41. Sybert, V.P., Dale, B.A., and Holbrook, K.A. (1988). Palmar-plantar keratoderma. A clinical, ultrastructural, and biochemical study. J. Am. Acad. Dermatol. *18*, 75–86.
42. Hatsell, S.J., Eady, R.A., Wennerstrand, L., Dopping-Hepenstal, P., Leigh, I.M., Munro, C., and Kelsell, D.P. (2001). Novel splice site mutation in keratin 1 underlies mild epidermolytic palmoplantar keratoderma in three kindreds. J. Invest. Dermatol. *116*, 606–609.
43. Reis, A., Hennies, H.C., Langbein, L., Digweed, M., Mischke, D., Drechsler, M., Schröck, E., Royer-Pokora, B., Franke, W.W., Sperling, K., et al. (1994). Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK). Nat. Genet. *6*, 174–179.
44. Küster, W., Reis, A., and Hennies, H.C. (2002). Epidermolytic palmoplantar keratoderma of Vörner: re-evaluation of Vörner's original family and identification of a novel keratin 9 mutation. Arch. Dermatol. Res. *294*, 268–272.
45. Kimonis, V., DiGiovanna, J.J., Yang, J.M., Doyle, S.Z., Bale, S.J., and Compton, J.G. (1994). A mutation in the V1 end domain of keratin 1 in non-epidermolytic palmar-plantar keratoderma. J. Invest. Dermatol. *103*, 764–769.
46. Gach, J.E., Munro, C.S., Lane, E.B., Wilson, N.J., and Moss, C. (2005). Two families with Greither's syndrome caused by a keratin 1 mutation. J. Am. Acad. Dermatol. *53*(*Suppl 1*), S225–S230.
47. Covello, S.P., Irvine, A.D., McKenna, K.E., Munro, C.S., Nevin, N.C., Smith, F.J., Uitto, J., and McLean, W.H. (1998). Mutations in keratin K9 in kindreds with epidermolytic palmoplantar keratoderma and epidemiology in Northern Ireland. J. Invest. Dermatol. *111*, 1207–1209.
48. Hamm, H., Happle, R., Butterfass, T., and Traupe, H. (1988). Epidermolytic palmoplantar keratoderma of Vörner: is it the most frequent type of hereditary palmoplantar keratoderma? Dermatologica *177*, 138–145.
49. Küster, W., and Becker, A. (1992). Indication for the identity of palmoplantar keratoderma type Unna-Thost with type Vörner. Thost's family revisited 110 years later. Acta Derm. Venereol. *72*, 120–122.
50. Bouadjar, B., Benmazouzia, S., Prud'homme, J.F., Cure, S., and Fischer, J. (2000). Clinical and genetic studies of 3 large, consanguineous, Algerian families with Mal de Meleda. Arch. Dermatol. *136*, 1247–1252.

Genomic Basis of Aromatase Excess Syndrome: Recombination- and Replication-Mediated Rearrangements Leading to *CYP19A1* Overexpression

Maki Fukami, Takayoshi Tsuchiya, Heike Vollbach, Kristy A. Brown, Shuji Abe, Shigeyuki Ohtsu, Martin Wabitsch, Henry Burger, Evan R. Simpson, Akihiro Umezawa, Daizou Shihara, Kazuhiko Nakabayashi, Serdar E. Bulun, Makio Shozu, and Tsutomu Ogata*

Context: Genomic rearrangements at 15q21 have been shown to cause overexpression of *CYP19A1* and resultant aromatase excess syndrome (AEXS). However, mutation spectrum, clinical consequences, and underlying mechanisms of these rearrangements remain to be elucidated.

Objective: The aim of the study was to clarify such unsolved matters.

Design, Setting, and Methods: We characterized six new rearrangements and investigated clinical outcome and local genomic environments of these rearrangements and of three previously reported duplications/deletions.

Results: Novel rearrangements included simple duplication involving exons 1–10 of *CYP19A1* and simple and complex rearrangements that presumably generated chimeric genes consisting of the coding region of *CYP19A1* and promoter-associated exons of neighboring genes. Clinical severities were primarily determined by the copy number of *CYP19A1* and the property of the fused promoters. Sequences at the fusion junctions suggested nonallelic homologous recombination, nonhomologous end-joining, and replication-based errors as the underlying mechanisms. The breakpoint-flanking regions were not enriched with GC content, palindromes, noncanonical DNA structures, or known rearrangement-associated motifs. The rearrangements resided in early-replicating segments.

Conclusions: These results indicate that AEXS is caused by duplications involving *CYP19A1* and simple and complex rearrangements that presumably lead to the usage of cryptic promoters of several neighboring genes. Our data support the notion that phenotypes depend on the dosage of *CYP19A1* and the characteristics of the fused promoters. Furthermore, we show that the rearrangements in AEXS are generated by both recombination- and replication-mediated mechanisms, independent of the known rearrangement-inducing DNA features or late-replication timing. Thus, AEXS represents a unique model for human genomic disorders. (***J Clin Endocrinol Metab* 98: E2013–E2021, 2013**)

Aromatase excess syndrome (AEXS; MIM no. 139300) is a rare autosomal dominant disorder that causes prepubertal- or peripubertal-onset gynecomastia, hypogonadotropic hypogonadism, advanced bone age, and short adult height in male patients (1, 2). Female patients are usually asymptomatic, although macromastia, irregular menses, and short stature have been reported in a few individuals (2). AEXS results from excessive expression of the aromatase gene *CYP19A1* on chromosome 15q21.2 (NM_000103) (1). *CYP19A1* comprises 11 noncoding exons 1 that function as tissue-specific promoters (exons I.1, IIa, I.8, I.4, I.5, I.7, 1f, I.2, I.6, I.3, and PII), and nine coding exons (exons 2–10) (3, 4). We and other groups have identified various chromosomal rearrangements at

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

Received June 13, 2013. Accepted September 19, 2013.
First Published Online September 24, 2013

For editorial see page 4676

* Author affiliations are shown at the bottom of the next page.

Abbreviations: AEXS, aromatase excess syndrome; CGH, comparative genomic hybridization.

15q21 in patients with AEXS (1, 2, 5). These rearrangements included duplications that encompassed seven of the 11 noncoding exons 1 of *CYP19A1* and deletions and inversions that generated chimeric genes consisting of coding exons of *CYP19A1* and promoter-associated exons of neighboring genes. Genotype-phenotype analysis has indicated that clinical severities primarily depend on the functional properties of the fused promoters. These findings provide a novel example of gain-of-function mutations resulting from submicroscopic genomic rearrangements.

Rearrangements in the human genome are known to be generated by recombination-based mechanisms, namely, nonallelic homologous recombination and nonhomologous end-joining, and by replication-based mechanisms (6–9). Of these, nonallelic homologous recombination results from unequal crossover between two homologous sequences, usually on the same but sometimes on different chromosomes (10). Nonallelic homologous recombination accounts for most of the recurrent simple deletions and duplications in the human genome and represents the most common abnormality involved in human genomic disorders (9–11). Nonhomologous end-joining occurs as a result of double-strand DNA breakage and subsequent ligation of the two broken DNA ends (12). Nonhomologous end-joining often underlies nonrecurrent simple deletions associated with short nucleotide stretches at the fusion junctions (9–12). Replication-based mechanisms are caused by aberrant template switching during replication and can produce both simple and complex rearrangements that carry microhomologies at the fusion junctions (8, 9, 13). Previous studies have indicated that nonallelic homologous recombination, nonhomologous end-joining, and replication-based mechanisms are facilitated by various local DNA features including high GC content and palindromes (10, 14–16). Highly similar sequences widely spread in the genome ("repetitive elements"), such as *Alu*, LINE1, and MIR, can mediate the occurrence of genomic rearrangements (12). Non-B structures, ie, DNA conformations that differ from the canonical Watson-Crick right-handed double helix, and specific short sequence motifs and tri/tetranucleotides have also been suggested as local genomic stimulants (14–22). Furthermore, replication timing of each chromosomal region appears to determine the frequency of rearrangements; nonallelic homologous recombination preferentially occurs in DNA segments that replicate in early S phase (early-replicating segments), whereas nonhomologous end-joining and replication-based errors frequently appear in late-replicating segments (23).

At present, the underlying mechanisms of the AEXS-associated rearrangements remain largely unknown. Although sequence analysis of the fusion junctions has indicated that nonallelic homologous recombination and nonhomologous end-joining—and possibly replication-based mechanisms as well—are involved in the formation of simple duplications and deletions in AEXS (5), the molecular basis of inversions remains to be determined. Here, we characterized the fine genomic structures of six rearrangements involved in AEXS. Furthermore, we investigated clinical consequences and local genomic environments of the six rearrangements and of three previously reported duplications/deletions.

Patients and Methods

Patients

This study consisted of six cases (cases 1–6) ascertained by prepubertal- or peripubertal-onset gynecomastia. Clinical findings of cases 1–6 are summarized in Table 1. Cases 1–4 are hitherto unreported. Cases 5 and 6 have been described previously, although the genomic structure remains to be determined (1, 2). Cases 1–3 and 5–6 had a 46,XY karyotype, whereas case 4 had a 46,XY inv (9) karyotype that is known as a normal variant. Case 2 had a brother with prepubertal-onset gynecomastia, a sister with premature thelarche, and a father and several paternal relatives with advanced bone age and/or short stature. Case 6 had a son with prepubertal-onset gynecomastia. There was no family history of AEXS in the remaining cases. This study was approved by the Institutional Review Board Committee at the National Center for Child Health and Development. Written informed consent was obtained from the patients and/or their parents.

Copy-number analyses

Leukocyte genomic DNA samples were obtained from cases 1–6, the parents and siblings of case 2, and the son of case 6. Genomic abnormalities involving *CYP19A1* exons and/or its flanking regions were examined by comparative genomic hybridization (CGH) using a custom-made oligoarray or a catalog human array (4 × 180K format, ID 030700 or G4449A; Agilent Technologies). The procedures were performed as described previously (5).

