J Tokyo Wom Med Univ
83 (Extra) E20～E24 (2013)

Original

Identification of a Duplication Breakpoint in the *DMD* Gene Using Array Comparative Genomic Hybridization

Takashi SAITO[1,2], Keiko ISHIGAKI[2], Terumi MURAKAMI[2], Takatoshi SATO[2], Sachiko KAJINO[2], Shin'ichi TAKEDA[1] and Makiko OSAWA[2]

[1]Department of Molecular Therapy, National Institute of Neuroscience, National Center of Neurology and Psychiatry
[2]Department of Pediatrics, Tokyo Women's Medical University School of Medicine

(Accepted December 25, 2012)

Mutations in the *DMD* gene cause Duchenne/Becker muscular dystrophy (DMD/BMD). Full characterization of the mutations, including the analysis of deletion or duplication breakpoints, is diagnostically and therapeutically beneficial. To identify duplication breakpoints, the array comparative genomic hybridization (aCGH) method was used, covering the whole *DMD* gene. For the previously identified duplication of exons 5–7, DMD-aCGH revealed the duplication at a higher resolution, and enabled detection of the breakpoint junction. The 29 kb duplication from intron 4 to 7, with its precise breakpoint junction sequence, was determined. Additionally we found a complicated rearrangement, fork stalling and template switching, in the breakpoint.

Key Words: Duchenne muscular dystrophy, duplication mutation, array comparative genomic hybridization, fork stalling and template switching

Introduction

Mutations in the *DMD* gene (MIM 310200) cause DMD/BMD. Deletions of one or more exons are most frequent, occurring in approximately two-thirds of all patients. Large duplications have been reported in approximately 6% of all DMD/BMD cases[1)2)]. Genetic testing of the *DMD* gene is the initial method for confirming the diagnosis. Multiplex ligation-dependent probe amplification (MLPA) analysis, which is used to examine every exon for deletion and/or duplication, has contributed to a marked improvement in the mutation detection rate[3)]. However, full characterization of the mutational spectrum, including the analysis of deletion and duplication breakpoints, is desirable for genetic counseling and eligibility assessment, ultimately leading to mutation-specific therapy, such as an antisense mediated exon-skipping[4)5)]. It is reported that the length of the flanking introns affects the dynamics of splicing; therefore, a determination of the breakpoint junction and the length of an intron is important in an antisense therapy[6)]. We planned an antisense mediated exon-skipping assay in the duplicated *DMD* gene; therefore, we performed a breakpoint analysis before the assay. Compared with the characterization of deletion breakpoints, duplication breakpoint analysis is more challenging. It usually requires a technically laborious long-range PCR to obtain a PCR fragment containing the duplication breakpoint[6)]. The aCGH method has been widely used to identify chromosomal copy number and structural changes, at a high resolution. To identify the *DMD* duplication breakpoint, we chose the CytoSure DMD Array, a commercially available "DMD-aCGH" with comprehensive coverage of the whole *DMD* gene, which was used in a previous study of copy number variations in *DMD* gene[7)]. This paper reports the rapid characterization of the duplication breakpoint, utilizing high-resolution aCGH, customized for the *DMD* gene.

