Retrotransposition disrupting EBP in a girl and her mother with X-linked dominant chondrodysplasia punctata

Retrotransposition disrupting *EBP* in a girl and her mother with X-linked dominant chondrodysplasia punctata

Takuya Hiraide[1,2], Yohei Masunaga[1], Akira Honda[3], Fumiko Kato[4], Tokiko Fukuda[4], Maki Fukami[5], Mitsuko Nakashima[2], Hirotomo Saitsu[2] and Tsutomu Ogata[1,2,6]

[1]Department of Pediatrics, Hamamatsu University School of Medicine, Hamamatsu, Japan
[2]Department of Biochemistry, Hamamatsu University School of Medicine, Hamamatsu, Japan
[3]Joint Research Center, Tokyo Medical University Ibaraki Medical Center, Ibaraki, Japan
[4]Hamamatsu Child Health and Developmental Medicine, Hamamatsu, Japan.
[5]Department of Molecular Endocrinology, National Research Institute for Child Health and Development, Tokyo, Japan
[6]Department of Pediatrics, Hamamatsu Medical Center, Hamamatsu, Japan

Correspondence
Tsutomu Ogata, Department of Pediatrics, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu 431-3192, Japan. E-mail: tomogata@hama-med.ac.jp

Funding Information
This work was supported by the grants from Japan Agency for Medical Research and Development (AMED) (JP19ek0109297 to H.S. and JP21ek0109549 to T.O.).

Abstract

X-linked dominant chondrodysplasia punctata (CDPX2) is a rare congenital disorder caused by pathogenic variants in *EBP* on Xp11.23. We encountered a girl and her mother with CDPX2-compatible phenotypes including punctiform calcification in the neonatal period of the girl, and asymmetric limb shortening and ichthyosis following the Blaschko lines in both subjects. Although Sanger direct sequencing failed to reveal a disease-causing variant in *EBP*, whole genome sequencing (WGS) followed by Manta analysis identified a ~ 4.5 kb insertion at *EBP* exon 2 of both subjects. The insertion was associated with the hallmarks of retrotransposition such as an antisense poly(A) tail, a target site duplication, and a consensus endonuclease cleavage site, and the inserted sequence harbored full-length SVA_F1 element with 5′- and 3′-transductions containing the Alu sequence. The results imply the relevance of retrotransposition to the human genetic diseases and the usefulness of WGS in the identification of retrotransposition.

Keywords:

X-linked dominant chondrodysplasia punctata, *EBP*, whole genome sequencing, retrotransposition

Introduction

X-linked dominant chondrodysplasia punctata (CDPX2; OMIM # 302960), also known as Conradi-Hünermann-Happle syndrome, is a rare congenital disorder characterized by punctiform calcification of the bones, rhizomelic shortness, transient congenital ichthyosis following the Blaschko lines, patchy alopecia, cataracts, and midface hypoplasia [1]. CDPX2 is caused by pathogenic variants in *EBP* encoding emopamil-binding protein for 3β-hydroxysteroid Δ^8,Δ^7-sterol isomerase involved in cholesterol biosynthesis [2]. Thus, the substrates for EBP such as 8-dehydrocholesterol are increased, and the products via EBP such as 7-dehydrocholesterol are decreased in CDPX2 (Supplementary Figure 1).

Retrotransposition events are an important underlying factor leading to genetic diversities [3]. Furthermore, they can cause genetic diseases, when transposable elements impair disease-related genes [4]. In the human genome, the long interspersed element 1 (LINE-1) encoding two proteins (ORF1p and ORF2p) is the sole active autonomous transposable element. ORF1p has a single-stranded nucleic acid binding activity and a chaperone activity, and ORF2p has a reverse transcriptase activity and an endonuclease activity [5–8]. Thus, LINE-1 can act *in cis* and *in trans* to mobilize transcribed RNA sequences to new genomic locations, thereby mediating retrotransposition of transcribed non-autonomous retrotransposons such as Alu and SVA (short interspersed element (SINE)-variable number tandem repeat-Alu element) as well as mobilization of mRNAs to generate processed pseudogenes [9–12]. SVA elements are active retrotransposons specific to hominid genomes [13]. The SVA F subfamily is specific to humans with SVA_F1 (MAST2) members characterized by the presence of exon 1 of *MAST2*, and accounts for at least 32% of the SVA_F insertions [14].

Here, we report a Japanese girl (the proband) and her mother with CDPX2 caused by retrotransposition disrupting *EBP*.

Clinical report

The pedigree of this family is shown in Figure 1A. The proband (III-2) was born at 38 weeks and 6 days of gestation to a 33-year-old mother (II-2) and a 34-year-old father (II-1) by