Department of Molecular Endocrinology (M.F., T.T., D.S., T.O.), National Research Institute for Child Health and Development, 157-8535 Tokyo, Japan; Department of Pediatrics (T.T.), Dokkyo Medical University Koshigaya Hospital, 343-8555 Koshigaya, Japan; Department of Pediatrics and Adolescent Medicine (H.V., M.W.), University Medical Center Ulm, 89081 Ulm, Germany; Metabolism and Cancer Laboratory (K.A.B., H.B., E.R.S.), Prince Henry's Institute, Monash Medical Centre, Clayton, 3168 VIC, Australia; Department of Pediatrics (S.A.), Hakodate Goryoukaku Hospital, 040-8611 Hakodate, Japan; Department of Pediatrics (S.O.), Kitasato University School of Medicine, 252-0375 Kanagawa, Japan; Department of Reproductive Biology (A.U.), Center for Regenerative Medicine, National Institute for Child Health and Development, 157-8535 Tokyo, Japan; Department of Maternal-Fetal Biology (K.N.), National Research Institute for Child Health and Development, 157-8535 Tokyo, Japan; Division of Reproductive Biology Research (S.E.B.), Department of Obstetrics and Gynecology, Feinberg School of Medicine, Northwestern University, Chicago, 60611 Illinois; Department of Reproductive Medicine (M.S.), Graduate School of Medicine, Chiba University, 260-8670 Chiba, Japan; and Department of Pediatrics (T.O.), Hamamatsu University School of Medicine, 431-3192 Hamamatsu, Japan

Table 1. Phenotypic and Endocrine Findings of Cases 1–6

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Genomic rearrangement	Duplication	Deletion	Complex	Complex	Complex	Complex
Age at examination, y	10	8 (18)[a]	15	13	17	36
Phenotypic findings						
Gynecomastia (Tanner stage)	**2–3**	**3**	**4–5**	**3–4**	**Severe**	**Severe**
Onset of gynecomastia, y	**7**	Unknown	**8**	**11**	**7**	**5**
Mastectomy	No	**Yes**	**Yes**	No	**Yes**	**Yes**
Testis, mL	6	N.E.	15	12	Normal	Normal
Pubic hair (Tanner stage)	None	None	3–4	4	N.D.	Normal
Facial hair	None	None	None	Scarce	Scarce	**None**
Final height	Unknown	Unknown	−0.9 SD	Unknown	**<1%ile**	**<1%ile**
Bone age, y[b]	**13.0**	**13.5**	N.E.	**18.0**	N.E.	N.E.
Fertility (spermatogenesis)	Unknown	Unknown	Yes	Unknown	Unknown	Yes
Endocrine findings[c]						
At diagnosis						
LH, mIU/mL	<0.1 (0.4–1.6) → 0.4 (10.9–20.6)[d]		2.4 (1.6–3.5)	1.3 (1.6–3.5) → 24.9 (21.7–39.5)[d]	4.3 (1.4–9.2)	1.7 (1.4–9.2)
FSH, mIU/mL	0.3 (1.7–4.2) → 1.6 (4.6–10.8)[d]		<1.0 (4.2–8.2)	0.6 (4.2–8.2) → 2.1 (11.2–17.3)[d]	2.7 (2.0–8.3)	1.5 (4.2–8.2)
T, ng/mL	0.06 (0.4–1.1) → 3.6 (>2.0)[e]	2.6 (2.8–7.0)	0.7 (2.8–7.0)	1.5 (2.8–7.0)	2.3 (2.8–7.0)	3.2 (2.8–7.0)
E_1, pg/mL				**111** (14–50)	**556** (15–32)	**903** (15–32)
E_2, pg/mL	**14** (<10)	**65** (10–35)	**406** (15–50)	**43** (2–30)	**392** (10–35)	**223** (10–35)
On AI treatment						
LH, mIU/mL	0.5 (0.4–1.6) → 7.3 (10.9–20.6)[d]	**44.8** (0.7–5.7)[f]	**4.7** (1.6–3.5)		8.9 (1.4–9.2)	2.9 (1.4–9.2)
FSH, mIU/mL	1.7 (1.7–4.2)→ 3.2 (4.6–10.8)[d]	**34.9** (2.0–8.3)[f]	2.5 (4.2–8.2)		5.6 (2.0–8.3)	5.6 (4.2–8.2)
T, ng/mL	0.9 (0.4–1.1)	**8.6** (2.8–7.0)	6.9 (2.8–7.0)		5.3 (2.8–7.0)	**10.7** (2.8–7.0)
E_1, pg/mL					**89** (15–32)	27 (15–32)
E_2, pg/mL	<10 (<10)	6 (10–35)	13 (15–50)		**59** (10–35)	**68** (10–35)
Reference	Present study	Present study	Present study	Present study	Ref. 1	Ref. 1

Abbreviations: AI, aromatase inhibitor; E_1, estrone; E_2, estradiol; N.D., not described; N.E., not examined. Abnormal clinical findings are boldfaced. Hormone values below the reference range (shown in parentheses) are underlined, and those above the reference range are boldfaced. Conversion factors to the SI unit: LH, 1.0 (IU/L); FSH, 1.0 (IU/L); E_1, 3.699 (pmol/L); E_2, 3.671 (pmol/L); and T, 3.467 (nmol/L).

[a] Physical examination and endocrine studies were carried out at 8 and 18 years of age, respectively.

[b] Assessed by the Tanner-Whitehouse 2 method standardized for Japanese or by the Greulich-Pyle method constructed for Caucasians.

[c] Evaluated by age-matched male reference data.

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

[e] Human chorionic gonadotropin stimulation tests (3000 IU/m^2, maximum 5000 IU im for 3 consecutive days; blood sampling on d 1 and 4).

[f] Increased levels of LH and FSH during AI treatment may be associated with low E_2 levels (24).

Characterization of the genomic structures of rearrangements

Breakpoints of the rearrangements were determined by direct sequencing of the PCR-amplified DNA fragments harboring the fusion junctions. PCRs were carried out using a number of primer pairs for various genomic positions around *CYP19A1*. The sequences of the primers utilized in the present study are available upon request. To confirm the formation of a chimeric gene in a case with a complex rearrangement, we performed RT-PCR using leukocyte mRNA and primers annealing to exon 2 of *CYP19A1* and exons of neighboring genes. The presence or absence of promoter-associated histone marks in the fused exons was analyzed using the UCSC genome browser (http://genome.ucsc.edu/).

Genotype-phenotype analysis

We performed genotype-phenotype analyses in cases 1–6 and in 18 patients identified in our previous study (5).

DNA sequences at the fusion junctions

To clarify the underlying mechanisms of the rearrangements, we examined the presence or absence of microhomologies and short nucleotide stretches at the fusion junctions. In addition, we searched for repeat elements around the breakpoints using Repeatmasker (http://www.repeatmasker.org).

Genomic environments around the breakpoints

We studied the frequencies of known rearrangement-inducing DNA features in the breakpoint-flanking regions. In silico analyses were carried out in the 300-bp regions at the proximal and distal sides of each breakpoint. We also examined control regions (n = 53) randomly selected at an interval of 1.5 Mb from the entire 15q (Supplemental Table 1, published on The Endocrine Society's Journals Online web site at http://jcem.endojournals.org). We calculated the average GC content using GEECEE (http://emboss.bioinformatics.nl/cgi-bin/emboss/geecee) and searched for palindromes using PALINDROME (http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::palindrome) and Non-B structures using Non-B DB (http://nonb.abcc.ncifcrf.gov). Examined Non-B structures included direct repeats, inverted repeats (cruciforms), mirror repeats, A-phased repeats, G-quadruplex repeats, short tandem repeats, and Z-DNA motifs (17). The presence or absence of the 10 specific sequence motifs and two tri/tetranucleotides implicated in rearrangements in various chromosomal regions (14, 18–22) were analyzed using Fuzznuc (http://emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc).

Replication timing analysis

We analyzed whether the rearrangements at 15q21 have occurred at a specific timing of S phase (23). Replication timing profiles of the approximately 10-Mb genomic interval around *CYP19A1* were evaluated using 92 datasets currently available in

Table 2. Genomic Rearrangements in Cases 1–6

	Rearrangement	Genomic Abnormality	Affected Genes[a]
Case 1	Simple	Simple duplication	*CYP19A1, TNFAIP8L3, AP4E1*
Case 2	Simple	Simple deletion	*CYP19A1, GLDN, DMXL2*
Case 3[b]	Complex	Multiple deletions?	*TMOD3, GLDN, DMXL2?*
Case 4	Complex	Multiple duplications and inversion	*CYP19A1, GLDN, SEMA6D*
Case 5	Complex	Multiple duplications, deletion, and inversion	*TMOD3, DMXL2, TMOD2, LYSMD2, SCG3*
Case 6	Complex	Multiple deletions and inversion	*CGNL1, CYP19A1*

[a] Genes involved in the deletion or duplication. Genes affected by copy-number-neutral inversions are not shown.

[b] Genomic structure of the rearrangement in case 3 remains to be characterized.

the ReplicationDomain database (http://www.replicationdomain.com/replication_timing.php).

Statistical analyses

Statistical significance of the average GC content between the breakpoint-flanking and control regions was analyzed by Student's *t* test. Differences in the frequencies of other rearrangement-inducing DNA features were examined by Fisher's exact probability test. $P < .05$ was considered significant.

Results

Copy-number alterations in cases 1–6

CGH analyses indicated heterozygous genomic rearrangements involving *CYP19A1* and/or its neighboring genes; ie, an approximately 0.4-Mb duplication involving *CYP19A1*, *TNFAIP8L3*, and *AP4E1* in case 1; an approximately 0.3-Mb deletion affecting *DMXL2*, *CYP19A1*, and *GLDN* in case 2; an approximately 80-kb deletion involving *TMOD3* and an approximately 250-kb deletion involving *DMXL2* and *GLDN* in case 3; an approximately 130-kb duplication involving *GLDN* and *CYP19A1* and an approximately 340-kb duplication involving *SEMA6D* at a position of approximately 3.6 Mb distant from *CYP19A1* in case 4; an approximately 370-kb duplication involving *TMOD3*, *TMOD2*, *LYSMD2*, *SCG3*, and *DMXL2*, and a 3- to 35-kb deletion between *DMXL2* and *GLDN* in case 5; and an approximately 3.5-kb deletion in the promoter region of *CYP19A1* in case 6 (Table 2 and Figure 1). The deletion in case 5 could not be narrowed down because of the absence of CGH probes around the breakpoints. The father and siblings of case 2 and the son of case 6 carried the same abnormalities as the probands.

Genomic structures of six rearrangements

We were able to characterize all fusion junctions in cases 1, 2, and 6 and one of the multiple junctions in cases 4 and 5 (Table 3, Supplemental Table 2, and Figure 2). The remaining breakpoints could not be determined due to the low quality of the DNA samples, the presence of long repetitive sequences around the breakpoints, or the complex structures of the rearrangements. In case 1, we identified a 387 622-bp tandem duplication involving six of the 11 exons 1 (exons I.7, 1f, I.2, I.6, I.3, and PII) and all coding exons of *CYP19A1*, together with all exons of *TNFAIP8L3* and *AP4E1*. In case 2, we detected a 303 624-bp deletion involving six of the *CYP19A1* exons 1 (exons I.1, IIa, I.8, I.4, I.5, and I.7), all exons of *GLDN*, and *DMXL2* exons 2–43. In case 4, we identified two duplications: an approximately 130-kb duplication encompassing all noncoding exons 1 and coding exons 2–3 of *CYP19A1* and *GLDN* exon 1, and an approximately 340-kb duplication involving *SEMA6D* exons 1–3. PCR products were obtained with a primer pair for *GLDN* intron 1 and *SEMA6D* intron 3 (P5 and P6 in Figure 2A), indicating that the approximately 3.6-Mb genomic interval harboring *GLDN* exon 1, all noncoding and coding exons of *CYP19A1*, and *SEMA6D* exons 4–20 was inverted. In addition, we analyzed mRNA of case 4 and detected a chimeric clone composed of *CYP19A1* exon 2 and *SEMA6D* noncoding exon 3 (Supplemental Figure 1). Thus, although we could not determine the fusion junctions of the duplication, these data imply that the rearrangement was caused by an inversion of an approximately 3.6-Mb region and a duplication of the telomeric part of the inverted DNA fragment. In case 5, we identified an approximately 370-kb duplication containing *TMOD3* exon 1, *DMXL2* exons 1–29, and all exons of *TMOD2*, *LYSMD2*, and *SCG3*. PCR products were obtained with a primer pair for *TMOD3* intron 1 and the downstream region of *GLDN* (P7 and P8), indicating that the approximately 370-kb region was duplicated and inserted into the genome in a reverse direction. The small deletion between *DMXL2* and *GLDN* detected by CGH could not be characterized because of the presence of long repetitive sequences around the breakpoints. In case 6, we identified a complex deletion–inversion–deletion rearrangement: a 202-bp deletion within *CGNL1* intron 1, an approximately 6.1-Mb inversion encompassing *CGNL1* exon 1, eight of the *CYP19A1* exons 1 (exons I.1, IIa, I.8, I.4, I.5, I.7, 1f, and I.2), and ≥25 genes, and a 3476-bp deletion within *CYP19A1* intron 1.

