski JR, Lane AH, Shanske AL, Stankiewicz P, Schere G. 2007. Two novel translocation breakpoints upstream of SOX9 define borders of the proximal and distal breakpoint cluster region in campomelic dysplasia. Clin Genet 71:67–75.

MacPherson RI, Skinner SA, Donnenfeld AE. 1989. Acampomelic campomelic dysplasia. Pediatr Radiol 20:90–93.

Mansour S, Hall CM, Pembery ME, Young ID. 1995. A clinical and genetic study of campomelic dysplasia. J Med Genet 32:145–420.

Maroteaux P, Spranger J, Opitz J, Kucera J, Lowry R, Schimke R, Kagan S. 1971. Le syndrome campomelique. Presse Med 22:1157–1162.

Meyer J, Südbeck P, Held M, Wagner T, Schmitz ML, Bricarelli FD, Eggermont E, Friedrich U, Haas OA, Kobelt A, Leroy JG, Van Maldergem L, Michel E, Mitulla B, Pfeiffer RA, Schinzel A, Schmidt H, Scherer G. 1997. Mutational analysis of the SOX9 gene in campomelic dysplasia and autosomal sex reversal: Lack of genotype/phenotype correlations. Hum Mol Genet 6:91–98.

Moog U, Jansen NJ, Scherer G, Schrander-Stumpel CT. 2001. Acampomelic campomelic syndrome. Am J Med Genet 104:239–245.

Pfeifer D, Kist R, Dewar K, Devon K, Lander ES, Brirren B, Korniszewski L, Back E, Scherer G. 1999. Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to *SO* X9: Evidence for an extended control region. Am J Hum Genet 65:111–124.

Pop R, Conz C, Lindenberg KS, Blesson S, Schmalenberger B, Briault S, Pfeifer D, Scherer G. 2004. Screening of the 1 Mb SOX9 5' control region by array CGH identifies a large deletion in a case of campomelic dysplasia with XY sex reversal. J Med Genet 41:e47.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, González JR, Gratacòs M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. 2006. Global variation in copy number in the human genome. Nature 444:444–454.

Thong MK, Scherer G, Kozlowski K, Haan E, Morris L. 2000. Acampomelic campomelic dysplasia with SOX9 mutation. Am J Med Genet 93: 421–425.

Genomics 93 (2009) 461–472

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Identification of the mouse paternally expressed imprinted gene *Zdbf2* on chromosome 1 and its imprinted human homolog *ZDBF2* on chromosome 2

Hisato Kobayashi [a,1], Kaori Yamada [a,1], Shinnosuke Morita [a], Hitoshi Hiura [a], Atsushi Fukuda [a], Masayo Kagami [b], Tsutomu Ogata [b], Kenichiro Hata [c], Yusuke Sotomaru [d], Tomohiro Kono [a,*]

[a] Department of BioScience, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan
[b] Department of Endocrinology and Metabolism, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya-ku, Tokyo 157-8535, Japan
[c] Department of Maternal-Fetal Biology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya-ku, Tokyo 157-8535, Japan
[d] Natural Science Center for Basic Research and Development, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan

ARTICLE INFO

Article history:
Received 4 December 2008
Accepted 30 December 2008
Available online 4 February 2009

Keywords:
Genomic imprinting
Imprinted gene
Zdbf2
Differentially methylated region
Mouse chromosome 1
Human chromosome 2

ABSTRACT

In mammals, both the maternal and paternal genomes are necessary for normal embryogenesis due to parent-specific epigenetic modification of the genome during gametogenesis, which leads to non-equivalent expression of imprinted genes from the maternal and paternal alleles. In this study, we identified a paternally expressed imprinted gene, *Zdbf2*, by microarray-based screening using parthenogenetic and normal embryos. Expression analyses showed that *Zdbf2* was paternally expressed in various embryonic and adult tissues, except for the placenta and adult testis, which showed biallelic expression of the gene. We also identified a differentially methylated region (DMR) at 10 kb upstream of exon 1 of the *Zdbf2* gene and this differential methylation was derived from the germline. Furthermore, we also identified that the human homolog (*ZDBF2*) of the mouse *Zdbf2* gene showed paternal allele-specific expression in human lymphocytes but not in the human placenta. Thus, our findings defined mouse chromosome 1 and human chromosome 2 as the loci for imprinted genes.

Introduction

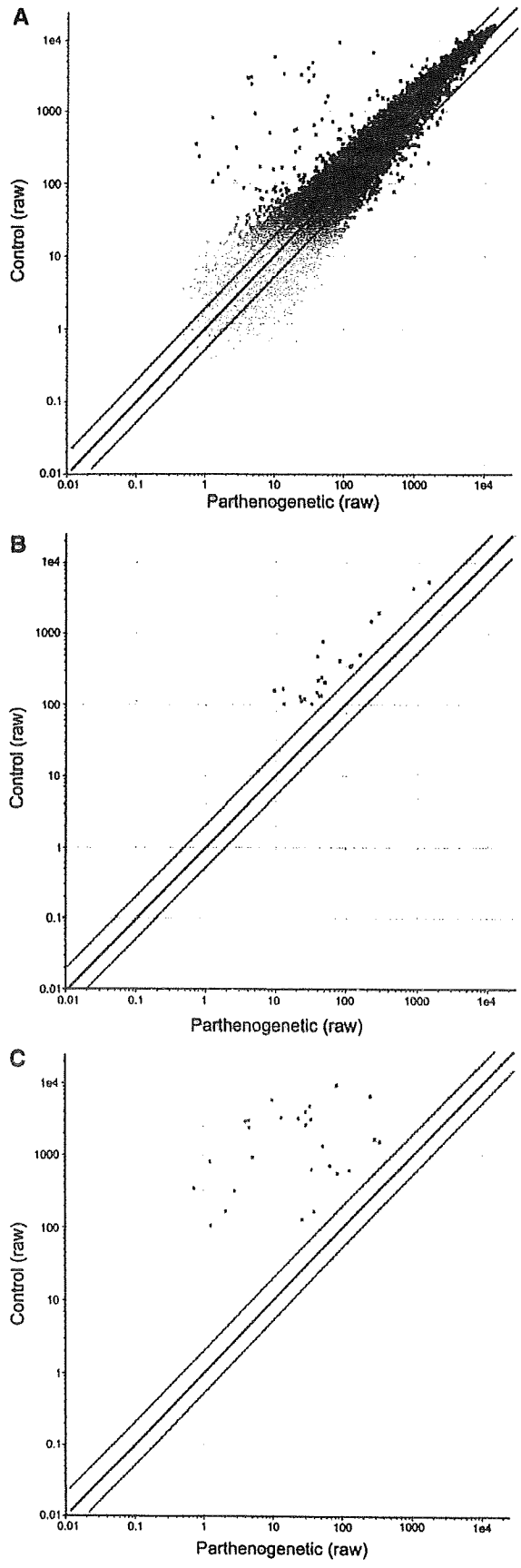
Genomic imprinting is an epigenetic gene-marking phenomenon in mammals, which leads to parent-of-origin-dependent monoallelic expression of certain genes, termed imprinted genes [1]. To date, approximately 80 imprinted genes have been identified in mice, the majority of which are present in 11 clusters including the Prader–Willi syndrome/Angelman syndrome and Beckwith–Wiedemann syndrome clusters. These clusters were assigned to 8 autosomal chromosomes, 2, 6, 7, 9, 11, 12, 15, and 17; whereas many solo imprinted genes have been identified on 5 chromosomes (numbers 2, 10, 14, 18, and 19). Many of these imprinted genes with expression patterns were well-conserved between mice and humans. These genes play an important role in fetal growth, development of particular somatic lineages, maternal behavior, tumorigenesis, and birth defects (MRC Mammalian Genetics Unit, Harwell, UK, http://www.mgu.har.mrc.ac.uk/research/imprinting/function.html).