Materials and Methods

1. Samples

As the reference normal human genome, TIG-119

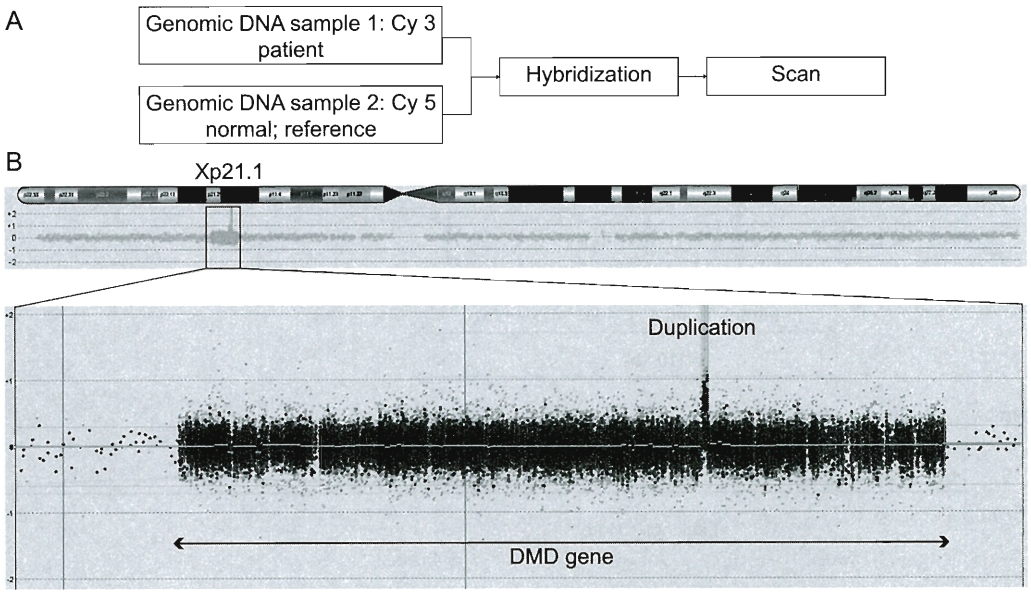

Fig. 1 Copy number changes in the *DMD* gene detected on the DMD-aCGH platform
(A) Overview of the DMD-aCGH procedure. (B) Each unique 60-mer oligonucleotide probe (gray dot) is represented by a data point along the X-axis, based on its physical position at the Xp21.1 *DMD* locus. In the magnified view of the Xp21.1 locus, probe (gray dot with black circle) copy number gains are shown at the duplication involving *DMD* exons 5-7.

fibroblasts were obtained from the Health Science Research Resource Bank (Osaka, Japan). As the DMD patient genome, GM04327 fibroblasts (clinically diagnosed DMD, exons 5–7 duplications identified by MLPA) were obtained from the Coriell Cell Repositories (Camden, NJ, USA). DNA was prepared from each of the fibroblast samples using the Wizard SV Genomic DNA Purification system (Promega, Fitchburg, WI, USA).

2. DMD-aCGH

The CytoSure DMD Array (Oxford Gene Technology, Oxford, UK) was used in this study. It comprises 44,000 probes of 60mer oligonucleotides that cover the whole *DMD* gene on a single array. The average probe spacing is 10 bp within the exons and 106 bp within the introns. The restriction digestion of genomic DNA, the labeling of the DNA, the hybridization of arrays with labeled target and the scanning of arrays was performed according to the manufacturer's instructions. Briefly, patient samples and normal reference samples were labeled with Cy5 and Cy3, respectively. After hybridization on the DMD-aCGH, the scanned fluorescent signal was analyzed (Fig. 1A). Data analysis was performed using CytoSure Interpret Software ver. 3.4.6 (Oxford Gene Technology).

3. Breakpoint PCR

Two intron 7 forward primers and two intron 4 reverse primers were designed based on the DMD-aCGH result (primer sequences are available on request). Each primer pair flanked the duplication breakpoint junction, and was expected to yield PCR products within the range of 1–15 kb. PCR was performed using KOD FX (Toyobo, Osaka, Japan), and the cycling program was set to yield 15 kb products with a program of 35 cycles of 98℃ for 10 sec, 60℃ for 30 sec, and 72℃ for 450 sec. One of the four primer pairs yielded a 3 kb PCR fragment. Primer walk sequencing was performed on the fragment (Operon Biotechnologies, Tokyo, Japan). Human genome sequence NCBI Build 36.1 was used as the reference genomic sequence.