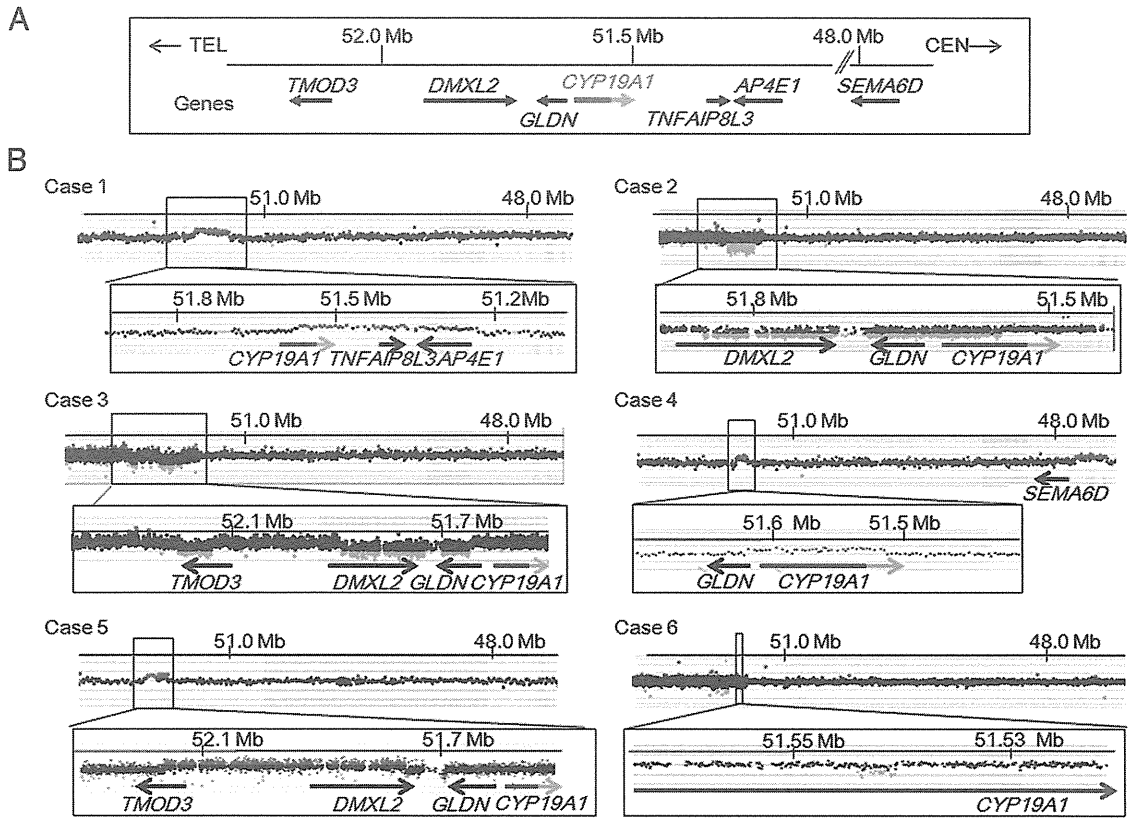

Figure 1. Copy-number analyses in cases 1–6. A, Schematic representation of the normal genomic structure around *CYP19A1*. The arrows indicate genomic positions and transcriptional direction of genes (5′→3′). For *CYP19A1*, the dark and light blue lines denote the genomic regions for noncoding exons 1 and coding exons 2–10, respectively. Genomic positions refer to Human Genome Database (hg19, build 37). Only genes around the fusion junctions are shown. B, CGH analyses in the six cases. The black, red, and green dots denote signals indicative of the normal, increased (>+0.5) and decreased (<−1.0) copy-numbers, respectively.

Phenotypic consequences of the six new and three previously reported rearrangements

We studied genotype-phenotype correlation in cases 1–6 and 18 previously reported patients (four patients from families A–B with simple duplications involving the *CYP19A1* promoter region, and 14 patients from families C–F with *DMXL2-CYP19A1* chimeric genes) (5). The results are summarized in Table 4. First, clinical severities were relatively mild in case 1 and patients from families A–B with simple duplications, obviously severe in cases 5 and 6 with complex rearrangements, and moderate in the remaining cases/patients with simple deletions or complex rearrangements. Second, among cases/patients with simple duplications, case 1 showed earlier onset of gyneco-

Table 3. Fusion Junctions in Cases 1–6

	No. of Fusion Junctions	No. of Fusion Junctions Characterized in This Study[a]	Sequences at the Fusion Junctions		
			Microhomology	Nucleotide Stretch	Predicted Mechanism
Case 1	1	1	Yes (4 bp)	Yes (2 bp)	RBM
Case 2	1	1	Yes (2 bp)	No	RBM
Case 3[b]	2?	0	Unknown	Unknown	RBM?
Case 4	5	1	Yes (20 bp)	No	RBM
Case 5	3	1	Yes (3 bp)	Yes (5 bp)	RBM
Case 6	2	2	Yes (3 bp)[c]/No	No	RBM

Abbreviation: RBM, replication-based mechanism.

[a] Several breakpoints could not be determined due to low quality of the DNA samples, the presence of long repetitive sequences around the breakpoints, or the complex structures of the rearrangements.

[b] Genomic structure of the rearrangement in case 3 remains to be characterized.

[c] Microhomology was observed at the telomeric junction.

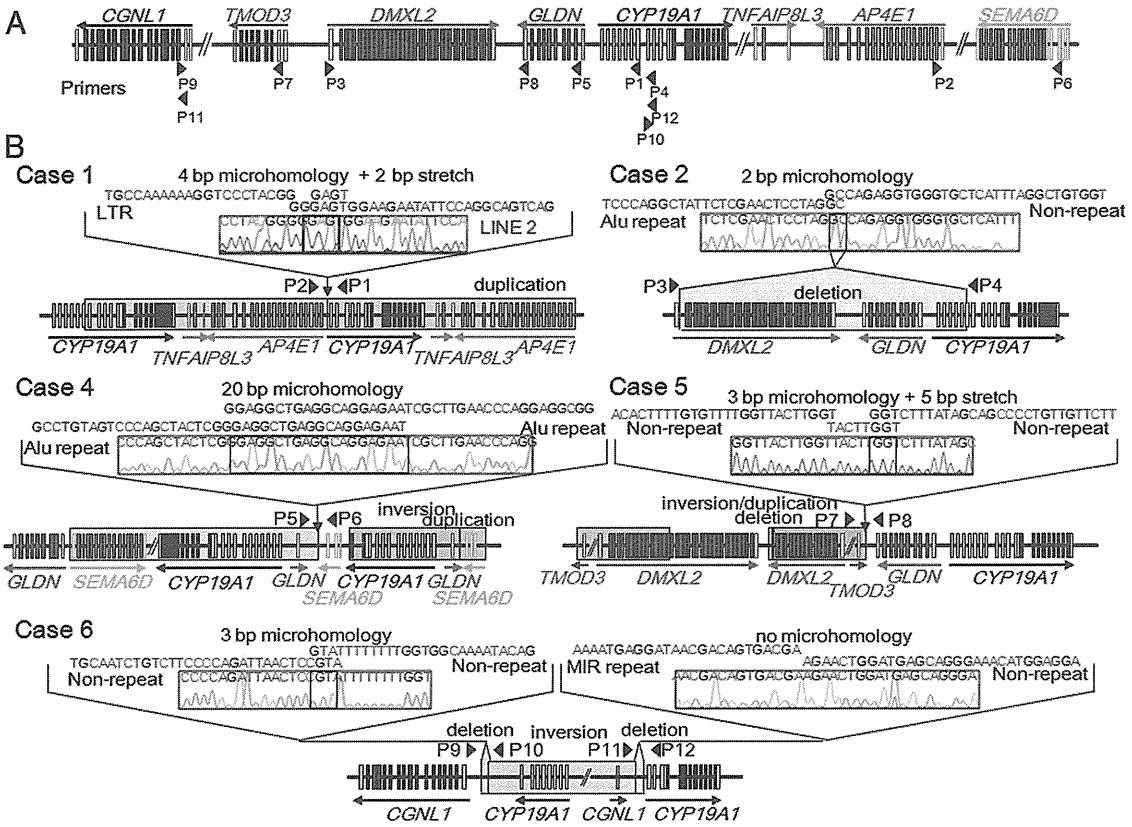

Figure 2. Fine genomic structures of the rearrangements. A, Schematic representation of the normal genomic structure. Arrowheads indicate the positions and the directions (5′→3′) of PCR primers utilized in this study (P1–P12). The open and color-painted boxes denote noncoding and coding exons, respectively. The sizes of the exons, introns, and primers are not drawn to scale. B, Schematic representation of the rearrangements and the DNA sequences at the fusion junctions. The red, blue, and green areas indicate duplications, deletions, and inversions, respectively. P1–P12 indicate the same PCR primers as shown in panel A. The fusion junctions of case 3 were not characterized. For case 4, the precise genomic position of the duplication remains to be clarified.

mastia and more severely advanced bone age than patients from families A–B. Third, among cases/patients with deletions, case 2 manifested milder gynecomastia than case 3 and patients from families C–F. Lastly, among cases/patients with deletions or complex rearrangements, cases 2–4 and patients from families C–F showed milder phenotypes than cases 5 and 6.

DNA sequences at the fusion junctions

We characterized fusion junctions of the rearrangements in cases 1, 2, and 4–6 and in patients from families A–F (Table 3, Supplemental Table 2, and Figure 2). The results indicated the following: 1) nonallelic homologous recombination for the recurrent simple deletions in patients from families D–F that took place between two homologous sequences; 2) nonhomologous end-joining for the nonrecurrent simple deletions in patients from family C that were associated with short nucleotide stretches at the fusion junction; and 3) replication-based mechanisms for the simple and complex aberrations in cases 1, 2, and 4–6, and in patients from families A–B that were accompanied by microhomologies at the fusion junction. Nine of the 18 breakpoints resided within repetitive elements, such as LINE 1, LINE 2, *AluJo*, *AluY*, and *AluSx3*.

Genomic environments around the breakpoints

The average GC content was similar between the breakpoint-flanking and control regions (Supplemental Tables 2 and 3). Furthermore, the frequencies of known rearrangement-inducing DNA features (12, 14, 18–22) did not significantly differ between the breakpoint-flanking and control regions, except for some non-B structures enriched around the breakpoints of the deletions in patients from families D–F (Supplemental Tables 2 and 3).