Gene imprinting is initiated by epigenetic modifications such as DNA methylation that occur in the parental germline. In mammals, DNA methylation occurs exclusively at the cytosine residues within cytosine–guanine (CpG) dinucleotides, which plays an important role in normal development [2]. Indeed, many imprinted genes have differentially methylated regions (DMRs), which exhibit parent-of-origin-dependent DNA methylation patterns. Some DMRs function as cis-acting imprinting control regions (ICRs) that exert a regional control on gene expression from an imprinted cluster. Knockout mice studies demonstrated that the de novo DNA methyltransferase Dnmt3a and its related protein Dnmt3L are required to establish the methylation imprints in both maternal and paternal germlines [3–5]. The maintenance methyltransferase Dnmt1 is then required to maintain the differential methylation and imprinted expression patterns in the embryo proper [6,7].

A number of studies have suggested that imprinted genes may have characteristic structural features. For example, it was reported that imprinted genes tend to have fewer and smaller introns [8]. Other reports have described that human and mouse imprinted gene regions contain a lower density of short interspersed transposable elements (SINEs) than non-imprinted regions [9,10]. Thus far, the presence of direct repeats near or within the DMRs has been identified as the potential feature of these regions [11]. However, these features are not observed with regard to all imprinted regions, and their functional relevance is controversial. In a previous study, the dimensions of 15 DMRs (12 maternally imprinted genes and 3 paternally imprinted genes) were measured, and it was revealed that paternally methylated DMRs contain fewer CpGs than maternally methylated DMRs [12]. Furthermore, a recent study has demonstrated that the Dnmt3a/Dnmt3L complex could preferentially methylate CpG site pairs that are 8–10 base pairs apart, and a similar periodicity was observed for the frequency of CpG sites in the 12 maternally imprinted regions [13].

* Corresponding author. Fax: +81 354772543.
E-mail address: tomohiro@nodai.ac.jp (T. Kono).
[1] These authors contributed equally to this work.

0888-7543/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2008.12.012



However, no other consensus sequence has been identified for DMRs, and the features that cause them to get preferentially methylated via the Dnmt3a/Dnmt3L complex in the germline remain unknown.

After the first imprinted gene, *Igf2*, was identified in a knockout study, many other methods have subsequently been used to identify imprinted genes in mice. In 1994, *U2afbp-rs* (*Zrsr1*) [14] and *Rasgrf1* [15] were identified as paternally expressed genes with the use of methylation-sensitive restriction enzyme sites as the restriction landmark in the restriction landmark genomic scanning (RLGS) method for screening methylated sites. Subsequently, Ishino's group identified 2 paternally expressed genes, namely, *Peg1* (*Mest*) and *Peg3*, and 2 maternally expressed genes, *Meg1* (*Grb10*) and *Meg3*, by comparison analysis of gene expression among in vitro fertilized parthenogenetic (containing only maternally derived chromosomes) and androgenetic (containing only paternally derived chromosomes) embryos, by suppression subtractive hybridization [16–19]. In 1997, the paternally expressed gene *Impact* was identified using allelic message display without positional cloning or production of parthenogenetic and androgenetic embryos [20]. Furthermore, the development of DNA microarray technology facilitated the identification of many imprinted genes by gene expression profiling. In 2000, Affymetrix GeneChip was used with in vitro fertilized and parthenogenetic embryos and, in 2002, RIKEN cDNA Microarray was used with parthenogenetic and androgenetic embryos to identify new imprinted genes [21,22]. Further, in 2006, Affymetrix GeneChip was used with uniparental disomies [23]. Recently, the imprinted *Mcts2* gene was identified using the sequence features of imprinted genes [24], indicating that bioinformatics analysis can contribute to the identification of novel imprinted genes. Although a recent study has estimated that there are 600 imprinted genes in mice [25], a complete global analysis for locating imprinted genes has not been performed. To elucidate the biological importance of genomic imprinting and other characteristics of imprinted genes, it is important to systematically identify the remaining imprinted genes.

In this study, we compared the gene expression profiles between parthenogenetic and in vitro fertilized embryos (control) by using the Affymetrix GeneChip probe array to identify novel imprinted genes. The control embryos containing both the maternal and paternal genomes exhibited normal expression patterns of both maternally and paternally expressed genes; however, the parthenogenetic embryos that contain 2 maternal genomes exhibited a significantly decreased expression of paternally expressed genes. On the basis of this information, we screened the imprinted gene candidates and confirmed the imprinted expression of these genes by using reverse transcriptase-polymerase chain reaction (RT-PCR) analysis. From this screening, we identified a paternally expressed imprinted gene, *Zdbf2*, on mouse chromosome 1. We also identified a DMR in the paternal allele methylated at 10 kb upstream of the predicted exon 1 of the *Zdbf2* gene. Furthermore, we demonstrated that the human homolog *ZDBF2*, which is mapped to chromosome 2, is also paternally expressed. The newly identified imprinted genes provide an opportunity to further investigate the function and mechanisms of the genomic imprinting machinery.

Results

Screening for new imprinted gene candidates by microarray analyses

In this study, we compared the global expression profiles of mouse parthenogenetic and in vitro fertilized (control) embryos by using the

Fig. 1. Comparison between the control and parthenogenetic samples. The red lines indicate equal expression levels between the 2 samples. The pink lines indicate a 2-fold change in the expression levels between the 2 samples. (A) Scatter plots of all the genes. (B) Scatter plots of candidate genes. (C) Scatter plots of known imprinted genes obtained by our screening. Partheno: parthenogenetic embryos ($n=4$); control: in vitro fertilized control embryos ($n=4$).

Table 1
Genes identified by microarray analysis between the control and parthenogenetic embryos

