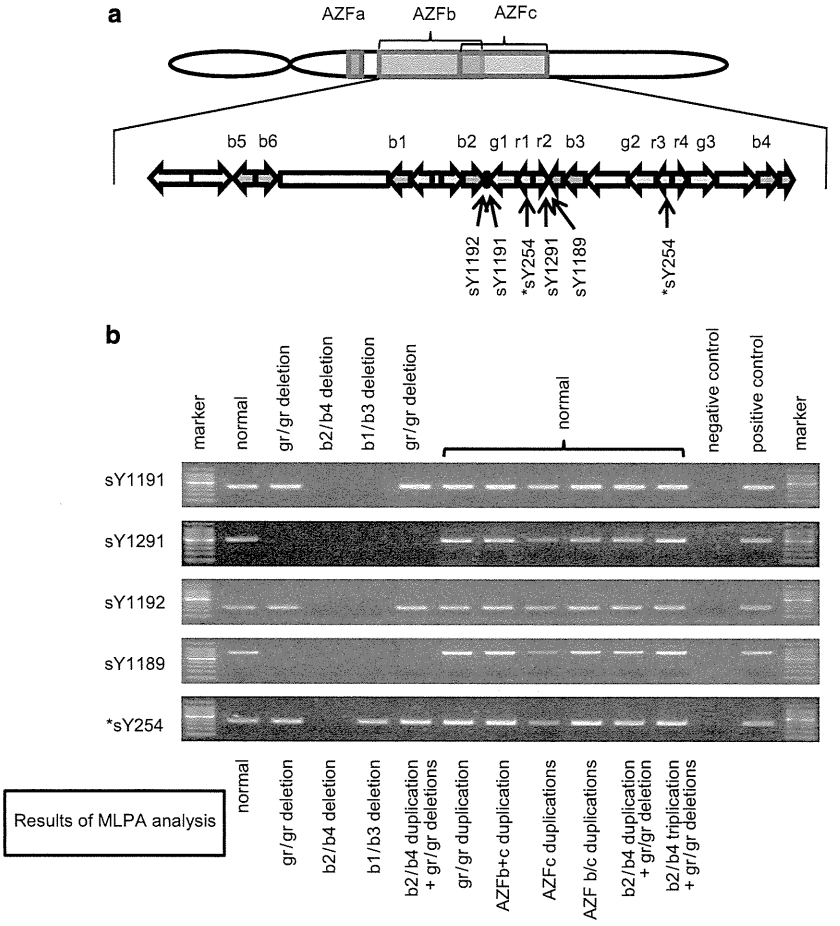


Figure 2 PCR analyses for sequence-tagged site (STS) markers. (**a**) Genomic positions of five STS markers. Of the five markers, sY254 is located at multiple sites (asterisks). (**b**) Representative results of STS-PCR. This method was able to identify four of the 11 copy-number variations identified by multiplex ligation-dependent probe amplification (MLPA). AZF, azoospermia factor. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

CNVs were more frequently identified in the patient group (33 of 56, 58.9%) than in the control group (25 of 65, 38.5%; $P=0.025$; Table 1). Although simple deletions and complex rearrangements were observed in the two groups at similar frequencies ($P=0.224$ and $P=0.331$, respectively), simple duplications were detected exclusively in the patient group ($P=0.043$). Multiple duplications were identified in one patient and one control male. The frequencies of copy-number gain of multi-copy genes were similar between the two groups (Supplementary Table 1).

DISCUSSION

We performed MLPA for 121 Japanese individuals and identified 11 types of CNVs in AZF regions. These CNVs were detected in ~60% of SF patients and in ~40% of control individuals. The total frequency of CNVs was higher than that reported in previous studies on the Japanese population.[11,23] Notably, 7 of the 11 CNVs identified by MLPA were undetectable by STS-PCR. In addition, MLPA was capable of characterizing a complex structure of the b2/b4 duplication–gr/gr deletions, which was assessed as a simple gr/gr deletion by STS-PCR. These results demonstrate the usefulness of MLPA in the identification and characterization of AZF-linked CNVs. As MLPA is a relatively simple method and requires only a small amount of genomic DNA,[19] it can be used for the molecular diagnosis of several clinical samples. MLPA analysis appears to be beneficial for patients with SF, because it has been suggested that detection of AZF-linked CNVs would help to predict the sperm recovery rate at testicular sperm extraction.[9]

The 11 CNVs identified in this study were widely distributed in AZF regions and included three rearrangements that have not been reported previously. Our findings provide further evidence for a high frequency and genetic heterogeneity of AFZ-linked CNVs. Notably, 10 of the 11 CNVs had their breakpoints within AZF-specific repeats, while the breakpoints of the multiple duplications in AZFc region remain to be determined. These data support the previously proposed notion that non-allelic homologous recombination has a critical role in the development of CNVs in AZF regions.[24]

Two matters are noteworthy for the frequencies of CNVs in the patient and control groups. First, the most common gr/gr deletion was detected in 23 of 56 patients (41.1%) and in 20 of 65 controls (30.8%). Previous studies have shown that the frequency of the gr/gr deletion in control subjects is variable among ethnic groups, ranging from 0.0% in the Dutch to 33.9% in the Japanese.[18] It was suggested that the gr/gr deletion is accompanied by multiple haplogroups of diverse pathogenicity, and this deletion in Japanese individuals is usually

accompanied by a functionally neutral haplogroup.[11,12] Our findings confirm that, in Japan, the gr/gr deletion represents a common variation that has a negligible effect on spermatogenesis. Second, although most of the 11 CNVs were identified in the patient and control groups at similar frequencies, simple duplications were detected exclusively in the patient group. These results are consistent with those of a previous study on the Taiwan Han Chinese population in which the risk of SF was associated with AZF duplications but not with deletions.[16,17] Copy-number gain of multi-copy genes, such as *DAZ* and *CDY1*, or conformational changes of Y chromosome may exert deleterious effects on spermatogenesis.[16,17] However, copy-number gain of the multi-copy genes would not be sufficient to cause SF, because two of our control individuals also had an increased copy-number of these genes (Table 1). Indeed, we found no significant difference in the frequencies of copy-number gain of multi-copy genes between the patient and control groups (Supplementary Table 1). Furthermore, *USP9Y*, a putative causative gene for SF,[25] was not affected in our patients.

Considering the small number of subjects in this study, the pathogenicity of AZF-linked CNVs, except for the gr/gr deletion, remains unclear. Furthermore, DNA samples of patients' relatives were not analyzed in this study. As previous studies have shown that *de novo* occurrence of CNVs in AZF regions is a relatively rare event,[1,22,26] molecular analysis of the fathers of SF patients with CNVs would provide critical information regarding the effect of the CNVs on spermatogenesis. Further studies, including MLPA analysis in large cohorts and familial analyses of CNV-positive individuals, will clarify the clinical significance of each CNV.

In conclusion, the results expanded our understanding of the frequency and genetic heterogeneity of AZF-linked CNVs. It appears that non-allelic homologous recombination underlies most, if not all, CNVs in AZF regions. In addition, we confirmed the functional neutrality of the gr/gr deletion in Japanese individuals. We found a possible association between AZF microduplications and the risk of SF, which needs to be evaluated in future studies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to the participants of this study. The study was supported by the Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology and from the Japan Society for the Promotion of Science and by the Grants from the Ministry of Health, Labor and Welfare (Grant Numbers: H26-062, 082, Tokyo, Japan) from National Center for Child Health and Development (Grant Numbers: 26-11, 26-19, Tokyo, Japan) and from Takeda foundation. The sponsors had no role in study design, in the collection, analysis or interpretation of data, in the writing of the report or in the decision to submit the article for publication.

1 Vogt, P. H., Edelmann, A., Kirsch, S., Henegariu, O., Hirschmann, P. & Kiesewetter, F. *et al.* Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum. Mol. Genet.* **5**, 933–943 (1996).
2 Giacco, D. L., Chianese, C., Sanchez-Curbelo, J., Bassas, L., Ruiz, P., Rajmil, O. *et al.* Clinical relevance of Y-linked CNV screening in male infertility: new insights based on the 8-year experience of a diagnostic genetic laboratory. *Eur. J. Hum. Genet.* **22**, 754–761 (2014).
3 Repping, S., Skaletsky, H., Lange, J., Silber, S., Veen, F., Oates, R. D. *et al.* Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure. *Am. J. Hum. Genet.* **71**, 906–922 (2002).
4 Kuroda-Kawaguchi, T., Skaletsky, H., Brown, L. G., Minx, P. J., Cordum, H. S., Waterston, R. H. *et al.* The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nat. Genet.* **29**, 279–286 (2001).
5 Lu, C., Jiang, J., Zhang, R., Wang, Y., Xu, M., Qin, Y. *et al.* Gene copy number alterations in the azoospermia-associated AZFc region and their effect on spermatogenic impairment. *Mol. Hum. Reprod.* **20**, 836–843 (2014).
6 Shen, Y., Yan, Y., Liu, Y., Zhang, S., Yang, D., Zhang, P. *et al.* A significant effect of the *TSPY1* copy number on spermatogenesis efficiency and the phenotypic expression of the gr/gr deletion. *Hum. Mol. Genet.* **22**, 1679–1695 (2013).
7 Stahl, P. S., Mielnik, A. N., Barbieri, C. E., Schlegel, P. N. & Paduch, D. A. Deletion or underexpression of they-chromosome genes CDY2 and HSFY is associated with maturation arrest in American men with nonobstructive azoospermia. *Asian J. Androl.* **14**, 676–682 (2012).
8 Kleiman, S. E., Yogev, L., Hauser, R., Botchan, A., Maymon, B. B., Schreiber, L. *et al.* Members of the CDY family have different expression patterns: CDY1 transcripts have the best correlation with complete spermatogenesis. *Hum. Genet.* **113**, 486–492 (2003).
9 Krausz, C., Hoefsloot, L., Simoni, M. & Tuttelmann, F. EAA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions: state-of-the-art 2013. *Andrology* **2**, 5–19 (2014).
10 Rosen, S. G., Marszalek, J. D., Irenze, K., Skaletsky, H., Brown, L. G. & Oates, R. D. *et al.* AZFc deletions and spermatogenic failure: a population-based survey of 20,000 Y chromosomes. *Am. J. Hum. Genet.* **91**, 890–896 (2012).
11 Carvalho, C. M. B., Zuccherato, L. W., Fujisawa, M., Shirakawa, T., Ribeiro-dos-Santos, A. K. C., Santos, S. E. B. *et al.* Study of AZFc partial deletion gr/gr in firtile and infertile Japanese males. *J. Hum. Genet.* **51**, 794–799 (2006).
12 Repping, S., Skaletsky, H., Brown, L., Daalen, S. K. M., Korver, C. M., Pyntikova, T. *et al.* Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. *Nat. Genet.* **35**, 247–251 (2003).
13 Fernandes, S., Paracchini, S., Meyer, L. H., Floridia, G., Tyler-Smith, C. & Vogt, P. H. A large AZFc deletion removes *DAZ3/DAZ4* and nearby genes from men in Y haplogroup N. *Am. J. Hum. Genet.* **74**, 180–187 (2004).
14 Lu, C., Zhang, F., Yang, H., Xu, M., Du, G., Wu, W. *et al.* Additional genomic duplications in AZFc underlie the b2/b3 deletion-associated risk of spermatogenic impairment in Han Chinese population. *Hum. Mol. Genet.* **20**, 4411–4421 (2011).
15 Krausz, C., Giachini, C., Xue, Y., O'Brya, M. K., Gromoll, J., Meyts, E. R. *et al.* Phenotypic variation within European carriers of the Y-chromosomal gr/gr deletion is independent of Y-chromosomal background. *J. Med. Genet.* **46**, 21–31 (2009).
16 Lin, Y., Hsu, L. C., Kuo, P., Huang, W. J., Chiang, H., Yeh, S. *et al.* Partial duplication at AZFc on the Y chromosome is a risk factor for impaired spermatogenesis in Han Chinese in Taiwan. *Hum. Mutat.* **28**, 486–494 (2007).
17 Ye, J., Ma, L., Yang, L., Wang, J., Wang, Y., Guo, H. *et al.* Partial AZFc duplications not deletions are associated with male infertility in the Yi population of Yunnan Province, China. *J. Zhejiang Univ. Sci. B* **14**, 807–815 (2013).
18 Giachini, C., Laface, I., Guarducci, E., Balercia, G., Forti, G. & Krausz, C. Partial AZFc deletions and duplications: clinical correlates in the Italian population. *Hum. Genet.* **124**, 399–410 (2008).
19 Stuppia, L., Antonucci, I., Palka, G. & Gatta, V. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. *Int. J. Mol. Sci.* **13**, 3245–3276 (2012).
20 Coughlin, C. R., Scharer, G. H. & Shaikh, T. H. Clinical impact of copy number variation analysis using high-resolution microarray technologies: advantages, limitations and concerns. *Genome Med.* **4**, 80 (2012).
21 Bunyan, D. J., Callaway, J. L. A. & Laddach, N. Detection of partial deletions of Y-chromosome AZFc in infertile men using the multiplex Ligation-dependent Probe Amplification Assay. *J. Reprod. Infertil.* **13**, 174–178 (2012).
22 Liu, X. H., Yan, L. Y., Lu, C. L., Li, R., Zhu, X. H. & Jin, H. Y. ART do not increase the risk of Y-chromosome microdeletion in 19 candidate genes at AZF regions. *Reprod. Fertil. Dev.* **26**, 778–786 (2013).
23 Kihaile, P. E., Yasui, A. & Shuto, Y. Prospective assessment of Y-chromosome microdeletions and reproductive outcomes among infertile couples of Japanese and African origin. *J. Exp. Clin. Assist. Reprod.* **2**, 9 (2005).
24 Carvalho, C. M. B., Zhang, F. & Lupski, J. R. Structural variation of the human genome: mechanisms, assays, and role in male infertility. *Syst. Biol. Reprod. Med.* **57**, 3–16 (2011).
25 Sun, C., Skaletsky, H., Birren, B., Devon, K., Tang, Z., Silber, S. *et al.* An azoospermic man with a de novo point mutation in the Y-chromosomal gene *USP9Y*. *Nat. Genet.* **23**, 429–432 (1999).
26 Cram, D. S., Ma, K., Bhasin, S., Arias, J., Pandjaitan, M., Chu, B. *et al.* Y chromosome analysis of infertile men and their sons conceived through intracytoplasmic sperm injection: vertical transmission of deletions and rarity of de novo deletions. *Fertil. Steril.* **74**, 909–915 (2000).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

