drastically compromised cortisol response to ACTH stimulation, and DSD (bilateral cryptorchidism in case 1, partial labial fusion in case 2, and mild clitoromegaly in case 3). Cases 2 and 3 also experienced adrenal crisis, whereas maternal virilization during pregnancy was not identified in cases 1–3. In addition, case 2 had right vesicoureteral reflux, and case 3 manifested imperforated anus. In cases 1–3, direct sequencing for leukocyte genomic DNA indicated apparent heterozygosity for the Japanese founder mutation p.R457H, and that for leukocyte cDNA demonstrated transcription failure of an apparently normal allele (7). Thus, although cases 1–3 were found to have compound heterozygosity for p.R457H and transcription failure, the cause of transcription failure remained to be clarified.

Primer and probe

The primers and probes used in the present study are shown in Supplemental Table 1 (published on The Endocrine Society's Journals Online web site at http://endo.endojournals.org).

Genome structure analysis

Oligoarray comparative genomic hybridization (CGH) was performed for leukocyte genomic DNA, using a custom-build oligo-microarray containing 39,169 probes for an approximately 8-Mb region around *POR* and 26,662 reference probes for a different genomic interval (2x105K format, design ID 022431) (Agilent Technologies, Palo Alto, CA). The procedure was as described in the manufacturer's instructions. To determine the deletion size and the junction structure, serial direct sequencing was performed for long PCR products obtained with primer pairs flanking the deleted region, and the obtained junction sequence was compared with the reference sequence at the NCBI Database (NT_007933.15). The presence or absence of repeat sequences around the breakpoints was examined with Repeatmasker (http://www.repeatmasker.org).

In silico analysis

In silico analysis was performed for CpG islands, evolutionally conserved sequences, and promoter-associated histone marks, using UCSC genome browser (http://genome.ucsc.edu/). Putative transcription factor binding sites were searched by TFSEARCH (http://mbs.cbrc.jp/research/db/TFSEARCH.html). In addition, because animal *Por* has been well studied in rats (9), conservation status of identified sites was examined using rat data. The transcription start site of *POR* exon 1U (+1) was determined on the basis of the *POR* cDNA sequence (NM_000941) obtained from the NCBI database.

Luciferase assays

A series of promoter-reporter constructs were generated by inserting PCR-amplified DNA fragments into PGL3-enhancer vector or pGL3-basic vector (Promega, Madison, WI). Deletion mutants were created by site-directed mutagenesis. Transient transfection was carried out using human embryonic kidney (HEK) 293 cells with endogenous SP families, because of their stable transfection efficiency and usefulness in *in vitro* functional studies for SP1 binding sites (13). HEK 293 cells were cultured in DMEM at 37 C, seeded in 12-well dishes, and transfected using Lipofectamine 2000 (Life Technologies, Carlsbad, CA) with 0.6 μg of the reporter plasmids. As an internal control for the transfection, 20 ng pRL-CMV vector (Promega) was used. In addition, transient transfection was also performed using *Drosophila* Schneider line 2 (SL2) cells (CRL-1963; American Type Culture Collection, Manassas, VA) that lack endogenous SP families. SL2 cells were grown in Schneider's medium at 25 C, seeded in six-well dishes, and transfected using calcium phosphate (14) with 1.0 μg of the reporter plasmid and a total of 50 ng of various combinations of the SP1 expression vector (pPAC-SP1) and an empty pPAC vector, as well as 50 ng of the SP3 expression vector (pPAC-SP3). As an internal control for the transfection, 50 ng pPAC-β-galactosidase vector was used. For both experiments using HEK 293 cells and SL2 cells, luciferase activities were determined at 48 h after the transfections.

Transfections were performed in triplicate within a single experiment, and the experiments were repeated three times. The results are expressed as mean ± SEM, and statistical significance was examined by the *t* test. $P < 0.05$ was considered significant.

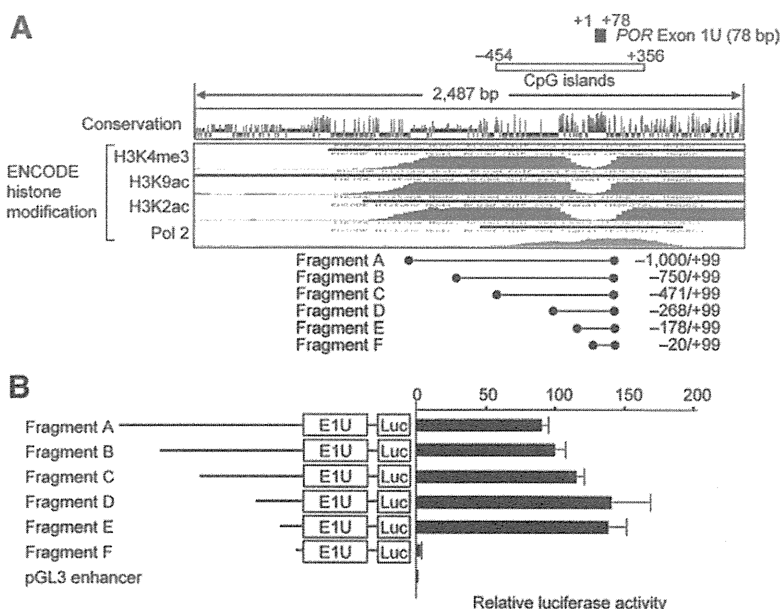

FIG. 2. Localization of the promoter region to a 178-bp segment just upstream of exon 1U. Panel A, *In silico* analysis in search of the promoter-compatible sequences. The transcription start site of *POR* exon 1U (+1) is based on the *POR* cDNA sequence at the NCBI database (NM_000941). The CpG-rich region spans from −454 to +356 bp. The ENCODE histone modification analysis indicates the presence of a highly conserved promoter-compatible sequence just upstream of exon 1U. The fragments A–F denote the DNA sequences used for the luciferase assays. Panel B, Luciferase reporter assays using the fragments A–F. The results are expressed as fold-change of the target vectors over the empty pGL3 enhancer vector (mean ± SEM). Transfections were performed in triplicate within a single experiment, and the experiments were repeated three times. Although the increase in the relative luciferase activity is significant for fragment A (92.6 ± 5.2, $P = 0.0006$), fragment B (101.6 ± 5.8, $P = 0.0006$), fragment C (106.0 ± 5.5, $P = 0.0004$), fragment D (137.7 ± 29.0, $P = 0.0009$), and fragment E (131.3 ± 13.4, $P = 0.0006$), it is not significant for fragment F (2.6 ± 1.1, $P = 0.25$).

DNA binding analysis

EMSA was performed as described previously (15). In brief, 10 μg of nuclear extracts of HEK 293 cells were incubated with ^{32}P-labeled oligonucleotides and unlabeled polydeoxyinosinic-deoxycytidylic acids and subjected to polyacrylamide gel electrophoresis (4%). For a competition experiment, a 200-fold molar excess of unlabeled competitor DNA was added. Supershift assay was performed by preincubating the nuclear extracts with anti-SP1 antisera (PEP2) and/or anti-SP3 antisera (D-20) (Santa Cruz Biotechnology, Santa Cruz, CA).

Methylation analysis

Bisulfite sequencing was performed for human leukocyte- and HEK 293-derived genomic DNA samples treated with the EZ DNA Methylation Kit (Zymo Research, Orange, CA) that converts all the cytosines except for methylated cytosines at the CpG dinucleotides into uracils and subsequently thymines. A 282-bp CpG-rich region containing SP1 binding sites just upstream of exon 1U was amplified with primer sets that hybridize to both methylated and unmethylated alleles because of absent CpG dinucleotides within the primer sequences. Subsequently, the PCR products were subcloned with the TOPO TA Cloning Kit (Life Technologies), and multiple clones were subjected to direct sequencing on the CEQ 8000 autosequencer (Beckman Coulter, Fullerton, CA).

Results

Identification and characterization of microdeletions in cases 1–3

Oligoarray CGH analysis indicated cryptic heterozygous deletions in cases 1–3 (Fig. 1). Furthermore, sequencing of the long PCR products harboring the fusion points revealed a 2,487-bp microdeletion (13,575,403–13,577,889 bp) encompassing exon 1U in case 1 and an identical 49,604-bp deletion (13,571,326–13,620,929 bp) involving exon 1U and exon 1 in cases 2 and 3. Thus, the 2,487-bp microdeletion on the noncoding upstream region was common to cases 1–3. The microdeletion in case

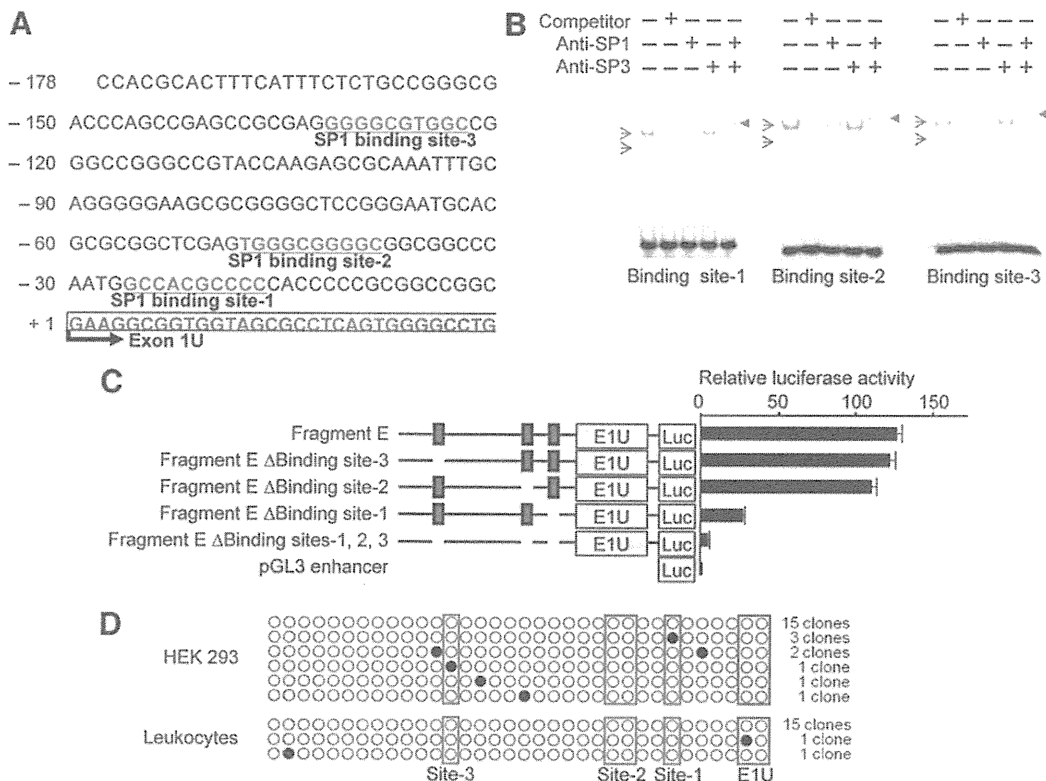

FIG. 3. Functional studies of the SP1 binding sites. Panel A, The three potential SP1 binding sites 1–3 at the position just upstream of exon 1U. The transcription start site of *POR* exon 1U (+1) is based on the *POR* cDNA sequence at the NCBI database (NM_000941). Panel B, EMSA showing positive bindings of SP1 and SP3 proteins to the SP1 binding sites 1–3. The *red arrows* indicate the strong bands derived from the SP1 protein binding to the probes containing the SP1 binding sites. These bands become weak, and supershifted bands (*red arrowheads*) are seen by adding anti-SP1. In addition, the *blue arrows* denote specific bands derived from the SP3 protein binding to the same probes. These bands become very weak by adding anti-SP3; the extremely faint supershifted bands are not visible in this figure. The band shift pattern is more obvious for SP1 protein than for SP3 protein. Panel C, Luciferase reporter assays using fragment E and its deletion mutants. The results are expressed as fold change of the target vectors over the empty pGL3 enhancer vector (mean ± SEM). Transfections were performed in triplicate within a single experiment, and the experiments were repeated three times. Although the relative luciferase activity is similar between Fragment E (121.8 ± 3.4) and ΔBinding site-3 (117.8 ± 3.1) ($P = 0.22$), it is significantly different between Fragment E and ΔBinding site-2 (105.7 ± 3.5) ($P = 0.015$), ΔBinding site-1 (25.8 ± 1.2) ($P = 0.0007$), and ΔBinding site-1, -2, and -3 (5.2 ± 0.5) ($P = 0.0004$). Panel D, Methylation analysis of the CpG-rich region. Each *circle* denotes a CpG island, and *filled* and *open circles* represent methylated and unmethylated cytosines, respectively. The CpG dinucleotides within the exon 1U are surrounded by *blue squares*, and those within the SP1 binding sites 1, 2, and 3 by *red squares*.