Control	Systematic	GenBank accession	Common name	Fold change	Control		Parthenogenetic		Chromosome
					Raw	Normalized	Raw	Normalized	
<500 raw	1423294_at	AW555393	***Mest***	96	9196.1	1	84.2	0.01	6
	1448152_at	NM_010514	***Igf2***	28.4	6508.5	1.01	253.6	0.04	7
	1426208_x_at	AF147785	***Plagl1***	97.3	5727.5	0.97	9.8	0.01	10
	1421144_at	NM_023879	*Rpgrip1*	5.5	5316.2	0.97	1460.6	0.18	14
	1433924_at	BM200248	***Peg3***	84.9	4702.7	1.02	34.5	0.01	7
	1425966_x_at	D50527	*Ubc*	5.5	4289.8	0.96	884.1	0.17	5
	1417355_at	AB003040	***Peg3***	88.2	3949.5	1.02	30	0.01	7
	1423506_a_at	AV218841	***Nnat***	97.7	3321.8	0.98	13.4	0.01	2
	1435383_x_at	AW743020	***Ndn***	90.5	3216.6	1.03	23.6	0.01	7
	1435382_at	AW743020	***Ndn***	78.7	3115.9	1.03	36	0.01	7
	1455792_x_at	AV124445	***Ndn***	99.6	2982	1	4.5	0.01	7
	1415923_at	NM_010882	***Ndn***	100.2	2871.3	1	4	0.01	7
	1437853_x_at	BB074430	***Ndn***	81.9	2585.1	1	30.3	0.01	7
	1417356_at	AB003040	***Peg3***	102.9	2394.6	1.03	4.6	0.01	7
	1417184_s_at	BC027434	*Hbb-b1*	7	1944	1.01	286.3	0.14	7
	1449939_s_at	NM_010052	***Dlk1***	6.5	1668.7	1.07	291.2	0.17	12
	AFFX-18SRNAMur/X00686_5_at	AFFX-18SRNAMur/X00686_5	*Pigt*[a]	8.9	1622	2.33	56.9	0.26	2
	1417714_x_at	NM_008218	*Hba-a1*	7.6	1470.9	1.15	222.8	0.15	11
	1415896_x_at	NM_013670	***Snrpn***	28	1336	1.01	53.4	0.04	7
	1435716_x_at	AI836293	***Snrpn***	102.7	916.7	1.04	5.1	0.01	7
	1421063_s_at	NM_033174	***Snrpn***	104.8	807.3	1.05	1.2	0.01	7
	1451058_at	AV017653	***Mcts2***	17.6	766.6	1.04	46.9	0.06	2
	1428111_at	AK003626	***Slc38a4***	16.9	710.8	1.01	67.8	0.06	15
	1448889_at	NM_027052	***Slc38a4***	31.6	629.1	1	36.6	0.03	15
	1415911_at	NM_008378	***Impact***	5.2	618.1	0.94	127.6	0.18	18
	1420688_a_at	NM_011360	***Sgce***	8.3	558.6	1.04	86.6	0.13	6
	1457356_at	BI793602	***Airn***	33.4	550	1	20.5	0.03	17
	1457781_at	BG063584	***Kcnq1***	33.4	537.2	0.97	19	0.03	7
	1455966_s_at	BG070110		5.5	502.8	1.01	156.6	0.18	[b]
	1458161_at	BM248551	***Kcnq1***	68.1	500.6	1	8.3	0.01	7
<400 raw	1456783_at	BB075402	*9330107J05Rik*	14.9	467.5	0.98	39.2	0.07	1
	1427797_s_at	BF580235		5.4	413.9	0.87	80.4	0.16	[b]
	1437213_at	BG070110		6	411.3	0.98	128.7	0.16	[b]
<300 raw	1452705_at	AK004611	*Pdxdc1*	5.5	355.2	0.59	117.4	0.11	16
	1455087_at	AV328498	***D7Ertd715e***	103	341.8	1.03	0.7	0.01	7
	1445966_at	BG075586	***Airn***	15	310.6	0.99	22.7	0.07	17
	1456139_at	BM124989	***Airn***	90.1	309.2	0.97	2.8	0.01	17
<200 raw	1451634_at	BC009123	***Airn***	19.8	271.1	1.01	19.8	0.05	17
	1427127_x_at	M12573	*Hspa1b*	5.1	240.5	0.98	45.3	0.19	17
	1436964_at	BB314814	***D7Ertd715e***	62.9	234.3	0.63	0.8	0.01	7
	1452646_at	AK003956	*Trp53inp2*	7.6	216	0.95	40.4	0.13	2
	1451386_at	BC027279	*Blvrb*	6.9	204.3	0.97	49.4	0.14	7
<100 raw	1458179_at	BB526903	***Airn***	36.6	169	1.01	6.1	0.03	17
	1424010_at	BC022666	*Mfap4*	16.3	167.1	1.04	12.6	0.06	11
	1450533_a_at	NM_009538	***Plagl1***	81.4	166.3	0.99	2.1	0.01	10
	1443007_at	AW545941	***Gnas***	8.9	162	0.94	19	0.11	2
	1439483_at	BI438039	*AI506816*	12	155.7	0.53	9.4	0.04	[b]
	1418632_at	BI694835	*Ube2h*	5.6	147.7	1.01	38.5	0.18	6
	1420978_at	NM_010938	*Nrf1*	5.1	132.4	0.9	44.6	0.17	6
	1442029_at	BM250850	***Kcnq1***	68	132.2	0.96	1.5	0.01	7
	1426009_a_at	BC003763	*Plp5k1b*	6.3	130.1	0.98	42.1	0.15	3
	1417217_at	NM_013779	***Magel2***	6.3	129.5	0.98	27	0.16	7
	1453164_a_at	AI596401	*Ptdss2*	9.8	126.2	0.99	22.1	0.1	7
	1453224_at	AW049828	*Zfand5*	8	117.9	0.99	25.8	0.12	19
	1450383_at	AF425607	*Ldlr*	5.7	110.3	0.84	23	0.15	9
	1420406_at	NM_013788	***Peg12***	71.3	105.4	0.97	1.3	0.01	7
	1434952_at	BI734783		5.7	101.8	0.95	32.2	0.17	8
	1429115_at	AK008077	*2010003O02Rik*	9.3	101.7	0.96	13	0.1	4

Genes indicated in bold are known imprinted genes.
[a] Was used as a control gene; therefore, it was excluded from the list of candidate genes.
[b] These sequences matched at more than 2 loci.

Affymetrix GeneChip assay to identify novel imprinted genes. Logically, the expression of an imprinted gene transcribed from the paternally inherited allele will be repressed in parthenogenetic embryos as compared to control embryos. In the microarray analysis, 4 parthenogenetic samples and 4 control samples were hybridized to the GeneChip (Fig. 1A). We used 2 approaches for screening candidates of paternally expressed genes. First, we excluded genes whose raw expression intensities were below 100 in the control embryos, because analysis of genes with low intensities would produce unreliable results. Second, we selected genes whose normalized data in parthenogenetic embryos was more than 5-fold lower than that of control embryos (Table 1). By this screening, we obtained 21 imprinted gene candidates (Fig. 1B), which seemed to be predominantly expressed by paternal allele, excluding 18 known imprinted genes (Fig. 1C). Of the 18 known imprinted genes obtained, 16 genes were known as paternally expressed genes. Therefore, the results of the microarray analysis were reasonable because there is down-regulation of the imprinted genes in parthenogenesis.

Identification of novel imprinted transcripts by RT-PCR

We investigated the polymorphisms in the candidate genes among C57BL/6, DBA/2, and JF1 mice in order to confirm that these candidates are true imprinted genes by allele-specific RT-PCR sequencing analysis. The polymorphism analyses of the candidate genes revealed polymorphisms in a total of 13 candidates: 2 candidates (GenBank accession numbers D50527 and NM_010938) between C57BL/6 and DBA/2 mice, 9 candidates (BB075402, AK004611, BC027279, BI694835, BC003763, AI596401, AF425607, BI734783, and AK008077) between C57BL/6 and JF1 mice, and 2 candidates (NM_008218 and BC027434) among C57BL/6, DBA/2, and JF1 mice (Table 2). To identify the alleles of these genes that were predominantly expressed, we performed RT-PCR sequencing of the candidate genes using BDF1 (C57BL6×DBA/2), DBF1 (DBA/2×C57BL/6), JBF1 (JF1×C57BL/6), and BJF1 (C57BL/6×JF1) mouse embryos at the 9.5-day-old stage (E9.5). All the primer sets for this analysis are listed in Supplemental Table 1. Allele-specific RT-PCR sequencing analysis showed that the BB075402 transcript was expressed only from the paternal allele (Fig. 2).