Replication timing of the 15q21 region

Replication timing analysis indicated that in most cell lines examined, the genomic region around *CYP19A1* is replicated during early S phase (Supplemental Figure 2).

Discussion

We characterized six genomic rearrangements in patients with AEXS (Supplemental Figure 3). In case 1, the tandem duplication seems to have enhanced the transcriptional efficiency of *CYP19A1* in native *CYP19A1*-expressing tissues by increasing the number of transcription start sites. In cases 2–6, the rearrangements are predicted to have

Table 4. Genotype-Phenotype Correlation in Cases 1–6 and Families A–F

Cases/Families[a]	Case 1	Families A and B	Case 2	Case 3[b], Families C–F	Case 4	Case 5	Case 6
Molecular defects							
Predicted mechanism for *CYP19A1* overexpression	Duplication of *CYP19A1* coding exons	Duplication of *CYP19A1* promoters	Chimeric gene formation	Chimeric gene formation	Chimeric gene formation	Chimeric gene formation	Chimeric gene formation
Genes involved in chimeric gene formation	None	None	*DMXL2*	*DMXL2*	*SEMA6D*	*TMOD3*	*CGNL1*
Copy-number of the *CYP19A1* exon 1.4[c]	Normal	Increased	Decreased	Normal	Increased[d]	Normal	Decreased
Clinical findings							
Onset of gynecomastia, y	7	10–13	Unknown	7–12	11	7	5
Gynecomastia (Tanner stage)	2–3	2–3	1–3[e]	3–5	3–4	Severe	Severe
Advanced bone age	Mild	Subtle	Moderate	Mild/ moderate	Severe	N.E.	N.E.

Abbreviation: N.E., not examined.

[a] Cases 1–6 were present cases, whereas families A–F were reported previously (5).

[b] Fine genomic structure of case 3 remains to be characterized.

[c] Exon 1.4 functions as the major promoter in extragonadal tissues.

[d] Duplicated exon 1.4 has been disconnected from the coding exons of *CYP19A1*.

[e] The patient and his father had gynecomastia of Tanner stages 3 and 1, respectively.

created chimeric genes consisting of coding exons of *CYP19A1* and promoter-associated exons of neighboring genes. Actually, the deletions in cases 2 and 3 appear to have permitted splicing between *DMXL2* exon 1 and *CYP19A1* exon 2, as has been shown in the patients with similar deletions (5). Furthermore, the inversion in case 4 was found to produce a chimeric gene consisting of exon 3 of *SEMA6D* and exon 2 of *CYP19A1* (Supplemental Figure 1), and the inversions in cases 5 and 6 have previously been shown to form *TMOD3*- and *CGNL1-CYP19A1* chimeric genes, respectively (2). In this regard, the rearrangements in cases 2–6 have brought not only exons 1 of other genes, but also their flanking regions of >10 kb, to lie near the coding region of *CYP19A1*. Because these flanking regions harbor several enhancer- and promoter-associated histone marks (H3K4Me1 and H3K4Me3) (Supplemental Figure 4), they appear to contain most, if not all, components of cis-regulatory elements. Thus, although we can not examine the actual expression pattern of the chimeric genes, these genes seem to be expressed in a wide range of tissues where the original genes are expressed. These results argue for a broad mutation spectrum of AEXS.

Such diverse genetic basis of AEXS would be relevant to phenotypic variations (Table 4). First, cases/patients with copy-number gains of *CYP19A1* showed milder phenotypes than those with chimeric genes. This is consistent with the limited tissue expression pattern of *CYP19A1* and broad expression patterns of other genes involved in the chimeric gene formation (5, 25). Second, among cases/patients with simple duplications, case 1 showed a more severe phenotype than patients from families A–B. This suggests that tandem duplications encompassing the transcriptional unit, ie, the promoter region plus the coding exons, permit more efficient aromatase protein production than tandem duplications encompassing the promoter region only. Third, among cases/patients with the same *DMXL2-CYP19A1* chimeric gene, case 2 manifested milder phenotypes than case 4 and patients from families C–F. These results can be explained by the difference in the number of *CYP19A1* exons 1, because six of *CYP19A1* exons 1 were deleted in case 2 and all exons 1 were preserved in the remaining cases/patients (Supplemental Figure 5). Fourth, case 4 with a *SEMA6D-CYP19A1* chimeric gene showed a milder phenotype than cases 5 and 6 with a *TMOD3*- and *CGNL1-CYP19A1* chimeric gene, respectively. This is consistent with a tissue expression pattern being broader in *TMOD3* and *CGNL1* than in *SEMA6D* (5, 25). Lastly, cases/patients with *DMXL2-CYP19A1* chimeric genes manifested milder phenotype than cases with a *TMOD3*- or *CGNL1-CYP19A1* chimeric gene. This would primarily be ascribed to the presence or absence of a translational start codon on the fused promoter-associated exons (Supplemental Figure 6). It is likely that *DMXL2-CYP19A1* chimeric mRNAs transcribed by the *DMXL2* promoter preferentially recognize the natural start codon on *DMXL2* exon 1 and undergo nonsense-mediated mRNA decay, and rather exceptional chimeric mRNAs utilize the start codon on *CYP19A1* exon 2 and produce the aromatase protein (5). Such a phenomenon would not be postulated for

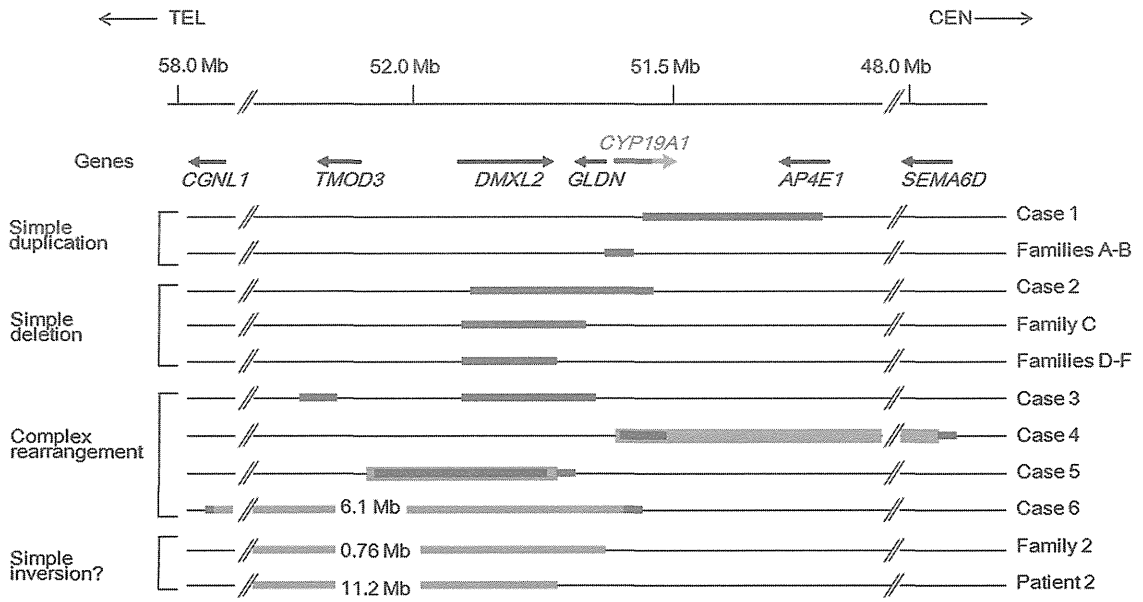

Figure 3. Schematic representation of the 11 rearrangements. Cases 1, 2, and 4–6 are from the present study, and patients from families A–F, patient 2, and patients from family 2 have been reported previously (1, 2, 5). The genomic abnormalities of case 3 were not characterized. The arrows indicate the positions and transcriptional direction of *CYP19A1* and its neighboring genes (5′→3′). Only genes around the fusion junctions are shown. The red, blue, and green lines indicate duplications, deletions, and inversions, respectively. For *CYP19A1*, the dark and light blue lines denote the genomic regions for the noncoding and coding exons, respectively. The inversions of family 2 and patient 2 may be complex rearrangements because copy-number analyses have not been performed in these cases.

the *TMOD3*- and *CGNL1-CYP19A1* chimeric mRNAs because of the absence of a translation start codon on exons 1 of *TMOD3* and *CGNL1*. Taken together, the present study suggests that phenotypic severity is primarily determined by the copy-number of *CYP19A1* and by the expression patterns and structural properties of the fused promoters. It should be pointed out, however, that this conclusion is based on the observation of only a limited number of patients. Phenotypic variation of the patients may be due to low penetrance of the clinical features.

To date, 11 genomic rearrangements have been identified in patients with AEXS (Figure 3). The 11 rearrangements are widely distributed on an approximately 9-Mb region and include simple duplications, deletions, and inversions, as well as complex rearrangements. Of these, the rearrangements in cases 1, 2, and 4–6 and in patients from families A–B are predicted to be replication-based errors (Supplemental Table 2 and Figure 2). Although the short nucleotide stretches at the fusion junctions in cases 1 and 5 may represent "information scars" characteristic of nonhomologous end-joining (9), the complex structures of the rearrangements would be consistent with replication-based mechanisms rather than end-joining (8). However, these rearrangements may result from microhomology-mediated end-joining (26). In contrast, the simple deletions in patients from family C and those in patients from families D–F are compatible with nonhomologous end-joining and nonallelic homologous recombination, respectively (Supplemental Table 2 and Figure 2). These results imply that the genomic region at 15q21 is vulnerable to both recombination- and replication-mediated errors.

In silico analyses revealed that deletions in families D–F due to nonallelic homologous recombination were associated with non-B structures and were located within an early-replicating segment of the genome, whereas the breakpoint-flanking regions of other rearrangements were independent of known rearrangement-inducing DNA features or late-replication timing. These data indicate that there are hitherto unidentified factors that facilitate nonhomologous end-joining and replication-based errors at 15q21. In this regard, it is noteworthy that nine of the 18 breakpoints resided within repetitive elements, and frequencies of *Alu*s (16%) and LINEs (22%) in the breakpoint-flanking regions were slightly higher than expected from the draft human genome (*Alu*, 9.9%; and LINE, 16.1%) (27). An increased number of repetitive sequences was found around the breakpoints of various rearrangements (14, 18, 19, 21), and Boone et al (28) have reported that a high concentration of *Alu* elements may predispose replication-based errors. The presence of various *Alu* family members (*AluJo*, *AluY*, and *AluSx3*) at the fusion junction of our cases supports the notion that moderate sequence similarity between *Alu* elements would be sufficient to provide substrates for replication-based errors (28). Further studies are necessary to clarify the role of repetitive sequences in the formation of rearrangements.

In summary, the present study implies a broad mutation spectrum of AEXS and supports the previously proposed

notion that clinical severities of AEXS are determined by the dosage of the promoter and coding regions of *CYP19A1* and by characters of the fused promoters. We show that rearrangements involved in AEXS can be attributed to nonallelic homologous recombination that is induced by repeats and/or by early-replication timing, and to nonhomologous end-joining and replication-based mechanisms that occur independently of known rearrangement-inducing DNA features or a late-replicating timing. Thus, AEXS represents a unique model for human genomic disorders.