1 occurred between an *Alu* element and a nonrepeat sequence and was associated with an addition of a 47-bp segment of unknown origin, whereas that in cases 2 and 3 occurred between two *Alu* elements with an overlap of a 13-bp segment.

Critical function of the SP1 binding sites

In silico analysis for the noncoding 2,487-bp region showed an 810-bp long CpG-rich region involving exon 1U, an approximately 350-bp long evolutionally conserved sequence-rich region encompassing exon 1U, and an approximately 1.3-kb region with promoter-associated histone marks (Fig. 2A). The TATA box was not identified. Thus, relative luciferase activity was examined for fragments A–F with various lengths of the candidate promoter region, localizing a critical sequence for the *POR* promoter to a 178-bp segment defined by fragment E and fragment F (Fig. 2B).

The 178-bp segment was found to harbor three SP1 binding sites, *i.e.* site 1 at the position −26/−17, site 2 at the position −48/−39, and site 3 at the position −132/−123 (Fig. 3A). The three binding sites were well conserved in rats. EMSA indicated specific binding of SP1 and SP3 proteins to the three binding sites, with the band shift pattern being more obvious for SP1 protein than for SP3 protein (Fig. 3B). Deletion of the binding site 1 and the binding site 2 significantly reduced the relative luciferase activity (by ~80 and ~15%, respectively), although deletion of the binding site 3 had no significant effect on the relative luciferase activity; furthermore, loss of the binding sites 1–3 virtually abolished the relative luciferase activity (Fig. 3C). The 282-bp segment containing the three SP1 binding sites was almost completely unmethylated (Fig. 3D).

Furthermore, relative luciferase activity was examined for a 170-bp fragment (−120/+50) harboring the SP1 binding site 1 and the SP1 binding site 2, using SL2 cells devoid of endogenous SP families. Relative luciferase activity was clearly increased in a dose-dependent manner by adding the *SP1* expression vector but was barely elevated by adding the *SP3* expression vector (Fig. 4).

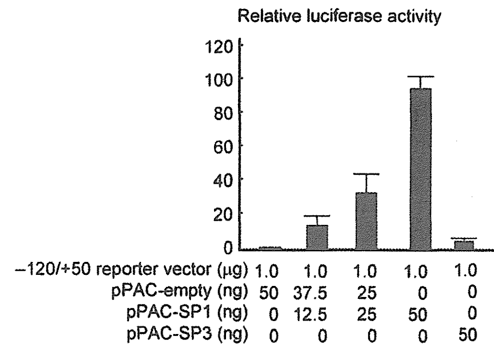

FIG. 4. Luciferase assays of a fragment containing the SP1 binding sites 1 and 2, using SL2 cells lacking endogenous SP families. The results are expressed as fold change of the target vectors over the empty pPAC vector (mean ± SEM). Transfections were performed in triplicate within a single experiment, and the experiments were repeated three times. The relative luciferase activity is significantly increased by adding the *SP1* expression vector of 12.5 ng (14.7 ± 4.4) ($P = 0.037$), 25.0 ng (31.8 ± 7.6) ($P = 0.035$), and 50 ng (95.8 ± 7.1) ($P = 0.0002$), although it is barely elevated by adding the *SP3* expression vector of 50 ng (5.2 ± 1.5) ($P = 0.054$).

Discussion

We identified two types of cryptic deletions, one involving exon 1U alone and the other encompassing exon 1U and exon 1, in three cases with PORD. The microdeletion in case 1 is explained by nonhomologous end joining that occurs between nonhomologous sequences and is frequently accompanied by an insertion of a short segment at the fusion point (16). The microdeletion in cases 2 and 3 is compatible with a repeat sequence mediated nonallelic intrachromosomal or interchromosomal recombination (16). Although cases 2 and 3 were apparently nonconsanguineous, it would not be unexpected that the same repeat-mediated genomic rearrangement took place in unrelated individuals. Notably, because the apparently normal allele in cases 1–3 was not transcribed (7), this implies that the 2,487-bp microdeletion common to cases 1–3 has affected the promoter function for *POR*. In this context, because approximately 12% of patients with PORD are known to be heterozygotes with one apparently normal *POR* allele (4), it might be possible that some, if not all, of them have similar microdeletions or other genetic aberrations affecting the *POR* transcription.

The present study revealed a pivotal role of the SP1 binding sites, especially the binding site 1, in the transcription of *POR*. This implies that the SP1 binding sites constitute an essential element of the *POR* proximal promoter. Indeed, SP1 binding sites as well as other noncore promoter elements are usually located in multiple copies within the proximal promoter region (~250 bp upstream of the transcription initiation site) of a ubiquitously expressed gene like *POR* (10). In this regard, several findings are noteworthy. First, the TATA box was apparently absent from the *POR* promoter region. This is compatible with the ubiquitous expression of *POR*, because the TATA box is usually identified in genes with a tissue-specific expression pattern (10). Second, the SP1 binding sites were highly conserved between the human and the rat. This finding, in conjunction with the previous data indicating absence of polymorphism for the three SP1 binding sites in 842 individuals (17), implies that the wild-type sequences of the SP1 binding sites are indispensable for the regulation of *POR* transcription. Third, the functional

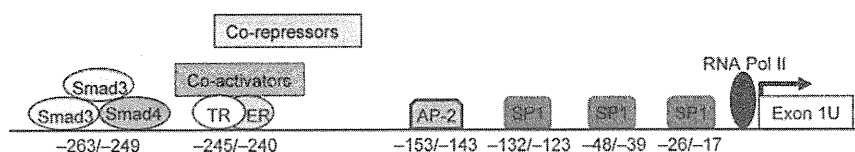

FIG. 5. Schematic representation indicating the binding sites for various factors in the proximal promoter region of *POR*. The diagram of the promoter upstream of −143 has been taken from Tee *et al.* (12). ER, Estrogen receptor; Pol II, polymerase II; TR, thyroid hormone receptor; AP-2, activator protein 2.

data using SL2 cells indicated a major role of SP1, rather than SP3, in the *POR* transcription. This is consistent with the notion that although both SP1 and SP3 can bind to the same cognate SP1 binding site, the DNA binding properties and regulatory functions are quite different between SP1 and SP3, depending on the promoter context and the cell type (18). Lastly, the SP1 binding sites were almost completely unmethylated. This argues for a transcriptionally active status of *POR*, because SP1 protein binding is known to be reduced when the CpG-rich region around the SP1 binding sites is methylated (19).

The proximal promoter region of *POR* has been studied previously (11, 12). Scott *et al.* (11) analyzed the 5′ region of *POR* coding exons by means of comparative genomics and characterized human *POR* exon 1U and its flanking sequences. Subsequently, Tee *et al.* (12) examined a 361-bp region around the transcription start site of exon 1U (−325/+36) using adrenal NCI-H295A and liver Hep-G2 cells and found a major trophic effect of thyroid hormone on *POR* expression primarily via thyroid hormone receptor-β as well as modulatory effects of thyroid hormone receptor-α, estrogen receptor-α, Smad3, and Smad4 on *POR* expression. The binding sites for these factors reside in a −263/−240 region upstream of the SP1 binding sites (Fig. 5). Furthermore, Tee *et al.* (12) screened functional alterations of polymorphisms within the 325-bp region, suggesting that the common −152C→A polymorphism may play a certain role in the genetic variation of steroid biosynthesis and drug metabolism. In this regard, whereas the −152C→A polymorphism resides on the AP-2 (activator protein 2) binding site, the functional difference of the polymorphism is obviously independent of the recruit of AP-2 (12). Thus, the underlying factors for the reduced activity of the −152A allele remain to be clarified.

Taken together, multiple regulatory elements have been identified in the proximal promoter region of *POR* (Fig. 5). Although the regulatory machinery has not yet been fully elucidated, we suggest that the presence of the SP1 binding sites has permitted the ubiquitous expression of *POR* and that the presence of other sites including thyroid hormone receptors is relevant to the variability in *POR* expression level among different tissues. In this regard, although the present study failed to identify the effects of the −263/−240 regulatory sequence identified by Tee *et al.* (12) (fragment D *vs.* fragment E in Fig. 2), this may be due to the difference in the cell type and/or in the promoter-luciferase construct used in the study by Tee *et al.* (+36) and in this study (+99). In addition, the hormonal effects on the *POR* transcription have not been examined in this study.

Finally, it would be useful to refer to clinical phenotypes of cases 1–3. In this context, we have previously compared clinical phenotype between Japanese PORD patients with homozygosity for the hypomorphic p.R457H mutation (group A) and those with compound heterozygosity for p.R457H and one apparently null mutation including nonsense and frameshift mutations (group B) and found that skeletal features are definitely more severe and adrenal dysfunction and 46,XY DSD are somewhat more severe in group B than in group A, whereas 46,XX DSD, maternal virilization during pregnancy, and anorectal and urinary anomalies are similarly identified in the two groups (5, 7). It is likely, therefore, that the residual POR activity reflected by the p.R457H dosage constitutes the underlying factor for clinical variability in some features but not in other features, probably due to the simplicity and complexity of POR-dependent metabolic pathways relevant to each phenotype. The clinical features of cases 1–3 are quite comparable to those of group B patients and, therefore, are consistent with transcription failure of one allele being a null mutation.

In summary, we identified microdeletions involving exon 1U and its upstream region in PORD patients, and revealed the critical function of the SP1 binding sites in the transcription of *POR*. Additional studies will permit to elucidate the regulatory machinery for *POR* expression.

Acknowledgments

We thank Dr. Timothy F. Osborne, University of California, Irvine, CA, and Dr. Guntram Suske, Philipps Universität, Marburg, Germany, for providing us with pPAC-β-galactosidase and pPAC-SP1/Sp3, respectively.

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

This work was supported by the Grant-in-Aid for Scientific Research on Innovative Areas (22132004) from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT); by the Grant-in-Aid for Scientific Research (S) (22227002) from

the Japan Society for the Promotion of Science (JSPS); by the Grants for Research on Intractable Diseases (H22-098) and for Health Research on Children, Youth, and Families (H21-005) from the Ministry of Health, Labor and Welfare; by the Grant from Japan Foundation for Pediatric Research (10-001); and by the Grant of National Center for Child Health and Development (23A-1).

Disclosure Summary: The authors have nothing to declare.

References

1. **Flück CE, Tajima T, Pandey AV, Arlt W, Okuhara K, Verge CF, Jabs EW, Mendonça 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. **Fukami M, Nagai T, Mochizuki H, Muroya K, Yamada G, Takitani K, Ogata T** 2010 Anorectal and urinary anomalies and aberrant retinoic acid metabolism in cytochrome P450 oxidoreductase deficiency. Mol Genet Metab 100:269–273
6. **Huang N, Pandey AV, Agrawal V, Reardon W, Lapunzina PD, Mowat D, Jabs EW, Van Vliet G, Sack J, Flück 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
7. **Fukami M, Nishimura G, Homma K, Nagai T, Hanaki K, Uematsu A, Ishii T, Numakura C, Sawada H, Nakacho M, Kowase T, Motomura K, Haruna H, Nakamura M, Ohishi A, Adachi M, Tajima T, Hasegawa Y, Hasegawa T, Horikawa R, Fujieda K, Ogata T** 2009 Cytochrome P450 oxidoreductase deficiency: identification and characterization of biallelic mutations and genotype-phenotype correlations in 35 Japanese patients. J Clin Endocrinol Metab 94: 1723–1731
8. **Shephard EA, Palmer CN, Segall HJ, Phillips IR** 1992 Quantification of cytochrome P450 reductase gene expression in human tissues. Arch Biochem Biophys 294:168–172
9. **O'Leary KA, Kasper CB** 2000 Molecular basis for cell-specific regulation of the NADPH-cytochrome P450 oxidoreductase gene. Arch Biochem Biophys 379:97–108
10. **Strachan T, Read AP** 2004 Human gene expression. In: Strachan T, Read AP, eds. Human molecular genetics. 3rd ed. New York: Garland Science; 275–313
11. **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
12. **Tee MK, Huang N, Damm I, Miller WL** 2011 Transcriptional regulation of the human P450 oxidoreductase gene: hormonal regulation and influence of promoter polymorphisms. Mol Endocrinol 25:715–731
13. **Blondet A, Gout J, Durand P, Bégeot M, Naville D** 2005 Expression of the human melanocortin-4 receptor gene is controlled by several members of the Sp transcription factor family. J Mol Endocrinol 34:317–329
14. **Di Nocera PP, Dawid IB** 1983 Transient expression of genes introduced into cultured cells of Drosophila. Proc Natl Acad Sci USA 80:7095–7098
15. **Yazawa T, Mizutani T, Yamada K, Kawata H, Sekiguchi T, Yoshino M, Kajitani T, Shou Z, Miyamoto K** 2003 Involvement of cyclic adenosine 5′-monophosphate response element-binding protein, steroidogenic factor 1, and Dax-1 in the regulation of gonadotropin-inducible ovarian transcription factor 1 gene expression by follicle-stimulating hormone in ovarian granulosa cells. Endocrinology 144:1920–1930
16. **Gu W, Zhang F, Lupski JR** 2008 Mechanisms for human genomic rearrangements. Pathogenetics 1:4
17. **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
18. **Li L, He S, Sun JM, Davie JR** 2004 Gene regulation by Sp1 and Sp3. Biochem Cell Biol 82:460–471
19. **Zhu WG, Srinivasan K, Dai Z, Duan W, Druhan LJ, Ding H, Yee L, Villalona-Calero MA, Plass C, Otterson GA** 2003 Methylation of adjacent CpG sites affects Sp1/Sp3 binding and activity in the p21(Cip1) promoter. Mol Cell Biol 23:4056–4065

Original Article

Sexual Development

Sex Dev 330122
DOI: 10.1159/000330122

Accepted: June 3, 2011
by M. Schmid
Published online: ■■■

© S. Karger AG, Basel
PROOF Copy for personal use only
ANY DISTRIBUTION OF THIS ARTICLE WITHOUT WRITTEN CONSENT FROM S. KARGER AG, BASEL IS A VIOLATION OF THE COPYRIGHT.