Expression analysis of mouse Zdbf2—a novel imprinted gene

According to the NCBI Entrez Gene database (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene), the 655-bp region of the BB075402 sequence corresponded to the 3′-untranslated region (UTR) of the *Zdbf2* (zinc finger, DNA binding factor type containing 2) gene, which contains 7 predicted exons (Fig. 3A). The *Zdbf2* gene was mapped to mouse chromosome 1C2 (Gene ID: 73884). First, we designed 3 specific primer sets (Z1, Z2, and Z3; Z1 primers were used in allele-specific RT-PCR sequencing analysis) for the 3 expressed sequence tags (ESTs) (BB075402, AK033878, and AK015271) that matched the predicted complete *Zdbf2* gene structure in order to confirm the expression levels of the gene (Fig. 3A). Incidentally, the AK033878 transcript (3113 bp) corresponds to the 3′-UTR region of the *Zdbf2* gene, and is registered as a candidate mouse imprinted transcript in the RIKEN database (http://fantom2.gsc.riken.jp/EICODB/imprinting/). The AK015271 transcript (984 bp) spliced fragment contains exons 1–7 of the *Zdbf2* gene. The microarray assay showed that the BB075402 transcript expression level in the parthenogenetic embryos was 8% (the expression level of the control was 100%) (Fig. 3B). The results of real-time PCR showed that the expression levels of the BB075402 transcript (Z1) and AK033878 (Z2) in the parthenogenetic embryos were 2% and 3%, respectively (Fig. 3C, D). Furthermore, results of the RT-PCR conducted for the AK015271 transcript (Z3) containing exons 5–7 showed a PCR band in the control embryo, but no such band was detected in the parthenogenetic embryo (Fig. 3E). Thus, we confirmed the repression of the *Zdbf2* gene in parthenogenesis. Second, we designed another primer set (Z4) for the translated region at exon 7 of the *Zdbf2* gene to confirm whether this gene is imprinted. We identified a single nucleotide polymorphism (SNP) in the Z4 region between B6 and JF1 mice. Further, allele-specific RT-PCR sequencing analysis using BJF1 and JBF1 embryos at E9.5 showed that the transcript at the Z4 region was paternally expressed (Fig. 3F). In addition, 5′-rapid amplification of cDNA ends (RACE) analysis of the *Zdbf2* gene, which was performed using BJF1 mouse embryos at E9.5, showed that the expressed transcript almost completely matched exons 1–7 of AK015271; however, exons 1 and 2 were partially matched, and exon 5 was 30 bp longer than that in the transcripts (Supplemental Fig. 1). These results suggest that the mouse *Zdbf2* gene transcript containing at least 7 exons was paternally expressed.

Expression analysis of mouse Zdbf2 in differential developmental stages and tissues

Next, we investigated the *Zdbf2* gene expression pattern in mice during 4 differential developmental stages: 15.5- and 18.5-day-old embryos (E15.5 and E18.5) and 1- and 9-week-old mice. Further, the pattern was investigated by RT-PCR analysis (at Z1 and Z4 regions) of various mouse tissues: the brain, tongue, heart, liver, lung, kidney, and muscle at all ages; the intestine and placenta in only embryos; the spleen in only 1- and 9-week-old mice; and the testis in only 9-week-old adult mice (Fig. 4). Results of the BB075402 (Z1) transcript analysis showed that gene expression was detected in almost all the major tissues from E15.5 and E18.5, except the liver and intestine, which did not show detectable expression in a few cases (Fig. 4A, Supplemental Fig. 1). Furthermore, allele-specific RT-PCR sequencing of BB075402 was performed for all the tissues (strong expression was detected) of BJF1 embryos at E15.5 and 9-week-old adult mice. Interestingly, although almost all the tissues were paternally expressed, placentas from the E15.5 embryos and adult testis exhibited biallelic expression (Fig. 4C,

Table 2
DNA polymorphism information of each paternally expressed candidate gene

Control	Gene	Nucleotide number	C57BL/6	DBA/2	JF1
< 500 raw	NM_023878	[a]			
	D50527	308	CCCTG	CCTTG	
	BC027434	232, 235	AAGAAAGT	AAAAAGGT	AAAAAGGT
	NM_008218	271	CGGTG	CGCTG	CGCTG
	BG070110	[a]			
	BB075402	589	TGAAA		TGGAA
< 400 raw	BF580235	[a]			
< 300 raw	AK004611	772	ACGTA		ACATA
	M12573	[a]			
< 200 raw	AK003956	[a]			
	BC027279	263	CCGTC		CCATC
< 100 raw	BC022666	[a]			
	BI438039	[b]			
	BI694835	154	TAAGA		TADGA
	NM_010938	1995	GAATG	GAGTG	
	BC003763	1894	GGACC		GGGCC
	AI596401	152	ATAGG		ATTGG
	AW049828	[b]			
	AF425607	395	CGATG		CGCTG
	BI734783	227	CCCAA		CCDAA
	AK008077	733	CATAG		CAAAG

All polymorphisms are shown in red.
[a] No polymorphisms were identified.
[b] No PCR bands were detected by RT-PCR.

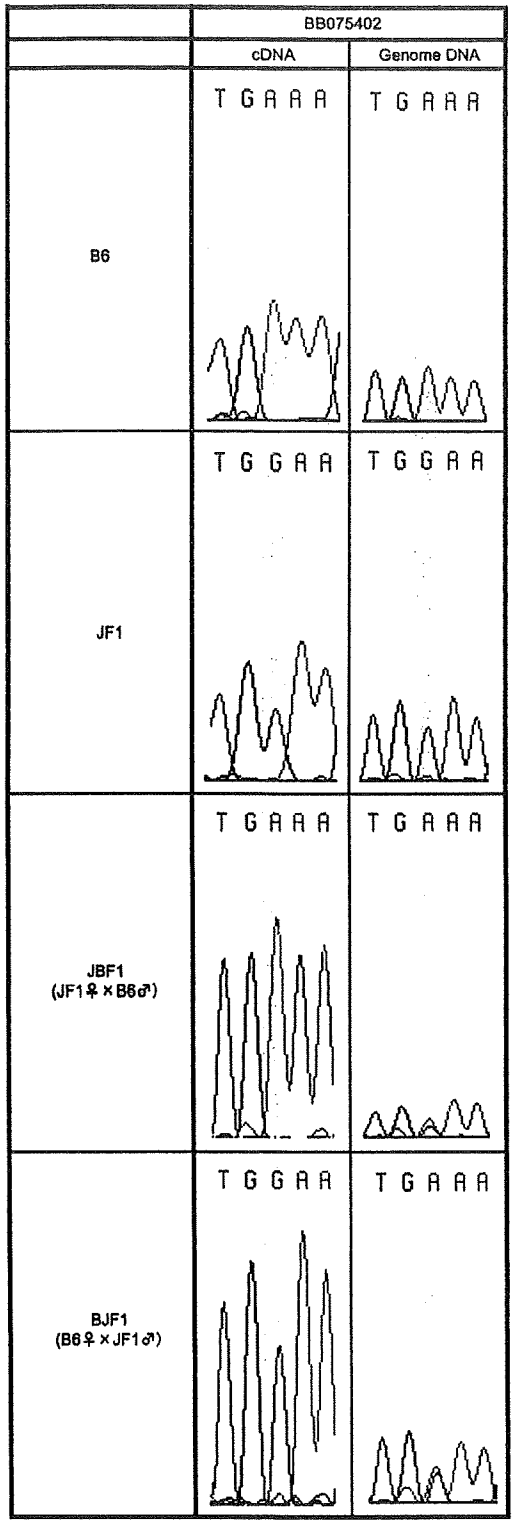


Fig. 2. Expression analysis of BB075402. Allele-specific RT-PCR sequencing analysis of BB075402 was performed with C57BL/6 (B6), JF1, BJF1, and JBF1 mouse embryos at E9.5 ($n=3$). The SNP of BB075402 is highlighted in yellow.