vaginal delivery. She had been found to have short left lower limb by fetal ultrasound studies. At birth, her length was 44.0 cm (–2.4 SD), her weight 2,522 g (–1.3 SD), and her occipitofrontal circumference 33.0 cm (–0.1 SD). Physical examination at birth revealed overt scaling ichthyosis following the Blaschko lines, erythematous skin, patchy alopecia, bilateral cataracts, and midface hypoplasia, together with shortened left lower limb (Figure 1B). The skin lesion ameliorated in a few months (Figure 1C), while hypopigmented skin lesions following the Blaschko lines remained clinically discernible (Figure 1D). Roentgenographic examinations were repeatedly performed, delineating skeletal features characteristic of CDPX2, such as epiphyseal stippling at the right knee joint in infancy, asymmetrical shortening of the lower extremities, hemivertebra of the fourth thoracic vertebra, scoliosis, and left acetabular dysplasia (Figure 1E–H). Serum sterol profile analysis was performed at 4.5 months of age by the previously described gas chromatography tandem mass spectrometry method with a minor modification [15, 16], indicating defective 3β-hydroxysteroid Δ^8,Δ^7-sterol isomerase activity (Supplementary Table 1). Cataract extraction was performed for the severe right side cataract at two months of age (artificial lens implant was planned to be considered in a later age), and cataract extraction followed by artificial lens implant was carried out for the left side cataract at five years of age. Since lower limb asymmetry became severe (~ 4 cm), she received surgical bone lengthening of the left lower limb at 3 and 5 years of age. Based on the above findings, she was diagnosed with CDPX2. On the last examination at 8 years and 2 months of age, she measured 105.0 cm (–3.8 SD), weighed 24.8 kg (–0.2 SD), and showed apparently normal intellectual development.

The mother (II-2) also manifested CDPX2-compatible features such as rhizomelic short stature with a height of 143.9 cm (–2.7 SD) and a weight of 43.9 kg (–1.1 SD), mildly short left lower limb, and hypopigmented skin lesions following the Blaschko lines, although she lacked patchy alopecia and cataracts. Allegedly, she exhibited marked linear ichthyosis in her infancy. Serum sterol profile analysis indicated impaired 3β-hydroxysteroid Δ^8,Δ^7-sterol isomerase activity (Supplementary Table 1).

The maternal parents (I-3 and I-4) were clinically normal, as were the 11-year-old elder sister (III-1), the father (II-1), and the paternal parents (I-1 and I-2) (Figure 1A).

Genetic studies

This study was approved by the Institutional Review Board Committee at Hamamatsu University School of Medicine, and was performed using peripheral blood samples after obtaining written informed consent.

Cytogenetic analysis showed a 46,XX karyotype, and Sanger direct sequencing revealed no discernible variant on exons and splice sites of *EBP2* in the affected III-2 and II-2. Thus, whole genome sequencing (WGS) was carried out, to explore a non-coding variant or a structural variant (SV) in *EBP*. WGS was commissioned to BGI Japan Corporation (Kobe, Japan), and data processing, variant calling, annotation, and filtering were performed as described previously [17]. SVs including medium-sized indels and large insertions were investigated by Manta [18] and visualized by Integrative Genomic Viewer (IVG) (https://www.igv.org). The character of repetitive elements was examined by RepeatMasker (http://www.repeatmasker.org/).

Consequently, WGS revealed an increased depth of a 16 bp segment at *EBP* exon 2 (NM_006579.3:c. 275_290), which was flanked by multiple clipped or discordant reads in the affected III-2 and II-2 (Figure 2A). This SV was detected as an insertion by Manta. Thus, we performed PCR amplification using the primers flanking the 16 bp segment, identifying a ~ 4.5 kb product in III-2 and II-2, but not in I-3 and I-4 (Figure 2B). Sequential sequencing of the ~ 4.5 kb PCR product revealed the presence of a 32 bp poly(T) tract, a poly(A) signal, an AluSx element, an SVA_F1 containing the *MAST2* exon 1 sequence, and a truncated AluSx element (5′→ 3′) between the duplicated 16 bp segments regarded as a target site duplication (TSD), as well as a consensus endonuclease cleavage site (5'-TCTTAT-3') on exon 2 (Figure 2C, Supplementary Figure 2) [12, 14, 19, 20], although the sequencing of the entire inserted region could not be performed due to GC-rich tandem repeats. These retrotransposons were inserted in an antisense orientation. The inserted sequence was found to share high homology with 10q24.2 (99,837,056–99,840,608) and 19q12 (29,897,923–29,902,184) (GRCh38) loci by BLAST search (https://genome.ucsc.edu/cgi-bin/hgBlat).

We also performed genotyping and X-inactivation analysis for the CAG repeat length

polymorphism at exon 1 of *AR* on Xq12 [21]. The results showed transmission of the X chromosome from I-3 (the maternal father) to II-2 and III-1, and random X-inactivation patterns in I-4, II-2, and III-2 (Supplementary Figure 3).

Discussion

WGS and subsequent studies successfully revealed a ~ 4.5 kb insertion disrupting *EBP* exon 2 in the affected III-2 and II-2, thereby confirming the diagnosis of CDPX2. The results argue for the value of WGS in the molecular diagnosis of human genetic diseases. Indeed, short read WGS has served to identify a broad range of disease-causing SVs, including CNVs, balanced translocations, inversions, and insertions, as well as pathogenic variants on the non-coding regions [22]. Although there is no gold standard tool for the detection of SVs, Manta is considered as one of the most useful analytic tools for detecting insertions. Furthermore, the present study also revealed that the insertion occurred in the X chromosome inherited from I-3 to II-2 as a *de novo* event.