Nagata *et al. Orphanet Journal of Rare Diseases* 2014, **9**:125
http://www.ojrd.com/content/9/1/125

ORPHANET JOURNAL OF RARE DISEASES

RESEARCH Open Access

Japanese founder duplications/triplications involving *BHLHA9* are associated with split-hand/foot malformation with or without long bone deficiency and Gollop-Wolfgang complex

Eiko Nagata[1†], Hiroki Kano[2†], Fumiko Kato[1], Rie Yamaguchi[1], Shinichi Nakashima[1], Shinichiro Takayama[3], Rika Kosaki[4], Hidefumi Tonoki[5], Seiji Mizuno[6], Satoshi Watanabe[7], Koh-ichiro Yoshiura[7], Tomoki Kosho[8], Tomonobu Hasegawa[9], Mamori Kimizuka[10], Atsushi Suzuki[11], Kenji Shimizu[11], Hirofumi Ohashi[11], Nobuhiko Haga[12], Hironao Numabe[13], Emiko Horii[14], Toshiro Nagai[15], Hiroshi Yoshihashi[16], Gen Nishimura[17], Tatsushi Toda[18], Shuji Takada[19], Shigetoshi Yokoyama[19,22], Hiroshi Asahara[19,20], Shinichiro Sano[1,21], Maki Fukami[21], Shiro Ikegawa[2] and Tsutomu Ogata[1*]

Abstract

Background: Limb malformations are rare disorders with high genetic heterogeneity. Although multiple genes/loci have been identified in limb malformations, underlying genetic factors still remain to be determined in most patients.

Methods: This study consisted of 51 Japanese families with split-hand/foot malformation (SHFM), SHFM with long bone deficiency (SHFLD) usually affecting the tibia, or Gollop-Wolfgang complex (GWC) characterized by SHFM and femoral bifurcation. Genetic studies included genomewide array comparative genomic hybridization and exome sequencing, together with standard molecular analyses.

Results: We identified duplications/triplications of a 210,050 bp segment containing *BHLHA9* in 29 SHFM patients, 11 SHFLD patients, two GWC patients, and 22 clinically normal relatives from 27 of the 51 families examined, as well as in 2 of 1,000 Japanese controls. Families with SHFLD- and/or GWC-positive patients were more frequent in triplications than in duplications. The fusion point was identical in all the duplications/triplications and was associated with a 4 bp microhomology. There was no sequence homology around the two breakpoints, whereas rearrangement-associated motifs were abundant around one breakpoint. The rs3951819-*D17S1174* haplotype patterns were variable on the duplicated/triplicated segments. No discernible genetic alteration specific to patients was detected within or around *BHLHA9*, in the known causative SHFM genes, or in the exome.

(Continued on next page)

* Correspondence: tomogata@hama-med.ac.jp
†Equal contributors
[1]Department of Pediatrics, Hamamatsu University School of Medicine, Hamamatsu 431-3192, Japan
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusions: These results indicate that *BHLHA9* overdosage constitutes the most frequent susceptibility factor, with a dosage effect, for a range of limb malformations at least in Japan. Notably, this is the first study revealing the underlying genetic factor for the development of GWC, and demonstrating the presence of triplications involving *BHLHA9*. It is inferred that a Japanese founder duplication was generated through a replication-based mechanism and underwent subsequent triplication and haplotype modification through recombination-based mechanisms, and that the duplications/triplications with various haplotypes were widely spread in Japan primarily via clinically normal carriers and identified via manifesting patients. Furthermore, genotype-phenotype analyses of patients reported in this study and the previous studies imply that clinical variability is ascribed to multiple factors including the size of duplications/triplications as a critical factor.

Keywords: *BHLHA9*, Split-hand/foot malformation, Long bone deficiency, Gollop-Wolfgang complex, Expressivity, Penetrance, Susceptibility, Japanese founder copy number gain

Introduction

Split-hand/foot malformation (SHFM), also known as ectrodactyly, is a rare limb malformation involving the central rays of the autopod [1,2]. It presents with median clefts of the hands and feet, aplasia/hypoplasia of the phalanges, metacarpals, and metatarsals, and syndactyly. SHFM results from failure to maintain the central portion of the apical ectodermal ridge (AER) in the developing autopod [1,2]. SHFM is divided into two forms: a non-syndromic form with limb-confined manifestations and a syndromic form with extra-limb manifestations [2]. Furthermore, non-syndromic SHFM can occur as an isolated abnormality confined to digits (hereafter, SHFM refers to this type) or in association with other limb abnormalities as observed in SHFM with long bone deficiency (SHFLD) usually affecting the tibia and in Gollop-Wolfgang complex (GWC) characterized by femoral bifurcation [1,2]. Both syndromic and non-syndromic forms are associated with wide expressivity and penetrance even among members of a single family and among limbs of a single patient [2].

SHFM and SHFLD are genetically heterogeneous conditions reviewed in ref. [2]. To date, SHFM has been identified in patients with heterozygous deletions or translocations involving the *DLX5–DLX6* locus at 7q21.2–21.3 (SHFM1) [3] (*DLX5* mutations have been detected recently), heterozygous duplications at 10q24 (SHFM3), heterozygous mutations of *TP63* at 3q27 (SHFM4), heterozygous deletions affecting *HOXD* cluster at 2q31 (SHFM5), and biallelic mutations of *WNT10B* at 12q31 (SHFM6); in addition, SHFM2 has been assigned to Xq26 by linkage analyses in a large Pakistani kindred [2]. Similarly, a genomewide linkage analysis in a large consanguineous family has identified two SHFLD susceptibility loci, one at 1q42.2–q43 (SHFLD1) and the other at 6q14.1 (SHFLD2); furthermore, after assignment of another SHFLD locus to 17p13.1–13.3 [4], duplications at 17p13.3 (SHFLD3) have been found in patients with SHFLD reviewed in ref. [2]. However, the GWC locus (loci) remains unknown at present.

The duplications at 17p13.3 identified to date are highly variable in size, and harbor *BHLHA9* as the sole gene within the smallest region of overlap [5-9]. *Bhlha9*/*bhlha9* is expressed in the limb bud mesenchyme underlying the AER in mouse and zebrafish embryos, and *bhlha9* knockdown has resulted in shortening of the pectoral fins in zebrafish [6]. Furthermore, *BHLHA9*-containing duplications have been identified not only in patients with SHFLD but also in those with SHFM and clinically normal family members [4-10]. These findings argue for a critical role of *BHLHA9* duplication in the development of SHFM and SHFLD, with variable expressivity and incomplete penetrance.

In this study, we report on *BHLHA9*-containing duplications/triplications with an identical fusion point and various haplotype patterns that were associated with a range of limb malformations including GWC, and discuss on characteristic clinical findings, genomic basis of Japanese founder copy number gains, and underlying factors for phenotypic variability.

Materials and methods

Patients/subjects

We studied 68 patients with SHFM (n = 55), SHFLD (n = 11), or GWC (n = 2), as well as 60 clinically normal relatives, from 51 Japanese families; the pedigrees of 27 of the 51 families and representative clinical findings are shown in Figure 1. All the probands 1–51 had a normal karyotype. Southern blot analysis for SHFM3 locus had been performed in 28 probands with SHFM, indicating 10q24 duplications in two of them [11]. Clinical features including photographs and roentgenograms of a proband with GWC and his brother with SHFLD (family 23 in Figure 1A) were as described previously [12]. The residencies of families 1–51 were widely distributed throughout Japan.