Association of Primary Ovarian Insufficiency with a Specific Human Leukocyte Antigen Haplotype (A*24:02-C*03:03-B*35:01) in Japanese Women

T. Ayabe[a, c] B. Ishizuka[d] T. Maruyama[b] M. Fukami[a] R. Yoshida[a] H. Uchida[b] Y. Yoshimura[b] T. Nagai[c] T. Ogata[a, e]

[a]Department of Molecular Endocrinology, National Research Institute for Child Health and Development, and [b]Department of Obstetrics and Gynecology, Keio University School of Medicine, Tokyo, [c]Department of Pediatrics, Dokkyo Medical University Koshigaya Hospital, Koshigaya, [d]Department of Obstetrics and Gynecology, St. Marianna University School of Medicine, Kawasaki, and [e]Department of Pediatrics, Hamamatsu University School of Medicine, Hamamatsu, Japan

Key Words

Association study · Haplotype analysis · Human leukocyte antigen · Primary ovarian insufficiency · Susceptibility

Abstract

Primary ovarian insufficiency (POI) is a heterogeneous condition defined by the triad of oligo/amenorrhea, elevated gonadotropins and estrogen deficiency in women under the age of 40 years. Although autoimmune abnormalities appear to be involved in the development of POI, there are only a few studies with respect to human leukocyte antigen (HLA). The objective of this study was to identify an HLA allele(s) and/or haplotype(s) constituting a susceptibility factor(s) for POI. We examined 83 Japanese women with apparently idiopathic isolated POI. For controls, Japanese HLA reference data registered in the HLA Laboratory were utilized. No significant association was found for a total of 94 alleles for HLA-A, B, C, DRB1, and DQB1 loci, after both stringent Bonferroni correction and less stringent Benjamini-Hochberg (B-H) correction for multiple comparisons. By contrast, of 86 haplotypes identified for MHC class I (HLA-A, B, and C) and 31 haplotypes detected for MHC class II (HLA-DRB1 and DQB1), a single haplotype (A*24:02-C*03:03-B*35:01) remained significant after Bonferroni and B-H corrections (frequency: 4.82% in women with POI and 1.06% in the control data; $p = 0.00049$). The results imply that a specific HLA haplotype (A*24:02-C*03:03-B*35:01) constitutes a susceptibility factor for apparently isolated POI in Japanese women.

Copyright © 2011 S. Karger AG, Basel

Primary ovarian insufficiency (POI), or premature ovarian failure, is a heterogeneous condition defined by the triad of oligo/amenorrhea, elevated gonadotropins and estrogen deficiency in women under the age of 40 years [Kalantaridou et al., 1998]. While POI is frequently observed in women with sex chromosome aberrations, it also occurs in women with normal karyotype [Laml et al., 2000]. In this regard, although underlying factors for POI remain to be elucidated in most women with normal karyotype, various genetic and environmental factors have been implicated in the development of POI in such women. Indeed, mutations of several genes including *BMP15* [Di Pasquale et al., 2004], *NOBOX* [Qin et al., 2009] and *NR5A1* (alias *SF-1* and *AD4BP*) [Lourenço et al., 2009] are known to cause POI, as well as premutations

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2011 S. Karger AG, Basel
1661–5425/11/0000–0000$38.00/0

Accessible online at:
www.karger.com/sxd

Dr. T. Ogata
Department of Molecular Endocrinology
National Research Institute for Child Health and Development
2-10-1 Ohkura, Setagaya, Tokyo 157-8535 (Japan)
Tel. +81 3 3416 0181, ext. 4920, E-Mail tomogata@nch.go.jp

of *FMR1* [Cronister et al., 1991]. In addition, several candidate genes such as *LHX8* [Qin et al., 2008] and *GDF9* [Kovanci et al., 2007] have been identified to date. Furthermore, endocrine disruptors, chemotherapy, radiation, smoking, and viral infections also constitute risk factors for POI [Morrison et al., 1975; Howell and Shalet, 1998; Sharara et al., 1998; Di Prospero et al., 2004].

Autoimmune abnormalities also appear to be involved in the development of POI in women with normal karyotype. Indeed, POI can occur as a part of autoimmune polyendocrine syndromes [Kauffman and Castracane, 2003]. Moreover, women with POI occasionally have autoimmune hypothyroidism, type 1 diabetes mellitus (T1DM), hypoadrenalism, myasthenia gravis and systemic lupus erythematosus and, conversely, women with autoimmune disorders occasionally exhibit POI [Moncayo-Naveda et al., 1989; Belvisi et al., 1993; Bakalov et al., 2002; Ryan and Jones, 2004]. Non-specific autoantibodies such as antinuclear antibodies are often identified in women with POI [Ishizuka et al., 1999], as are organ-specific autoantibodies such as anti-thyroid autoantibodies [Belvisi et al., 1993] and steroid-producing cell autoantibodies [Betterle et al., 1993].

Human leukocyte antigen (HLA) of the major histocompatibility complex (MHC) plays an essential role in the human immune system. Consistent with this, specific HLA alleles or haplotypes are known to constitute susceptibility factors for several autoimmune disorders such as T1DM [Todd et al., 1988]. For POI, however, there are only a few studies with respect to HLA, and there is no haplotype association study. Thus, we performed HLA allele and haplotype analyses in Japanese women with POI. Here, we focused on apparently idiopathic isolated POI as a relatively homogeneous group, because POI with associated autoimmune abnormalities may represent genetically heterogeneous disorders.

Subjects and Methods

Subjects

This study consisted of 83 Japanese women with apparently idiopathic isolated POI who satisfied the following selection criteria: (1) lack of somatic abnormalities; (2) absence of clinically discernible autoimmune diseases; (3) no history of chemotherapy or radiation; (4) 46,XX karyotype in all the ≥30 lymphocytes examined; (5) no demonstrable mutations in *BMP15, NOBOX, NR5A1, LHX8*, and *GDF9*, and (6) no *FMR1* premutations. Two women were familial cases with a similarly affected sister or mother, and the remaining 81 women were sporadic cases. Organ-specific autoantibodies such as anti-thyroid autoantibodies as well as non-specific antinuclear antibodies were not detected in 60 patients examined, although they were not examined in the remaining 23 patients. The menarchial age ranged from 10 to 15 years (median, 13 years) (menarchial age in normal Japanese girls, 12.25 ± 1.25 years), and the age of POI onset (amenorrhea persisting ≥6 months) ranged from 13 to 39 years (median, 31 years). Serum FSH was above 40 IU/l and estradiol was undetectable or below the normal range on at least 2 occasions. Serum testosterone was undetectable, and polycystic ovary was excluded by abdominal ultrasound studies.

Thus, the following patients with overt autoimmune diseases were excluded from the present study: (1) ten patients with Hashimoto thyroiditis accompanied by anti-thyroid peroxidase autoantibodies; (2) one patient with Basedow disease; (3) one patient with Sjögren syndrome; (4) two patients with systemic lupus erythematosus, and (5) two patients with anti-phospholipid antibody syndrome.

Human Leukocyte Antigen Analysis

This study was approved by the Institutional Review Board Committees of the investigators' affiliations. After obtaining written informed consent, genotyping was performed for HLA-A, B, C, DRB1, and DQB1 loci by the polymerase chain reaction-based sequence-specific oligonucleotide probe method, using leukocyte genomic DNA samples. Haplotype inference was performed for MHC class I (HLA-A, B, and C) and class II (DRB1 and DQB1) by the maximum likelihood method using expectation maximization algorithm [Excoffier and Slatkin, 1995] implemented in the software LDSUPPORT [Kitamura et al., 2002]. For controls, we utilized Japanese HLA reference data registered in the HLA Laboratory (http://www.hla.or.jp/) that have been obtained from 14,631 control subjects for HLA-A, B, and DRB1 loci, 8,240 control subjects for HLA-C locus and MHC class I haplotype, and 2,934 control subjects for HLA-DQB1 locus and MHC class II haplotype. The sex ratio was nearly equal in the control subjects, and HLA allele and haplotype patterns have been shown to be comparable between men and women control subjects (http://www.hla.or.jp/).

Statistical Analysis

Allele and haplotype frequencies were compared between women with POI and control subjects by the Fisher's exact probability test, using R environment version 2.7.1 (http://www.r-project.org/). The odds ratio and the 95% confidence interval were also calculated using the same environment. To address the problem of multiple comparisons, the Bonferroni correction and the Benjamini-Hochberg (B-H) correction were employed in this study [Shaffer, 1995]. The Bonferroni correction is the most stringent method that sets the corrected significance level by dividing the empirical significance level (0.05) by the number of tests, and corrected p values by multiplying the observed p values by the number of tests. The B-H correction is a subsequently developed less stringent method that defines the significance level by multiplying the Bonferroni significance level by the frequency order of observed p values, and corrected p values by dividing the Bonferroni corrected p values by the frequency order of observed p values. In both corrections, observed p values are regarded as significant when corrected p values remain below the empirical significance level (0.05). Since individual HLA genotyping data are not available in the control reference data, it was impossible to perform the permutation test that is frequent-

Table 1. Representative results of the association studies between POI and each HLA allele (p < 0.05)

Allele	Frequency, %		Statistical analysis			Bonferroni correction		Benjamini-Hochberg correction		
	women with POI	control subjects	p value	odds ratio	95% CI	significance level	corrected p	significance level	corrected p	frequency order of p value
B*27:07	0.60	0.00	0.0056	530.4	21.51–13,080	0.0017	0.16	0.0017	0.16	1
B*35:03	0.60	0.00	0.0056	530.4	21.51–13,080	0.0017	0.16	0.0017	0.16	1
B*51:02	1.20	0.19	0.034	6.60	1.59–27.29	0.0017	1.00	0.0069	0.25	4
B*55:04	1.20	0.13	0.0071	9.38	2.24–39.21	0.0017	0.21	0.0052	0.07	3
C*15:05	0.60	0.02	0.043	24.96	2.77–224.70	0.0033	0.65	0.0033	0.64	1
DRB1*04:01	3.01	0.94	0.019	3.29	1.34–8.07	0.0023	0.42	0.0023	0.41	1
DRB1*12:02	4.22	1.82	0.045	2.38	1.11–5.09	0.0023	0.99	0.0068	0.33	3
DRB1*14:01	6.63	3.36	0.035	2.04	1.11–3.78	0.0023	0.77	0.0046	0.38	2

CI = Confidence interval.

ly utilized to estimate significant p values in multiple comparisons [Shaffer, 1995].

The power (1 – β [type II error]: the probability of rejecting a false null hypothesis) was estimated using GDesignPlus (StaGen Co., Ltd.). The power of >50% is usually required to approve the results obtained in a pilot study like the present study, while the power of >80% is usually necessary to admit the results obtained in a replication study [Cohen, 1988]. Furthermore, false positive report probability (FPRP), the probability of no true association between a genetic variant and disease given a statistically significant finding, was also calculated using α (the significance level), power (1 – β), and π (assumed frequency of susceptibility allele or haplotype; in this study, π was assumed as 0.02) [Wacholder et al., 2004]. The FPRP of <5% is usually regarded as verifying the quality of the obtained data.

Results

HLA genotyping identified 16 alleles for the HLA-A locus, 29 alleles for HLA-B, 15 alleles for HLA-C, 22 alleles for HLA-DRB1, and 12 alleles for HLA-DQB1 in women with POI (online suppl. table 1; for all suppl. material, see www.karger.com/doi/10.1159/000330122). While low p values (p < 0.05) were identified for 8 of these alleles (table 1), none of them were significant after the Bonferroni and B-H corrections.