Supplemental Fig. 2). Meanwhile, expression at the Z4 region was detected in the brain, tongue, heart, lung, intestine, kidney, muscle, and placenta in E15.5 embryos and in the brain, tongue, kidney, muscle, and placenta in E18.5 embryos (Fig. 4B). Allele-specific RT-PCR sequencing of the Z4 region showed similar results to that of BB075402 (data not shown). These results strongly suggested that the transcript of the mouse *Zdbf2* gene is paternally expressed in almost all expressed tissues, but is biallelically expressed in only the placenta and testis.

Parent-of-origin-specific methylation of the mouse Zdbf2 gene

Many imprinted genes are epigenetically regulated by epigenetic mechanisms such as DNA methylation. DNA methylation occurs exclusively at cytosine residues within CpG dinucleotides. DMRs have been identified in CpG-rich regions (CpG islands) around imprinted genes in the maternal and paternal genomes, and it has been demonstrated that these regions function as ICRs. To explore and understand the regulation of the imprinted *Zdbf2* gene, we analyzed the DNA methylation status by using the bisulfite sequencing method [26].

The genomic DNA sequence of the mouse *Zdbf2* gene was derived from the mouse BAC clone RP23-434D24 (GenBank accession number AL669947, Fig. 5A). First, we identified 3 putative CpG islands within the AL669947 sequence using EMBOSS CpGplot (http://www.ebi.ac.uk/emboss/cpgplot/), and termed them CG1, CG2, and CG3 (Fig. 5A). These CpG islands were defined as a 200-bp stretch of DNA with a GC content of over 50% and an observed CpG/expected CpG (Obs/Exp CpG) ratio greater than 0.6. The CG1 region was observed 25 kb upstream of the *Zdbf2* gene. CG2 contains the promoter and exon 1 of the *Zdbf2* gene. We identified some SNPs in the CG1 and CG3 regions but not in CG2 by direct sequencing of samples from C57BL/6 and JF1 mice (Table 3). The CG3 contains 29 copies of the cytosine-rich 18-bp direct repeat and is included in exon 7 (Supplemental Fig. 3). To examine the differential methylation between paternal and maternal alleles in these CpG islands, we subjected them to bisulfite sequencing analyses for the CG1 and CG3 regions using 9.5-day-old in vitro fertilized embryos (BJF1 mice) and for the CG2 region using 9.5-day-old parthenogenetic and androgenetic embryos. The results showed that all the analyzed regions were almost unmethylated in both the alleles (Fig. 5B–D).

Second, by changing the CpG island criteria (minimum length, 100 bp; GC content, >50%; Obs/Exp CpG, >0.6), we identified a relatively CpG-rich region 10 kb upstream of the *Zdbf2* gene, between the CG1 and CG2 regions, and also identified SNPs in this region between C57BL/6 and JF1 mice (Table 3). We performed bisulfite sequence analysis for this region similar to the 3 CpG islands (Fig. 5E). Interestingly, paternal allele-specific methylation was detected in the CpG-rich region. Furthermore, bisulfite sequencing analyses of oocytes and sperms revealed that this region was hypomethylated in oocytes but highly methylated in sperms. These results showed that this CpG-rich region was methylated on the paternal alleles and that the methylation was derived from germline, similar to *H19* DMR, *Dlk1-Gtl2* IG-DMR, and *Rasgrf1* DMR, which are known as paternal methylation imprints [27–29]; thus, we termed this region *Zdbf2* DMR.

Human homologous ZDBF2 is paternally expressed in lymphocytes but not in the placenta

Most known imprinted genes have been identified among mammalian species, especially in mice and humans, but species-specific imprinting has been reported in some genes, like *Igf2r* and *Impact* [30,31]. To verify the imprinting status of the human homolog *ZDBF2*, we examined an SNP in the *ZDBF2* gene by direct sequencing of cDNA from human tissues. On the basis of the NCBI Entrez human gene database, the human homolog *ZDBF2* gene containing 5 exons was mapped to chromosome 2q33.3 (Gene ID: 57683), and no imprinted genes were identified on human chromosome 2. We designed a primer set within the last exon of *ZDBF2* (exon 5) including the SNP (reference SNP ID: rs10932150). We then