Ethical approval and samples

This study was approved by the Institutional Review Board Committees of Hamamatsu University School of

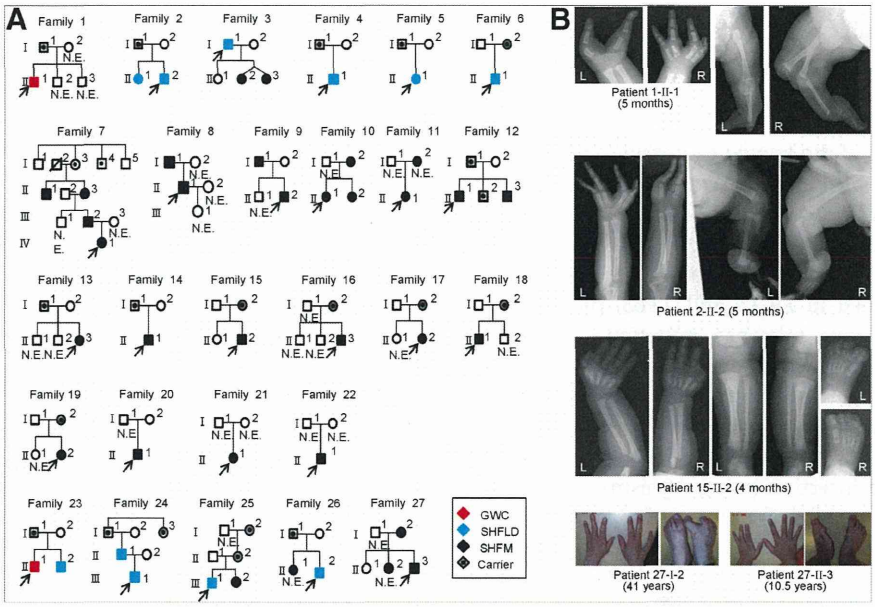


Figure 1 Clinical summary. A. Pedigrees of 27 Japanese families with duplications (families 1–22) and triplications (families 23–27) of a ~200 kb region involving *BHLHA9*. The duplications/triplications are associated with GWC, SHFLD, SHFM, or normal phenotype (carriers). N.E.: Not examined molecularly. **B**. Representative clinical findings. Each patient is indicated by a family-generation-individual style and corresponds to the patient/subject shown in Figure 1A and Additional file 5. The top panel: GWC with right bifid femur; the second panel: SHFLD with bilateral tibial deficiencies, the third panel: SHFM with polydactyly; and the bottom panel: SHFM.

Medicine, RIKEN, and National Center for Child Health and Development, and was performed using peripheral leukocyte samples after obtaining written informed consent for the molecular analysis and the publication of genetic and clinical data after removing information for personal identification (e.g., name, birthday, and facial photograph) from the adult subjects (^{3} 20 years) or from the parents of the child subjects (below 20 years). Furthermore, informed assent was also obtained from child subjects between 6–20 years.

Samples and primers

The primers utilized in this study are summarized in Additional file 1.

Molecular studies

Sanger sequencing, fluorescence *in situ* hybridization (FISH), microsatellite genotyping, Southern blotting, and bisulfite sequencing-based methylation analysis were performed by the standard methods, as reported previously [13]. Quantitative real-time PCR (qPCR) analysis was carried out by the SYBR Green methods on StepOnePlus system, using *RNaseP* as an internal control (Life Technologies). Genomewide oligonucleotide-based array comparative genomic hybridization (CGH) was performed with a catalog human array (4 × 180 K format, ID G4449A) according to the manufacturer's instructions (Agilent Technologies), and obtained copy number variants/polymorphisms were screened with Agilent Genomic Workbench software using the Database of Genomic Variants (http://dgv.tcag.ca/dgv/app/home). Sequencing of a long region encompassing *BHLHA9* was performed with the Nextera XT kit on MiSeq (Illumina), using SAMtools v0.1.17 software (http://samtools.sourceforge.net/). Exome sequencing was performed as described previously [14].

Assessment of genomic environments around the fusion points

Repeat elements around the fusion point were searched for using Repeatmasker (http://www.repeatmasker.org). Rearrangement-inducing DNA features were investigated for 300 bp regions at both the proximal and the distal sides of each breakpoint, using GEECEE (http://emboss.bioinformatics.nl/cgi-bin/emboss/geecee) for calculation of the average GC content, PALINDROME (http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::palindrome) and Non-B DB (http://nonb.abcc.ncifcrf.gov) for the examination of the palindromes and non-B (non-canonical) structures, and Fuzznuc (http://emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc) for the assessment of rearrangement-associated sequence motifs and tri/tetranucleotides [15-20]. For controls, we examined 48 regions of 600 bp long selected at an interval of 1.5 Mb from the entire chromosome 17.

Statistical analysis

The statistical significance of the frequency was analyzed by the two-sided Fisher's exact probability test.

Results

Sequence analysis of the known causative/candidate genes

We performed direct sequencing for the previously known causative genes (*DLX5*, *TP63*, and *WNT10B*) reviewed in ref. [2] in the probands 1–51. Although no pathologic mutation was identified in *DLX5* and *TP63*, the previously reported homozygous missense mutation of *WNT10B* (c.944C > T, p.R332W) [21] was detected in the proband 48 with SHFM who was born to healthy consanguineous parents heterozygous for this mutation. In addition, while no variation was detected in *DLX5* and *WNT10B*, rs34201045 (4 bp insertion polymorphism) in *TP63* [21] was detected with an allele frequency of 61%.

We also examined *BHLHA9*, because gain-of-function mutations of *BHLHA9* as well as *BHLHA9*-harboring duplications may lead to limb malformations. No sequence variation was identified in the 51 probands.

Array CGH analysis

Array CGH analysis was performed for the probands 1–51, showing increased copy numbers at 17p13.3 encompassing *BHLHA9* (SHFLD3) in the probands 1–27 from families 1–27 (Figure 1A). Furthermore, heterozygous duplications at 10q24 (SHFM3) were detected in the probands 49–51, i.e., a hitherto unreported patient with paternally inherited SHFM (his father also had the duplication) and the two patients who had been indicated to have the duplications by Southern blot analysis [11]. No copy number alteration was observed at other SHFM/SHFLD loci in the probands 1–27 and 49–51. In the remaining probands 28–48, there was no copy number variation that was not registered in the Database of Genomic Variants.

Identical fusion points in *BHLHA9*-containing duplications/triplications

The array CGH indicated that the increased copy number regions at 17p13.3 were quite similar in the physical size in the probands 1–27 and present in three copies in the probands 1–22 and in four copies in the probands 23–27 (Figure 2A). Thus, FISH analysis was performed using 8,259 bp PCR products amplified from this region, showing two signals with a different intensity that was more obvious in the probands 23–27 (Figure 2A).

We next determined the fusion points of the duplications/triplications (Figure 2B). PCR products of 2,195 bp long were obtained with P1/P2 primers in the probands 1–27, and the fusion point was determined by direct sequencing for 418 bp PCR products obtained with P3/P4 primers. The fusion point was identical in all the probands 1–27; it resided on intron 1 of *ABR* and intron 1 of *YWHAE*, and was associated with a 4 bp microhomology.

Then, we performed qPCR analysis for a 214 bp region harboring the fusion point, using P5/P6 primers (Figure 2C and Additional file 2). The fusion point was present in a single copy in the probands 1–22 and in two copies in the probands 23–27. The results showed that the identical genomic segment harboring *BHLHA9* was tandemly duplicated in the probands 1–22 and triplicated in the probands 23–27. According to GRCh37/hg19 (http://genome.ucsc.edu/), the genomic segment was 210,050 bp long.

We also performed array CGH and qPCR for the fusion point in 15 patients other than the probands and 47 clinically normal relatives from the 27 families (Figures 1 and 2C). The duplications/triplications were identified in all the 15 patients. Thus, in a total of 42 patients, duplications/triplications were found in 29 SHFM patients, 11 SHFLD patients, and two GWC patients. Furthermore, the duplications/triplications were also present in 22 of the 47 clinically normal relatives. In particular, they were invariably identified in either of the clinically normal parents when both of them were examined; they were also present in other clinically normal relatives in families 7, 12, 24, and 25.

Since the above data indicated the presence of duplications/triplications in clinically normal subjects, we performed qPCR for the fusion point in 1,000 Japanese controls. The fusion point was detected in a single copy in two subjects (Subjects 1 and 2 in Figure 2C). We also performed array CGH in 200 of the 1,000 controls including the two subjects, confirming the duplications in the two subjects and lack of other copy number variations, including deletions involving *BHLHA9*, which were not registered in the Database of Genomic Variants in the 200 control subjects. The frequency of duplications/triplications was significantly higher in the probands than in the control subjects (27/51 vs. 2/1,000, $P = 3.5 \times 10^{-37}$).

Various haplotype patterns on the duplicated/triplicated segments

We carried out genotyping for rs3951819 (A/G SNP on *BHLHA9*) and *D17S1174* (CA repeat microsatellite locus) on the genomic segment subjected to duplications/triplications (Figure 2A), and determined rs3951819-*D17S1174* haplotype patterns. Representative results are shown in Figure 2D, and all the data are available on request. Various haplotype patterns were identified on the single, the duplicated, and the triplicated segments, and the [A-14] haplotype was most prevalent on the duplicated/triplicated segments (Table 1). While the distribution of CA repeat lengths on the single segments was discontinuous, similar discontinuous distribution was

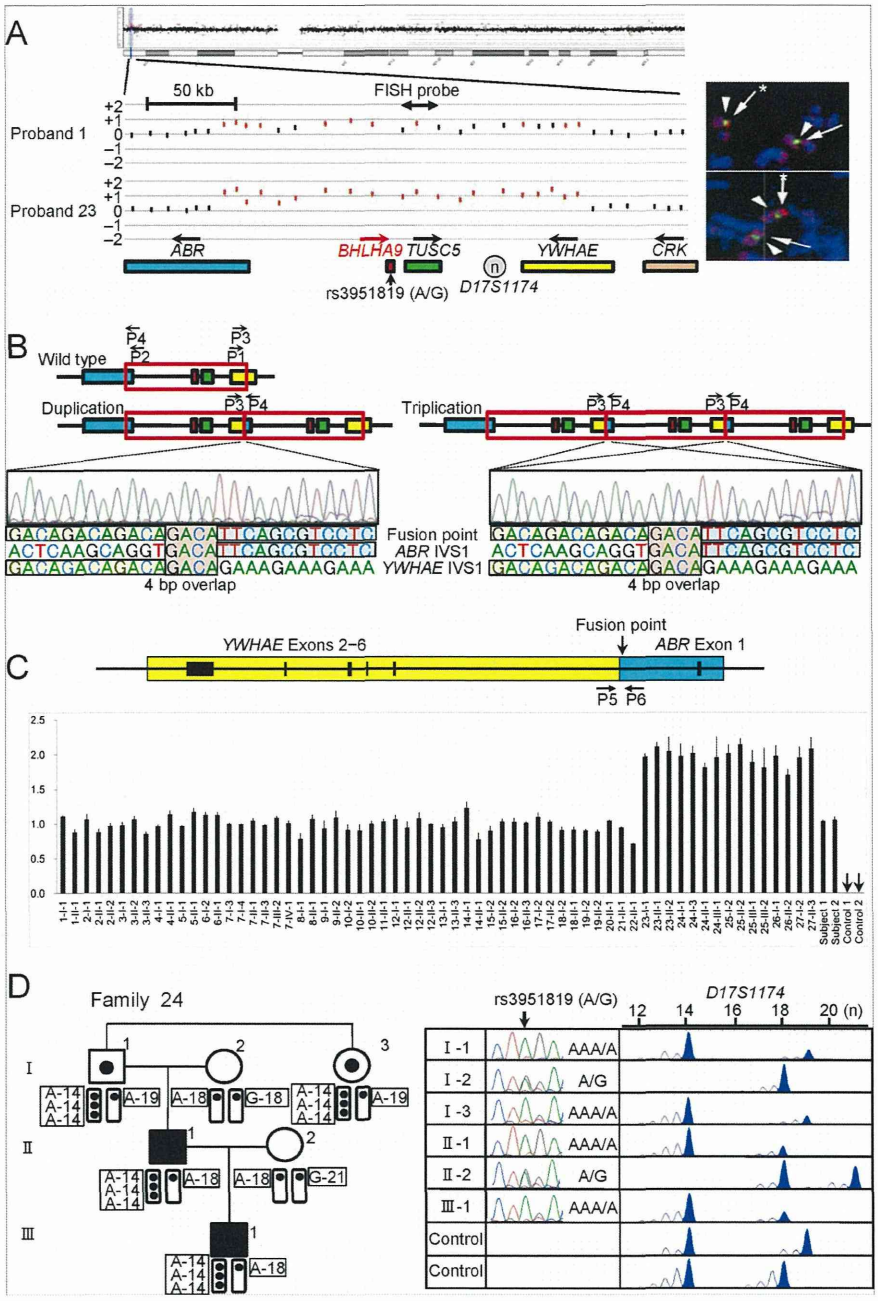


Figure 2 (See legend on next page.)