The ~ 4.5 kb inserted sequence contained a poly(T) tract and was flanked by 16 bp segment from *EBP* exon 2 on its both sides. The poly(T) tract is derived from an antisense non-template poly(A) tail induced by a polyadenylation signal (5'-AATAAA-3'), and the duplicated 16 bp segment has the property of TSD [12]. Furthermore, a consensus endonuclease cleavage site was identified in *EBP* exon 2, where endonucleases cleave genomic DNA between T and A (5'-TCTT/AT-3') [8]. The liberated 3′ hydroxyl residue then serves as a primer used by ORF2p through a mechanism referred to as target-site priming reverse transcription [12]. The presence of such an antisense poly(A) tail, a TSD, and a consensus endonuclease cleavage site is characteristic of retrotransposition, and implies that the inserted sequence was templated from the RNA species and integrated by LINE-1 endonuclease and reverse transcriptase activities [12]. In this regard, it would be worth pointing out that the increased depth for the TSD in association with clipped or discordant reads served as the hallmark for the detection of retrotransposition in the WGS analysis.

The inserted sequence contained the full-length SVA_F1 and the 5′- and 3′-transductions containing the Alu sequence. For the 5'-transductions of the incorporated

SVA, the transcription of the SVA is regulated by a relatively strong external promoter rather than a relatively weak internal promoter, so that the upstream sequence is incorporated via the upstream transcription start site [14, 23], and the transcript passes through the SVA's own poly(A) signal and proceeds to the downstream poly(A) signal of the inserted 3'-transduction [24]. Thus, insertions containing 3'-transductions typically contain two poly(A) tracts, which is also a hallmark of LINE-1 mediated retrotransposition [25]. Since the Alu element is an excellent substrate for the reverse transcription by LINE-1 ORF2p protein, the presence of Alu in the newly acquired sequence would have made SVA a better target for LINE-1 mediated retrotransposition [14, 26].

The inserted sequence was found to show high homology with 10q24.2 and 19q12 loci which harbor an SVA_F1 element with 5′- and 3′-transductions containing the Alu sequence. The source locus of the inserted SVA_F1 sequence identified in this study is assessed to be 19q12 based on the size of the 5'-transduced Alu sequence [14]. In this regard, the locus on 10q24.2 is known to be one of the most active source elements from which multiple SVA_F1 insertions are derived, because SVA_F1 insertions generated from 10q24.2 contain common characteristic 3' transductions including Alu and poly(A) signals, as well as variably truncated 5' transductions, as revealed in this study [14, 19, 20]. It is likely, therefore, that the 19p12 sequence, which was generated by the insertion of 10q24.2 sequence followed by truncation during the evolution, was inserted into *EBP* of the mother by retrotransposition.

The skin lesion resolved in a few months in the proband, as has been reported previously [1]. This implies that *EBP* is required for the skin viability, so that cutaneous cells with inactive normal *EBP* and active abnormal *EBP* were gradually eliminated because of lethality, resulting in scar-like skin appearance and the predominance of cutaneous cells with active normal *EBP* and inactive abnormal *EBP* [27]. In this regard, since the two X chromosomes were randomly inactivated in the leukocyte genomic DNA samples of III-2 and II-2, this would suggest that *EBP* plays no critical role in the leukocyte viability, so that leukocytes with inactive normal *EBP* and active abnormal *EBP* remained viable, as well as those with active normal *EBP* and inactive abnormal *EBP*.

In summary, we identified a ~ 4.5 kb insertion into *EBP* via the LINE-1 mediated

retrotransposition event in two familial patients with CDPX2. The results imply that retrotransposition is an important mechanism for genetic diseases and is worth considering as a possible cause for genetic diseases, especially for those of an unknown etiology.

Acknowledgments

We would like to thank the patients for participating in this work. We also thank Ms. Aya Kitamoto, and Mr. Naoki Adachi for their technical support.

Conflict of Interest

The authors have nothing to disclose.

References

1. Happle R. X-linked dominant chondrodysplasia punctata. Review of literature and report of a case. Hum Genet. 1979;53:65-73.
2. Braverman N, Lin P, Moebius FF, Obie C, Moser A, Glossmann H, et al. Mutations in the gene encoding 3 beta-hydroxysteroid-delta 8, delta 7-isomerase cause X-linked dominant Conradi-Hünermann syndrome. Nat Genet. 1999;22:291-294.
3. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. Nat Rev Genet. 2002;3:370-379.
4. Hancks DC, Kazazian HH Jr. Roles for retrotransposon insertions in human disease. Mob DNA. 2016;7:9.
5. Hohjoh H, Singer MF. Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. EMBO J. 1997;16:6034-6043.
6. Martin SL, Bushman FD. Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Mol Cell Biol. 2001;21:467-475.
7. Mathias SL, Scott AF, Kazazian HH Jr, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. Science. 1991;254:1808-1810.
8. Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell. 1996;87:905-916.
9. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet. 2003;35:41-48.
10. Ostertag EM, Goodier JL, Zhang Y, Kazazian HH Jr. SVA elements are nonautonomous retrotransposons that cause disease in humans. Am J Hum Genet. 2003;73:1444-1451.
11. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. Nat Genet. 2000;24:363-367.
12. Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, et al. Human L1 retrotransposition: cis preference versus trans complementation. Mol Cell Biol. 2001;21:1429-1439.
13. Wang H, Xing J, Grover D, Hedges DJ, Han K, Walker JA, et al. SVA elements: a hominid-specific retroposon family. J Mol Biol. 2005;354:994-1007.