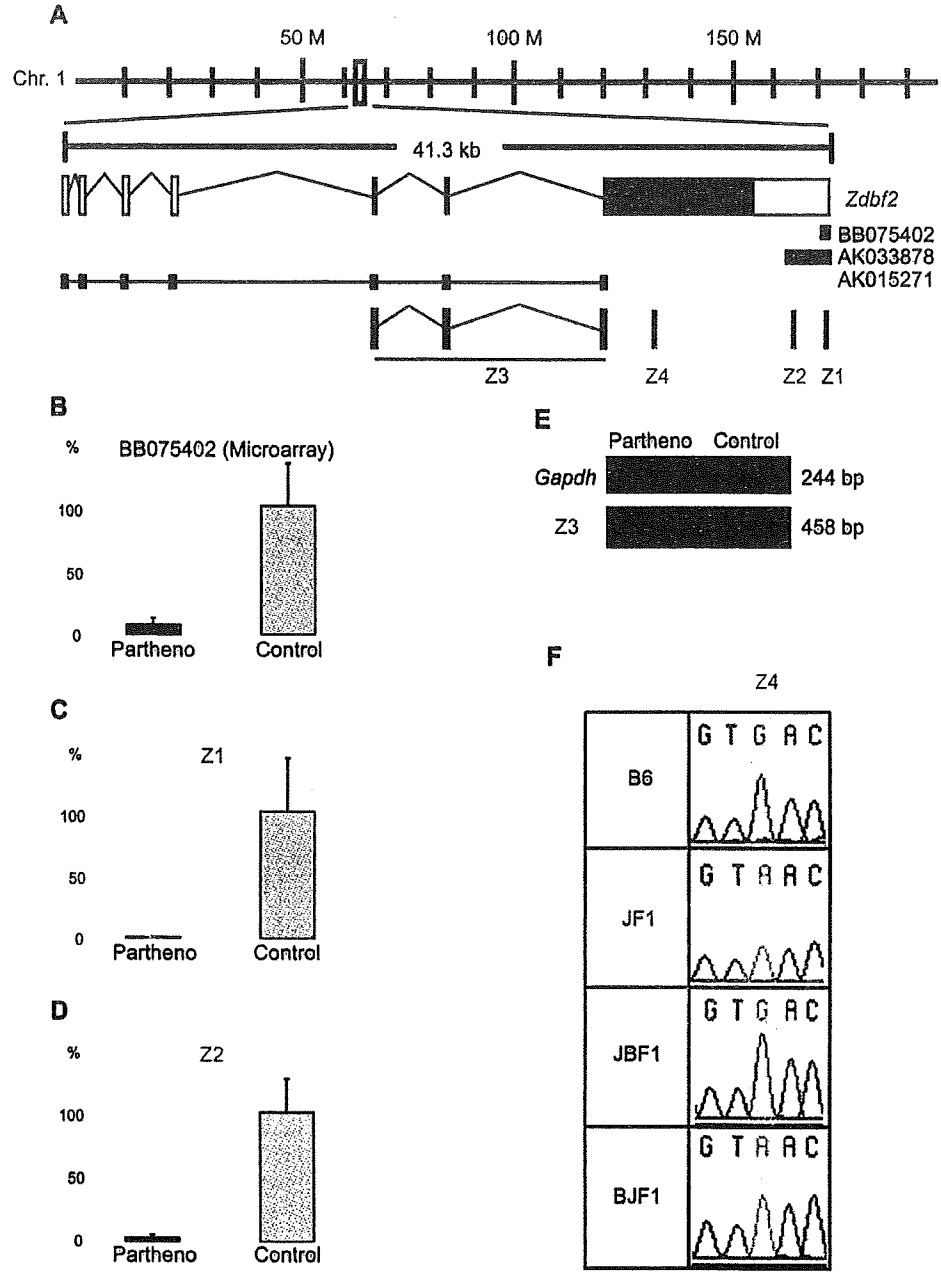


Fig. 3. Expression analysis of a novel paternally expressed gene, *Zdbf2*, on mouse chromosome 1. (A) Map of the *Zdbf2* gene on chromosome 1. Three EST positions and the primer positions of the Z1–Z4 regions are shown in the map. (B) Expression analysis of BB075402 by the microarray method. Partheno: parthenogenetic embryos ($n=4$); control: in vitro fertilized control embryos ($n=4$). Expression analysis of the Z1 (BB075402) (C) and Z2 regions (AK033878) (D) by real-time RT-PCR in the parthenogenetic ($n=3$) and control embryos ($n=3$). Values are expressed as mean ± s.e.m. (E) Expression analysis of the Z3 region (AK0152712) by RT-PCR in the control and parthenogenetic embryos. (F) Allele-specific RT-PCR sequencing analysis of the Z4 region with B6, JF1, BJF1, and JBF1 mouse embryos at E9.5 ($n=3$). The SNP of Z4 is highlighted in yellow.

performed RT-PCR analysis using human lymphocytes from a child who was heterozygous for a G/A polymorphism at the rs10932150 site, and only allele A from the father was detected (Fig. 6A, B). Furthermore, we examined the SNP in the human placenta, which was also heterozygous and, surprisingly, both G and A alleles were detected (Fig. 6C). These results revealed that the human *ZDBF2* gene is paternally expressed in lymphocytes but biallelically expressed in the placenta. This placenta-specific gene escape from imprinting is similar to that observed in the imprinted mouse *ZDBF2* gene. These results demonstrated that an imprinted locus is present on human chromosome 2.

Discussion

This study aimed to identify novel imprinted genes by comprehensive comparison of mouse gene expression. We successfully identified a imprinted gene, *Zdbf2*, which was mapped to mouse chromosome 1C2, and its imprinted human homolog, *ZDBF2*, which was mapped to human chromosome 2q33.3. The discovery of the imprinted *Zdbf2* gene may provide an opportunity to identify additional imprinted genes in the vicinity of this gene because of the clustering tendency of imprinted genes. Currently, the function of the *Zdbf2*/*ZDBF2* gene is unknown. Meanwhile, previous studies

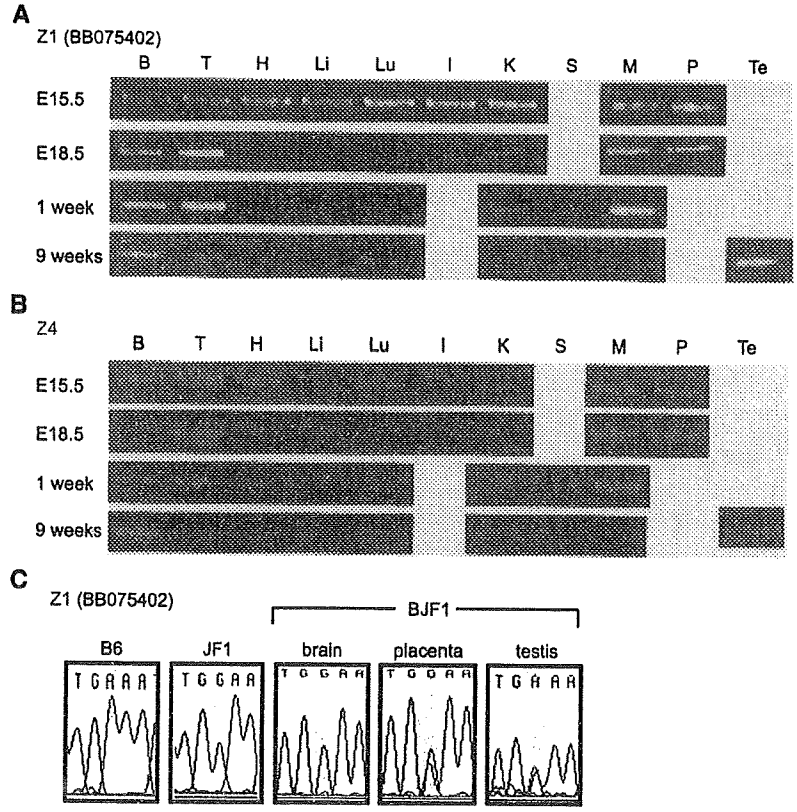


Fig. 4. Stage- and tissue-specific expression of the *Zdbf2* gene. Expression analysis by RT-PCR of Z1 (A) and Z4 (B) in the tissues (B = brain, T = tongue, H = heart, Li = liver, Lu = lung, I = intestine, K = kidney, S = spleen, M = muscle, P = placenta, and Te = testis; $n = 3$, respectively) of BJF1 mouse embryos at 15.5 and 18.5 days (E15.5 and E18.5) and 1- and 9-week-old adult mice (1 week and 9 weeks). (C) Allele-specific RT-PCR sequencing analysis of Z1 (BB075402) using B6 and JF1 mouse embryos at E9.5, the brain and placenta from BJF1 embryos at E15.5, and the testis from adult BJF1 mice ($n = 2$). The SNP of BB075402 is highlighted in yellow.

reported that maternal and paternal uniparental isodisomies for human chromosome 2 were responsible for various abnormalities [32–35]. Investigating the functions of the *Zdbf2* gene and other imprinted genes may provide further information about the imprinting disorders and mechanism.

In this study, we compared gene expression profiles between parthenogenetic and in vitro fertilized embryos (control) using the Affymetrix GeneChip probe array. We obtained 18 known imprinted genes and 21 candidates of paternally expressed genes. Some of the known imprinted genes included paternally expressed genes that were previously identified using subtraction hybridization in 9- to 10-day-old parthenogenetic and control embryos, similar to those used in our study [16,17,21,36], or embryonic fibroblast cell lines [37]. Of the obtained known imprinted genes, *Kcnq1* and *Gnas* are known as maternally expressed genes. These genes were accompanied by a paternally expressed antisense gene (*Kcnq1ot1*) or an alternative gene form (*Gnasxl*); therefore, these transcripts might be hybridized to the array [38,39]. Although we obtained several known imprinted genes from this screening, we could not obtain all the known imprinted genes. The reasons may include the decreased detection capability of tissue-specific imprinted gene expression because of RNA isolation from whole embryos or immature organ formation of each sample.