Haplotype estimation indicated 86 haplotypes for MHC class I and 31 haplotypes for MHC class II in women with POI (online suppl. table 2). Low p values (p < 0.05) were identified for 34 of these haplotypes (table 2), and one of them (A*24:02-C*03:03-B*35:01) remained significant after the Bonferroni and B-H corrections. This haplotype was invariably present in a heterozygous condition, and accounted for 4.82% of haplotypes in women with POI (the 2nd most frequent haplotype in this group) and 1.06% of haplotypes in the control subjects (the 15th most common haplotype in this group). The power was calculated as 72.2%, and the FPRP was 3.8%. Notably, each of the A*24:02, C*03:03, and B*35:01 alleles constituting the specific haplotype was fairly common in both women with POI and control subjects and was identified with similar frequencies (online suppl. table 1). There was no haplotype that was regarded as significant by the B-H correction but not by the Bonferroni correction.

Discussion

We studied Japanese patients with apparently idiopathic isolated POI who had no associated overt autoimmune diseases. This selection was performed to reduce genetic heterogeneity among patients. Indeed, POI in patients with Sjögren syndrome and systemic lupus erythematosus may represent specific groups of POI, as well as POI in patients with single gene diseases such as autoimmune polyendocrine syndromes [Kauffman and Castracane, 2003]. Similarly, thyroid disease-associated POI may also exhibit a specific group of POI. Thus, we excluded such patients from this study. It should be pointed out, however, that not all patients with POI received examination of autoantibodies, and that several patients may have subclinical autoimmune diseases or may develop autoimmune diseases at later ages.

The present study revealed significant association between apparently isolated POI and a specific HLA haplotype (A*24:02-C*03:03-B*35:01) in Japanese women. In this regard, since the frequency of each allele for the specific haplotype was similar between women with POI and

Table 2. Representative results of the association studies between POI and each HLA haplotype ($p < 0.05$)

Haplotype	Frequency, %		Statistical analysis			Bonferroni correction		Benjamini-Hochberg correction		
	women with POI	control subjects	p value	odds ratio	95% CI	significance level	corrected p	significance level	corrected p	frequency order of p value
A*02:01-C*01:02-B*54:01	2.41	0.73	0.036	3.37	1.23–9.23	0.00058	1.00	0.01	0.13	23
A*02:01-C*03:03-B*55:04	1.20	0.07	0.0083	16.74	3.72–75.40	0.00058	0.71	0.0017	0.24	3
A*02:01-C*03:04-B*40:02	2.41	0.66	0.026	3.74	1.36–10.28	0.00058	1.00	0.01	0.13	17
A*02:01-C*08:01-B*40:02	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*02:01-C*08:01-B*48:01	1.81	0.16	0.0029	11.65	3.49–38.88	0.00058	0.25	0.0012	0.12	2
A*02:06-C*01:02-B*15:01	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*02:06-C*01:02-B*40:06	0.60	0.01	0.03	49.93	4.50–553.7	0.00058	1.00	0.01	0.14	18
A*02:06-C*07:02-B*67:01	0.60	0.02	0.049	24.96	2.77–224.7	0.00058	1.00	0.02	0.16	27
A*02:07-C*01:02-B*54:01	0.60	0.02	0.049	24.96	2.77–224.7	0.00058	1.00	0.02	0.16	27
A*02:10-C*01:02-B*15:01	0.60	0.01	0.03	49.93	4.50–553.7	0.00058	1.00	0.01	0.14	18
A*11:01-C*03:04-B*15:01	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*11:01-C*03:04-B*51:01	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*11:01-C*07:02-B*40:02	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*11:01-C*08:03-B*15:01	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*24:02-C*01:02-B*46:01	2.41	0.80	0.047	3.06	1.12–8.37	0.00058	1.00	0.02	0.16	26
A*24:02-C*01:02-B*56:03	0.60	0.01	0.03	49.93	4.50–553.7	0.00058	1.00	0.01	0.14	18
A*24:02-C*03:03-B*35:01	4.82	1.06	0.00049	4.75	2.30–9.81	0.00058	0.042	0.00058	0.042	1
A*24:02-C*04:01-B*27:07	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*24:02-C*08:01-B*07:05	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*24:02-C*08:01-B*40:01	0.60	0.02	0.049	24.96	2.77–224.7	0.00058	1.00	0.02	0.16	27
A*26:01-C*01:02-B*15:01	0.60	0.02	0.049	24.96	2.77–224.7	0.00058	1.00	0.02	0.16	27
A*26:01-C*03:04-B*40:01	1.20	0.18	0.049	6.69	1.58–28.22	0.00058	1.00	0.01	0.17	25
A*26:01-C*07:02-B*39:02	0.60	0.01	0.03	49.93	4.50–553.7	0.00058	1.00	0.01	0.14	18
A*26:01-C*08:01-B*35:01	1.20	0.17	0.036	7.17	1.69–30.34	0.00058	1.00	0.01	0.13	24
A*26:01-C*12:02-B*52:01	1.20	0.21	0.0499	5.90	1.41–24.77	0.00058	1.00	0.02	0.14	31
A*26:01-C*14:02-B*51:01	1.81	0.26	0.01	7.20	2.21–23.48	0.00058	0.86	0.0093	5.5E-02	16
A*29:01-C*15:05-B*15:02	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*31:01-C*03:04-B*51:02	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*31:01-C*07:02-B*67:01	0.60	0.01	0.03	49.93	4.50–553.7	0.00058	1.00	0.01	0.14	18
A*32:01-C*15:02-B*35:03	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
A*33:03-C*07:02-B*67:01	0.60	0.00	0.01	298.70	12.12–7,366	0.00058	0.86	0.0023	0.22	4
DRB1*04:01-DQB1*03:01	3.01	1.06	0.046	2.91	1.15–7.33	0.0016	1.00	0.0048	0.48	3
DRB1*12:02-DQB1*03:01	4.22	1.70	0.034	2.54	1.16–5.55	0.0016	1.00	0.0032	0.53	2
DRB1*14:06-DQB1*06:02	0.60	0.00	0.0039	106.40	4.31–2,623	0.0016	0.12	0.0016	0.12	1

CI = Confidence interval.

the control subjects, this would imply that an additive or a synergic effect of the alleles constituting the specific haplotype increases susceptibility to the development of POI. Thus, this study exemplifies the importance of haplotype analysis in the investigation of genetic susceptibility.

The mechanism(s) by which this specific haplotype raises the susceptibility to POI remains to be clarified. It may be possible, however, that this haplotype has some ovary-specific effect such as the production of hitherto unknown autoantibodies against an ovary-specific tissue(s), because this haplotype was identified as a susceptibility factor for the apparently isolated POI, and because there is no study suggesting a positive association between this haplotype and non-ovarian autoimmune disorders. In support of this notion, it is known in T1DM that pancreatic β-cell-specific insulinoma-associated antigen 2 autoantibodies are preferentially detected in patients with a specific HLA allele [Qu and Polychronakos, 2009].

HLA association studies have previously been performed for POI in non-Japanese populations. Walfish et al. [1983] found that the prevalence of DR3 antigen was significantly higher in 19 women with POI than in 80 control women. Anasti et al. [1994] reported that the frequency of DR4 antigen was significantly higher in 102

women with POI than in the reference data obtained from 1,927 subjects, whereas it was similar between the 102 women with POI and 102 control women. Jaroudi et al. [1994] identified no significant difference in the frequency of HLA loci between 37 women with POI and 100 control women. In these studies, however, HLA typing was performed by classic serological methods rather than refined genotyping methods, and haplotype analysis was not performed, while the Bonferroni correction was utilized to cope with multiple comparisons. Thus, while a possible association of POI with DR3 and/or DR4 may be suggested, this matter awaits further investigations. In this context, it may be worth pointing out that the present study failed to identify a positive association between POI and DR3 (DRB1*03) or DR4 (DRB1*04) (online suppl. table 1), although this may primarily be due to the ethnic difference, as has been shown in T1DM [Thomson et al., 2007].

Several points should be made with respect to the present study. First, the specific HLA haplotype accounts for only a minor portion (4.82%) of haplotypes identified in women with POI. Second, affected women may have some unidentified pathologic cause(s) for POI, such as mutations of hitherto unknown causative genes. Third, there may be some hidden polygenic and environmental differences between women with POI and the control subjects. Finally, it remains to be determined whether similar results can be reproduced in other studies. These matters imply that the etiologies of POI still remain elusive and await further investigations in most women with POI.

Despite the above caveats, the present study provides a useful clue to clarify the underlying factors for the development of POI. In summary, we propose that a specific HLA haplotype (A*24:02-C*03:03-B*35:01) constitutes a susceptibility factor for apparently isolated POI in Japanese women.

Acknowledgements

We would like to thank Professor Naoyuki Kamatani and Dr. Shigeo Kamitsuji, Department of Statistical Genetics, StaGen Co., Ltd., for their critical support in the genetic statistical analysis. This study was supported by Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) [22132004 to T.O.]; by Grants for Research on Intractable Diseases [H22-098 to T.O.]; by Grants for Health Research on Children, Youth and Families from the Ministry of Health, Labor and Welfare [H21-005 to T.O.]; and by Grant of National Center for Child Health and Development [20C-2 to T.O.].

References

Anasti JN, Adams S, Kimzey LM, Defensor RA, Zachary AA, Nelson LM: Karyotypically normal spontaneous premature ovarian failure: evaluation of association with the class II major histocompatibility complex. J Clin Endocrinol Metab 78:722–723 (1994).

Bakalov VK, Vanderhoof VH, Bondy CA, Nelson LM: Adrenal antibodies detect asymptomatic auto-immune adrenal insufficiency in young women with spontaneous premature ovarian failure. Hum Reprod 17:2096–2100 (2002).

Belvisi L, Bombelli F, Sironi L, Doldi N: Organ-specific autoimmunity in patients with premature ovarian failure. J Endocrinol Invest 16:886–892 (1993).

Betterle C, Rossi A, Dalla-Pria S, Artifoni A, Pedini B, et al: Premature ovarian failure: autoimmunity and natural history. Clin Endocrinol 39:35–43 (1993).

Cohen J: Statistical Power Analysis for the Behavioral Sciences, 2nd ed. (Lawrence Erlbaum Associates, New Jersey 1988).

Cronister A, Schreiner R, Wittenberger M, Amiri K, Harris K, Hagerman RJ: Heterozygous fragile X female: historical, physical, cognitive, and cytogenetic features. Am J Med Genet 38:269–274 (1991).

Di Pasquale E, Beck-Peccoz P, Persani L: Hypergonadotropic ovarian failure associated with an inherited mutation of human bone morphogenetic protein-15 *(BMP15)* gene. Am J Hum Genet 75:106–111 (2004).

Di Prospero F, Luzi S, Iacopini Z: Cigarette smoking damages women's reproductive life. Reprod Biomed Online 8:246–247 (2004).

Excoffier L, Slatkin M: Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol 12: 921–927 (1995).

Howell S, Shalet S: Gonadal damage from chemotherapy and radiotherapy. Endocrinol Metab Clin North Am 27:927–943 (1998).

Ishizuka B, Kudo R, Amemiya A, Yamada H, Matsuda T, Ogata T: Anti-nuclear antibodies in patients with premature ovarian failure. Hum Reprod 14:70–75 (1999).

Jaroudi KA, Arora M, Sheth KV, Sieck UV, Willemsen WN: Human leukocyte antigen typing and associated abnormalities in premature ovarian failure. Hum Reprod 9:2006–2009 (1994).

Kalantaridou SN, Davis SR, Nelson LM: Premature ovarian failure. Endocrinol Metab Clin North Am 27:989–1006 (1998).

Kauffman RP, Castracane VD: Premature ovarian failure associated with autoimmune polyglandular syndrome: pathophysiological mechanisms and future fertility. J Womens Health (Larchmt) 12:513–520 (2003).

Kitamura Y, Moriguchi M, Kaneko H, Morisaki H, Morisaki T, et al: Determination of probability distribution of diplotype configuration (diplotype distribution) for each subject from genotypic data using the EM algorithm. Ann Hum Genet 66:183–193 (2002).

Kovanci E, Rohozinski J, Simpson JL, Heard MJ, Bishop CE, Carson SA: Growth differentiating factor-9 mutations may be associated with premature ovarian failure. Fertil Steril 87:143–146 (2007).

Laml T, Schulz-Lobmeyr I, Obruca A, Huber JC, Hartmann BW: Premature ovarian failure: etiology and prospects. Gynecol Endocrinol 14:292–302 (2000).

Lourenço D, Brauner R, Lin L, De Perdigo A, Weryha G, et al: Mutations in *NR5A1* associated with ovarian insufficiency. N Engl J Med 360:1200–1210 (2009).