14. Damert A, Raiz J, Horn AV, Löwer J, Wang H, Xing J, et al. 5'-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. Genome Res. 2009;19:1992-2008.
15. Honda A, Yamashita K, Miyazaki H, Shirai M, Ikegami T, Xu G, et al. Highly sensitive analysis of sterol profiles in human serum by LC-ESI-MS/MS. J Lipid Res. 2008;49:2063-2073.
16. Honda A, Miyazaki T, Ikegami T, Iwamoto J, Yamashita K, Numazawa M, et al. Highly sensitive and specific analysis of sterol profiles in biological samples by HPLC-ESI-MS/MS. J Steroid Biochem Mol Biol. 2010;121:556-564.
17. Hiraide T, Nakashima M, Ikeda T, Tanaka D, Osaka H, Saitsu H. Identification of a deep intronic POLR3A variant causing inclusion of a pseudoexon derived from an Alu element in Pol III-related leukodystrophy. J Hum Genet. 2020;65:921-925.
18. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics. 2016;32:1220-1222.
19. Bantysh OB, Buzdin AA. Novel family of human transposable elements formed due to fusion of the first exon of gene MAST2 with retrotransposon SVA. Biochemistry (Mosc). 2009;74:1393-1399.
20. Hancks DC, Ewing AD, Chen JE, Tokunaga K, Kazazian HH Jr. Exon-trapping mediated by the human retrotransposon SVA. Genome Res. 2009;19:1983-1991.
21. Miyamoto S, Nakashima M, Ohashi T, Hiraide T, Kurosawa K, Yamamoto T, et al. A case of de novo splice site variant in SLC35A2 showing developmental delays, spastic paraplegia, and delayed myelination. Mol Genet Genomic Med. 2019;7:e814.
22. Neerman N, Faust G, Meeks N, Modai S, Kalfon L, Falik-Zaccai T, et al. A clinically validated whole genome pipeline for structural variant detection and analysis. BMC Genomics. 2019;20:545.
23. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, et al. Human l1 retrotransposition is associated with genetic instability in vivo. Cell. 2002;110:327-338.
24. Moran JV, DeBerardinis RJ, Kazazian HH Jr. Exon shuffling by L1 retrotransposition.

Science. 1999;283:1530-1534.

25. Szak ST, Pickeral OK, Landsman D, Boeke JD. Identifying related L1 retrotransposons by analyzing 3' transduced sequences. Genome Biol. 2003;4:R30.

26. Cost GJ, Feng Q, Jacquier A, Boeke JD. Human L1 element target-primed reverse transcription in vitro. EMBO J. 2002;21:5899-5910.

27. Steijlen PM, van Geel M, Vreeburg M, Marcus-Soekarman D, Spaapen LJ, Castelijns FC, et al. Novel EBP gene mutations in Conradi-Hünermann-Happle syndrome. Br J Dermatol. 2007;157:1225-1229.

Figure legends

Figure 1. Representative clinical findings.

A. The pedigree of this family.

B–D. Photographs of the proband (III-2).

E–H. Roentgenograms the proband (III-2). The arrowhead represents hemivertebra of the fourth thoracic vertebra.

Figure 2. Summary of molecular studies.

A. IVG screenshot of WGS pair-end reads in the proband, showing an increased depth of 16 bp segment regarded as a target site duplication (TSD) and clipped and discordant reads clustered at the boundaries of the 16 bp segment. Discordant reads shown with different colors have partners aligned to different chromosomes with aberrant alignments, and this suggests the presence of inserted sequences. The arrows indicate the positions of PCR primers (forward and reverse) utilized for the PCR amplification shown in Figure 2B.

B. PCR products obtained with primers shown in Figure 2A. A ~ 4.5 kb PCR product is delineated in III-2 (the proband) and II-2 (the mother), but not in the maternal father (I-3) and mother (I-4). NC, negative control.

C. Schematic presentation showing retrotransposition disrupting *EBP* exon 2. The inserted region is flanked by TSD and contains a 32-bp poly(T) tract, a poly(A) signal, an AluSx element, an SVA_F1 element containing *MAST2* exon 1 sequence (indicated with a red rectangle), and a truncated AluSx element. Consensus ECS (endonuclease cleavage site) is present on exon 2 (highlighted with light blue background).

Figure 1.