Results

In the chromosomal overview, high-density probes were mapped to the Xp21.2 region. Magnification of the region indicated an at least 24 kb gain aberration from intron 4 to intron 7, consistent with the result of MLPA analysis (Fig. 1B). We selected four representative probes near the breakpoint junction. Probes 1 and 2 were from intron 4, and

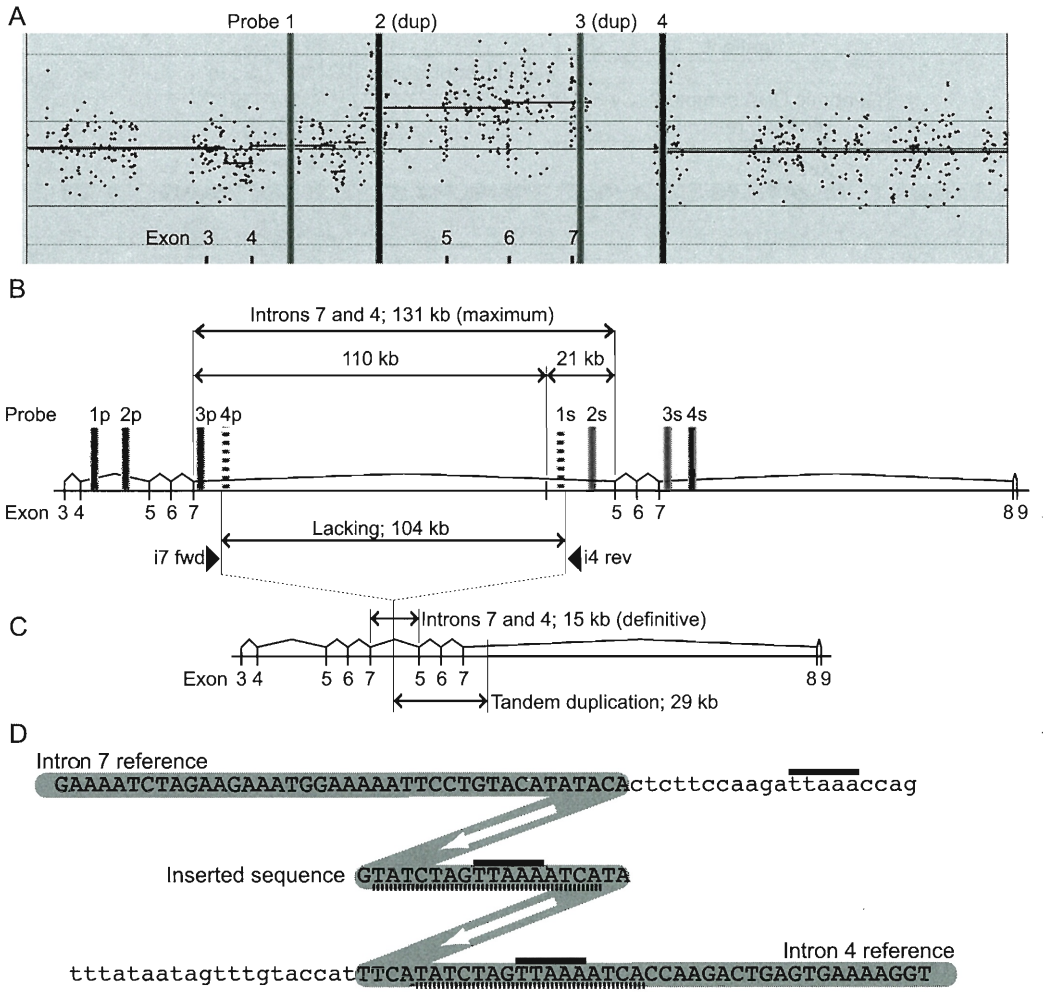

Fig. 2 Detection of breakpoint junction in the exons 5-7 duplication
(A) Each probe (black dot) is displayed horizontally along with its genomic position and vertically by the signal. Probes 1-4 (gray vertical bar) were speculated to be positioned near the breakpoint junction. (B) In a tandem duplication model, probes 1-4 were mapped twice (1-4p: primary, 1-4s: secondary). Probe 4p and 1s (gray dashed vertical bar) were speculated to be in the region missing from the genome. Breakpoint detection primers: an intron 7 forward primer (i7 fwd) and an intron 4 reverse primer (i4 rev) were designed between probes 3p and 4p, and 1s and 2s, respectively. (C) The rearranged concatenation from intron 7 to intron 4 was 15 kb. The duplication size was 29 kb. (D) Bases written in upper case with a gray background are the sequence of the breakpoint junction. "TTAAA" (solid line) is common between introns 7 and 4, and is inserted into the breakpoint. "TATCTAGTTAAAATCA" (dashed line) is common between the inserted sequence and the breakpoint of intron 4.