Moncayo-Naveda H, Moncayo R, Benz R, Wolf A, Lauritzen C: Organ-specific antibodies against ovary in patients with systemic lupus erythematosus. Am J Obstet Gynecol 160: 1227–1229 (1989).
Morrison JC, Givens JR, Wiser WL, Fish SA: Mumps oophoritis: a cause of premature menopause. Fertil Steril 26:655–659 (1975).
Qin Y, Zhao H, Kovanci E, Simpson JL, Chen ZJ, Rajkovic A: Analysis of *LHX8* mutation in premature ovarian failure. Fertil Steril 89: 1012–1014 (2008).
Qin Y, Shi Y, Zhao Y, Carson SA, Simpson JL, Chen ZJ: Mutation analysis of *NOBOX* homeodomain in Chinese women with premature ovarian failure. Fertil Steril 91 (Suppl):1507–1509 (2009).
Qu HQ, Polychronakos C: The effect of the MHC locus on autoantibodies in type 1 diabetes. J Med Genet 46:469–471 (2009).
Ryan MM, Jones HR Jr: Myasthenia gravis and premature ovarian failure. Muscle Nerve 30: 231–233 (2004).
Shaffer JP: Multiple hypothesis testing. Annu Rev Psychol 46:561–584 (1995).
Sharara FI, Seifer DB, Flaws JA: Environmental toxicants and female reproduction. Fertil Steril 70:613–622 (1998).
Thomson G, Valdes AM, Noble JA, Kockum I, Grote MN, et al: Relative predispositional effects of HLA class II DRB1-DQB1 haplotypes and genotypes on type 1 diabetes: a meta-analysis. Tissue Antigens 70:110–127 (2007).
Todd JA, Acha-Orbea H, Bell JI, Chao N, Fronek Z, et al: A molecular basis for MHC class II-associated autoimmunity. Science 240:1003–1009 (1988).
Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N: Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 96:434–442 (2004).
Walfish PG, Gottesman IS, Shewchuk AB, Bain J, Hawe BS, Farid NR: Association of premature ovarian failure with HLA antigens. Tissue Antigens 21:168–169 (1983).

Aromatase Excess Syndrome: Identification of Cryptic Duplications and Deletions Leading to Gain of Function of *CYP19A1* and Assessment of Phenotypic Determinants

Maki Fukami, Makio Shozu, Shun Soneda, Fumiko Kato, Akemi Inagaki, Hiroshi Takagi, Keiichi Hanaki, Susumu Kanzaki, Kenji Ohyama, Tomoaki Sano, Toshinori Nishigaki, Susumu Yokoya, Gerhard Binder, Reiko Horikawa, and Tsutomu Ogata

Department of Molecular Endocrinology (M.F., S.S., F.K., T.O.), National Research Institute for Child Health and Development, Tokyo 157-8535, Japan; Department of Reproductive Medicine (M.S.), Graduate School of Medicine, Chiba University, Chiba 206-8670, Japan; Department of Diabetes and Endocrinology (A.I., H.T.), Nagoya Second Red Cross Hospital Nagoya 466-8650, Japan; Department of Women's and Children's Family Nursing (K.H.) and Division of Pediatrics and Perinatology (S.K.), Tottori University, Yonago 683-8503, Japan; Department of Pediatrics (K.O., T.S.), Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Chuo 408-3898, Japan; Department of Pediatrics (T.N.), Osaka Police Hospital, Osaka 543-0035, Japan; Department of Medical Subspecialties (S.Y., R.H.), National Medical Center for Children and Mothers, Tokyo 157-8535, Japan; and Pediatric Endocrinology Section (G.B.), University Children's Hospital, Tuebingen 72076, Germany

Context: Aromatase excess syndrome (AEXS) is a rare autosomal dominant disorder characterized by gynecomastia. Although cryptic inversions leading to abnormal fusions between *CYP19A1* encoding aromatase and its neighboring genes have been identified in a few patients, the molecular basis remains largely unknown.

Objective: The objective of the study was to examine the genetic causes and phenotypic determinants in AEXS.

Patients: Eighteen affected males from six families participated in the study.

Results: We identified three types of heterozygous genomic rearrangements, *i.e.* a 79,156-bp tandem duplication involving seven of 11 noncoding *CYP19A1* exons 1, a 211,631-bp deletion involving exons 2–43 of *DMXL2* and exons 5–10 of *GLDN*, and a 165,901-bp deletion involving exons 2–43 of *DMXL2*. The duplicated exon 1 functioned as transcription start sites, and the two types of deletions produced the same chimeric mRNA consisting of *DMXL2* exon 1 and *CYP19A1* coding exons. The *DMXL2* exon 1 harbored a translation start codon, and the *DMXL2/CYP19A1* chimeric mRNA was identified in only 2–5% of *CYP19A1*-positive transcripts. This was in contrast to the inversion-mediated chimeric mRNA that had no coding sequence on the fused exon 1 and accounted for greater than 80% of *CYP19A1*-positive transcripts. *CYP19A1* was expressed in a limited number of tissues, whereas its neighboring genes involved in the chimeric mRNA formation were expressed widely.

Conclusions: This study provides novel mechanisms leading to gain of function of *CYP19A1*. Furthermore, it appears that clinical severity of AEXS is primarily determined by the tissue expression pattern of relevant genes and by the structural property of promoter-associated exons of chimeric mRNA. (***J Clin Endocrinol Metab* 96: E1035–E1043, 2011**)

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.
Copyright © 2011 by The Endocrine Society
doi: 10.1210/jc.2011-0145 Received January 18, 2011. Accepted March 22, 2011.
First Published Online April 6, 2011

Abbreviations: AEXS, Aromatase excess syndrome; CGH, comparative genomic hybridization; E_2, estradiol; FISH, fluorescence *in situ* hybridization; hCG, human chorionic gonadotropin; LCL, lymphoblastoid cell line; NMD, nonsense-mediated mRNA decay; RACE, rapid amplification of cDNA ends; SF, skin fibroblast; T, testosterone.

Aromatase is a cytochrome P450 enzyme that plays a crucial role in the estrogen biosynthesis (1). It catalyzes the conversion of Δ4-androstendione into estrone and that of testosterone (T) into estradiol (E_2) in the placenta and ovary as well as in other tissues such as the fat, skin, bone, and brain (1). It is encoded by *CYP19A1* consisting of at least 11 noncoding exons 1 and nine coding exons 2–10 (Supplemental Fig. 1, published on The Endocrine Society's Journals Online web site at http://jcem.endojournals.org) (2, 3). Each exon 1 is accompanied by a tissue-specific promoter and is spliced alternatively onto a common splice acceptor site at exon 2, although some transcripts are known to contain two of the exons 1, probably due to a splice error (2, 4). Of the 11 exons 1, exon I.4 appears to play a critical role in the regulation of estrogen biosynthesis in males because this exon contains a major promoter for extragonadal tissues including the skin and fat (2).

Excessive *CYP19A1* expression causes a rare autosomal dominant disorder known as aromatase excess syndrome (AEXS) (5–8). AEXS is characterized by pre- or peripubertal onset gynecomastia, advanced bone age from childhood to the pubertal period, and short adult height in affected males (5–8). Affected females may show several clinical features such as macromastia, precocious puberty, irregular menses, and short adult height (6–8). In this regard, previous studies have identified four heterozygous cryptic inversions around *CYP19A1* in patients with AEXS (5, 8). Each inversion results in the formation of a chimeric gene consisting of a noncoding exon(s) of a neighboring gene (*CGNL1*, *MAPK6*, *TMOD3*, or *TLN2*) and coding exons of *CYP19A1*. Because this condition is predicted to cause aberrant *CYP19A1* expression in tissues in which each neighboring gene is expressed, such inversions have been regarded to be responsible for AEXS (5, 8).

However, such inversions have been revealed only in a few patients with AEXS, and, despite extensive studies, no other underlying genetic mechanisms have been identified to date (6, 8–10). Here we report novel genomic rearrangements in AEXS and discuss primary phenotypic determining factors in AEXS.

Patients and Methods

Patients

This study was approved by the Institutional Review Board Committee at the National Center for Child Health and Development and was performed after obtaining informed consent. We examined 18 male patients aged 8–69 yr (cases 1–18) from six unrelated families A-F (Fig. 1A). The probands were ascertained by bilateral gynecomastia (Fig. 1B) and the remaining 12 males by familial studies. Ten other males allegedly had gynecomastia. There were four obligatory carrier females.

Phenotypic assessment showed pre- or peripubertal onset gynecomastia in all cases, small testes and fairly preserved masculinization in most cases, obvious or relative tall stature in childhood and grossly normal or relative short stature in adulthood, and age-appropriate or mildly advanced bone ages (Table 1) (for detailed actual data, see Supplemental Table 1). Such clinical features, especially gynecomastia, tended to be milder in cases 1–4 from families A and B than in the remaining cases from families C-F. Fertility or spermatogenesis was preserved in all adult cases (≥20 yr). In addition, the obligatory carrier females from families B and D had apparently normal phenotype, and such females from families E and F exhibited early menarche (9.0 yr) and short adult stature (−2.8 SD), respectively.

Blood endocrine studies revealed that LH values were grossly normal at the baseline and variably responded to GnRH stimulation, whereas FSH values were low at the baseline and responded poorly to GnRH stimulation, even after preceding GnRH priming (Table 1) (for detailed actual data, see Supplemental Table 1) (see also Fig. 1C for the cases aged ≥15 yr). Δ^4-Androstendione, T, and dihydrotestosterone values were low or normal. A human chorionic gonadotropin (hCG) test indicated relatively low but normal T responses in five young cases. In most cases, estrone values were elevated, E_2 values were normal or elevated, and E_2/T ratios were elevated. These endocrine data were grossly similar among cases 1–18.

Aromatase inhibitor (anastrozole, 1 mg/d) was effective in all the four cases treated (Supplemental Table 1) (see also Fig. 1C for cases aged ≥15 yr). Gynecomastia was mitigated within 6 months of treatment, and endocrine data were ameliorated within 1 month of treatment.

Primers

Primers used in this study are shown in Supplemental Table 2.

CYP19A1 mRNA levels and aromatase activities

We analyzed relative mRNA levels of *CYP19A1* and catalytic activities of aromatase in skin fibroblasts (SF) and lymphoblastoid cell lines (LCL). mRNA were extracted by a standard method and were subjected to RT-PCR using a high capacity RNA-to-cDNA kit (Life Technologies, Carlsbad, CA). A relative amount of *CYP19A1* mRNA against *B2M* was determined by the real-time PCR method using the Taqman gene expression assay on ABI PRISM 7500fast (Life Technologies) (assay no. Hs00903411_m1 for *CYP19A1* and Hs99999907_m1 for *B2M*). PCR was performed in triplicate. Aromatase activity was determined by a tritium incorporation assay (11). In brief, the samples were incubated with androstenedione-2-^{3}H for 2 h, and $^3H\,H_2O$ in the supernatant of the culture media was measured with a scintillation counter LSC-5100 (Aloka, Tokyo, Japan).

Sequence analysis of *CYP19A1*

Leukocyte or SF genomic DNA samples from the six probands and additional four male patients (Fig. 1A) were PCR amplified for the coding exons 2–10 and their flanking splice sites of *CYP19A1*. Subsequently the PCR products were subjected to direct sequencing from both directions on CEQ 8000 autosequencer (Beckman Coulter, Fullerton, CA).