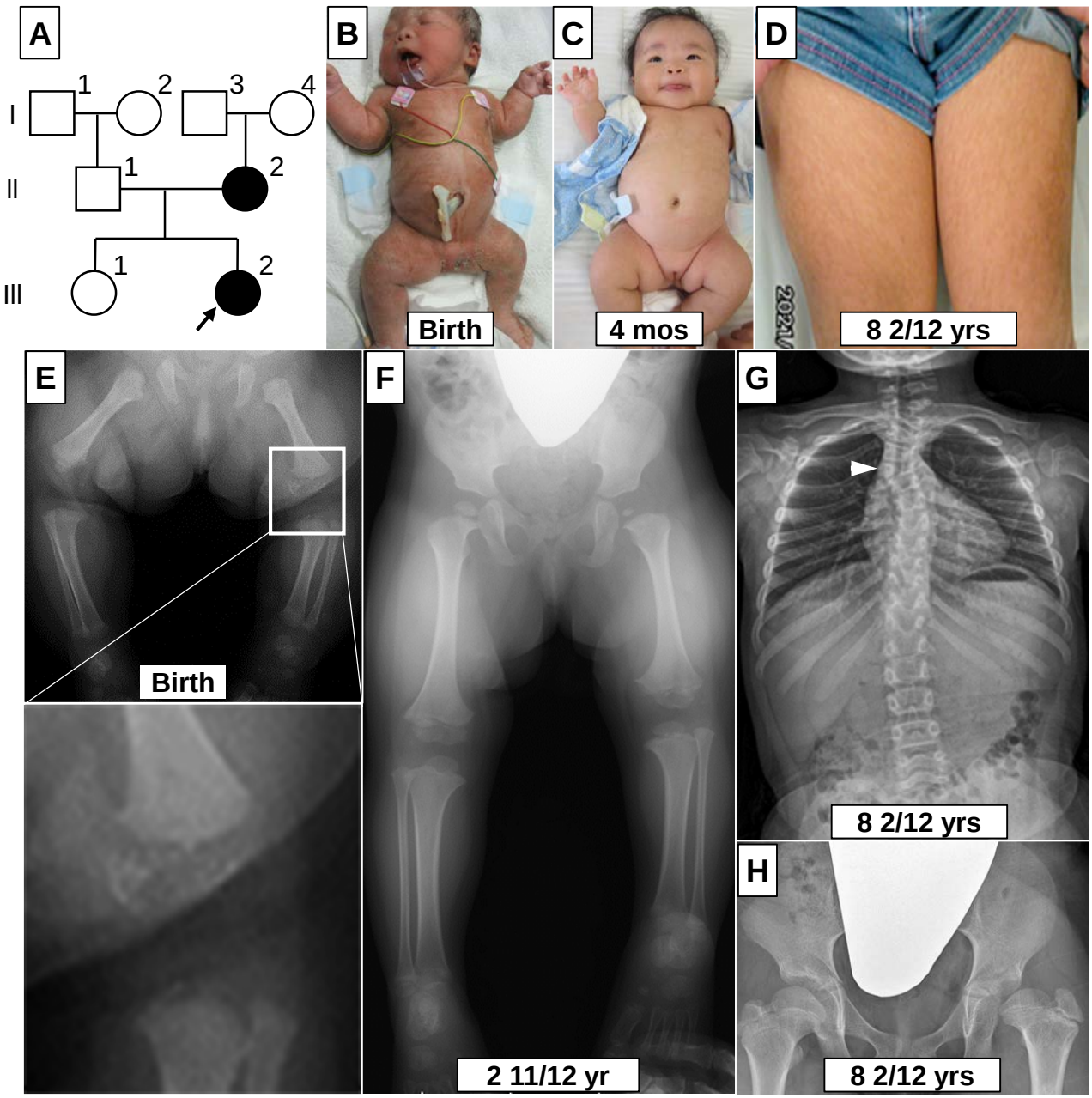
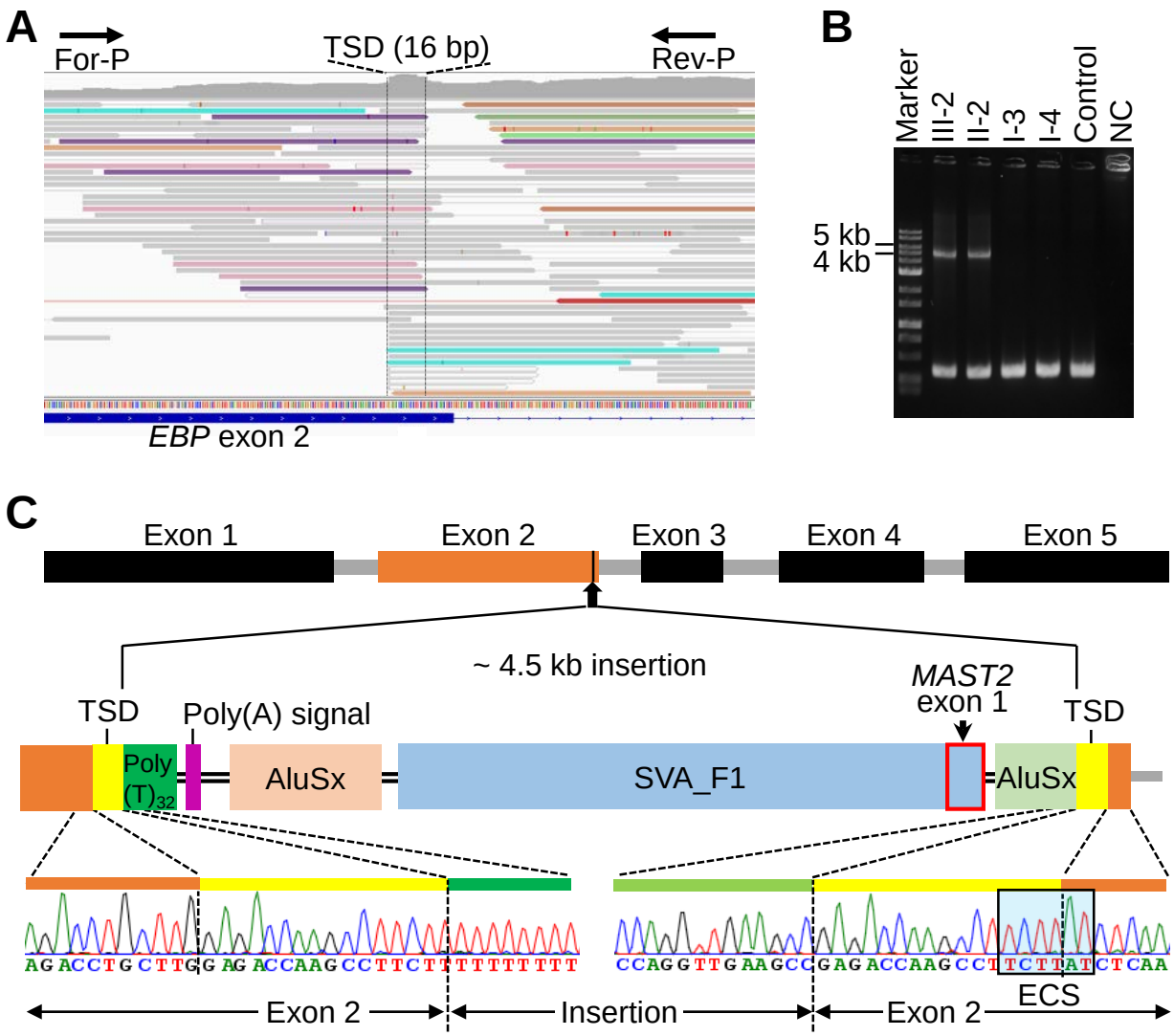