probes 3 and 4 were from intron 7. From the mean signal ratio of these probes, we speculated that probes 1 and 4 were positioned in the non-duplicated region and probes 2 and 3 were positioned in the duplicated region (Fig. 2A). The tandem duplication model was used; the tandem duplication of introns 7 and 4 was a maximum of 131 kb, and probes 1–4 were mapped twice (1–4p: primary, 1–4s: secondary). Probes 4p and 1s were speculated to be in the region missing from the tandem intron; therefore, the 104 kb region between probes 4p and 1s was excluded from design of the breakpoint primers. We designed an intron 7 forward primer between probes 3p and 4p, and an intron 4 reverse primer between probes 1s and 2s (Fig. 2B). A primer pair flanking the breakpoint junction yielded a 3 kb PCR product, which revealed that the rearranged concatenation from intron 7 to in-

tron 4 was 15 kb (Fig. 2C). The sequence analysis of the 3 kb PCR fragment revealed the genotype; arr Xp21.1 (32,731,239 – 32,760,260) ×2, and the duplication size was 29 kb. In the breakpoint, we identified a 19 bp insertion "GTATCTAGTTAAAATCATA". (Fig. 2D). "TTAAA" is common between intron 4 and 7. The inserted sequence of "TCTAGTTAAA ATCA" is also found in intron 4. This result suggested that the fork stalling and template switching (FoSTeS) mechanism had mediated the genomic rearrangement in this case.

Discussion

The characteristics of a breakpoint provide an insight into complex chromosomal rearrangements[6]. In determining a breakpoint junction, the correct design of PCR primers covering the breakpoint regions is essential. In this case, if the conventional long-range PCR approach was chosen, candidate primer regions would be spread throughout the duplicated region, and a trial-and-error approach would be necessary for successful PCR. The DMD-aCGH technique narrowed the candidate primer region from 131 kb to 17 kb, which dramatically accelerated the design of primers, quickly achieved successful PCR, and provided a definitive sequence of the breakpoint.

In the *DMD* gene, the duplications are evenly spread throughout the gene. An exception is the duplication of exon 2, which is the most common single-exon duplication[2]. In the Leiden Muscular Dystrophy Database (http://www.dmd.nl/), we found only 17 cases of exons 5–7 duplication, compared to 113 cases of exon 2 duplication[8]. Among the 17 cases, only one case had a reported breakpoint determined by custom-designed aCGH, and was genotyped as arr Xp21.1 (32,725,684 – 32,765,708) ×2; the minimum duplication size was 40.0 kb[9]. Compared with our result, the duplication size of this case was 11.0 kb longer, and the intron size was 5.5 kb longer for introns 4 and 7. A study of 11 cases of exon 2 duplication showed that the breakpoints in intron 1 were relatively scattered, whereas the 10 breakpoints in intron 2 were clustered in the first 40 kb region[10]. The apparent clustering of breakpoints within intron 2 suggested a breakpoint hotspot in that particular intron; however, analysis of the two cases of exons 5–7 duplication provided no apparent bias in breakpoint distribution within intron 4 and 7.

The breakpoint analysis suggested that the FoSTeS mechanism had mediated the genomic arrangement in this case. FoSTeS is proposed as a microhomology-mediated replication error mechanisms and has been found previously in a *DMD* gene rearrangement[11)12)]. The revealed rearrangement was a tandem duplication, but contained 5 bp of microhomology and 16 bp of replication from intron 4. A large-scale breakpoint analysis study demonstrated that microhomology-mediated processes, including FoSTeS, account for 28% of observed rearrangements[13].

Conclusion

The DMD-aCGH platform was used to detect genomic rearrangements involving the *DMD* gene. It revealed a 29 kb duplication from intron 4 to 7 and a complicated FoSTeS rearrangement at the breakpoint.

Declarations

The authors declare no financial conflict of interest.