Acknowledgments

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

This work was supported by the Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science, and Technology; by the Grant-in-Aid for Scientific Research and for Challenging Exploratory Research from the Japan Society for the Promotion of Science; by the Grant for Research on Intractable Diseases from the Ministry of Health, Labor, and Welfare; by grants from the National Center for Child Health and Development; and by grants from the Takeda Foundation and the Daiichi-Sankyo Foundation of Life Science.

Disclosure Summary: The authors have nothing to disclose.

References

1. **Shozu M, Sebastian S, Takayama K, et al.** Estrogen excess associated with novel gain-of-function mutations affecting the aromatase gene. *N Engl J Med.* 2003;348:1855–1865.
2. **Demura M, Martin RM, Shozu M, et al.** Regional rearrangements in chromosome 15q21 cause formation of cryptic promoters for the CYP19 (aromatase) gene. *Hum Mol Genet.* 2007;16:2529–2541.
3. **Bulun SE, Takayama K, Suzuki T, Sasano H, Yilmaz B, Sebastian S.** Organization of the human aromatase p450 (CYP19) gene. *Semin Reprod Med.* 2004;22:5–9.
4. **Demura M, Reierstad S, Innes JE, Bulun SE.** Novel promoter I.8 and promoter usage in the CYP19 (aromatase) gene. *Reprod Sci.* 2008; 15:1044–1053.
5. **Fukami M, Shozu M, Soneda S, et al.** Aromatase excess syndrome: identification of cryptic duplications and deletions leading to gain of function of CYP19A1 and assessment of phenotypic determinants. *J Clin Endocrinol Metab.* 2011;96:E1035–E1043.
6. **Lee C, Iafrate AJ, Brothman AR.** Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. *Nat Genet.* 2007;39:S48–S54.
7. **Lupski JR, Stankiewicz P.** Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet.* 2005;1:e49.
8. **Hastings PJ, Ira G, Lupski JR.** A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* 2009;5:e1000327.
9. **Gu W, Zhang F, Lupski JR.** Mechanisms for human genomic rearrangements. *Pathogenetics.* 2008;1:4.
10. **Shaw CJ, Lupski JR.** Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet.* 2004;13:R57–R64.
11. **Conrad DF, Bird C, Blackburne B, et al.** Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. *Nat Genet.* 2010;42:385–391.
12. **Chen JM, Cooper DN, Férec C, Kehrer-Sawatzki H, Patrinos GP.** Genomic rearrangements in inherited disease and cancer. *Semin Cancer Biol.* 2010;20:222–233.
13. **Colnaghi R, Carpenter G, Volker M, O'Driscoll M.** The consequences of structural genomic alterations in humans: genomic disorders, genomic instability and cancer. *Semin Cell Dev Biol.* 2011; 22:875–885.
14. **Froyen G, Belet S, Martinez F, et al.** Copy-number gains of HUWE1 due to replication- and recombination-based rearrangements. *Am J Hum Genet.* 2012;91:252–264.
15. **Wang G, Zhao J, Vasquez KM.** Methods to determine DNA structural alterations and genetic instability. *Methods.* 2009;48:54–62.
16. **Kurahashi H, Inagaki H, Ohye T, Kogo H, Kato T, Emanuel BS.** Palindrome-mediated chromosomal translocations in humans. *DNA Repair (Amst).* 2006;5:1136–1145.
17. **Cer RZ, Donohue DE, Mudunuri US, et al.** Non-B DB v2.0: a database of predicted non-B DNA-forming motifs and its associated tools. *Nucl Acids Res.* 2013;41:D94–D100.
18. **Verdin H, D'haene B, Beysen D, et al.** Microhomology-mediated mechanisms underlie non-recurrent disease-causing microdeletions of the FOXL2 gene or its regulatory domain. *PLoS Genet.* 2013;9: e1003358.
19. **Carvalho CM, Zhang F, Liu P, et al.** Complex rearrangements in patients with duplications of MECP2 can occur by fork stalling and template switching. *Hum Mol Genet.* 2009;18:2188–2203.
20. **Kornreich R, Bishop DF, Desnick RJ.** α-Galactosidase A gene rearrangements causing Fabry disease. Identification of short direct repeats at breakpoints in an Alu-rich gene. *J Biol Chem.* 1990;265: 9319–9326.
21. **Vissers LE, Bhatt SS, Janssen IM, et al.** Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. *Hum Mol Genet.* 2009; 18:3579–3593.
22. **Liu P, Carvalho CM, Hastings PJ, Lupski JR.** Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev.* 2012;22:211–220.
23. **Koren A, Polak P, Nemesh J, et al.** Differential relationship of DNA replication timing to different forms of human mutation and variation. *Am J Hum Genet.* 2012;91:1033–1040.
24. **Patry G, Jarvi K, Grober ED, Lo KC.** Use of the aromatase inhibitor letrozole to treat male infertility. *Fertil Steril.* 2009;92:829.e1–e2.
25. **Nagase T, Kikuno R, Ishikawa K, Hirosawa M, Ohara O.** Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res.* 2000;7:143–150.
26. **Lieber MR.** The mechanism of human nonhomologous DNA end joining. *J Biol Chem.* 2008;283:1–5.
27. **Venter JC, Adams MD, Myers EW, et al.** The sequence of the human genome. *Science.* 2001;291:1304–1351.
28. **Boone PM, Liu P, Zhang F, et al.** Alu-specific microhomology-mediated deletion of the final exon of SPAST in three unrelated subjects with hereditary spastic paraplegia. *Genet Med.* 2011;13:582–592.
29. **Pope BD, Tsumagari K, Battaglia D, et al.** DNA replication timing is maintained genome-wide in primary human myoblasts independent of D4Z4 contraction in FSH muscular dystrophy. *PLoS One.* 2011;6:e27413.
30. **Ryba T, Battaglia D, Chang BH, et al.** Abnormal developmental control of replication-timing domains in pediatric acute lymphoblastic leukemia. *Genome Res.* 2012;22:1833–1844.
31. **Pope BD, Chandra T, Buckley Q, et al.** Replication-timing boundaries facilitate cell-type and species-specific regulation of a rearranged human chromosome in mouse. *Hum Mol Genet.* 2012;21: 4162–4170.

Ehlers–Danlos Syndrome Associated with Glycosaminoglycan Abnormalities

10

Noriko Miyake, Tomoki Kosho, and Naomichi Matsumoto

Abstract

Ehlers–Danlos syndrome (EDS) is a genetically and clinically heterogeneous group of connective tissue disorders that typically present with skin hyperextensibility, joint hypermobility, and tissue fragility. The major cause of EDS appears to be impaired biosynthesis and enzymatic modification of collagen. In this chapter, we discuss two types of EDS that are associated with proteoglycan abnormalities: the progeroid type of EDS and dermatan 4-*O*-sulfotransferase 1 (D4ST1)-deficient EDS. The progeroid type of EDS is caused by mutations in *B4GALT7* or *B3GALT6*, both of which encode key enzymes that initiate glycosaminoglycan (GAG) synthesis. D4ST1-deficient EDS is caused by mutations in *CHST14*, which encodes an enzyme responsible for post-translational modification of GAG. The clinical and molecular characteristics of both types of EDS are described in this chapter.

Keywords

Ehlers–Danlos syndrome (EDS) • Progeroid type • *B4GALT7* • *B3GALT6* • Xylosylprotein beta 1,4-galactosyltransferase, polypeptide 7 • UDP-Gal:βGal β 1,3-galactosyltransferase polypeptide 6 • Dermatan 4-*O*-sulfotransferase 1 (D4ST1)-deficient EDS • *CHST14*

N. Miyake (✉) • N. Matsumoto
Department of Human Genetics, Yokohama City University Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama, 236-0004, Japan
e-mail: nmiyake@yokohama-cu.ac.jp; naomat@yokohama-cu.ac.jp

T. Kosho
Department of Medical Genetics, Shinshu University School of Medicine, 3-1-1 Asahi, Matsumoto, 390-8621, Japan
e-mail: ktomoki@shinshu-u.ac.jp

J. Halper (ed.), *Progress in Heritable Soft Connective Tissue Diseases*, Advances in Experimental Medicine and Biology 802, DOI 10.1007/978-94-007-7893-1_10,

Abbreviations

CHST14	Carbohydrate (*N*-Acetylgalactosamine 4-*O*) Sulfotransferase 14
D4ST1	Dermatan 4-*O*-sulfotransferase 1
EDS	Ehlers–Danlos Syndrome
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	*N*-Acetylgalactosamine
GlcA	Glucuronic Acid
IdoA	Iduronic Acid
PG	Proteoglycan
Xyl	Xylose

10.1 Introduction

Ehlers–Danlos syndrome (EDS) is a heterogeneous connective tissue disorder that affects as many as 1 in 5,000 individuals. It is characterized by joint and skin laxity, and tissue fragility [44]. In a revised classification, Beighton et al. classified EDS into six major types and several minor types [2]. The major causes of EDS are thought to include abnormal collagen biosynthesis through dominant-negative effects, haploinsufficiency of mutant procollagen α-chains, or deficiencies in collagen processing enzymes [29]. Abnormal glycosaminoglycan (GAG) synthesis and incorrect post-translational modification of GAG in proteoglycans (PGs) were recently identified in the progeroid type of EDS (EDS, progeroid form; MIM#130070, MIM#615349) and dermatan 4-*O*-sulfotransferase 1 (D4ST1)-deficient EDS (EDS, musculocontractural type; MIM#601776), respectively. In this chapter, the clinical and molecular characteristics of both types of EDS are described.

10.2 Background

Glycosylation is the addition of a sugar chain (a glycan) to a protein (generating a glycoprotein) or lipid (generating a glycolipid). More than 40 human disorders are thought to be caused by abnormal glycosylation [15, 19]. PGs are composed of core proteins and one or more glycans with modifications. PGs are present in the extracellular matrix and have important diverse biological functions [5]. PG synthesis is initiated by the sequential addition of four monosaccharides (xylose [Xyl], two molecules of galactose [Gal] and glucuronic acid [GlcA]), known as a linker tetrasaccharide, to the serine residue of the core protein backbone (Fig. 10.1a). Additional sugar chains are extended from the linker tetrasaccharide by the addition of repeated disaccharides (usually consisting of 50–150 disaccharides in vivo). Afterwards, some sugars are modified by a series of epimerases (epimerization) and sulfotransferases (sulfation).

GAGs are long unbranched polysaccharides consisting of repeating disaccharide units. GAGs are highly negatively charged because of the acidic sugar residues and/or sulfation. Consequently, GAG can change its conformation, attract cations, and bind water. Hydrated GAG gels enable joints and tissues to absorb large pressure changes, providing tissue elasticity. Post-translational modifications such as epimerization, sulfation, and acetylation/deacetylation result in the formation of diverse motifs in the GAG chains, which can bind to a large variety of ligands. Therefore, GAG chains play important roles in regulating growth factor signaling, cell adhesion, proliferation, differentiation, and motility [3, 5, 45].