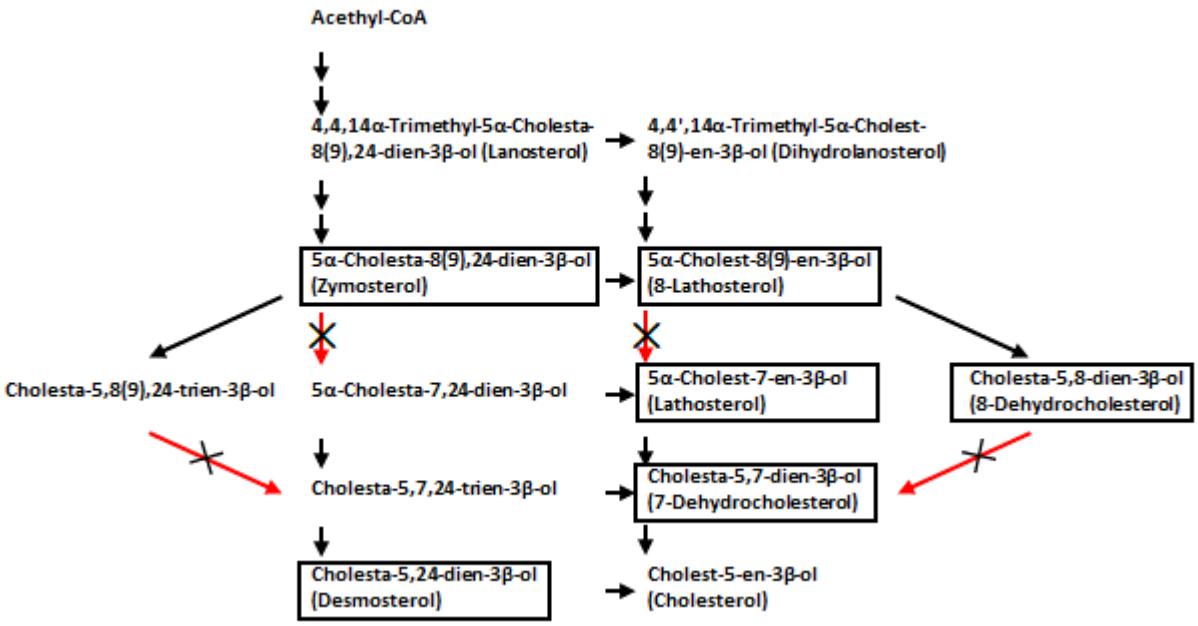