Genome structure analysis

Oligonucleotide array-based comparative genomic hybridization (CGH) analyses were carried out using a custom-built

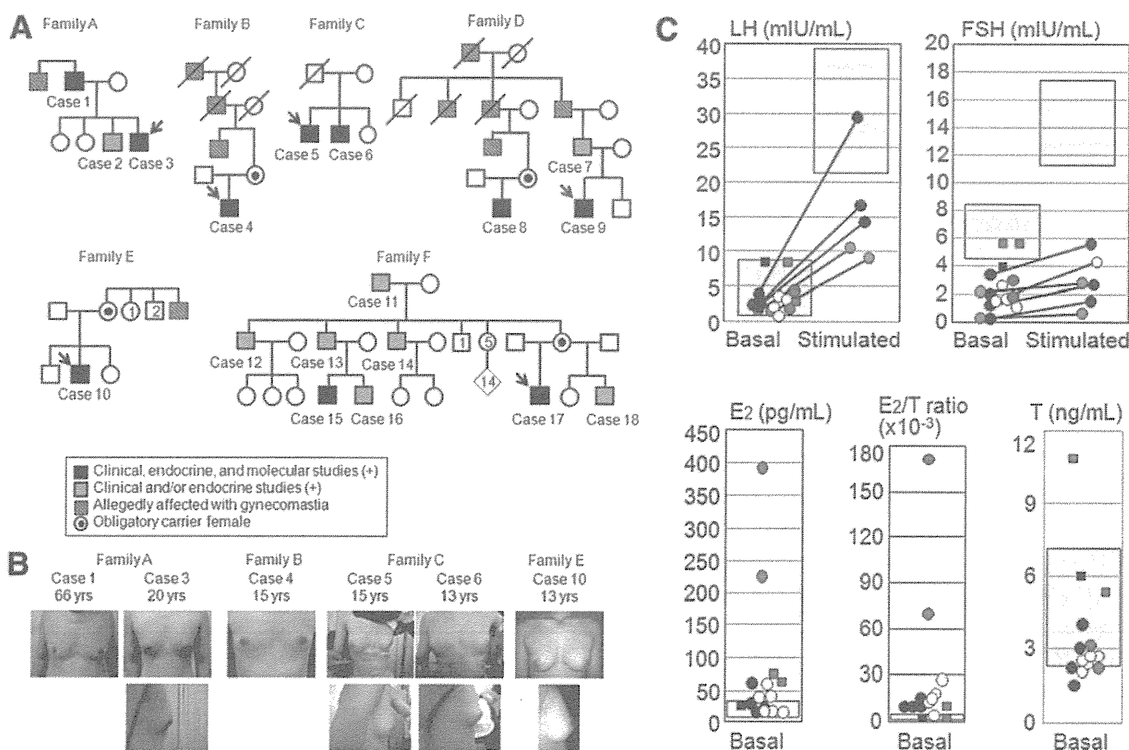

FIG. 1. Summary of clinical data. A, Pedigrees of six families with patients exhibiting AEXS-compatible phenotype. Families A–E are of Japanese origin, and family F is of German origin. Cases from families A–D were hitherto unreported, whereas those from families E and F have previously been described as having AEXS phenotypes (6, 8). B, Gynecomastia of six cases. C, Endocrine data in cases 15 yr of age or older. The *black*, *white*, and *red colors* represent the data in cases of the duplication, the deletion, and the inversion types, respectively; the *blue color* indicates the data of GnRH test after GnRH priming in two cases of the duplication type. The data at the time of diagnosis are denoted by *circles*, and those on aromatase inhibitor (anastrozole) treatment (1 mg/d in the duplication and the deletion types and 2–4 mg/d in the inversion types) are depicted by *squares*. The *light purple areas* represent the normal reference ranges.

oligo-microarray containing 90,000 probes for the 15q11.2-q26.3 region and approximately 10,000 reference probes for other chromosomal region (2 × 105K format, design identification 026533) (Agilent Technologies, Palo Alto, CA). The procedure was as described in the manufacturer's instructions. Fluorescence *in situ* hybridization (FISH) analysis was performed for lymphocyte or SF metaphase spreads, using long PCR products (FISH probes 1 and 2) for rearranged regions and CEP 15 probe for *D15Z4* used as an internal control (Abbott, Abbott Park, IL). The FISH probes 1 and 2 were labeled with digoxigenin and detected by rhodamine antidigoxigenin, and the CEP 15 probe was detected according to the manufacturer's protocol.

Characterization of the duplications and deletions

The duplication junctions were determined by direct sequencing for standard PCR products obtained with a variety of combinations of primers hybridizing to different positions within the *CYP19A1* exons 1 region. The deletion junctions were identified by direct sequencing of the long PCR products obtained with primer pairs flanking the deletions. The sizes of duplications and the deletions were determined by comparing obtained sequences with NT_010194 sequences at the National Center for Biotechnology Information Database (http://www.ncbi.nlm.nih.gov/; Bethesda, MD). The presence or absence of repeat sequences around the breakpoints was examined with Repeatmasker (http://www.repeatmasker.org).

For mRNA analysis, we preformed 5′-rapid amplification of cDNA ends (RACE) using a SMARTER RACE cDNA amplification kit (Takara Bio, Ohtsu, Japan). For both duplications and deletions, first PCR was carried out using the forward primer mix provided in the kit (Universal primer A mix) and an antisense reverse primer specific to *CYP19A1* exon 3 (RACE Rev). Second PCR was carried out for diluted products of the first PCR, using the nested forward primer of the kit (Nested universal primer A) and a reverse primer for *CYP19A1* exon 2 (Nested Rev). For duplications, furthermore, second PCR was also performed using various combinations of primers hybridizing to each *CYP19A1* exon 1. Subsequently PCR products were subcloned into TOPO cloning vector (Life Technologies) and subjected to direct sequencing. Then, the obtained sequences were examined with BLAST Search (National Center for Biotechnology Information). The presence or absence of promoter-compatible sequences was analyzed with the University of California, Santa Cruz, genome browser (http://genome.ucsc.edu/).

Relative mRNA levels of *CYP19A1* and its neighboring genes

We investigated relative mRNA levels of *CYP19A1* and *DMXL2* as well as those of *CGNL1*, *MAPK6*, *TMOD3*, and *TLN2* involved in the previously reported cryptic inversions (5, 8) in various human tissues. In this experiment, cDNA of SF and LCL were obtained from control males, and the remaining human cDNA samples were purchased from Life Technologies or Takara Bio. Relative quantification of mRNA against *TBP* was carried out using Taqman gene expression assay kit

TABLE 1. Summary of clinical studies in male patients with aromatase excess syndrome[a]

	Present study						Previous studies			
	Family A	Family B	Family C	Family D	Family E	Family F	Family 1	Family 2	Sporadic	
Cases	Cases 1–3	Case 4	Cases 5–6	Cases 7–9	Case 10	Cases 11–18	Two cases[b]	Proband[c]	Patient 1	Patient 2
Mutation type	Duplication	Duplication	Deletion	Deletion	Deletion	Deletion	Inversion	Inversion	Inversion	Inversion
Phenotypic findings										
Gynecomastia	Yes (mild)	Yes (mild)	Yes (moderate)	Yes (moderate)	Yes (moderate)	Yes (moderate)	Yes (severe)	Yes (severe)	Yes (severe)	Yes (severe)
Pubertal defect	Yes (mild)	Yes (mild)	Yes (mild)	No	No	Yes (mild)	N.D.	Yes (mild)	No	N.D.
Short adult height	No	No	N.D.	No	N.D.	No	Yes	N.D.	Yes	N.D.
Spermatogenesis	Preserved	N.D.	N.D.	Preserved	N.D.	Preserved	Preserved	N.D.	N.D.	N.D.
Endocrine findings										
LH (basal)	Normal	Normal	Normal	Normal/low	Normal	Normal/low	Normal	Normal/low	Normal	N.E.
LH (GnRH stimulated)[d]	Low	Normal	High	Normal	Normal	Normal	N.E.	Low	N.E.	N.E.
FSH (basal)	Low	Low	Low	Low	Low	Normal/low	Normal/low	Low	Low	N.E.
FSH (GnRH stimulated)[d]	Low	Low	Low	Low	Low	Low	N.E.	Low	N.E.	N.E.
T (basal)	Normal/low	Normal	Normal/low	Normal/low	Normal	Normal/low	Normal	Normal/low	Low	N.E.
T (hCG stimulated)[e]	N.E.	N.E.	Normal	Normal	Normal	Normal	N.E.	Normal	N.E.	N.E.
E_1 (basal)	High	High	N.E.	High	High	High	High	High	High	N.E.
E_2 (basal)	Normal	High	High	Normal	High	Normal/high	High	High	High	N.E.
E_2 to T ratio	High	High	High	High	High	High	High	High	High	N.E.

E_1, Estrone; N.D., not determined; N.E., not examined.

[a] Detailed actual data are shown in Supplemental Table 1.

[b] A father-son pair.

[c] The sister has macromastia, large uterus, and irregular menses; the parental phenotype has not been described.

[d] GnRH 100 μg/m^2 (maximum 100 μg) bolus iv; blood sampling at 0, 30, 60, 90, and 120 min.

[e] hCG 3000 IU/m^2 (maximum 5000 IU) im for 3 consecutive days; blood sampling on d 1 and 4.

(assay no. Hs00903411_m1 for *CYP19A1*; Hs00324048_m1 for *DMXL2*; Hs00262671_m1 for *CGNL1*; Hs00833126_g1 for *MAPK6*; Hs00205710_m1 for *TMOD3*; Hs00322257_m1 for *TLN2*; and Hs99999910_m1 for *TBP*). The experiments were carried out three times.

Results

CYP19A1 mRNA levels and aromatase activities

Although relative mRNA levels of *CYP19A1* and catalytic activities of aromatase were grossly similar between LCL of case 3 (family A), case 4 (family B), and case 5 (family C) and those of control subjects, they were significantly higher in SF of case 3 (family A), case 4 (family B), case 9 (family D), and case 10 (family E) than in those of control subjects (Fig. 2).

Sequence analysis of *CYP19A1*

Direct sequencing showed no mutation in *CYP19A1* coding exons 2–10 of the 10 cases examined.

Genome structure analysis

CGH analysis revealed heterozygous cryptic duplications involving most of the *CYP19A1* exons 1 region in cases from families A and B, heterozygous cryptic deletions involving most of *DMXL2* and part of *GLDN* in cases from family C, and heterozygous cryptic deletions involving most of *DMXL2* in cases from families D–F (Fig. 3A). FISH analysis supported the duplications and confirmed the deletions.

Characterization of the cryptic duplications

Aberrant PCR products were obtained with the P2 primer (which amplifies a segment between exon I.1 and exon IIa with the P1 primer) and the P3 primer (which amplifies a segment between exon I.2 and exon I.6 with the P4 primer), and sequencing of the PCR products showed the same tandem duplication involving seven of the 11 exons 1 of *CYP19A1* in cases from families A and B (Fig. 3B). The duplicated region was 79,156-bp long, and the

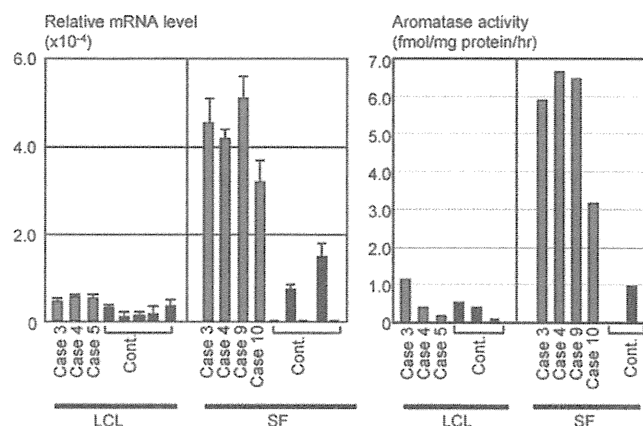

FIG. 2. Relative *CYP19A1* mRNA levels against *B2M* and catalytic activities of aromatase.