GAGs can be divided into two groups: (1) galactosaminoglycans such as chondroitin sulfate (CS) and dermatan sulfate (DS), and (2) glucosaminoglycans such as hyaluronic acid, keratan sulfate, heparan sulfate, and heparin [42]. Two types of glycosylation are known: *O*-glycosylation and *N*-glycosylation (Fig. 10.2a). Most GAGs (except for keratan sulfate and hyaluronic acid) are *O*-glycans that bind to the glycan via an oxygen molecule in the serine or threonine residue of the core protein (Fig. 10.2a). Notably, failure to add the first or second galactose residue of the tetrasaccharide results in the progeroid type of EDS (Fig. 10.1b, c).

The CS and DS GAGs are produced via the same pathway (Fig. 10.3a). In this pathway, after the linker tetrasaccharide attaches to the serine

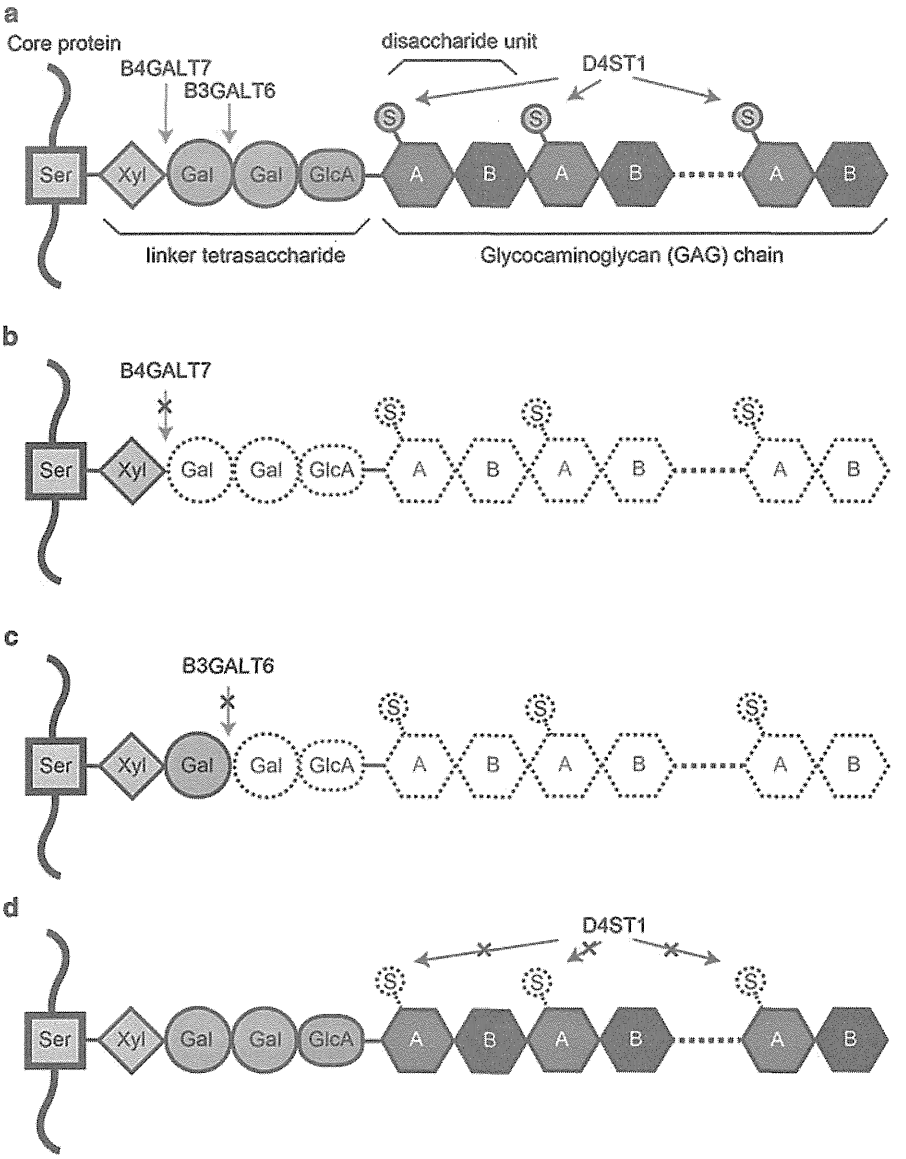

Fig. 10.1 Proteoglycan biosynthesis and its defects in two types of EDS. (**a**) Normal state. The serine residue (Ser) of the core protein and the GAG chain are bound via a linker tetrasaccharide. In CS, the disaccharides are composed of N-acetylgalactosamine (GalNAc) [position A] and glucuronic acid [position B]. In DS, the disaccharides are composed of GalNAc [position A] and Iduronic acid (IdoA) [position B]. B4GALT7 and B3GALT6 add the first and second galactose (Gal) to the xylose of the linker tetrasaccharide (*green arrows*). D4ST1 then adds the active sulfate to the 4-*O* position of GalNAc (*red arrows*) on DS. (**b**, **c**) Progeroid type of EDS. The impaired B4GALT7 cannot elongate the glycan chain from the first galactose (**b**). The impaired B3GALT6 cannot add the second galactose and the following glycan chain (**c**). (**d**) D4ST1-deficient EDS. The impaired/inactive D4ST1 cannot add the sulfate to GalNAc. *Gal* galactose, *GlcA* glucuronic acid, *S* active sulfate, *Ser* serine, *Xyl* xylose

residue of the core protein, GalNAc (*N*-acetyl galactosamine) transferase I elongates the glycan branch to create CS/DS. The enzyme C5-carboxy epimerase transforms glucuronic acid (GlcA) to iduronic acid (IdoA), which is specific for dermatan/DS (Fig. 10.3a). DS actually exists in a CS/DS hybrid state, containing GlcA–GalNAc and IdoA–GalNAc disaccharides (Figs. 10.2b

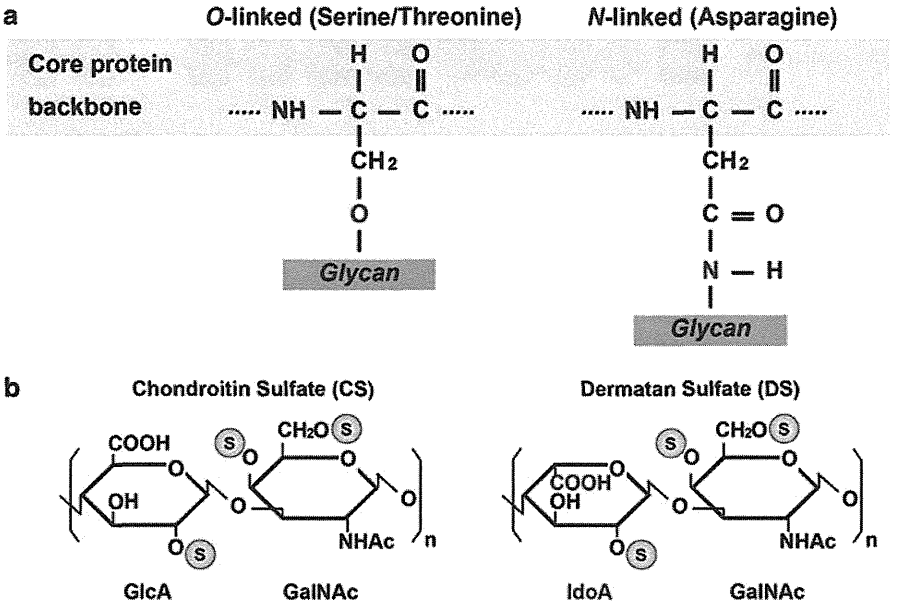

Fig. 10.2 Chemical structures of proteoglycan and disaccharides. (**a**) Chemical structure of *O*-linked and *N*-linked glycan. *O*-linked glycan can be linked via the *O*-element of serine or threonine. The diagram shows linking for serine. (**b**) Chemical structures of the disaccharide units of CS (*left*) and DS (*right*)

and 10.3a) [12]. Dermatan 4-*O*-sulfotransferase 1(D4ST1) specifically transfers an active sulfate to the 4-*O* position on the GalNAc residue of dermatan. The transfer of the active sulfate is impaired in D4ST1-deficient EDS (Figs. 10.1d and 10.3b).

10.3 The Progeroid Type of EDS (type 1: MIM#130070, type2: MIM#615349)

Alternative Names (MIM#130070)
Xylosylprotein 4-β-galactosyltransferase deficiency
XGPT deficiency
Galactosyltransferase I deficiency

10.3.1 Clinical Manifestations

Hernandez et al. reported five unrelated males in 1979, 1981, and 1986 representing a distinct variant of EDS. These males presented with a progeroid facial appearance, mild intellectual disability, and multiple nevi, in addition to hyperextensibility and fragility of skin, a high propensity for bruising, and joint hypermobility (particularly of the digits) [16–18]. A wrinkled face, curly and fine hair, scant eyebrows/eyelashes, telecanthus, periodontitis, multiple caries, low set/prominent ears, pectus excavatum, winged scapulae, and pes planus were observed in all five patients. Cryptorchidism and inguinal hernia were also noticed in four of the patients. Interestingly, the occurrence of the disorder in all of these patients was sporadic and the ages of their fathers were relatively advanced (33–55 years old). These characteristics prompted Hernandez et al. to speculate that the syndrome is caused by a de novo mutation [16].

In 1987, Kresse et al. reported a Danish male patient who was born to non-consanguineous healthy parents [26]. This patient presented with the clinical features observed in the original five patients, as well as a triangular head with a tiny face, frontal bossing, mid-face hypoplasia, a broad nasal bridge, prominent deep-set eyes, a small mouth, dental anomalies, low-set ears,