(See figure on previous page.)
Figure 2 Identification and characterization of the duplications/triplications involving *BHLHA9* at chromosome 17p13.3. A. Array CGH and FISH analyses in proband 1 and proband 23 with GWC. In array CGH analysis, the black and the red dots denote the normal and the increased copy numbers, respectively. Since the log2 signal ratios for a ~200 kb region encompassing *BHLHA9* are around +0.5 in the proband 1 and around +1.0 in the proband 23, this indicates the presence of three and four copies of this region in the two probands, respectively. In FISH analysis, two red signals with an apparently different density are detected by the 8,289 bp PCR probe (the stronger signals are indicated with asterisks). The green signals derive from an internal control probe (CEP17). The arrows on the genes show transcriptional directions. Rs3951819 (A/G) resides within *BHLHA9*. **B**. Determination of the fusion point. The fusion has occurred between intron 1 of *ABR* and intron 1 of *YWHAE*, and is associated with a 4 bp (GACA) microhomology. P1–P4 show the position of primers. **C**. Quantitative real-time PCR analysis. The upper part denotes the fusion point. P5 & P6 show the position of primers. The lower part shows the copy number of the fusion point in patients/subjects with duplications/triplications (indicated by a family-generation-individual style corresponding to that in Figure 1 and Additional file 5). Subject-1 and subject-2 denote the two control subjects with the duplication, and control-1 and control-2 represent normal subjects without the duplication/triplication. **D**. The rs3951819 (A/G SNP)–*D17S1174* (CA repeat number) haplotype patterns in family 24. Assuming no recombination between rs3951819 and *D17S1174*, the haplotype patterns of the family members are determined as shown here. The haplotype patterns of the remaining families have been interpreted similarly.

also observed in the Japanese general population (see Additional file 3).

Genomic environments around the breakpoints

The breakpoint on *YWHAE* intron 1 resided on a simple *Alu* repeat sequence, and that on *ABR* intron 1 was present on a non-repetitive sequence. There was no low copy repeat around the breakpoints. Comparison of the frequencies of known rearrangement-inducing DNA features between 600 bp sequences around the breakpoints and those of 48 regions selected at an interval of 1.5 Mb from chromosome 17 revealed that palindromes, several types of non-B DNA structures, and a rearrangement-associated sequence motif were abundant around the breakpoint on *YWHAE* intron 1 (see Additional file 4).

Clinical findings of families 1–27

Clinical assessment revealed several notable findings. First, duplications/triplications were associated with SHFM, SHFLD, GWC, or normal phenotype, with inter- and intra-familial clinical variability (Figure 1A). Second, in the 42 patients, split hand (SH) was more prevalent than split foot (SF) (41/42 vs. 17/42, $P = 6.2 \times 10^{-9}$), and long bone defect (LBD) was confined to lower extremities (0/42 vs. 13/42, $P = 4.1 \times 10^{-5}$) (Table 2 and Additional file 5). Third, there was no significant sex difference in the ratio between patients with limb malformations and patients/carriers with duplications/triplications (26/38 in males vs. 16/26 in females, $P = 0.60$) (Table 2 and Additional file 5). Fourth, the ratio of LBD positive families was significantly higher in triplications than in duplications (4/5 vs. 16/22, $P = 0.047$) (Figure 1A and Table 2). Fifth, while the duplications/triplications were transmitted from patients to patients, from carriers to patients, and from a carrier to a carrier (from I-1 to II-2 in family 12), transmission from a patient to a carrier was not identified (Figure 1A); it should be pointed out, however, that molecular analysis in a clinically normal child born to an affected parent was possible only in a single adult subject (II-1 in family 27), and that molecular analysis in clinically

Table 1 The rs3951819 (A/G SNP) – *D17S1174* (CA repeat number) haplotype

Patterns of the 210,050 bp segment subjected to copy number gains	
Haplotype pattern	**Family**
<Single segment>	
[A-14]	1, 5, 9, 15, 17, 19, 23, 26
[A-16]	12
[A-18]	3, 14, 15, 24, 25, 26
[A-19]	2, 6, 13, 19, 20, 24, 25, 27
[A-21]	5, 23
[G-12]	17
[G-14]	2, 3, 6, 12, 13, 19, 26
[G-18]	3, 5, 17, 18, 24, 25
[G-19]	9, 12, 18, 20, 25
[G-21]	1, 9, 19, 24, 27
[A-14] or [G-14]	16
[A-18] or [G-18]	4
[A-19] or [G-19]	4
[A-21] or [G-21]	16
<Duplicated segments>	
[A-14] + [A-14]	5, 12, 13, 14, 15, 20
[A-14] + [A-18]	1
[A-14] + [G-18] or [G-14] + [A-18]	2, 3, 4, 6, 9, 16, 17
[A-14] + [G-18] or [A-14] + [G-19]	18
[A-14] + [G-14] or [G-14] + [G-14]	19
<Triplicated segments>	
[A-14] + [A-14] + [A-14]	23, 24
[A-14] + [A-14] + [G-14]	25
[A-14] + [A-19] + [A-19]	26
[A-14] + [G-18] + [G-18] or [G-14] + [A-18] + [G-18]	27

The haplotype patterns written in the left column have been detected in at least one patient/subject in the families described in the right column.
Genotyping could not be performed in several patients/subjects who had been repeatedly examined previously, because of the extremely small amount of DNA samples that were virtually used up in the sequencing and array CGH analyses.

Table 2 Summary of clinical findings in patients/carriers with duplications/triplications involving *BHLHA9*

	SHFM (+) patients			LBD (+) patients			Patient ratio*			LBD (+) families		
	SH	SF	*P*-value	U-LBD	L-LBD	*P*-value	Male	Female	*P*-value	Trip	Dup	*P*-value
This study	41/42	17/42	6.2×10^{-9}	0/42	13/42	4.1×10^{-5}	26/38	16/26	0.60	4/5	16/22	0.047
Previous studies	63/84	23/84	8.6×10^{-10}	11/91	42/91	5.7×10^{-7}	68/114	31/79	5.7×10^{-3}	...	...	...
Sum	104/126	40/126	1.1×10^{-16}	11/133	55/133	3.0×10^{-10}	94/152	47/105	7.6×10^{-3}	...	...	...

SHFM: split-hand/foot malformation; SH: split hand; SF: split foot; LBD: long bone deficiency; U: upper; L: lower; Trip: triplication; and Dup: duplication.
In the previous studies, patients without detailed phenotypic description and those of unknown sex have been excluded (3–9).
*The ratio between patients with limb malformations and patients/carriers with duplications/triplications, i.e. the number of patients over the number of patients plus carriers.

normal children <20 years old was possible only in two subjects (II-2 in family 12 and II-1 in family 15). Lastly, limb malformation was inherited in an apparently autosomal dominant manner (from patients to patients), or took place as an apparently *de novo* event or as an apparently autosomal recessive trait (from clinically normal parents to a single or two affected children) (Figure 1A).

Attempts to identify a possible modifier(s)

The variable expressivity and incomplete penetrance in families 1–27 suggest the presence of a possible modifier (s) for the development of limb malformations. Thus, we performed further molecular studies in patients/subjects in whom DNA samples were still available, and compared the molecular data between patients with SHFM and those with SHFLD for the assessment of variable expressivity and between SHFM, SHFLD, or total patients and carriers for the evaluation of incomplete penetrance.

We first examined the possibility that the modifier(s) resides within or around *BHLHA9* (see Additional file 6). There was no *BHLHA9* mutation in all the 21 examined probands with SHFM, SHFLD, or GWC, as described in the section of "Sequence analysis of the known causative/candidate genes". The rs3951819 A/G SNP pattern on the duplicated/triplicated segments was apparently identical between patients and carriers (e.g. Figure 2D), and the frequency of A/G allele on the normal chromosome 17 was similar between SHFM and SHFLD patients and between SHFM, SHFLD, or total patients and carriers (see Additional file 7). The results of other known SNPs on *BHLHA9* (rs185242872, rs18936498, and rs140504068) were not informative, because of absence or extreme rarity of minor alleles. Furthermore, in SHFM families 7, 12, and 18, sequencing of a 7,406 bp region encompassing *BHLHA9* and Southern blot analysis using five probes and *Mfe*I-, *Ssp*I-, and *Sac*I-digested genomic DNA revealed no variation specific to the patients, and methylation analysis for a CpG rich region at the upstream of *BHLHA9* delineated massive hypomethylation in all the patients/carriers examined.

Next, we examined the possibility that a variant(s) of known causative genes constitutes the modifier(s). Since rs34201045 in *TP63* was identified in the mutation analysis, we compared rs34201045 genotyping data between the 27 probands and the 15 carriers. The allele and genotype frequencies were similar between SHFM and SHFLD patients and between SHFM, SHFLD, or total patients and carriers (see Additional file 8).

We finally performed exome sequencing in SHFM families 13 and 17–19. However, there was no variation specific to the patients. In addition, re-examination of the genomewide array CGH data showed no discernible copy number variation specific to the patients.

Discussion

BHLHA9 overdosage and clinical characteristics

We identified duplications/triplications of a ~ 200 kb genomic segment involving *BHLHA9* at 17p13.3 in 27 of 51 families with SHFM, SHFLD, or GWC. To our knowledge, this is the first study revealing the underlying genetic factor for the development of GWC, and demonstrating the presence of triplications involving *BHLHA9* that were suggested but not confirmed in the previous studies [5,9]. Furthermore, this study indicates that *BHLHA9*-containing duplications/triplications are the most frequent underlying factor for the development of limb malformations at least in Japan. Notably, SHFLD and GWC with LBD were significantly more frequent in patients with triplications than in those with duplications, and the duplications/triplications were identified in clinically normal familial members and in the general population. These findings imply that increased *BHLHA9* copy number constitutes a strong susceptibility, rather than a causative, factor with a dosage effect for the development of a range of limb malformations. Since *Bhlha9* is expressed in the developing ectoderm adjacent to the AER rather than the AER itself in mouse embryos [6], *BHLHA9* appears to play a critical role in the limb development by interacting with the AER. While the duplications/triplications identified in this study included *TUSC5* and generated an *ABR-YWHAE* chimeric gene (Figure 2C), *TUSC5* duplication and the chimeric gene formation are not common findings in the previously reported patients with duplications at 17p13.3 and SHFM and/or SHFLD [5-9]. In addition, none of *Tusc5*, *Abr*, and *Ywhae* is specifically expressed in the developing mouse limb buds [22] (A Transcriptome Atlas Database

for Mouse Embryo of Eurexpress Project, http://www.eurexpress.org/ee/project/).