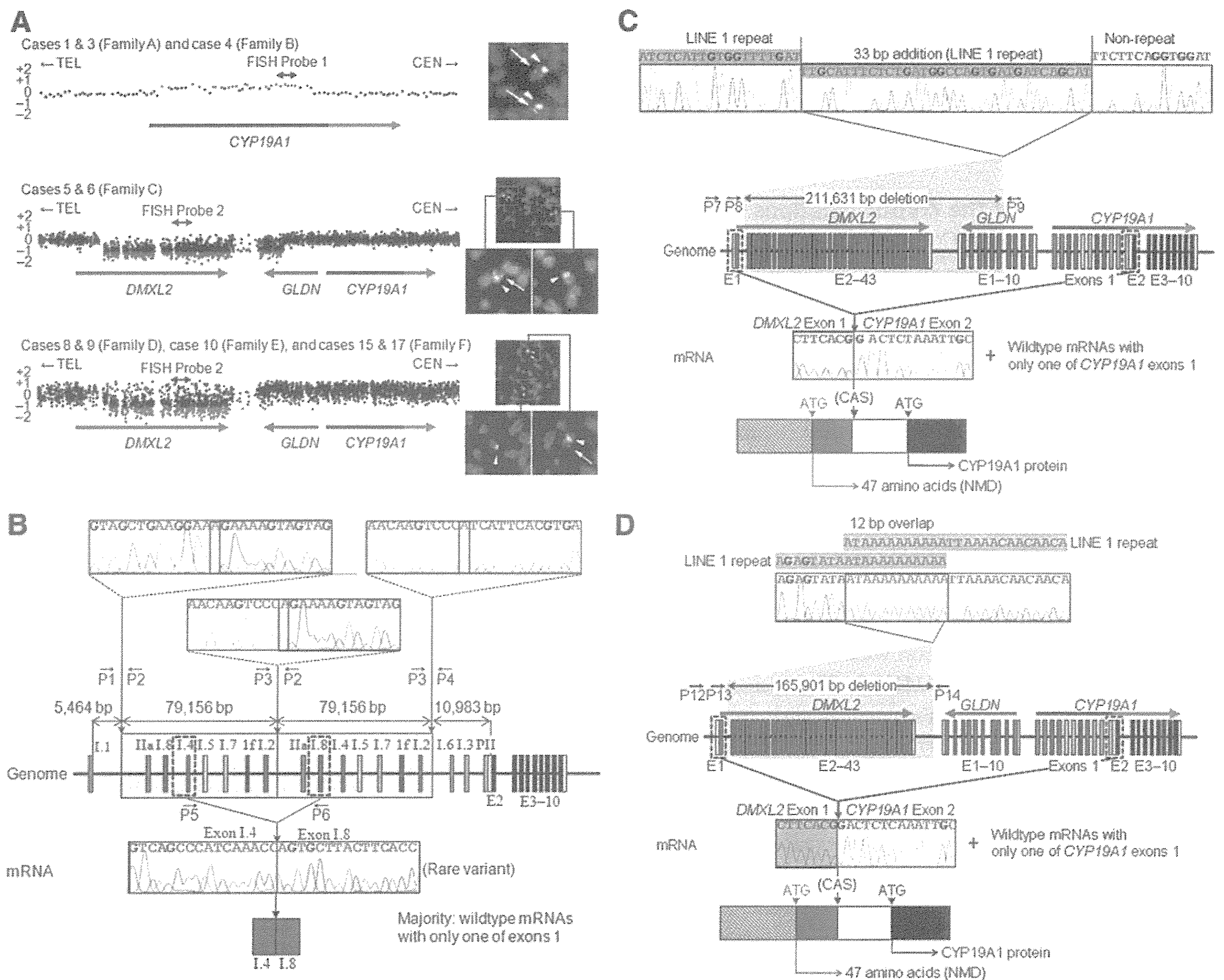

FIG. 3. Summary of molecular studies. For *CYP19A1*, the *dark* and *light blue lines* represent the genomic regions for noncoding exons 1 and coding exons 2–10, respectively. A, Oligoarray CGH and FISH analyses. In CGH analysis, the *black*, *red*, and *green dots* denote signals indicative of the normal, the increased (>+0.5), and the decreased (<−1.0) copy numbers, respectively. In FISH analysis, *two red signals* with an apparently different density are identified in cases from families A and B by FISH probe 1, whereas only a *single red signal* is found in cases from families C–F by FISH probe 2. The *green signals* are derived from the internal control probe. B, Schematic representation of the tandem duplication shared in common by cases 1 and 3 from family A and case 4 from family B. Genome, The junction sequence of the tandem duplication (*yellow boxes*) is shown, together with the original normal sequences at the 5′- and the 3′-ends of the duplicated region. The sequences highlighted with *light green* and *light orange* are identical, and 1 bp (A) is shared at the junction point (highlighted with *light yellow*). mRNA, The sequence of a rare clone is shown. The 3′-end of exon I.4 is connected with the 5′-end of exon I.8. C, Schematic representation of the deletion in sibling cases 5 and 6 from family C. Genome, The junction sequence of the deletion (*a gray area*) is shown. The fusion has occurred between a LINE 1 repeat sequence (*highlighted with blue*) at intron 1 of *DMXL2* and a nonrepeat sequence at intron 4 of *GLDN* and is accompanied by an addition of a 33-bp segment with a LINE 1 repeat sequence. mRNA, The sequence of a rare chimeric gene transcript is shown. *DMXL2* exon 1 consisting of a noncoding region (*a red striped box*) and a coding region (*a red box*) is spliced onto the common acceptor site (CAS) of *CYP19A1* exon 2 comprising an untranslated region (*a white box*) and a coding region (*a black box*). Thus, this transcript has two translation initiation codons (ATG), although the mRNA destined to produce a 47-amino acid protein from the ATG on *DMXL2* exon 1 is predicted to undergo NMD. D, Schematic representation of the deletion shared in common by cases 8 and 9 from family D, case 10 from family E, and cases 15 and 17 from family F. Genome, The junction sequence of the deletion (*a gray area*) is shown. The fusion has occurred between a LINE 1 repeat sequence (*highlighted with blue*) at intron 1 of *DMXL2* and that at a downstream region of *DMXL2*, with an overlap of a 12-bp segment. mRNA, The sequence of a chimeric gene transcript is delineated. The mRNA structure is the same as that described in the legend for Fig. 3C.

fusion occurred between nonrepeat elements with an overlap of one nucleotide.

All the 5′-RACE products (>500 clones) obtained from LCL and SF of case 3 (family A) and case 4 (family B) were found to be associated with a single exon 1, as observed in control materials. However, PCR amplifications for the 5′-RACE products with a variety of combinations of primers hybridizing to each exon 1 and subsequent sequencing of the PCR products revealed the presence of a chimeric clone consisting of exon I.4 at the 5′ side and exon I.8 at

the 3′ side in both LCL and SF (Fig. 3B). Although such a chimeric clone would have been produced by a splice error, this indicated that duplicated exon 1.4 at the distal nonphysiological position functioned as a transcription start site.

Characterization of the cryptic deletions

In cases from family C, long PCR products were obtained with the P7 primer and the P9 primer, and the deletion junction was determined by direct sequencing with the P8 primer (Fig. 3C). The deleted region was 211,631-bp long and involved exons 2–43 of *DMXL2* and exons 5–10 of *GLDN*. The two breakpoints resided within a LINE 1 repeat sequence and a nonrepeat sequence respectively, and a 33-bp segment with a LINE 1 repeat sequence was inserted to the fusion point. In cases from families D–F, long PCR products were obtained by sequential amplifications with the P12 primer and the P14 primer and with the P13 primer and the P14 primer, and an identical deletion was identified by direct sequencing with the P13 primer (Fig. 3D). The deletion was 165,901-bp long and involved exons 2–43 of *DMXL2*. The fusion occurred between two LINE 1 repeat sequences with an overlap of a 12-bp segment.

Sequence analysis of the 5′-RACE products obtained from LCL of cases 5 and 6 (family C) and from SF of case 9 (family D) and case 10 (family E) revealed the presence of a few clones with *DMXL2* exon 1 (2–5%), together with multiple clones with a single wild-type *CYP19A1* exon 1 (Fig. 3, C and D). Such a chimeric mRNA clone was absent from control materials. Furthermore, *DMXL2* exon 1 was found to be accompanied by a promoter-compatible sequence (Supplemental Fig. 2). This indicated a cryptic usage of *DMLX2* exon 1 as an alternative *CYP19A1* transcription start site in cases with deletions. Notably, because of the presence of the translation start codon on *DMXL2* exon 1, mRNAs of the *DMLX2/CYP19A1* chimeric genes are predicted to produce two proteins, *i.e.* CYP19A1 protein and an apparently nonfunctional 47-amino acid protein with a termination codon on *CYP19A1* exon 2, when the translation started from the initiation codons on *CYP19A1* exon 2 and on *DMLX2* exon 1, respectively. Furthermore, mRNA destined to yield the 47-amino acid protein is predicted to undergo nonsense-mediated mRNA decay (NMD) because it satisfies the condition for the occurrence of NMD (12).

Relative mRNA levels of *CYP19A1* and its neighboring genes

CYP19A1 showed a markedly high expression in the placenta and a relatively weak expression in a limited number of tissues including hypothalamus and ovary. By contrast, *DMXL2* was expressed in a range of tissues with some degree of variation as well as *CGNL1*, *MAPK6*, *TMOD3*, and *TLN2* (Fig. 4).

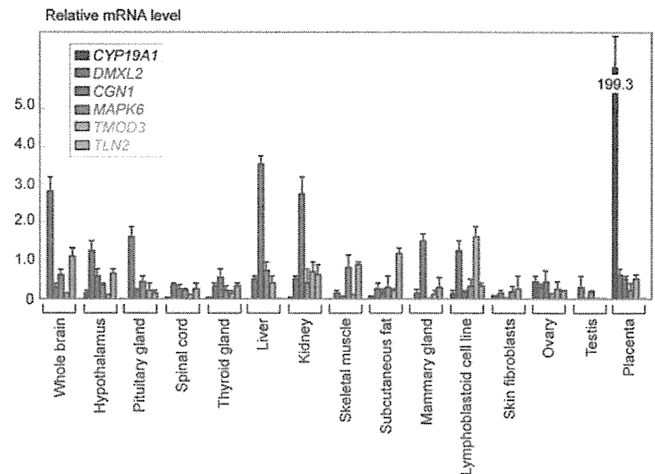

FIG. 4. Expression patterns of *CYP19A1* and the five neighboring genes involved in the chimeric gene formation. Relative mRNA levels against *TBP* are shown.

Discussion

We identified cryptic duplications of the *CYP19A1* promoter region and deletions of the *CYP19A1* upstream region in cases with AEXS. The tandem duplications would have caused *CYP19A1* overexpression because of an increased number of the wild-type transcription start sites. Indeed, because a rare mRNA variant with exon I.4 and exon I.8 was identified, this implies that duplicated exons 1 at the distal nonphysiological position can also function as transcription start sites. Similarly, the deletions would have caused *CYP19A1* overexpression because of a cryptic usage of *DMXL2* exon 1 with a putative promoter function as an extra transcription start site for *CYP19A1*. Indeed, because a few clones with *DMXL2* exon 1 and *CYP19A1* exon 2 were identified, this confirms the formation of a *DMXL2/CYP19A1* chimeric gene. Thus, our results suggest for the first time that duplications of a physiological promoter and deletions of an upstream region can cause overexpression of a corresponding gene and resultant human genetic disease.

Such cryptic genomic rearrangements can be generated by several mechanisms. The tandem duplication in families A and B would be formed by a replication-based mechanism of fork stalling and template switching that occurs in the absence of repeat sequences and is associated with microhomology (13). The deletion in family C is explained by nonhomologous end joining that takes place between nonhomologous sequences and is frequently accompanied by an insertion of a short segment at the fusion point (13).

The deletion in families D–F is compatible with a repeat sequence mediated nonallelic intrachromosomal or interchromosomal recombination (13). Thus, in conjunction with the previously identified four cryptic inversions that are also explainable by fork stalling and template switching or nonallelic recombination (8), genomic sequence around *CYP19A1* appears to harbor particular motifs that are vulnerable to replication and recombination errors.

To date, three types of cryptic genomic rearrangements have been identified in patients with AEXS, *i.e.* duplication type, deletion type (two subtypes), and inversion type (four subtypes) (Fig. 5). Here, although the deletion and the inversion types are associated with heterozygous impairment of neighboring genes (deletion or disconnection between noncoding exon(s) and coding exons), the phenotypes of patients are well explained by excessive *CYP19A1* activity alone. Thus, haploinsufficiency of these neighboring genes would not have a major clinical effect.

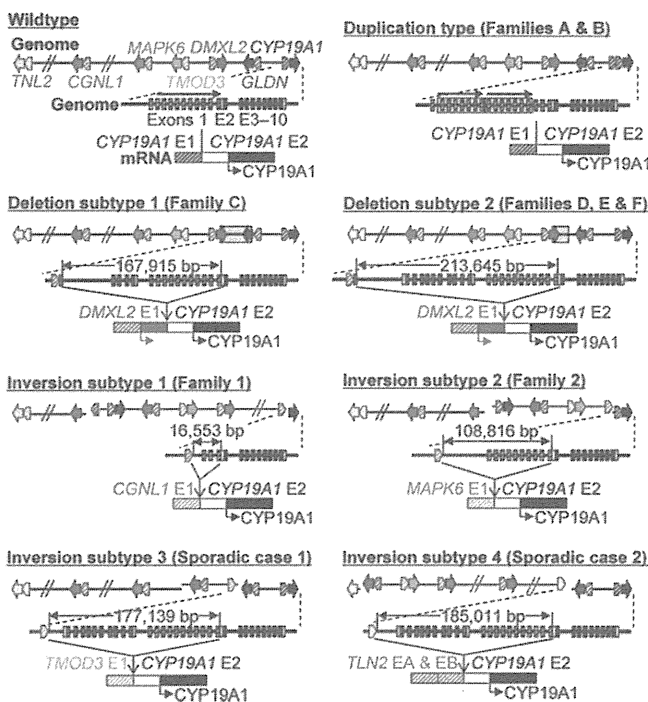

FIG. 5. Schematic representation of the rearranged genome and mRNA structures. The *white* and *black boxes* of *CYP19A1* exon 2 show untranslated region and coding region, respectively (for details, see Supplemental Fig. 1). For the duplication type and the deletion subtypes, see Fig. 3, C and D, for details. For genome, the *striped* and *painted arrows* indicate noncoding and coding exons, respectively (5'→3'). The inverted genomic regions are delineated in *blue lines*. For mRNA, *colored striped boxes* represent noncoding regions of each gene. For *TLN2*, exons A and B correspond to the previously reported exons 1 and 2 (8); because current exon 1 in the public database indicates the first coding exon, we have coined the terms exons A and B for the noncoding exons. The deletion and inversion types are associated with heterozygous impairment of neighboring genes [deletion or disconnection between noncoding exon(s) and the following coding exons]. The inversion subtype 1 is accompanied by inversion of eight of the 11 *CYP19A1* exons 1, and the inversion subtype 2 is associated with inversion of the placenta-specific *CYP19A1* exon I.1.