A
1
2
3
4
I
1
2
II
1
2
III
B
Birth
C
4 mos
D
8 2/12 yrs
E
Birth
F
2 11/12 yr
G
8 2/12 yrs
H
8 2/12 yrs

Figure 2.


A
For-P
TSD (16 bp)
Rev-P
EBP exon 2
B
Marker
III-2
II-2
I-3
I-4
Control
NC
5 kb
4 kb
C
Exon 1
Exon 2
Exon 3
Exon 4
Exon 5
~ 4.5 kb insertion
TSD
Poly(A) signal
MAST2
exon 1
TSD
Poly
(T)32
AluSx
SVA_F1
AluSx
AGACCTGCTTGGAGACCAAGCCTTCTTTTTTTTTT
CCAGGTTGAAGCCGAGACCAAGCCTTCTTATCTCAA
ECS
Exon 2
Insertion
Exon 2

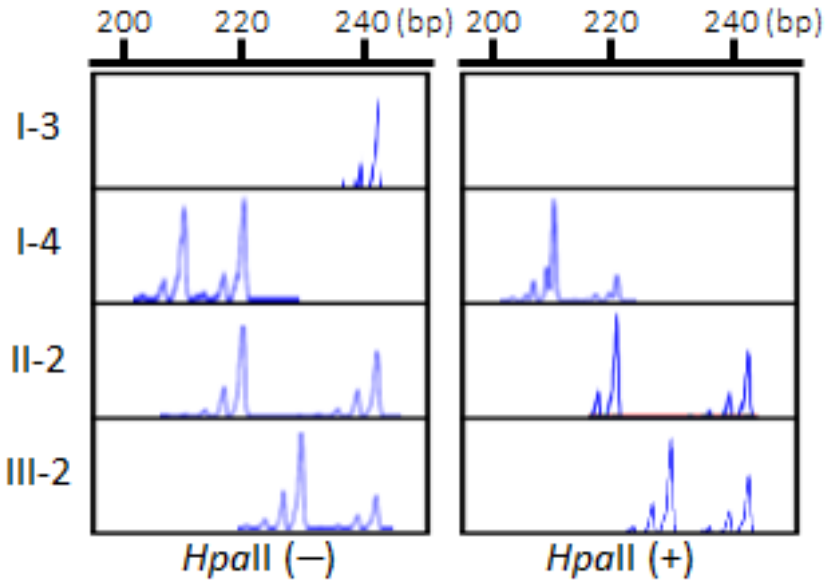


Supplementary Figure 1. Simplified cholesterol biosynthetic pathway. The metabolic conversions mediated by 3β-hydroxysteroid Δ^8,Δ^7-sterol isomerase encoded by *EBP* are shown with red arrows, and these conversions are impaired in CDPX2 (shown with crosses). The sterol intermediates measured in the proband and her mother are surrounded with rectangles.

AGTCGTGACCACATGGCTGTTGTCAGGTCGTGCTGCGGTTGTCCCATTGGGGACTTGGCGGCGACTGTCCCTGTG
CTGGTTTGCAGTGTGTGGGTTCATTCACCTGGTGATCGAGGGCTGGTTCGTTCTCTACTACGAAGACCTGCTTGG
AGACCAAGCCTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTAAAGGACATATCATATTTATTCATAC
ACATGCTGGAATTATTGGTGCAGACATTTAAATACATTTTCTTTGAGAAAGTCCTTTTTTTTTTTTTTTTTGAT
GGAGTTTCCCTCTTGTTGCCCAGGCTGGAGTGCAATGGTGCAATCTCAGCTCACAACAACCTCTGCCTCCTGGGT
TCAAGCAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATGCACCACCACGCCCAGCTAATTTTT
TTTATTTTTAGTAGAGACGGGGTTTCTCCGTGTTGGTCAGGCTGGTCTTGAACTCCTGATCTCAGGTGATCTGCC
CACCTTGGCCTGCCACAGTGCTGGGATTACAGTCGTGAGCCACCACAGCTGGCCTGGGAAAGTCCATTCTTTTTT
TTCTTTTTTTTTTTTTTTTTTAAATTTATTTTTTTATTGATAATTCTTGGGTGTTTCTCACAGAGGGGGATTTGGC
AGGGTCATGGGACAATAGTGGAGGGAAGGTCAGCAGATAAACAAGTGAACAAAGGTCTCTGGTTTTCCTAGGCAG
AGGACCCCGCGGCCTTCCGCAGTGTTTGTGTCCCTGATTACTTGAGATTAGGGATTGGTGATGACTCTTAACGAG
CATGCTGCCTTCAAGCATCTGTTTAACAAAGCACATCTTGCACTGCCCTTAATCCATTTAACCCTGAGTGGACAC
AGCACATGTTTCAGAGAGCACAGGGTTGGGGGTAAGGTCACAGATCAACAGGATCCCAAGGCAGAGGAATTTTTC
TTAGTGCAGAACAAAATGAAAAGTCTCCCATGTCTACTTCTTTCTACACAGACACGGCAACCATCCGATTTCTCA
ATCTTTTCCCCACCTTTCCCGCCTTTCTATTCCACAATGCCGCCATTGTCATCCTGGCCCGTTCTCAATGAGCTG
TTGGGCACACCTCCCAGACGGGGTGGTGGCCGGGCAGAGGGGCTCCTCACTTCCCAGTAGGGGCGGCCGGGCAGA
GGCGCCCCTCACCTCCCGGACGGGGCGGCTGGCCGGGCAGGGGGGCTGACCCCCCCCACCTCCCTCCCGGACGGG
GCGGCTGGCCGGGCGGGGGGCTGACCCCCCAACCTCCCTCCCGGACGGGGCGGCGGGCGGAGAGGCTCCTTACTT
----------(GC-rich tandem repeats)----------
CTCAGACGGGGCAGCTGCCGGGCGGAGGGGCTCCTCACTTCTCAGACGGGGTGGTTGCCAGGCAGAGGGTCTCCT
CACTTCTCAGACGGGGCGGCCGGGCAGAGACGCTCCTCACCTCCCAGACGGGGTCTCGGCCGGGCAGAGGCGCTC
CTCACATCCCAGATGGGGCGGAGGGGCAGAGGCGCTCCCCACATCTCAGACGATGGGCGGCCGGGCAGAGACGCT
CCTCACTTCCTAGATGTGATGGCGGCTGGGAAGAGGCGCTCCTCACTTCCTAGATGGGATGGCGGCCGGGCGGAG
ACGCTCCTCACTTTCCAGACTGGGCAGCCAGGCAGAGGGGCTCCTCACATCCCAGACGATGGGCGGCCAGGCAGA
GACACTCCTCACTTCCCAGACGGGGTGGCAGCCGGGCAGAGGCTGCAATCTCGGCACTTTGGGAGGCCAAGGCAG
GCGGCTGCTCCTTGCCCTCGGGCCCCGCGGGGCCCGTCCGCTCCTCCAGCCGCTGCCTCCCGGGCGGCGCTCGCC
GGCGCGGCGGCAAAGACTGAGACAGCTCCGCTGCCCGCTGAACTCCATCCTCCCGGCGGTCGGGCGGCGGCGGCT
GCGGTCGGTCGCGGCAGCGGCTCCGCTTCATATCTGCAGCTGGGGCCCGCGGGCGTCAGCGCCGCGACTGTCCTG
GCTCCGCACTGCCCCGGGCCGCAGCGCAGCCGCGCCAACCACCAGCCGCGGCCACCATGGCCGGACGGGCTCCCT
AAGCCACCGACCCCAGCCCGCGGCGCCTTCGACCCTTCTGGGGCCTCCGGCGCCGCGACCTCCTCTGCCTGAAAT
TTCTTTTTTCTTTTCCTTTTATTTTATTTTATTTTTTGAGACGGAGTCTTGCTCTGTTGTCTGGGTGGAGTGCAG
TGGTGCAATCTCGGCTCACTGCAACCTCTGCCTCCCAGGTTGAAGCCGAGACCAAGCCTTCTTATCTCAACTCTg
tgagtcctgatttctttcatatgctgtgggatgggatttgctgggcagggatcggcttgcatgtttacctatcca
cctattcttcttccattgatttattttaaatttttatttaatcttttaaaaaatttattataggccgggcacagt
ggctcacgcttgtaatcccagcactttgggaggccgaggcaggtggatcacttgaggtcaggaatttgagaccag
cctggtcaa