Several clinical findings are noteworthy in patients/subjects with duplications/triplications. First, SH was more frequent than SF in this study as well as in the previous studies, and LBD was confined to lower extremities in this study and was more frequent in lower extremities than in upper extremities in the previous studies (Table 2) [4-10]. This implies that *BHLHA9* overdosage exerts differential effects on the different parts of limbs. Second, while limb malformations were similarly identified between males and females in this study, they were more frequently observed in males than in females in the previous studies (Table 2) [4-10]. In this regard, it has been reported that testosterone influences the digital growth pattern as indicated by the lower second to fourth digit length ratio in males than in females [23-25], and that Caucasian males have higher serum testosterone values and lower second to fourth digit length ratios than Oriental males [26,27]. Such testosterone effects on the digital growth pattern with ethnic difference may explain why male dominant manifestation was observed in the previous studies primarily from Caucasian countries and was not found in this study. Lastly, LBD was more prevalent in patients with triplications than in those with duplications. This suggests that LBD primarily occurs when the effects of *BHLHA9* overdosage are considerably elevated.

Genomic basis of the Japanese founder copy number gains

The duplications/triplications were associated with the same fusion point and variable haplotype patterns. Since there was no sequence homology or low-copy repeats around the breakpoints, it is unlikely that such duplications/triplications were recurrently produced in different individuals by non-allelic homologous recombination (NAHR) [17,20]. Instead, it is assumed that a Japanese founder duplication took place in a single ancestor, and was spread with subsequent triplication and modification of the haplotype patterns.

The most likely genomic basis of the Japanese duplications/triplications is illustrated in Additional file 9. Notably, a 4 bp (GACA) microhomology was identified at the duplication fusion point (Figure 2B). A microhomology refers to two to five nucleotides common to the sequences of the two breakpoints, and is found as an overlapping sequence at the join point [16,19,20]. This suggests that the Japanese founder duplication was generated by replication-based mechanisms such as fork stalling and template switching (FoSTeS) and microhomology-mediated break-induced replication (MMBIR), because the presence of such a microhomology is characteristic of FoSTeS/MMBIR [17-20]. Indeed, such a simple tandem duplication with a microhomology can be produced by one time FoSTeS/MMBIR [17-20], although it could also be generated by non-homologous end-joining (NHEJ) [17]. Since the [A-14] haplotype was most prevalent on the duplicated/triplicated segments, it is inferred that a genomic rearrangement occurred in an ancestor with the [A-14] haplotype, yielding the founder duplication with the [A-14] + [A-14] haplotype. Furthermore, the presence of multiple stimulants for genomic rearrangements around the breakpoint on *YWHAE* intron 1 would have facilitated the generation of the founder duplication. In particular, non-B structures are known to stimulate the occurrence of both replication-based FoSTeS/MMBIR and double-strand breaks and resultant NHEJ [17,28,29], although the relative importance of each non-B DNA structure is largely unknown.

Subsequent triplication and haplotype modification can develop from the Japanese founder duplication through unequal interchromatid and interchromosomal recombinations [17,20]. Indeed, a tandem triplication with the [A-14] + [A-14] + [A-14] haplotype can be generated by unequal exchange between sister chromatids with the [A-14] + [A-14] haplotype, and various haplotype patterns are yielded by unequal interchromosomal exchanges involving the duplicated or triplicated segments. Furthermore, the haplotype variation would be facilitated by unequal exchanges between sister chromatids harboring duplications/triplications with various haplotype patterns and by the further unequal interchromosomal exchanges.

Underlying factors for the phenotypic variability

The duplications/triplications were accompanied by limb malformations with variable expressivity and incomplete penetrance. Although this may suggest the presence of a possible modifier(s) for the development of limb malformations, such a modifier(s) was not detected. In particular, while patient-to-carrier transmission of duplications/triplications was not identified in this study, even patient-to-carrier-to-patient transmission has been reported in three pedigrees [5,6,10]. Such transmission pattern with incomplete penetrance characterized by skipping of a generation is apparently inexplicable by assuming a modifier (s) interacting with *BHLHA9* or independent of *BHLHA9* on the duplication/triplication positive chromosome 17, on the normal chromosome 17, or on other chromosomes (Figure 3, Models A, B, and C, see also the legends in Figure 3).

In this regard, it is noteworthy that the development of limb malformations is obviously dependent on the size of genomic segment subjected to copy number gains. Actually, limb malformation has occurred in only one of 21 large duplications encompassing *BHLHA9* (average 1.55 Mb, mean 1.12 Mb) and in 29 of 80 small duplications encompassing *BHLHA9* (average 244 kb, mean 263 kb) ($P = 5.9 \times 10^{-3}$) [8]. Consistent with this, the patients with large and

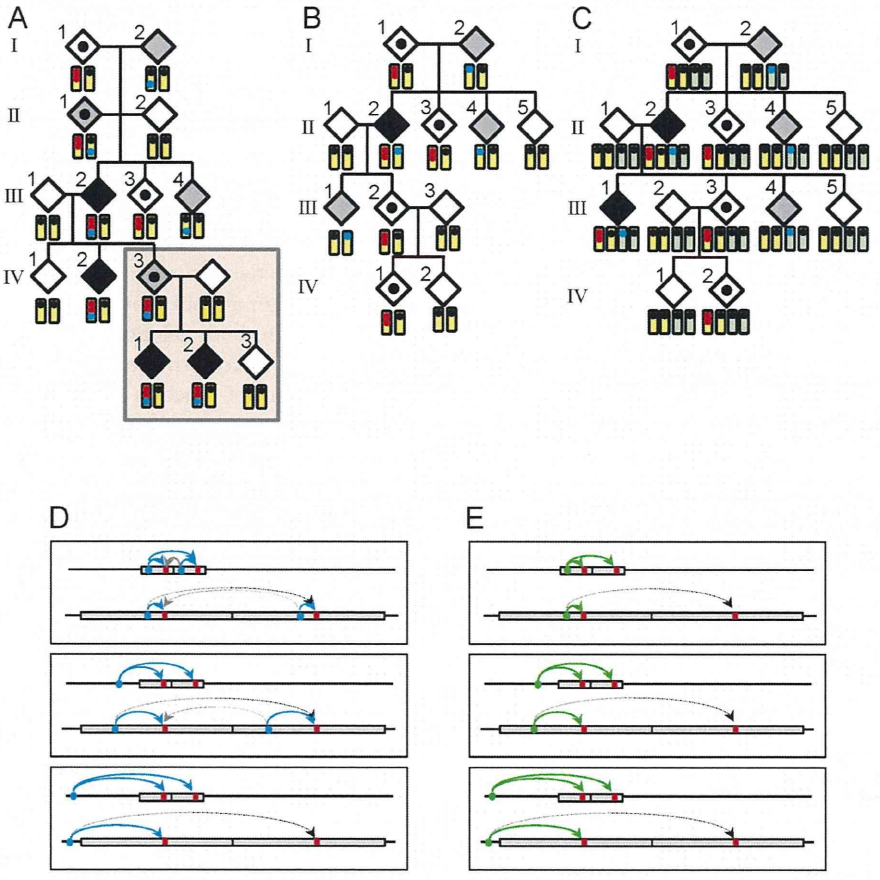


Figure 3 Models for a modifier(s) and effects of the duplication size. In models **A–C**, the yellow bars show chromosome 17, and the light green bars indicate other chromosomes. The two red dots represent the duplication at 17p13.3, and the blue dots indicate a putative modifier(s). Black painted diamonds represent limb malformation positive patients, dot-associated and gray painted diamonds indicate clinically normal carriers with the duplications and the modifier(s) respectively, and white painted diamonds denote clinically normal subjects without both the duplications and the modifier(s). **A**. This model assumes that co-existence of the duplication and a *cis*-acting modifier(s) causes limb malformation. If co-existence of the duplication and the *cis*-acting modifier(s) is associated with incomplete penetrance, this can explain all the transmission patterns observed to date, including the patient-to-carrier transmission and the presence of ≥ 2 affected children. **B**. This model postulates that the presence of a *cis*-acting modifier(s) on the normal chromosome 17 leads to limb malformation by enhancing the expression of the single *BHLHA9*, together with duplicated *BHLHA9* on the homologous chromosome. **C**. This model postulates that co-existence of the duplication at 17p13.3 and a modifier(s) on other chromosome causes limb malformation. In models **D–E**, the red bars represent *BHLHA9*, the blue circles indicate a physiological cis-regulatory element for *BHLHA9*, and the green circles indicate a non-physiological modifier(s) for *BHLHA9*. **D**. The physiological cis-regulatory element may be duplicated or non-duplicated, depending on its position relative to the size of the duplications. *BHLHA9* expression can be higher in small duplications than large duplications. **E**. The non-physiological modifier(s) can be transferred to various positions of the duplication positive chromosome 17, depending on the recombination places (see Model A). *BHLHA9* expression can be higher in small duplications than large duplications irrespective of the position of the modifier(s).

small duplications were ascertained primarily due to developmental retardation and limb malformation, respectively [8]. It is likely that a physiological cis-regulatory element for *BHLHA9* (e.g., an enhancer) can frequently but not invariably work on both of the duplicated *BHLHA9* when the duplication size is small but is usually incapable of working on duplicated *BHLHA9* when the duplication size is large, probably because of the difference in the chromatin structure (see Model D in Figure 3). Similar findings have also been reported in other genes. For example, small (~150 kb) and relatively small (600–800 kb) duplications involving a putative testis-specific enhancer(s) for *SOX9* have caused 46,XX testicular and ovotesticular disorders of sex development respectively, whereas large duplications (~2 Mb) involving the enhancer(s) have permitted normal ovarian development in 46,XX individuals [30].

Thus, a plausible explanation may be that a range of limb malformations emerge when the effects of *BHLHA9* overdosage exceed the threshold for the development of SHFM, SHFLD, or GWC, depending on the conditions of

other genetic and environmental factors including the size of duplications/triplications as an important but not definitive factor. One may argue that this notion is inconsistent with the apparent anticipation phenomenon that is suggested by the rare patient-to-carrier transmission and the frequent carrier-to-patient transmission of the duplications/triplications, because no specific factor(s) exaggerating the development of limb malformations is postulated in the next generation. However, the skewed transmission pattern would primarily be ascribed to ascertainment bias rather than anticipation [31]. Indeed, while clinically normal parents of disease positive children would frequently be examined for the underlying genetic factor(s) of the children, clinically normal children born to disease positive parents would not usually be studied for such factor(s), as exemplified in this study. Similarly, the frequent patient-to-patient transmission of the duplications/triplications would also be ascribed to ascertainment bias, because molecular studies would preferentially be performed in such families. Nevertheless, the apparently autosomal dominant inheritance pattern of limb malformations in several families may still suggest the relevance of a non-physiological *cis*-acting modifier(s) (see Models A and E in Figure 3). It is possible that such a modifier(s), once transferred onto the duplication/triplication positive chromosome 17, is usually co-transmitted with the duplications/triplications, leading to a specific condition in which the effects of *BHLHA9* overdosage frequently but not invariably exceed the threshold for the development of limb malformations in offsprings with the duplications/triplications.