For the deletion and inversion types, two factors should be considered. One factor is expression patterns of each chimeric gene. In this regard, the five genes involved in the formation of chimeric genes are widely expressed, with some degree of variation (Fig. 4). Furthermore, *in silico* analysis revealed promoter-compatible sequences around exon 1 of *DMXL2*, *CGN1*, *MAPK6*, and *TMOD3* in multiple cell types, although such sequences remain to be identified for noncoding exons of *TLN2* (Supplemental Fig. 2). These findings imply that the chimeric genes show wide expression patterns because expression patterns of chimeric genes would follow those of the original genes.

The other factor is expression dosage of each chimeric gene. In this context, the *DMXL2/CYP19A1* chimeric mRNA was identified only in 2–5% of transcripts from SF, whereas the *CGNL1/CYP19A1* chimeric mRNA and the *TMOD3/CYP19A1* chimeric mRNA accounted for 89–100% and 80% of transcripts from SF, respectively (no data for the *MAPK6/CYP19A1* and the *TLN2/CYP19A1* chimeric genes) (5). This difference is obviously inexplicable by the relative expression level in SF that is grossly similar between *DMXL2* and *TMOD3* and is quite low for *CGNL1* (Fig. 4). In this regard, it is notable that a translation start codon and a following coding region are present on exon 1 of *DMXL2* (Fig. 5). It is likely that *DMXL2/CYP19A1* chimeric mRNA transcribed by the *DMXL2* promoter preferentially recognized the natural start codon on *DMXL2* exon 1 and underwent NMD and that rather exceptional chimeric mRNAs, which recognized the start codon on *CYP19A1* exon 2, were identified by 5'-RACE. By contrast, such a phenomenon would not be postulated for the inversion-mediated chimeric mRNA because of the absence of a translation start codon on the fused exon 1 of *CGNL1* and *TMOD3* (as well as exon 1 of *MAPK6* and exons A and B of *TLN2*) (Fig. 5). For the *CGNL1/CYP19A1* chimeric gene, furthermore, the physical distance between *CGNL1* exon 1 and *CYP19A1* exon 2 is short, and whereas a splice competition may be possible between exon 1 of neighboring genes and original *CYP19A1* exons 1, eight of 11 *CYP19A1* exons 1 including exon I.4 functioning as the major promoter in SF have been disconnected from *CYP19A1*-coding exons by inversion. These structural characters would have also contributed to the efficient splicing between *CGNL1* exon 1 and *CYP19A1* exon 2 (14). In this context, although the *CGNL1/CYP19A1* chimeric gene is associated with functional loss of eight *CYP19A1* exons 1 and the resultant reduction of *CYP19A1* expression in *CYP19A1*-expressing tissues, overall aromatase activity would be increased

by the wide expression of the chimeric gene. These structural properties would primarily explain the difference in the expression dosage of chimeric mRNA between the deletion and the inversion types.

It is inferred, therefore, that the duplication type simply increases *CYP19A1* transcription in native *CYP19A1*-expressing tissues, whereas the deletion and the inversion types cause relatively mild and severe *CYP19A1* overexpression in a range of tissues, respectively. These notions would grossly explain why clinical features of affected males and carrier females and endocrine profiles of affected males are apparently milder in the duplication and the deletion types than in the inversion type and why clinical findings were ameliorated with 1 mg/d of anastrozole in the duplication and the deletion types and with 2–4 mg/d of anastrozole in the inversion type. In addition, the different expression pattern between *CYP19A1* and *DMXL2* may explain, in terms of autocrine and/or paracrine effects, why phenotypic features such as gynecomastia tended to be more severe in the deletion type than in the duplication type under similar endocrine profiles.

Furthermore, several findings are notable in this study. First, a similar degree of FSH-dominant hypogonadotropic hypogonadism is present in the three types, with no amelioration of FSH responses to GnRH stimulation after GnRH priming in two cases with the duplication. This suggests that a relatively mild excess of circulatory estrogens, as observed in the duplication and the deletion types, can exert a strong negative feedback effect on FSH secretion, primarily at the pituitary, as has been suggested previously (15–19). Second, although basal T values appear to be mildly and similarly compromised in the three types, age-matched comparison suggests that T responses to hCG stimulation are apparently normal in the duplication and the deletion types and somewhat low in the inversion type. These data, although they remain fragmentary, would primarily be compatible with fairly preserved LH secretion in the three types and markedly increased estrogen values in the inversion type because T production is under the control of LH (1), and excessive estrogens compromise testicular steroidogenic enzyme activity (20, 21). Lastly, although testis volume appears somewhat small, fertility (spermatogenesis) is normally preserved in the three types. This would be consistent with the FSH-dominant hypogonadotropic hypogonadism because FSH plays only a minor role in male fertility (spermatogenesis) (22). Indeed, males with mutations of *FSHR* encoding FSH receptor as well as mice lacking *FSHB* or *FSHR* can be fertile (23, 24).

The results of this study are contrastive to those of the previous studies. In the previous studies, inversions only have been identified, and each inversion is specific to each family or patient (8). By contrast, in this study, the identical duplication was found in two Japanese families A and B, and the same deletion (subtype 2 in Fig. 5) was shared by three Japanese and one Caucasian families D–F, despite apparent nonconsanguinity. This may be explained by assuming that patients with severe phenotype were preferentially examined in the previous studies, whereas those with the AEXS phenotype were analyzed in this study without ascertainment bias. Furthermore, because phenotypes are milder in the duplication and the deletion types than in the inversion type, this may have permitted the spread of the duplication and the deletion types, but not the inversion type, as the founder abnormalities. This notion predicts that the duplication and the deletion types would be identified by examining patients with mild AEXS phenotype.

In summary, the present study shows that AEXS can be caused by duplications of the physiological promoters and microdeletions of the upstream regions of *CYP19A1* and that phenotypic severity is primarily determined by the tissue expression pattern of *CYP19A1* and the chimeric genes and by structural properties of the fused exons. Most importantly, the present study provides novel models for the gain-of-function mutations leading to human genetic disease.

Acknowledgments

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

Present address for T.O.: Department of Pediatrics, Hamamatsu University School of Medicine, Hamamatsu 431-3192, Japan.

This work was supported by Grants for Research on Intractable Diseases (H22-035 and H22-098) from the Ministry of Health, Labor, and Welfare; Grants-in-Aid for Scientific Research (B) (20390265) and (S) (22227002) from the Japan Society for the Promotion of Science; and Grant-in-Aid for Scientific Research on Innovative Areas (22132004) from the Ministry of Education, Culture, Sports, Science, and Technology.

Disclosure Summary: The authors have nothing to declare.

References

1. **Bhasin S** 2008 Testicular disorders. In: Kronenberg HM, Melmed M, Polonsky KS, Larsen PR, eds. Williams textbook of endocrinology. 11th ed. Philadelphia: Saunders; 645–699
2. **Bulun SE, Takayama K, Suzuki T, Sasano H, Yilmaz B, Sebastian S**

2004 Organization of the human aromatase p450 (CYP19) gene. Semin Reprod Med 22:5–9
3. **Demura M, Reierstad S, Innes JE, Bulun SE** 2008 Novel promoter I. 8 and promoter usage in the CYP19 (aromatase) gene. Reprod Sci 15:1044–1053
4. **Harada N, Utsumi T, Takagi Y** 1993 Tissue-specific expression of the human aromatase cytochrome P-450 gene by alternative use of multiple exons 1 and promoters, and switching of tissue-specific exons 1 in carcinogenesis. Proc Natl Acad Sci USA 90:11312–11316
5. **Shozu M, Sebastian S, Takayama K, Hsu WT, Schultz RA, Neely K, Bryant M, Bulun SE** 2003 Estrogen excess associated with novel gain-of-function mutations affecting the aromatase gene. N Engl J Med 348:1855–1865
6. **Binder G, Iliev DI, Dufke A, Wabitsch M, Schweizer R, Ranke MB, Schmidt M** 2005 Dominant transmission of prepubertal gynecomastia due to serum estrone excess: hormonal, biochemical, and genetic analysis in a large kindred. J Clin Endocrinol Metab 90: 484–492
7. **Martin RM, Lin CJ, Nishi MY, Billerbeck AE, Latronico AC, Russell DW, Mendonca BB** 2003 Familial hyperestrogenism in both sexes: clinical, hormonal, and molecular studies of two siblings. J Clin Endocrinol Metab 88:3027–3034
8. **Demura M, Martin RM, Shozu M, Sebastian S, Takayama K, Hsu WT, Schultz RA, Neely K, Bryant M, Mendonca BB, Hanaki K, Kanzaki S, Rhoads DB, Misra M, Bulun SE** 2007 Regional rearrangements in chromosome 15q21 cause formation of cryptic promoters for the CYP19 (aromatase) gene. Hum Mol Genet 16:2529–2541
9. **Tiulpakov A, Kalintchenko N, Semitcheva T, Polyakov A, Dedov I, Sverdlova P, Kolesnikova G, Peterkova V, Rubtsov P** 2005 A potential rearrangement between CYP19 and TRPM7 genes on chromosome 15q21.2 as a cause of aromatase excess syndrome. J Clin Endocrinol Metab 90:4184–4190
10. **Stratakis CA, Vottero A, Brodie A, Kirschner LS, DeAtkine D, Lu Q, Yue W, Mitsiades CS, Flor AW, Chrousos GP** 1998 The aromatase excess syndrome is associated with feminization of both sexes and autosomal dominant transmission of aberrant P450 aromatase gene transcription. J Clin Endocrinol Metab 83:1348–1357
11. **Bellino FL, Osawa Y** 1977 Localization of estrogen synthetase in the chorionic villus fraction after homogenization of human term placenta. J Clin Endocrinol Metab 44:699–707
12. **Kuzmiak HA, Maquat LE** 2006 Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges. Trends Mol Med 12:306–316
13. **Gu W, Zhang F, Lupski JR** 2008 Mechanisms for human genomic rearrangements. Pathogenetics 1:4
14. **Castillo-Davis CI, Mekhedov SL, Hartl DL, Koonin EV, Kondrashov FA** 2002 Selection for short introns in highly expressed genes. Nat Genet 31:415–418
15. **Shaw ND, Histed SN, Srouji SS, Yang J, Lee H, Hall JE** 2010 Estrogen negative feedback on gonadotropin secretion: evidence for a direct pituitary effect in women. J Clin Endocrinol Metab 95:1955–1961
16. **Belgorosky A, Guercio G, Pepe C, Saraco N, Rivarola MA** 2009 Genetic and clinical spectrum of aromatase deficiency in infancy, childhood and adolescence. Horm Res 72:321–330
17. **Alexander DC, Miller WL** 1982 Regulation of ovine follicle-stimulating hormone β-chain mRNA by 17β-estradiol *in vivo* and *in vitro*. J Biol Chem 257:2282–2286
18. **Mercer JE, Clements JA, Funder JW, Clarke IJ** 1988 Luteinizing hormone-β mRNA levels are regulated primarily by gonadotropin-releasing hormone and not by negative estrogen feedback on the pituitary. Neuroendocrinology 47:563–566
19. **Raven G, de Jong FH, Kaufman JM, de Ronde W** 2006 In men, peripheral estradiol levels directly reflect the action of estrogens at the hypothalamo-pituitary level to inhibit gonadotropin secretion. J Clin Endocrinol Metab 91:3324–3328
20. **Moger WH** 1980 Direct effects of estrogens on the endocrine function of the mammalian testis. Can J Physiol Pharmacol 58:1011–1022
21. **Strauss L, Kallio J, Desai N, Pakarinen P, Miettinen T, Gylling H, Albrecht M, Mäkelä S, Mayerhofer A, Poutanen M** 2009 Increased exposure to estrogens disturbs maturation, steroidogenesis, and cholesterol homeostasis via estrogen receptor α in adult mouse Leydig cells. Endocrinology 150:2865–2872
22. **Kumar TR, Wang Y, Lu N, Matzuk MM** 1997 Follicle stimulating hormone is required for ovarian follicle maturation but not male fertility. Nat Genet 15:201–204
23. **Tapanainen JS, Aittomäki K, Min J, Vaskivuo T, Huhtaniemi IT** 1997 Men homozygous for an inactivating mutation of the follicle-stimulating hormone (FSH) receptor gene present variable suppression of spermatogenesis and fertility. Nat Genet 15:205–206
24. **Layman LC, McDonough PG** 2000 Mutations of follicle stimulating hormone-β and its receptor in human and mouse: genotype/phenotype. Mol Cell Endocrinol 161:9–17