5′ – *EBP* exon 2 (+ strand) – Target site duplication – Non-templated PolyA (– strand) –Polyadenylation signal – AluSx (– strand) – SVA_F1 (– strand) – Truncated AluSx (– strand) – Target site duplication – *EBP* exon 2 (+ strand) – *EBP* intron 2 (+ strand) – 3′

Supplementary Figure 2. DNA sequence of the ~4.5 kb region inserted between the duplicated 16 bp segments (target site duplication). The inserted region contains a 32-bp poly(T) tract, a poly(A) signal, an AluSx element, an SVA_F1 element, and a truncated AluSx element. The red underline indicates the exon 1 sequence of *MAST2*, and the blue rectangle denotes LINE-1 endonuclease cleavage site.



Supplementary Figure 3. Genotyping and X-inactivation analysis for the CAG repeat length polymorphism at exon 1 of *AR* on Xq12. PCR products are obtained from both active and inactive X chromosomes before *Hpa*II digestion and from inactive X chromosomes alone after *Hpa*II digestion. The X chromosome of I-3 is inherited by II-2 and III-2. X-inactivation ratio is calculated as 38%:62% in the proband (III-2), 50%:50% in the mother (II-2), and 79%:21% in the grandmother (I-4), after compensation of the unequal amplification between two peaks that is consistent with short products being more easily amplified than long products.

Supplementary Table 1. Serum sterol metabolite values (μg/mL) and ratios of substrates/products converted by 3β-hydroxysteroid Δ^8,Δ^7-sterol isomerase.

	Proband (4.5 mos)	Mother (33 yrs)	CDPX2 patient[a] (child)	Control children (n=37)[b] (8.6 ± 1.5 yrs)	Control adults (n=19)[c] (adults)
Zymosterol	1.97	1.35	2.30	0.26 ± 0.10	0.63 ± 0.36
8-Lathosterol	9.86	9.63	28.21	0.70, 3.59	0.77, 2.36
8-Dehydrocholesterol (8-DHC)	12.27	6.02	4.13	1.38 ± 0.34	2.49 ± 1.44
Desmosterol	0.81	0.84	0.77	0.71 ± 0.16	0.69 ± 0.27
Lathosterol	3.64	4.50	6.59	2.93 ± 1.19	6.12 ± 4.87
7-Dehydrocholesterol (7-DHC)	1.05	1.26	0.13	1.09 ± 0.34	3.81 ± 1.48
8-DHC/7-DHC ratio	11.68	4.78	31.77	1.26[d]	0.65[d]

The metabolic map for cholesterologenesis is shown in Supplementary Figure 1.

[a] A previously identified patient with CDPX2.

[b] Unpublished data obtained from 37 children, except for 8-Lathosterol values measured in two children only.

[c] Published data obtained from 19 adults [1], except for unpublished 8-Lathosterol values measured in two adults only.

[d] The ratios of the mean values.

The control values are described in mean ± SD.

Reference

1. Honda A, Yamashita K, Miyazaki H, Shirai M, Ikegami T, Xu G, et al. Highly sensitive analysis of sterol profiles in human serum by LC-ESI-MS/MS. J Lipid Res 2008;49:2063-73.