Remarks

Several matters should be pointed out in the present study. First, in contrast to diverse duplication sizes in non-Japanese populations [5-9], the size of the genomic segment subjected to duplications/triplications was identical in this study. Since families 1–27 were derived from various places of Japan, there is no selection bias in terms of a geographic distribution. Rather, since the small duplications/triplications identified in this study were not associated with developmental retardation, it is likely that they spread throughout Japan primarily via carriers with normal fitness and were found via patients with limb malformations. Obviously, this notion does not exclude the possible presence of other types of duplications/triplications at 17p13.3 in Japan. Second, except for the duplications/triplications at 17p13.3, we could reveal a homozygous *WNT10B* mutation (SHFM6) only in a single SHFM family and chromosome 10q24 duplications (SHFM3) only in three SHFM families. Thus, underlying factors are still unknown in the remaining 20 families, although tiny deletions and/or duplications affecting the known SHFM loci might have been overlooked because of the low resolution of the array. In addition, although all the probands had a normal karyotype, there might be cryptic translocations and/or inversions involving the known SHFM loci. Third, no deletion of *BHLHA9* was identified in the 51 probands and in the 200 control subjects. This argues against the relevance of *BHLHA9* haploinsufficiency to limb malformations, and coincides with the Japanese founder duplication being produced by a replication-mediated mechanism rather than an interchromatid/interchromosomal (but not an intrachromatid) NAHR that can lead to both deletions and duplications as a mirror image [17]. Furthermore, it remains to be determined (i) whether gain-of-function mutations (and possibly loss-of-function mutations as well) of *BHLHA9* are identified in patients with limb malformations, (ii) whether duplications/triplications involving *BHLHA9* underlie limb malformations other than SHFM, SHFLD, and GWC, and (iii) whether *BHLHA9*-containing duplications/triplications are also the most frequent underlying factors for limb malformations in non-Japanese populations.

Conclusions

The results imply that (i) duplications/triplications involving *BHLHA9* at chromosome 17p13.3 constitute a strong susceptibility factor for the development of a range of limb malformations including SHFM, SHFLD, and GWC; (ii) the Japanese founder duplication was generated by a replication-based mechanism and spread with subsequent triplication and haplotype modification through recombination-based mechanisms; and (iii) clinical variability appears to be due to multiple factors including the size of duplications/triplications. Thus, the present study provides useful information on the development of limb malformations.

Additional files

Additional file 1: Table S1. Primers utilized in this study.

Additional file 2: Figure S1. Real-time PCR analysis.

Additional file 3: Figure S2. *D17S1174* analysis in 200 Japanese control subjects, showing discontinuous distribution of the CA repeat numbers, as observed in the Japanese families with limb malformations.

Additional file 4: Table S2. *In silico* analysis for specific structures around the breakpoint-flanking regions and control regions.

Additional file 5: Table S3. Phenotypes in patients/subjects with increased copy number of *BHLHA9*.

Additional file 6: Figure S3. Genomic region encompassing *BHLHA9* examined in this study.

Additional file 7: Table S4. Polymorphism analysis of rs3951819 (A/G SNP) in *BHLHA9*.

Additional file 8: Table S5. Polymorphism analysis of rs34201045 (4 bp insertion) in *TP63*.

Additional file 9: Figure S4. Genomic basis of the Japanese founder copy number gain.

Abbreviations

AER: Apical ectodermal ridge; CEP17: Centromere of chromosome 17; CGH: Comparative genomic hybridization; Dup: Duplication; FoSTeS: Fork stalling and template switching; GWC: Gollop-Wolfgang complex; L: Lower; LBD: Long bone defect; MMBIR: Microhomology-mediated break-induced replication; NAHR: Non-allelic homologous recombination; N.E.: Not examined; NHEJ: Non-homologous end-joining; qPCR: Quantitative real-time PCR; SF: Split foo17)t; SH: Split hand; SHFLD: SHFM with long bone deficiency; SHFM: Split-hand/foot malformation; Trip: Triplication; U: Upper.

Competing interests

The authors have nothing to declare.

Authors' contributions

Molecular analysis using human samples was performed by EN, HK, FK. RY, SN, SW, KY, TT, SS, MF, and TT, ST, and SY; clinical assessment and blood sampling by RK, HT, SM, TK, TH, MK, AS, KS, HO, NH, HN, EH, TN, HY, GN, and TO; design of this study and interpretations of the data by HA, SI, and TO; and paper writing by TO. All authors read and approved the final manuscript.

Acknowledgements

This work was supported in part by Grants-in-Aid for Scientific Research on Innovative Areas [22132004-A01] from the Ministry of Education, Culture, Sports, Science and Technology, by Grant for Research on Intractable Diseases from the Ministry of Health, Labor and Welfare [H24-048], and by Grants from National Center for Child Health and Development [23A-1, 24–7]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author details

[1]Department of Pediatrics, Hamamatsu University School of Medicine, Hamamatsu 431-3192, Japan. [2]Laboratory of Bone and Joint Diseases, Center for Integrative Medical Sciences, RIKEN, Tokyo, Japan. [3]Department of Orthopedic Surgery, Tokyo, Japan. [4]Division of Medical Genetics, National Center for Child Health and Development, Tokyo, Japan. [5]Section of Clinical Genetics, Department of Pediatrics, Tenshi Hospital, Sapporo, Japan. [6]Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai, Japan. [7]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan. [8]Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan. [9]Department of Pediatrics, Keio University School of Medicine, Tokyo, Japan. [10]Department of Orthopedics, National Rehabilitation Center for Disabled Children, Tokyo, Japan. [11]Division of Medical Genetics, Saitama Children's Medical Center, Saitama, Japan. [12]Department of Rehabilitation Medicine, University of Tokyo, Tokyo, Japan. [13]Department of Genetic Counseling, Graduate School of Humanities and Sciences, Ochanomizu University, Tokyo, Japan. [14]Department of Orthopedic Surgery, Japanese Red Cross Nagoya Daiichi Hospital, Nagoya, Japan. [15]Department of Pediatrics, Dokkyo Medical University Koshigaya Hospital, Koshigaya, Japan. [16]Division of Medical Genetics, Tokyo, Japan. [17]Department of Pediatric Imaging, Tokyo Metropolitan Children's Medical Center, Tokyo, Japan. [18]Division of Neurology/Molecular Brain Science, Kobe University Graduate School of Medicine, Kobe, Japan. [19]Department of Systems Biomedicine, National Research Institute for Child Health and Development, Tokyo, Japan. [20]Department of Systems Biomedicine, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan. [21]Department of Molecular Endocrinology, National Research Institute for Child Health and Development, Tokyo, Japan. [22]Present address: Laboratory of Metabolism Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA.

Received: 15 April 2014 Accepted: 22 July 2014
Published online: 21 October 2014

References

1. Duijf PH, van Bokhoven H, Brunner HG: **Pathogenesis of split-hand/split-foot malformation.** *Hum Mol Genet* 2003, **12**:R51–60.
2. Gurrieri F, Everman DB: **Clinical, genetic, and molecular aspects of split-hand/foot malformation: an update.** *Am J Med Genet A* 2013, **161A**:2860–2672.
3. Lango Allen H, Caswell R, Xie W, Xu X, Wragg C, Turnpenny PD, Turner CL, Weedon MN, Ellard S: **Next generation sequencing of chromosomal rearrangements in patients with split-hand/split-foot malformation provides evidence for DYNC1/1 exonic enhancers of DLX5/6 expression in humans.** *J Med Genet* 2014, **51**:264–267.
4. Lezirovitz K, Maestrelli SR, Cotrim NH, Otto PA, Pearson PL, Mingroni-Netto RC: **A novel locus for split-hand/foot malformation associated with tibial hemimelia (SHFLD syndrome) maps to chromosome region 17p13.1-17p13.3.** *Hum Genet* 2008, **123**:625–631.
5. Armour CM, Bulman DE, Jarinova O, Rogers RC, Clarkson KB, DuPont BR, Dwivedi A, Bartel FO, McDonell L, Schwartz CE, Boycott KM, Everman DB, Graham GE: **17p13.3 microduplications are associated with split-hand/foot malformation and long-bone deficiency (SHFLD).** *Eur J Hum Genet* 2011, **19**:1144–1151.
6. Klopocki E, Lohan S, Doelken SC, Stricker S, Ockeloen CW, Soares Thiele de Aguiar R, Lezirovitz K, Mingroni Netto RC, Jamsheer A, Shah H, Kurth I, Habenicht R, Warman M, Devriendt K, Kordass U, Hempel M, Rajab A, Mäkitie O, Naveed M, Radhakrishna U, Antonarakis SE, Horn D, Mundlos S: **Duplications of BHLHA9 are associated with ectrodactyly and tibia hemimelia inherited in non-Mendelian fashion.** *J Med Genet* 2012, **49**:119–125.
7. Petit F, Andrieux J, Demeer B, Collet LM, Copin H, Boudry-Labis E, Escande F, Manouvrier-Hanu S, Mathieu-Dramard M: **Split-hand/foot malformation with long-bone deficiency and BHLHA9 duplication: two cases and expansion of the phenotype to radial agenesis.** *Eur J Med Genet* 2013, **56**:88–92.
8. Curry CJ, Rosenfeld JA, Grant E, Gripp KW, Anderson C, Aylsworth AS, Saad TB, Chizhikov VV, Dybose G, Fagerberg C, Falco M, Fels C, Fichera M, Graakjaer J, Greco D, Hair J, Hopkins E, Huggins M, Ladda R, Li C, Moeschler J, Nowaczyk MJ, Ozmore JR, Reitano S, Romano C, Roos L, Schnur RE, Sell S, Suwannarat P, Svaneby D, *et al*: **The duplication 17p13.3 phenotype: analysis of 21 families delineates developmental, behavioral and brain abnormalities, and rare variant phenotypes.** *Am J Med Genet A* 2013, **161A**:1833–1852.
9. Luk HM, Wong VC, Lo IF, Chan KY, Lau ET, Kan AS, Tang MH, Tang WF, She WM, Chu YW, Sin WK, Chung BH: **A prenatal case of split-hand malformation associated with 17p13.3 triplication - a dilemma in genetic counseling.** *Eur J Med Genet* 2014, **57**:81–84.
10. Petit F, Jourdain AS, Andrieux J, Baujat G, Baumann C, Beneteau C, David A, Faivre L, Gaillard D, Gilbert-Dussardier B, Jouk PS, Le Caignec C, Loget P, Pasquier L, Porchet N, Holder-Espinasse M, Manouvrier-Hanu S, Escande F: **Split hand/foot malformation with long-bone deficiency and BHLHA9 duplication: report of 13 new families.** *Clin Genet* 2014, **85**:464–469.
11. Kano H, Kurosawa K, Horii E, Ikegawa S, Yoshikawa H, Kurahashi H, Toda T: **Genomic rearrangement at 10q24 in non-syndromic split-hand/split-foot malformation.** *Hum Genet* 2005, **118**:477–483.
12. Matsuyama J, Mabuchi A, Zhang J, Iida A, Ikeda T, Kimizuka M, Ikegawa S: **A pair of sibs with tibial hemimelia born to phenotypically normal parents.** *J Hum Genet* 2003, **48**:173–176.
13. Kagami M, Sekita Y, Nishimura G, Irie M, Kato F, Okada M, Yamamori S, Kishimoto H, Nakayama M, Tanaka Y, Matsuoka K, Takahashi T, Noguchi M, Tanaka Y, Masumoto K, Utsunomiya T, Kouzan H, Komatsu Y, Ohashi H, Kurosawa K, Kosaki K, Ferguson-Smith AC, Ishino F, Ogata T: **Deletions and epimutations affecting the human 14q32.2 imprinted region in individuals with paternal and maternal upd(14)-like phenotypes.** *Nat Genet* 2008, **40**:237–242.
14. Iida A, Okamoto N, Miyake N, Nishimura G, Minami S, Sugimoto T, Nakashima M, Tsurusaki Y, Saitsu H, Shiina M, Ogata K, Watanabe S, Ohashi H, Matsumoto N, Ikegawa S: **Exome sequencing identifies a novel INPPL1 mutation in opsismodysplasia.** *J Hum Genet* 2013, **58**:391–394.
15. Cer RZ, Donohue DE, Mudunuri US, Temiz NA, Loss MA, Starner NJ, Halusa GN, Volfovsky N, Yi M, Luke BT, Bacolla A, Collins JR, Stephens RM: **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.
16. 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.
17. Gu W, Zhang F, Lupski JR: **(2008) Mechanisms for human genomic rearrangements.** *Pathogenetics* 2008, **1**:4.
18. Vissers LE, Bhatt SS, Janssen IM, Xia Z, Lalani SR, Pfundt R, Derwinska K, de Vries BB, Gilissen C, Hoischen A, Nesteruk M, Wisniowiecka-Kowalnik B, Smyk