References

1) **Hu XY, Ray PN, Murphy EG et al**: Duplicational mutation at the Duchenne muscular dystrophy locus: its frequency, distribution, origin, and phenotypegenotype correlation. Am J Hum Genet **46**: 682–695, 1990
2) **White S, Kalf M, Liu Q et al**: Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. Am J Hum Genet **71**: 365–374, 2002
3) **Janssen B, Hartmann C, Scholz V et al**: MLPA analysis for the detection of deletions, duplications and complex rearrangements in the dystrophin gene: potential and pitfalls. Neurogenetics **6**: 29–35, 2005
4) **Aartsma-Rus A, Janson AA, Van Ommen GJ et al**: Antisense-induced exon skipping for duplications in Duchenne muscular dystrophy. BMC Med Genet **8**: 43, 2007
5) **Saito T, Nakamura A, Aoki Y et al**: Antisense PMO found in dystrophic dog model was effective in cells from exon 7-deleted DMD patient. PLoS One **5**: e12239, 2010
6) **Gualandi F, Rimessi P, Trabanelli C et al**: Intronic breakpoint definition and transcription analysis in DMD/BMD patients with deletion/du-

plication at the 5' mutation hot spot of the dystrophin gene. Gene **370**: 26–33, 2006
7) **Ankala A, Kohn JN, Hegde A et al**: Aberrant firing of replication origins potentially explains intragenic nonrecurrent rearrangements within genes, including the human DMD gene. Genome Res **22**: 25–34, 2012
8) **White SJ, Den Dunnen JT**: Copy number variation in the genome; the human DMD gene as an example. Cytogenet Genome Res **115**: 240–246, 2006
9) **Del Gaudio D, Yang Y, Boggs BA et al**: Molecular diagnosis of Duchenne/Becker muscular dystrophy: enhanced detection of dystrophin gene rearrangements by oligonucleotide array-comparative genomic hybridization. Hum Mutat **29**: 1100–1107, 2008
10) **White SJ, Aartsma-Rus A, Flanigan KM et al**: Duplications in the DMD gene. Hum Mutat **27**: 938–945, 2006
11) **Lee JA, Carvalho CM, Lupski JR**: A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. Cell **131**: 1235–1247, 2007
12) **Oshima J, Magner DB, Lee JA et al**: Regional genomic instability predisposes to complex dystrophin gene rearrangements. Hum Genet **126**: 411–423, 2009
13) **Kidd JM, Graves T, Newman TL et al**: A human genome structural variation sequencing resource reveals insights into mutational mechanisms. Cell **143**: 837–847, 2010

アレイ CGH を用いた DMD 遺伝子の重複変異領域の断端解析

[1]独立行政法人国立精神・神経医療研究センター神経研究所遺伝子疾患治療研究部
[2]東京女子医科大学医学部小児科学

齊藤（サイトウ）　崇（タカシ）[1,2]・石垣（イシガキ）　景子（ケイコ）[2]・村上（ムラカミ）てるみ（テルミ）[2]・佐藤（サトウ）　孝俊（タカトシ）[2]
梶野（カジノ）　幸子（サチコ）[2]・武田（タケダ）　伸一（シンイチ）[1]・大澤（オオサワ）眞木子（マキコ）[2]

DMD 遺伝子の変異は Duchenne/Becker 型筋ジストロフィーを引き起こす．同遺伝子の変異領域の断端（breakpoint）解析は，同疾患に対する遺伝カウンセリングならびにエクソン・スキッピングなどの変異特異的な治療法の検討において有用な情報となる．断端解析を行うにあたっては，欠失変異では短いゲノム PCR 産物の有無により欠失を推測できるのに対し，重複変異では技術的に高度な long-range PCR が要求される．我々は重複変異における断端解析を簡易化するために，DMD 遺伝子全領域をカバーするアレイ CGH（DMD-aCGH）の有用性を検討した．MLPA 法によりあらかじめエクソン 5–7 の重複変異が同定されたサンプルを用いて検証したところ，DMD-aCGH によりゲノム上の重複の全長と断端の位置が容易に推測可能であった．最終的にはイントロン 4 から 7 にかけての 29 kb の重複領域，ならびに断端接合部の位置と塩基配列を同定できた．切断点の解析から，本例における重複の発生機構として FoSTeS（Fork Stalling and Template Switching）の関与が示唆された．