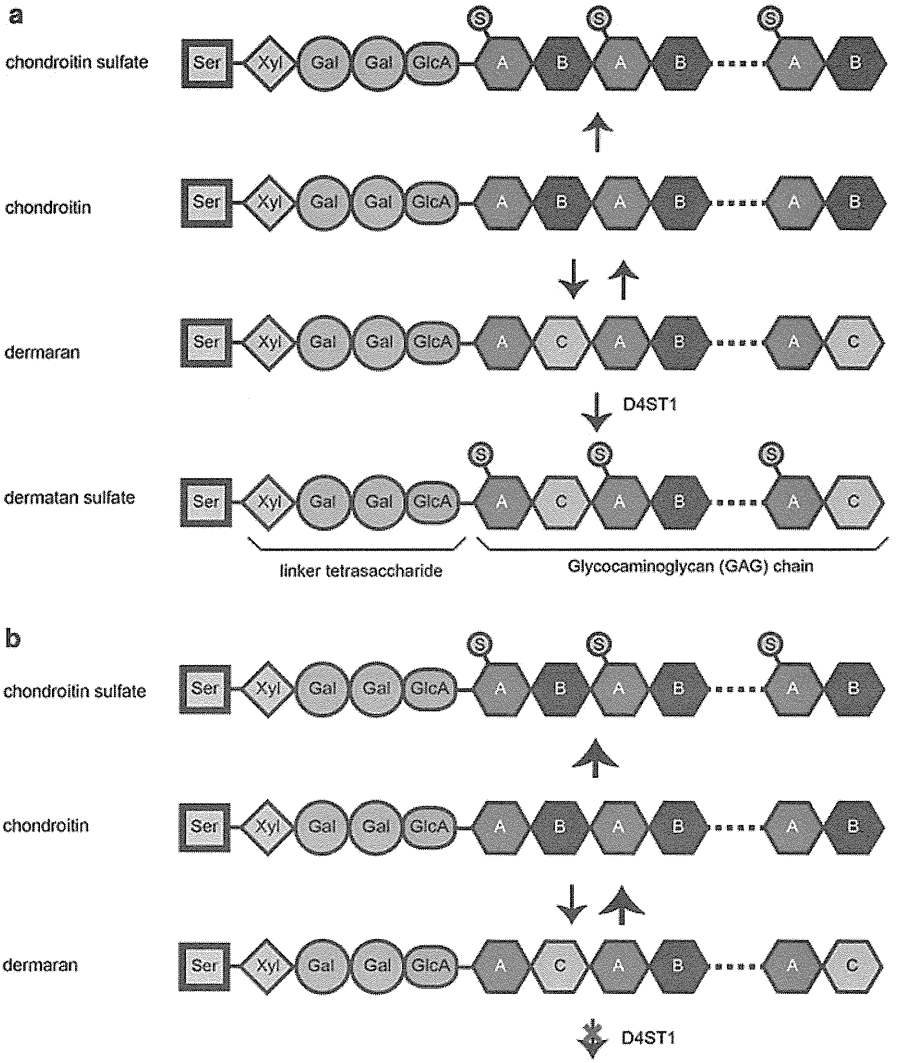

Fig. 10.3 Effects of D4ST1 defects on the biosynthesis of CS and DS. (**a**) The starting structure is chondroitin with a repeating disaccharide consisting of GalNAc [position A] and GlcA [position B]. Sulfation by 6-*O*-GalNAc sulfotransferase and 4-*O*-GalNAc sulfotransferase creates CS from chondroitin. To produce DS, first, C5-carboxy epimerase replaces GlcA with IdoA [position C]. This process is bidirectional as indicated by the bi-directional arrows. Then, D4ST1 adds sulfates to dermatan creating DS and prevents back epimerization. DS is often detected as a CS/DS hybrid. (**b**) In D4ST1-deficient EDS, back epimerization from IdoA to GlcA occurs. Consequently, neither DS nor dermatan are detected in fibroblasts derived from patients. *Gal galactose*, *GlcA* glucuronic acid, *S* active sulfate, *Ser* serine, *Xyl* xylose

short stature, osteopenia of all bones, dysplasia of some bones, and hypotonia. In 2004, Faiyaz-Ul-Haque et al. reported two patients from a large consanguineous Qatari family. The clinical features of both Qatari patients and the Danish patient seemed to be different from those of the original five patients [14].

10.3.2 Genetic Information

10.3.2.1 B4GALT7

In 1999, two different research groups [1, 33] identified compound heterozygous mutations of gene for xylosylprotein beta 1,4-galactosyltransferase, polypeptide 7 (*B4GALT7*, NM_007255.2),

c.557C>A (p.Ala186Asp) and c.617T>C (p. Leu206Pro) in the Danish patient reported by Kresse in 1987 [26]. The two Qatari patients from a large consanguineous family were analyzed in 2004 [14]. Based on the hypothesis of autosomal recessive inheritance, haplotype analysis using microsatellite markers for the limited candidate loci delineated a homozygous region from *D5S469* and *D5S2111*, which harbors *B4GALT7* [14]. A homozygous missense mutation (c.808C>T, p.Arg270Cys) in *B4GALT7* was identified. Interestingly, the clinical phenotype of the Qatari patients was milder than that of the Danish one.

B4GALT7 was cloned by Okajima et al. [34]. The gene consists of six coding exons with a 948-bp open reading frame. This gene encodes xylosylprotein β-1,4-galactosyltransferase, polypeptide 7 (B4GALT7; aliases: galactosyltransferase I, XGPT1, and XGALT1), which is 327 amino acids long and its molecular weight is 37.4 kDa. B4GALT7 is a type II transmembrane protein localized in the Golgi apparatus, and is a key initiator of GAG synthesis as it attaches the first galactose of the linker tetrasaccharide of PGs (Fig. 10.1a, b).

10.3.2.2 B3GALT6

In 2013, Nakajima et al. have identified compound heterozygous mutations of *B3GALT6* (NM_080605.3) in three patients with progeroid form of EDS [32]. This intronless gene has a 990-bp open reading frame and encodes UDP-Gal:βGal β 1,3-galactosyltransferase polypeptide 6 (alternatively galactosyltransferase -II: GalT-II), which is 329 amino acids long and its molecular weight is 37.1kDa. It is also the type II transmembrane protein localized in the Golgi apparatus, and it attaches the second galactose of the tetrasaccharide linker of PGs (Fig. 10.1a, c). So far, two missense (c.16C>T, p.Arg6Trp and c.925T>A, p.Ser309Thr), two frameshift deletions (c.353delA, p.Asp118Alafs*160 and c.588delG, p.Arg197Alafs*81) and one in-frame deletion (c.415_423del, p.Met139Ala141del) were reported in this type of EDS [32].

10.3.3 Biochemical Characteristics

10.3.3.1 B4GALT7

Kresse et al. reported that their patient's fibroblasts produced only PG chain-free core proteins (molecular weight: 46 and 44 kDa) whereas control fibroblasts produced normal PG chains [26]. Additionally, the GAG-free core protein in that patient contained unsubstituted xylose residues (Fig. 10.1b).

Okajima et al. measured the enzyme activity of exogenously expressed proteins (wild type, p. Ala186Asp, p.Leu206Pro) in XGalT-1/B4GALT7-deficient CHO cells [33]. In total cell lysates, the enzyme activity of the p.Ala186Asp mutant was approximately 50 % lower than that of the wild-type protein, whereas the activity of the p. Leu206Pro mutant was almost undetectable. Interestingly, the wild-type and p.Ala186Asp proteins were localized in the Golgi apparatus whereas the p.Leu206Pro mutant existed in the cytoplasm. The α-helix disrupted by p.Leu206Pro may alter the protein's conformation, thus impairing intracellular trafficking and enzyme activity [33].

B4GALT7 activity in fibroblasts from another patient with a homozygous mutation, c.808C>T (p. Arg270Cys), was also lower than that of controls [40]. The extracellular matrix around the B4GALT7^{Arg270Cys} mutant fibroblasts was disorganized without banded fibrils. Furthermore, the proliferation of B4GALT7^{Arg270Cys} fibroblasts was significantly reduced to 45 % of the level of control fibroblasts [40].

Bui et al. measured galactosyltransferase activity of B4GALT7 mutants expressed in CHO pgsB-618 cells using 4-methylumbelliferyl-β-D-xylopyranoside as acceptor substrate. The enzyme activities of the p.Arg270Cys, p. Ala186Asp, and p.Leu206Pro mutants were decreased to 60, 11, and 0 % (undetectable) of that of the wild-type enzyme [4]. It has been reported that the clinical features of patients with the homozygous p.Arg270Cys mutation appear to be milder than those of patients with compound heterozygous mutations, including p. Ala186Asp or p.Leu206Pro, supporting the different effects of these mutations.

10.3.3.2 B3GALT6

Nakajima et al. measured the galactosyltransferase activity of *B3GALT6* in vitro using soluble-FLAG-tagged proteins for wild-type and mutant (p.Ser309Thr) which was observed common in two families and revealed the enzyme activity of the mutant protein was significantly decreased compared to the wild-type [32].

10.4 D4ST1-Deficient EDS (MIM#601776)

Alternative Names
Ehlers–Danlos syndrome, type VIB, formerly
Ehlers–Danlos syndrome, Kosho type
Ehlers–Danlos syndrome, musculocontractural type
Adducted thumbs, clubfoot, and progressive joints and skin laxity syndrome
Adducted thumb-clubfoot syndrome (ATCS)
Dündar syndrome
Arthrogryposis, distal, with peculiar faces and hydronephrosis

10.4.1 Clinical Manifestations

The kyphoscoliosis type of EDS (formerly known as, EDS type VI) is characterized by generalized joint laxity, severe muscular hypotonia and scoliosis at birth, scleral fragility, and rupture of the ocular globe [2]. This disorder is essentially caused by lysyl hydroxylase deficiency (EDS type VIA); other patients with similar clinical manifestations but without lysyl hydroxylase deficiency were classified as EDS type VIB.

In 2005, Kosho et al. reported two unrelated patients with fragile and hyperextensible skin, a high propensity for bruising, generalized joint laxity, kyphoscoliosis, and the major features of EDS VI, as well as a characteristic craniofacial appearance, and multiple congenital contractures [25]. Lysyl hydroxylase deficiency was excluded in these patients by analysis of the urinary deoxypyridinoline:pyridinoline ratio, and the patients were tentatively classified as EDS VIB. Kosho et al. subsequently reported on four additional unrelated patients and concluded that the patients represented a new type of EDS [23]. Notably, all six patients had homozygous or compound heterozygous mutations in *CHST14* [31]. Loss-of-function mutations in *CHST14* were independently found in 11 patients from four families with a rare arthrogryposis syndrome known as "adducted thumb-clubfoot syndrome (ATCS)" [9–11, 21, 43] and in three patients from two families who were originally classified as suffering from EDS VIB [27]. Malfait et al. suggested that these patients had the same disorder, which they termed "musculocontractural EDS" [27]. Shimizu et al. described the clinical characteristics of two additional patients together with a review of all of the patients reported at that time; their findings support the notion that the three independently identified conditions represent a single type of EDS [41]. Conversely, Janecke et al. claimed that the disorder should not be categorized as a type of EDS because of the presence of atypical clinical features, including facial dysmorphism, multiple congenital contractures, visceral anomalies, and impaired biosynthesis of DS as a cause of the disorder, and proposed the term DS-deficient adducted thumb-clubfoot syndrome [20]. In their response, Kosho et al. provided clinical and etiological evidence from which the disorder could be categorized as a type of EDS, because of the presence of all major features of EDS, including connective tissue fragility which required special and appropriate management of these patients. Decorin-mediated impaired assembly of collagen fibrils was the primary cause of progressive connective tissue fragility in this type [24]. Therefore, Kosho et al. proposed that the term D4ST1-deficient EDS (adducted thumb-clubfoot syndrome) was appropriate for this syndrome [24]. The current OMIM (http://www.ncbi.nlm.nih.gov/omim) registration of this disorder is EDS, musculocontractural type.