Of the 21 paternally expressed gene candidates, we identified polymorphisms among C57BL/6, DBA/2, and JF1 mice with 13 candidates. We used RT-PCR analysis and identified that of these, the BB075402 transcript was expressed only from the paternal allele. Meanwhile, many paternally expressed gene candidates exhibited biallelic expression in RT-PCR analysis. The differential expression of such genes between parthenogenetic and control embryos could be explained using 2 reasons. First, since 9.5-day-old parthenogenetic embryos exhibited delayed development as compared to the controls at the same stage of development, stage-specific genes might have been selected in this screening. Second, disruption of the imprinted gene expression in parthenogenesis might affect the expression of the non-imprinted genes, which were detected as false imprinted genes. These arguments were described in the discussion section of the previous studies [22,40]. The remaining candidate genes in which no polymorphisms were detected need to be further evaluated to determine whether they are true imprinted genes. Further investigation with other reciprocal crosses would be useful for identifying polymorphisms between the strains.

The BB075402 transcript was registered as a RIKEN mouse EST obtained from adult male diencephalons, and our study demonstrated expression of the BB075402 transcript (Z1 region) in the brain. Though almost all the major tissues showed clear expression of this transcript during embryogenesis, the tissues expressing this gene were limited to the brain, tongue, and muscle, and the testis after birth. The 3′-region of the AK033878 transcript corresponds to the BB075402 transcript, and both transcripts showed decreased expression in parthenogenetic embryos. Moreover, a part of the AK015271 transcript (Z3 region), which contains exons 1–7 of the mouse *Zdbf2* gene, was observed to show decreased expression in parthenogenetic embryos as determined by RT-PCR. We could further perform allelic expression analysis on the translated region (Z4 region shown in Fig. 3) of the mouse *Zdbf2* gene, and the results revealed that the

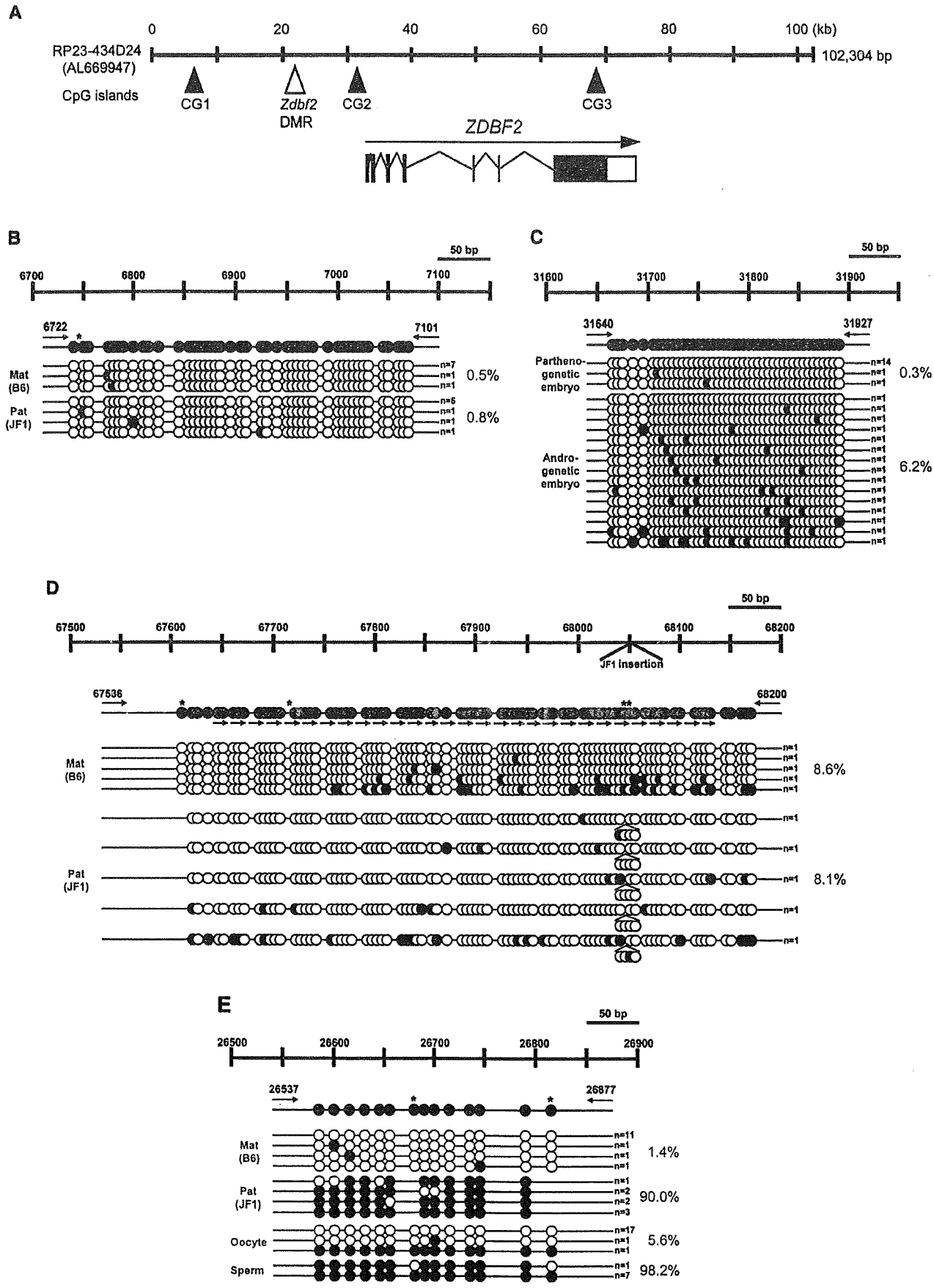

A
0
20
40
60
80
100 (kb)
RP23-434D24
(AL669947)
102,304 bp
CpG islands
CG1
Zdbf2
DMR
CG2
CG3
ZDBF2
B
50 bp
6700
6800
6900
7000
7100
6722
7101
Mat
(B6)
Pat
(JF1)
0.5%
0.8%
C
50 bp
31600
31700
31800
31900
31640
31927
Partheno-
genetic
embryo
Andro-
genetic
embryo
0.3%
6.2%
D
50 bp
67500
67600
67700
67800
67900
68000
68100
68200
JF1 insertion
67536
68200
Mat
(B6)
Pat
(JF1)
8.6%
8.1%
E
50 bp
26500
26600
26700
26800
26900
26537
26877
Mat
(B6)
Pat
(JF1)
Oocyte
Sperm
1.4%
90.0%
5.6%
98.2%

Table 3
DNA polymorphism information on each CpG-rich region surrounding the mouse *Zdbf2* gene

Region	Nucleotide number (AL669947)	C57BL/6	JF1
CG1	6753	GGAAG	GGGAG
Zdbf2 DMR	26681	CTCGA	CTTGA
	26819	TCCGA	TCTGA
CG2	a		
CG3	67619	GCGCT	GCCCT
	67716	GCCCC	GCACC
	68030	CGCGC	CGGGC
	68053	CC----------CG	CCCGTGCCCCCCGCGCGCCCCG

All polymorphisms are shown in red.
[a] No polymorphisms were identified.

translated region was paternally expressed similar to the BB075402 transcript. These results strongly indicated that these transcripts were the same products expressed solely from the paternal allele. It appeared that exons 5–6 of *Zdbf2* are also paternally expressed; however, it remained unknown as to whether the large (13 kb) transcript of the *Zdbf2* gene is expressed. Therefore, we tested 5′-RACE using mouse embryos at E15.5, and we could confirm that the expressed splice form (exons 1–7) was almost identical to the AK015271 transcript. Meanwhile, Z1 and Z4 primers did not detect expression in all the same tissues. For example, in embryonic liver, expression of the Z1 region was detectable but the Z4 region was not. Though this may be caused by the presence of different forms of the *Zdbf2* transcript in different tissues, to reveal (tissue-specific) splice variants of *Zdbf2* gene is required in future.