M, Brunner HG, Cheung SW, van Kessel AG, Veltman JA, Stankiewicz P: **Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture.** *Hum Mol Genet* 2009, **18**:3579–3593.
19. 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.
20. 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, **2011**(22):875–885.
21. Ugur SA, Tolun A: **Homozygous WNT10b mutation and complex inheritance in Split-Hand/Foot Malformation.** *Hum Mol Genet* 2008, **17**:2644–2653.
22. Oort PJ, Warden CH, Baumann TK, Knotts TA, Adams SH: **Characterization of Tusc5, an adipocyte gene co-expressed in peripheral neurons.** *Mol Cell Endocrinol* 2007, **276**:24–35.
23. Manning JT, Scutt D, Wilson J, Lewis-Jones DI: **The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen.** *Hum Reprod* 1998, **13**:3000–3004.
24. Manning JT, Trivers RL, Singh D, Thornhill R: **The mystery of female beauty.** *Nature* 1999, **399**:214–215.
25. Williams TJ, Pepitone ME, Christensen SE, Cooke BM, Huberman AD, Breedlove NJ, Breedlove TJ, Jordan CL, Breedlove SM: **Finger-length ratios and sexual orientation.** *Nature* 2000, **404**:455–456.
26. Heald AH, Ivison F, Anderson SG, Cruickshank K, Laing I, Gibson JM: **Significant ethnic variation in total and free testosterone concentration.** *Clin Endocrinol* 2003, **58**:262–266.
27. Manning JT, Stewart A, Bundred PE, Trivers RL: **Sex and ethnic differences in 2nd to 4th digit ratio of children.** *Early Hum Dev* 2004, **80**:161–168.
28. Wang G, Vasquez KM: **Non-B DNA structure-induced genetic instability.** *Mutat Res* 2006, **598**:103–119.
29. Zhao J, Bacolla A, Wang G, Vasquez KM: **Non-B DNA structure-induced genetic instability and evolution.** *Cell Mol Life Sci* 2010, **67**:43–62.
30. Benko S, Gordon CT, Mallet D, Sreenivasan R, Thauvin-Robinet C, Brendehaug A, Thomas S, Bruland O, David M, Nicolino M, Labalme A, Sanlaville D, Callier P, Malan V, Huet F, Molven A, Dijoud F, Munnich A, Faivre L, Amiel J, Harley V, Houge G, Morel Y, Lyonnet S: **Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development.** *J Med Genet* 2011, **48**:825–830.
31. Fraser FC: **Trinucleotide repeats not the only cause of anticipation.** *Lancet* 1997, **350**:459–460.

doi:10.1186/s13023-014-0125-5
Cite this article as: Nagata *et al.*: Japanese founder duplications/triplications involving *BHLHA9* are associated with split-hand/foot malformation with or without long bone deficiency and Gollop-Wolfgang complex. *Orphanet Journal of Rare Diseases* 2014 **9**:125.

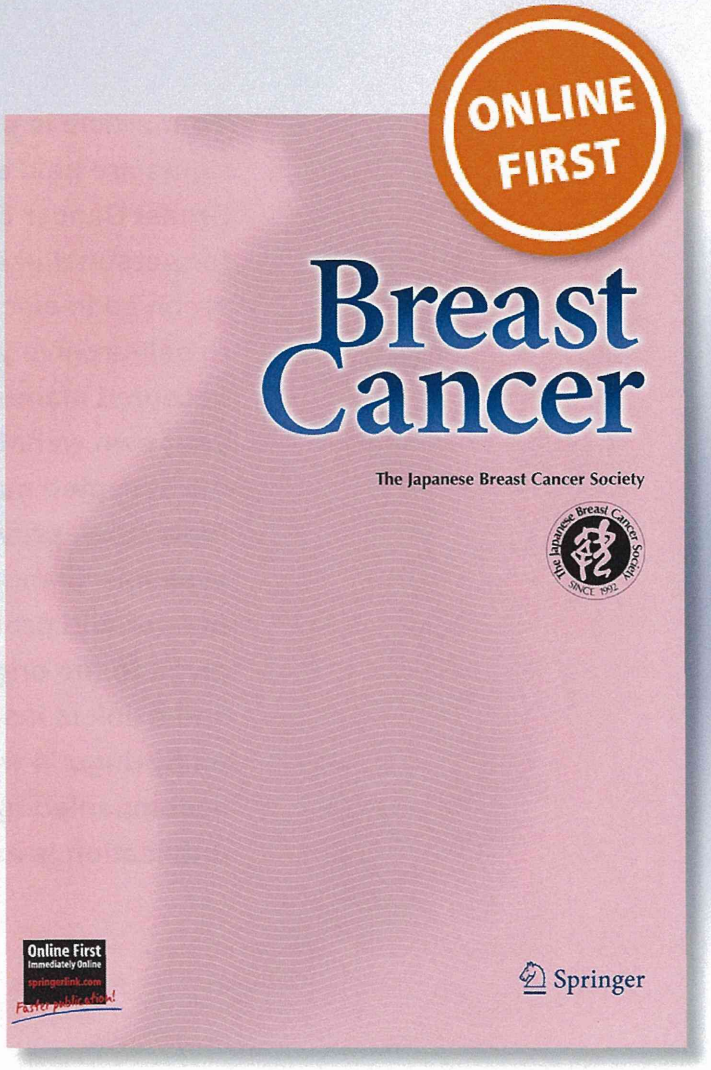
Lack of genomic rearrangements involving the aromatase gene CYP19A1 *in breast cancer*

Maki Fukami, Junichi Suzuki, Kazuhiko Nakabayashi, Ryo Tsunashima, Tsutomu Ogata, Makio Shozu & Shinzaburo Noguchi

Breast Cancer

ISSN 1340-6868

Breast Cancer
DOI 10.1007/s12282-013-0471-5



Springer

Breast Cancer
DOI 10.1007/s12282-013-0471-5

RAPID COMMUNICATION

Lack of genomic rearrangements involving the aromatase gene *CYP19A1* in breast cancer

Maki Fukami · Junichi Suzuki · Kazuhiko Nakabayashi · Ryo Tsunashima · Tsutomu Ogata · Makio Shozu · Shinzaburo Noguchi

Received: 25 February 2013 / Accepted: 15 April 2013

Abstract Increased intratumoral expression of aromatase, the key enzyme for estrogen biosynthesis, is predicted to be of critical importance in the development of breast cancer. Recently, several germline rearrangements at 15q21 have been shown to cause overexpression of the aromatase gene *CYP19A1* and resulting aromatase excess syndrome. To determine whether submicroscopic genomic rearrangements at 15q21 are involved in aromatase overexpression in breast cancer tissues, we investigated copy-number alterations in genomic DNA obtained from 44 tumor samples. Comparative genomic hybridization analysis identified no deletion or duplication at 15q21 in the 44 samples. These results, in conjunction with previous data, indicate that aromatase overexpression in breast cancer tissues is likely to result from a promoter switch of *CYP19A1* and/or accumulation of *CYP19A1*-expressing cells, rather than from cryptic transactivation of *CYP19A1* because of genomic rearrangements at 15q21.

Keywords Breast cancer · Aromatase · *CYP19A1* · Gene expression · Genomic rearrangement

Abbreviations
AEXS Aromatase excess syndrome
CGH Comparative genomic hybridization

Introduction

Estrogens stimulate the proliferation of breast cancer cells, possibly by transactivating growth factor genes or oncogenes [1, 2]. Although estrogens are usually synthesized in a variety of tissues, including ovaries, skin, and fat, they can also be produced locally in breast cancer tissues [1, 3, 4]. Previous studies have revealed enhanced expression of aromatase, the key enzyme for estrogen biosynthesis, in breast cancer epithelial cells and in immature fibroblasts around the malignant cells [1, 4, 5]. Aromatase overexpression in breast cancer tissues is predicted to be critically important in tumor development [1, 5].

Recently, we and other groups have described an autosomal dominant disorder referred to as aromatase excess syndrome (AEXS), which is characterized by gynecomastia resulting from increased expression of the aromatase gene *CYP19A1* [6–8]. The molecular bases of AEXS include tandem duplications at 15q21.2 involving the promoter region of *CYP19A1* and a variety of deletions and

M. Fukami (✉) · J. Suzuki · T. Ogata
Department of Molecular Endocrinology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya, Tokyo 157-8535, Japan
e-mail: fukami-m@ncchd.go.jp

J. Suzuki
Department of Pediatrics and Child Health, Nihon University School of Medicine, Tokyo, Japan

K. Nakabayashi
Department of Maternal-Fetal Biology, National Research Institute for Child Health and Development, Tokyo, Japan

R. Tsunashima · S. Noguchi
Department of Breast and Endocrine Surgery, Osaka University Graduate School of Medicine, Osaka, Japan

T. Ogata
Department of Pediatrics, Hamamatsu University School of Medicine, Hamamatsu, Japan

M. Shozu
Department of Reproductive Medicine, Graduate School of Medicine, Chiba University, Chiba, Japan

Published online: 30 April 2013

Table 1 Breast cancer samples examined in this study

Sample	ER[a]	PR[a]	HER2[b]	Menopausal status	Age (years)
1	Positive	Positive	Negative	Postmenopausal	78
2	Positive	Positive	Negative	Postmenopausal	63
3	Positive	Positive	Negative	Postmenopausal	51
4	Positive	Positive	Negative	Postmenopausal	56
5	Positive	Positive	Negative	Postmenopausal	55
6	Positive	Positive	Negative	Postmenopausal	65
7	Positive	Positive	Negative	Postmenopausal	58
8	Positive	Positive	Negative	Postmenopausal	71
9	Positive	Positive	Negative	Postmenopausal	64
10	Positive	Positive	Negative	Postmenopausal	66
11	Positive	Positive	Negative	Postmenopausal	59
12	Positive	Positive	Negative	Postmenopausal	81
13	Positive	Positive	Negative	Postmenopausal	61
14	Positive	Positive	Negative	Postmenopausal	68
15	Positive	Positive	Negative	Postmenopausal	65
16	Negative	Negative	Negative	Postmenopausal	67
17	Negative	Negative	Negative	Premenopausal	49
18	Negative	Negative	Positive	Postmenopausal	68
19	Negative	Negative	Negative	Postmenopausal	70
20	Negative	Negative	Positive	Premenopausal	42
21	Positive	Positive	Negative	Postmenopausal	73
22	Positive	Positive	Negative	Postmenopausal	56
23	Positive	Positive	Negative	Postmenopausal	56
24	Positive	Positive	Negative	Postmenopausal	57
25	Positive	Positive	Negative	Postmenopausal	56
26	Positive	Positive	Negative	Postmenopausal	70
27	Positive	Positive	Negative	Postmenopausal	71
28	Positive	Positive	Negative	Postmenopausal	80
29	Positive	Positive	Negative	Postmenopausal	60
30	Positive	Positive	Negative	Postmenopausal	80
31	Positive	Positive	Negative	Postmenopausal	72
32	Positive	Positive	Negative	Postmenopausal	64
33	Positive	Positive	Negative	Postmenopausal	65
34	Positive	Positive	Negative	Postmenopausal	80
35	Positive	Positive	Negative	Postmenopausal	59
36	Positive	Positive	Negative	Postmenopausal	60
37	Positive	Positive	Negative	Postmenopausal	62
38	Positive	Positive	Negative	Postmenopausal	63
39	Positive	Positive	Negative	Postmenopausal	58
40	Positive	Positive	Negative	Postmenopausal	69
41	Positive	Positive	Negative	Postmenopausal	65
42	Positive	Positive	Negative	Postmenopausal	70
43	Positive	Positive	Negative	Postmenopausal	62
44	Positive	Positive	Negative	Postmenopausal	69