To date, descriptions of 26 patients (12 males, 14 females) from 17 families have been published [9–11, 21, 23, 25, 27, 30, 31, 41, 43, 46,

Table 10.1 Classification of Ehlers-Danlos syndrome

	Prevalence/patient number	Inheritance	Causative gene
Major types			
Classical type	1/20,000	AD	*COL5A1, COL5A2*
Hypermobility type	1/5,000–20,000	AD	Unknown[a]
Vascular type	1/50,000–250,000	AD	*COL3A1*
Kyphoscoliosis type	1/100,000	AR	*PLOD1*
Arthrochalasia type	30	AD	*COL1A1, COL1A2*
Dermatosparaxis type	8	AR	*ADAMTS2*
Other types			
Brittle cornea syndrome	11	AR	*ZNF469*
EDS-like syndrome due to tenascin-XB deficiency	10	AR	*TNXB*
Progeroid form	7	AR	*B4GALT7, B3GALT6*
Cardiac valvular form	4	AR	*COL1A2*
EDS-like spondylocheirodysplasia	8	AR	*SLC39A13*
D4ST1-deficient EDS (DD-EDS)	22	AR	*CHST14*

AD, autosomal dominant; *AR*, autosomal recessive; *COL5A1*, collagen, type V, alpha 1; *COL5A2*, collagen, type V, alpha 2; *COL3A1*, collagen, type III, alpha 1; *PLOD1*, procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1; *COL1A1*, collagen, type I, alpha 1; *COL1A2*, collagen, type I, alpha 2; *ADAMTS2*, ADAM metallopeptidase with thrombospondin type 1 motif, 2; *ZNF469*, zinc finger protein 469; *TNXB*, tenascin XB; *B4GALT7*, xylosylprotein beta 1,4-galactosyltransferase, polypeptide 7; *B3GALT6*, UDP-Gal:βGal β 1,3-galactosyltransferase polypeptide 6; *COL1A2*, collagen, type I, alpha 2; *SLC39A13*, solute carrier family 39 (zinc transporter), member 13; *CHST14*, carbohydrate (*N*-acetylgalactosamine 4-*0*) sulfotransferase 14

[a] *TNXB* mutations in a small subset of patients

48, 49]. This syndrome is characterized by a unique set of clinical features consisting of progressive systemic manifestations, including tissue fragility (e.g., skin hyperextensibility and fragility, progressive spinal and foot deformities, and large subcutaneous hematomas) and various malformations (e.g., facial features, congenital eye/heart/gastrointestinal defects, congenital multiple contractures). We have summarized the main clinical features of this syndrome in each organ in Table 10.1

10.4.1.1 Craniofacial Features

The characteristic craniofacial features apparent at birth or during early infancy include a large fontanelle, hypertelorism, short and down slanting palpebral fissures, blue sclerae, a short nose with hypoplastic columella, low-set and rotated ears, a high or cleft palate, a long philtrum, a thin upper-lip vermilion, a small mouth, and micro-retrognathia (Fig. 10.4a, b). Slender and asymmetrical facial shapes with a protruding jaw are generally observed from school age onwards (Fig. 10.4c, d).

10.4.1.2 Skeletal Features

Congenital multiple contractures, particularly adduction-flexion contractures of the thumbs and talipes equinovarus, are the main skeletal features (Fig. 10.4e, g, h). Fingers with a "tapering", "slender", and "cylindrical" shape are also common (Fig. 10.4f). Aberrant finger movement was described in three patients. Four patients had tendon abnormalities, including anomalous insertion of the flexor muscles, which probably caused the congenital contractures. Spinal deformities (e.g., scoliosis and kyphoscoliosis) and talipes deformities (e.g., planus and valgus) (Fig. 10.4i) occurred and progressed during childhood. Marfanoid habitus, recurrent joint dislocations, and pectus deformities (e.g., flat and thin, excavatum, and carinatum) were also evident. Bone mineral density was decreased in five patients and normal in two. Urine concentrations of the N-telopeptide of collagen type I, an osteoclast marker, were increased in three patients, whereas serum bone-specific alkaline phosphatase concentrations, an osteoblast marker, were normal in three, suggesting that

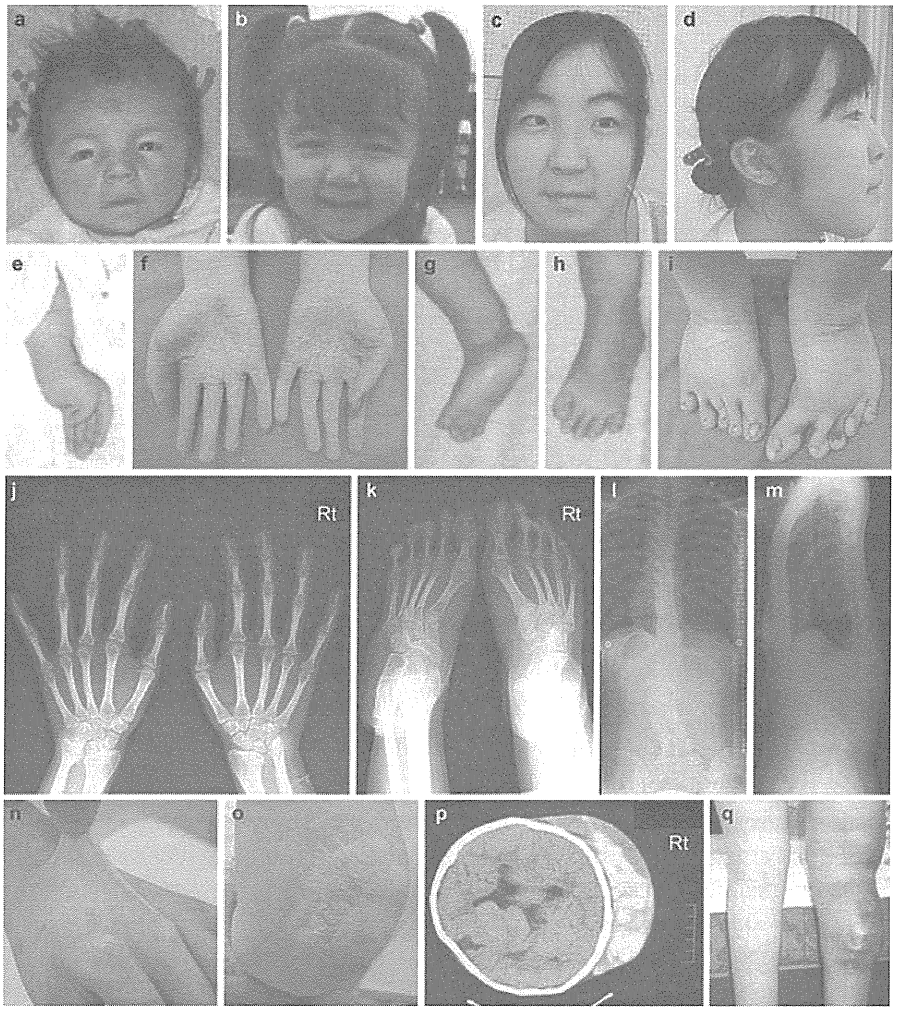

Fig. 10.4 Clinical photographs of patients with D4ST1-deficient EDS. (**a–d**) Facial features of a patient at 23 days (**a**), 3 years (**b**), and 16 years (**c**, **d**) of age. (**e**, **f**) Images of the hand in a patient with an adducted thumb at 1 month of age (**e**) and cylindrical fingers at 19 years of age (**f**). (**g–i**) Images of the foot in a patient with bilateral clubfeet at 1 month of age (**g**, **h**) and progressive talipes deformities (planus and valgus) at 19 years of age (**i**). (**j–m**) Radiographs of a 16-year-old patient show diaphyseal narrowing of the phalanges and metacarpals (**j**, **k**) and kyphoscoliosis with tall vertebral bodies (**l**, **m**). (**n**, **o**) Cutaneous features of a 19-year-old patient with hyperextensibility (**n**), atrophic scars, and fistula formation (**o**). (**p**) A massive cranial subcutaneous hematoma in the head of a 6-year-old patient after falling onto the floor. (**q**) A subcutaneous hematoma in the leg of a 16-year-old patient (All figures were originally published in Kosho et al. [23] except Fig. 10.4p, which was published in Kosho et al. [25])

increased osteoclast activity but normal osteoblast activity could cause osteopenia or osteoporosis. Radiologically, diaphyseal narrowing of the phalanges and metacarpals was noted in six patients (Fig. 10.4j, k). Talipes valgus and planus or cavum, with diaphyseal narrowing of the phalanges and metatarsals, were noted in six patients. Tall vertebral bodies were noted in five patients (Fig. 10.4l, m).

10.4.1.3 Cutaneous Features

Cutaneous features were apparent in most patients, including hyperextensibility to redundancy (Fig. 10.4n), a high propensity for bruising,

fragility leading to atrophic scars (Fig. 10.4o), acrogeria-like fine palmar creases or wrinkles, hyperalgesia to pressure, and recurrent subcutaneous infections with fistula formation. The palmar creases increased and became deeper with age.

10.4.1.4 Cardiovascular Features

Large subcutaneous hematomas were common, and frequently required intensive treatment, including hospital admission, blood transfusion, and surgical drainage (Fig. 10.4p, q). The lesions were thought to be caused by the rupture of a subcutaneous artery or vein. Bleeding time was prolonged in two patients (9 min and 11 min) and was normal in three. Intranasal administration of 1-desamino-8-D-arginine vasopressin prevented the development of large subcutaneous hematomas after trauma [49]. Four patients had congenital heart defects including an atrial septal defect in three, a patent ductus arteriosus in one, and coarctation of the aorta in one. Five patients had cardiac valve abnormalities including one who underwent surgery for infectious endocarditis, which was probably caused by aortic valve or mitral valve regurgitation.

10.4.1.5 Respiratory Features

Three adult patients developed pneumothorax or hemopneumothorax requiring chest tube drainage.

10.4.1.6 Gastrointestinal Features

Numerous gastrointestinal abnormalities were reported, including diverticular perforation in two adult patients, constipation in seven patients, abdominal pain in two patients, and other disorders in one patient (common mesentery, absence of the gastrocolic omentum with a spontaneous volvulus of small intestine, gastric ulcer, and malrotation with duodenal obstruction).

10.4.1.7 Genitourinary Features

Urological complications included nephrolithiasis or cystolithiasis in five patients, hydronephrosis in three, a dilated or atonic bladder with recurrent urinary tract infection in two, and a horseshoe kidney in one. Cryptorchidism was observed in eight male patients, including one who underwent orchiopexy because of hypogonadism in adulthood. Poor breast development was noted in five adolescent or adult patients. No pregnant females have been reported.

10.4.1.8 Ophthalmologic Features

Various ophthalmological complications have been reported, including strabismus in 12 patients, refractive errors in nine, glaucoma or elevated intraocular pressure in six, microcornea or microphthalmia in three, and retinal detachment in three.

10.4.1.9 Hearing Impairment

Six patients had hearing impairments, including for high-pitched sounds in three.

10.4.1.10 Growth

Patients showed mild prenatal growth retardation as the mean birth length was −0.5 standard deviations (SD), the mean birth weight was −0.6 SD, and the mean birth occipitofrontal circumference (OFC) was −0.2 SD. Postnatal growth was also mildly impaired, as the patients were generally slender with relative macrocephaly. The mean height was −0.9 SD, the mean weight was −1.5 SD, and the mean OFC was −0.2 SD.

10.4.1.11 Development and Neuromuscular Features

Gross motor developmental delay was observed in 14 patients, as the median age of independent walking was 2 years 1 month. Two patients, aged 15 years and 32 years, could not walk unassisted. An underlying myopathic process was observed in two patients. Mild intellectual disability was apparent in four patients. One patient had a global psychomotor delay at 1.5 years of age, but his intellectual quotient was approximately 90 at the age of 7 years 2 months. Brain imaging showed ventricular enlargement and/or asymmetry in seven patients, absence of the left septum pellucidum in one patient, and a short corpus callosum, mildly prominent Sylvian fissures, and periventricular nodular heterotopias. Two patients had spinal cord tethering.