According to the database, the mouse *Zdbf2* gene encodes a 2493-amino acid protein with a predicted mass of 270 kDa. In our study, we demonstrated that the human *ZDBF2* gene was paternally expressed in lymphocytes. Furthermore, the human unidentified gene-encoded (HUGE) protein database of the Kazusa DNA Research Institute provides the expression profile of the human *ZDBF2* gene based on RT-PCR and enzyme-linked immuno-sorbent assay (ELISA) [41]. It is shown that the expression level of this gene in the brain and muscle is higher than that in other tissues such as those in the heart, lung, liver, kidney, pancreas, and spleen (http://www.kazusa.or.jp/huge/gfpage/KIAA1571/). Interestingly, our study showed that mouse *Zdbf2* gene expression was detected in only the brain, tongue, and muscle through the embryo to the adult stages. Thus, the expression profile of human *ZDBF2* was similar to that of mouse *Zdbf2*, which was investigated in our study. Furthermore, analyses of imprinted expression patterns showed that biallelic expression of mouse and human homologs was detected in placental tissues despite the paternal allele-specific expression observed in almost all other tissues (except for the testis) The possibility of placental tissues containing maternal materials was not completely excluded, however, the other imprinted gene (H19) showed normal imprinted expression pattern in both human and mouse placental tisses (data not shown). These facts indicate that the regulation mechanism of *Zdbf2*/*ZDBF2* gene expression is well conserved between mice and humans. Further comparison analyses of *Zdbf2*/*ZDBF2* gene products may provide hints for revealing those functions. For example, additional homologous *Zdbf2* anchors among chimpanzees, rats, dogs, horses, and chickens have been identified (Gene ID: 470622, 501153, 488490, 100068542, and 424100). These facts indicated that this gene plays an evolutionally conserved role among at least these organisms. There were no reports of the imprinted genes showed biallelic expression specifically in placenta and testis, like *Zdbf2* gene. By contract, some imprinted genes shows placenta-specific imprinted expression, and one of them, *Mash2*/*Ascl2* gene is essential for placental development [42]. The elucidation of the function of *Zdbf2* gene may explain the reason of the placenta- and testis-specific escape imprinting.

As previously noted, imprinted genes were regulated by parent-of-origin-specific DNA methylation in the DMR in *cis*. On analyzing the DNA methylation status at 4 CpG-rich regions around the mouse *Zdbf2* gene, we identified a paternal allele-specific methylated region, *Zdbf2* DMR, which is 10 kb upstream of the *Zdbf2* gene. Thus far, similar DMRs have been found in only 3 imprinted loci—*H19-Igf2*, *Dlk1-Gtl2*, and *Rasgrf*—and it has been demonstrated that these DMRs function as the ICRs controlling the neighboring imprinted genes. In particular, methylation of *H19* DMR and *Dlk1-Gtl2* IG-DMR acts as paternal methylation imprinting and prevents parthenogenesis [43,44]. Interestingly, 2 paternally expressed genes, *Igf2* and *Dlk1*, were included in the 18 known imprinted genes that were obtained from our microarray screening. This indicates that hypomethylation of *H19* DMR and IG-DMR inhibits paternally expressed genes on both maternal alleles in a parthenogenetic embryo, and that the methylation of *Zdbf2* DMR may also regulate the paternally expressed *Zdbf2* gene and the hitherto undiscovered neighboring imprinted genes. Moreover, the methylation of *Zdbf2* DMR might be established in gonocytes because the other 3 paternal methylation imprints are established in this stage [45–49], but the details of this remain unknown. The regulatory mechanism of the *Zdbf2* gene and the role of DNA methylation at the *Zdbf2* DMR should be clarified. The identification of novel paternally methylated DMRs was important and valuable, because only 3 cases showing such methylation patterns were reported, even though over 10 maternally methylated DMRs were reported. Further characterization of *Zdbf2* DMR is required to demonstrate the mechanisms by which the paternally methylated DMRs were methylated (targeted) by DNA methyltransferases. Although the real role of the repeat element in genomic imprinting [50,51] is still unknown, we identified a direct repeat sequence on the *Zdbf2* gene, similar to the other imprinted genes. In conclusion, we successfully defined mouse chromosome 1 and human chromosome 2 as the imprinted loci. Our findings provide a new platform for further identification of new imprinted genes and new insight into control of parental gene expression.

Materials and methods

Extraction of total RNA from parthenogenetic, androgenetic, and control embryos for microarray analysis

Parthenogenetic and androgenetic embryos were prepared as described previously [52]. Briefly, parthenogenetic embryos were constructed by stimulating unfertilized BDF1 eggs (C57BL6 × DBA/2; Clea Japan, Tokyo, Japan) with strontium chloride solution, which contains cytochalasin B to prevent extrusion of the second polar body. Androgenetic embryos were produced by in vitro fertilization of enucleated oocytes from BDF1 mice. Pronuclear transfer (from male BDF1 mice) was performed to produce diploid androgenetic embryos. Control biparental embryos were produced by the in vitro fertilization of non-manipulated oocytes. These embryos were introduced into an

Fig. 5. Methylation profiles of *Zdbf2* DMR and CpG islands surrounding the mouse *Zdbf2* gene. (A) Genomic structure of the mouse *Zdbf2* gene. The filled vertical arrowheads indicate the positions of each CpG island, i.e., CG1, CG2, and CG3. The open vertical arrowhead indicates the position of the relatively CpG-rich region, i.e., *Zdbf2* DMR. The filled vertical arrowheads indicate the positions of each CpG island—CG1, CG2, and CG3. The methylation status of (B) CG1, (C) CG2, and (D) CG3 in the 9.5-day-old in vitro fertilized (BJF1 mice; $n = 2$), parthenogenetic, and androgenetic embryos ($n = 2$). Coordinates are from GenBank Sequence AL669947. The 29 continuous arrows (horizontal) indicate direct repeats within the CG3 region. (E) Methylation of *Zdbf2* DMR in the 9.5-day-old in vitro fertilized (BJF1 mice) embryos, oocytes, and sperm. The open and closed circles indicate methylated and unmethylated CpGs. The asterisks represent polymorphism positions between B6 and JF1 mice. The maternal and paternal alleles were distinguished by polymorphisms, if present, between C57BL6 (B6) and JF1 mice. Mat, maternal (B6) allele; Pat, paternal (JF1) allele. n: number of DNA clones.