ER estrogen receptor; *PR* progesterone receptor; *HER2* human epidermal growth factor receptor 2

[a] ER and PR were analyzed by use of immunohistochemical assays. The results were defined as positive when more than 10 % of the tumor cells were positively stained

[b] HER2 amplification was analyzed by fluorescence in situ hybridization analysis. Signal intensities >2.0 were assessed as positive

inversions at 15q21.1-21.3 that create chimeric genes consisting of coding exons of *CYP19A1* and the promoter-associated exons of other genes. Identification of highly heterogeneous rearrangements in patients with AEXS indicates that the chromosomal region around *CYP19A1* is particularly vulnerable to genomic abnormalities [8]. Therefore, submicroscopic rearrangements at 15q21 generated in germ cells or in somatic cells may underlie

aromatase overexpression in breast cancer tissues. Here, we performed copy-number analysis of the 15q21 region in 44 genomic DNA samples obtained from breast cancer tissues.

Materials and methods

Tumor tissue samples

This study was approved by the Institutional Review Board Committees at Osaka University and the National Center for Child Health and Development. After obtaining written informed consent, breast cancer tissues were obtained at surgery from 44 Japanese females aged 42–80 years (Table 1). The samples were frozen in liquid nitrogen. All 44 tumors were invasive ductal carcinoma. The samples were examined for levels of expression of estrogen receptor and progesterone receptor, and for amplification of human epidermal growth factor receptor 2, as described elsewhere [9] (Table 1).

Comparative genomic hybridization analysis

Genomic DNA was extracted from tumor tissues by use of a DNeasy blood and tissue kit (Qiagen, Maryland, USA). The DNA samples were subjected to comparative genomic hybridization (CGH) analyses by use of a custom-made oligoarray (8 × 60 k format; Agilent Technologies, Palo Alto, CA, USA). The array contained approximately 16,000 probes for the 1.5 Mb region at 15q21 (chr15:51,000,000–52,500,000; hg 19, build 37), together with several reference probes for other genomic regions. The procedure was conducted in accordance with the manufacturer's instructions. A genomic DNA sample obtained from leukocytes of an unaffected Japanese 58-year-old female was used as a control.

Results

CGH analysis

Comparative genomic hybridization analysis indicated a normal copy-number for exons 1–10 of *CYP19A1* in the 44 samples (Fig. 1a). Furthermore, no deletion or duplication was detected in the 1.5 Mb genomic region at 15q21 (Fig. 1a).

Discussion

Comparative genomic hybridization analysis identified no copy-number alterations at 15q21 in 44 breast cancer tissues. It is worth remarking that the method used in this

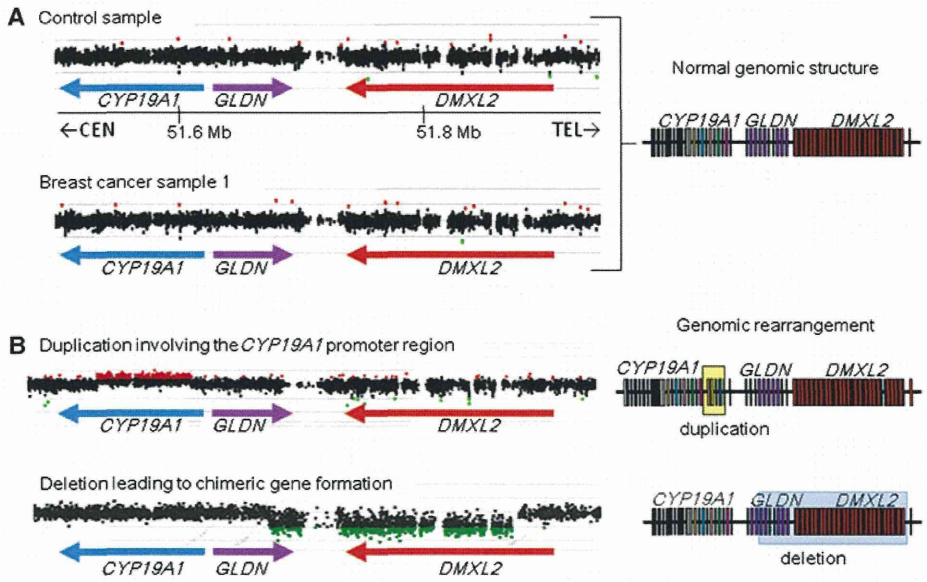


Fig. 1 **a** Comparative genomic hybridization (CGH) analysis (*left panel*) and schematic representation of the normal genomic structure in the control and breast cancer samples (*right panel*). The *black*, *red*, and *green dots* denote signals indicative of normal, increased (>+0.5), and reduced (<−1.0) copy-numbers, respectively. The *arrows* indicate genomic positions of *CYP19A1*, *GLDN*, and *DMXL2* (5′ → 3′). The positions correspond to the human genome reference assembly (UCSC genome browser, February 2009, hg19, build 37). **b** Genomic abnormalities previously identified in patients with aromatase excess syndrome. These germline rearrangements have been implicated in *CYP19A1* overexpression [8]. The results of CGH analysis (*left panel*) and schematic representation of the rearranged genome (*right panel*) are shown

study is capable of identifying all the duplications and deletions that have previously been associated with AEXS (Fig. 1b) [8], although copy-number-neutral inversions may not be detected by CGH analysis. These results imply that copy-number gains of *CYP19A1* and formation of chimeric genes arising from cryptic rearrangements are not the major factors underlying aromatase overexpression in breast cancer tissues. Furthermore, lack of genomic rearrangements in the 15q21 region in the 44 samples suggests that this region would not be a hotspot for somatic rearrangements, especially considering that breast cancer tissues are characterized by prominent genomic instability [10].

Previous studies have attributed aromatase overexpression in breast cancer tissues to alternative use of the non-coding exons 1 of *CYP19A1* [5, 11]. Of the 11 exons 1, exons 1.3, II, and 1.7 were found to predominantly control expression of *CYP19A1* in breast cancer tissues, whereas exon 1.4 functioned as the major promoter in normal breast tissues. This promoter switch seemed to result from the perturbed production of signal molecules in tumor tissues. For example, overproduction of prostaglandins has been associated with activation of exons 1.3 and II [1, 5, 11]. Furthermore, accumulation of immature adipose fibroblasts with strong expression of *CYP19A1* has also been proposed to contribute to aromatase excess in breast cancer tissues [5].

It should be noted that our conclusion is based on results from 44 samples only. Furthermore, *CYP19A1* mRNA levels were not analyzed in this study. In this regard, it is known that although increased aromatase activity is a common feature in breast cancer tissues, mRNA levels of *CYP19A1* vary among samples [12, 13]. Therefore, the absence of genomic rearrangements in our 44 samples may be ascribed to low expression of *CYP19A1* in these tissues. Indeed, genomic rearrangements at 15q21 may be hidden in a specific group of breast cancers with overexpression of *CYP19A1*. Further CGH and mRNA analysis for several tumor samples is necessary to clarify the presence or absence of an association between genomic rearrangements at 15q21 and intratumoral estrogen overproduction.

In summary, this study, in conjunction with previous data, implies that aromatase overexpression in breast cancer tissues is likely to result from a promoter switch of *CYP19A1* induced by trans-acting factors and/or accumulation of *CYP19A1*-expressing cells, rather than from cryptic transactivation of *CYP19A1* because of genomic rearrangements at 15q21.

Acknowledgments This work was supported by the Grant-in-Aid for Scientific Research on Innovative Areas (22132004) from the Ministry of Education, Culture, Sports, Science and Technology, by the Grant-in-Aid for Scientific Research (B) (23390249) 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, and by the Grants from National Center for Child Health and Development, from Takeda foundation, and from Daiichi-Sankyo Foundation of Life Science.

Conflict of interest The authors declare that no conflict of interests exists.

References

1. Bulun SE, Lin Z, Imir G, Amin S, Demura M, Yilmaz B, et al. Regulation of aromatase expression in estrogen-responsive breast and uterine disease: from bench to treatment. Pharmacol Rev. 2005;57:359–83.
2. Kulendran M, Salhab M, Mokbel K. Oestrogen-synthesising enzymes and breast cancer. Anticancer Res. 2009;29:1095–109.
3. Bulun SE, Chen D, Lu M, Zhao H, Cheng Y, Demura M, et al. Aromatase excess in cancers of breast, endometrium and ovary. J Steroid Biochem Mol Biol. 2007;106:81–96.
4. Sasano H, Miki Y, Nagasaki S, Suzuki T. In situ estrogen production and its regulation in human breast carcinoma: from endocrinology to intracrinology. Pathol Int. 2009;59:777–89.
5. Bulun SE, Lin Z, Zhao H, Lu M, Amin S, Reierstad S, et al. Regulation of aromatase expression in breast cancer tissue. Ann N Y Acad Sci. 2009;1155:121–31.
6. Demura M, Martin RM, Shozu M, Sebastian S, Takayama K, Hsu WT, et al. Regional rearrangements in chromosome 15q21 cause formation of cryptic promoters for the CYP19 (aromatase) gene. Hum Mol Genet. 2007;16:2529–41.
7. Shozu M, Sebastian S, Takayama K, Hsu WT, Schultz RA, Neely K, et al. Estrogen excess associated with novel gain-of-function mutations affecting the aromatase gene. N Engl J Med. 2003;348:1855–65.
8. Fukami M, Shozu M, Soneda S, Kato F, Inagaki A, Takagi H, 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:E1035–43.
9. Naoi Y, Kishi K, Tanei T, Tsunashima R, Tominaga N, Baba Y, et al. Development of 95-gene classifier as a powerful predictor of recurrences in node-negative and ER-positive breast cancer patients. Breast Cancer Res Treat. 2011;128:633–41.
10. Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R, et al. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer. 2006;6:96.
11. Khan SI, Zhao J, Khan IA, Walker LA, Dasmahapatra AK. Potential utility of natural products as regulators of breast cancer-associated aromatase promoters. Reprod Biol Endocrinol. 2011;9:91.
12. Bulun SE, Price TM, Aitken J, Mahendroo MS, Simpson ER. A link between breast cancer and local estrogen biosynthesis suggested by quantification of breast adipose tissue aromatase cytochrome P450 transcripts using competitive polymerase chain reaction after reverse transcription. J Clin Endocrinol Metab. 1993;77:1622–8.
13. Bulun SE, Simpson ER. Breast cancer and expression of aromatase in breast adipose tissue. Trends Endocrinol Metab. 1994;5:113–20.