Forensic Science International: Genetics 19 (2015) 81–85

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Case report

Concurrent copy number variations on chromosome 8 and 22 combined with mutation at FGA locus revealed in a parentage testing case

Yaran Yang[a,1], He Ren[b,1], Wei Chen[a], Bingbing Xie[a,c], Yan Wang[d], Yan Shi[e], Chong Chen[e], Chen Li[f], Le Yi[f], Xiangdong Fang[a], Jiangwei Yan[a,*]

[a] CAS Key Laboratory of Genome Sciences and Information, Beijing Institute ofGenomics, Chinese Academy of Sciences, Beijing, China
[b] Beijing Police Colle, Beijing, China
[c] University of Chinese Academy of Sciences, Beijing, China
[d] BaYi Children's Hospital, General Military Hospital of Beijing PLA, Beijing, China
[e] Beijing Tongda Shoucheng Institute of Forensic Science, Beijing, China
[f] Beijing Microread Genetics Co., Ltd, Beijing, China

ARTICLE INFO

Article history:
Received 13 March 2015
Received in revised form 22 June 2015
Accepted 2 July 2015

Keywords:
Copy number variations
Paternity testing
Short tandem repeats
Mutation
Tri-allele

ABSTRACT

Copy number variations (CNVs) are one of the major sources of human genetic diversity and are associated with rare genomic disorders as well as complex traits and diseases. A copy number variation was observed at the D8S1179 locus during routine STR based parentage testing, in which the child exhibited three alleles, "13, 15, 16", with the putative father a homozygous "15" and the mother homozygous "13". In addition, in the same testing case, there was a one-step mutation at the STR locus FGA, in which the putative father was a "22, 24", the mother was a "22, 25", and the child was a "22, 23". After further investigations by re-amplified with different primer sets, clone-based sequencing, karyotype analysis and whole-genome SNP analysis, the results showed that the child had the CNVs at chromosome 8q24.3 and 22q11.21. In conclusion, for parentage testing cases encountered with tri-allele patterns, more testings, such as cloning sequencing, karyotyping, or even whole genome analysis, as well as more appropriate statistical estimations might be conducted to further confirm or exclude the relationship.

1. Introduction

The human genome shows extensive copy number variations (CNVs) and the presence of variable numbers may range from about one kilobase to several megabases in size [1]. CNVs can be caused by structural rearrangements of the genome such as deletions, duplications, inversions, and translocations. Since CNVs correspond to relatively large regions of the genome that have been deleted (fewer than the normal number) or duplicated (more than the normal number) on certain chromosomes, the genotyping results of the genetic markers located in these regions may be affected and lead to null allele, tri-allele and peak height imbalance in short tandem repeats (STRs) markers routinely used in parentage testing. Therefore, false exclusion and incorrect interpretations may occur in these scenarios.

In this study, we report a paternity testing case with two concurrent CNVs at Chromosome 8q24.3 and 22q11.21 from the child and a one-step mutation of FGA between the father and the child. These genetic variables caused abnormal genotyping results for the corresponding STR loci.

2. Materials and methods

2.1. STRs analysis and clone-based sequencing confirmation

Blood samples from the putative father, the child (male) and the mother were collected with informed consent. The research protocol was approved by the ethical review committees of the Beijing Institute of Genomics (Protocol name: a study on the copy number variations revealed in a parentage testing case no. 2014032, date adopted: September 15, 2014). DNA was extracted by the Chelex-100 resin method [2] for all the STR amplifications.

* Corresponding author. Fax.: +86 10 8409 7720.
E-mail address: yanjw@big.ac.cn (J. Yan).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.fsigen.2015.07.002
1872-4973/

Genomic DNA for SNPs Chip was extracted from peripheral blood samples by QIAamp DNA Blood Midi kit (QIAGEN), according to the manufacturer's protocol. The quantity of the recovered DNA was determined by Qubit® Quantitation System (Invitrogen, CA, USA), according to the manufacturer's specifications.

First, 1 ng DNA was amplified with commercial STR typing kit GoldenEye™ 20A (Peoplespot Incorporation, China) which included 19 autosomal STR loci [3]. The same amount of DNA was also re-amplified with another commercial STR typing kit AmpFlSTR Identifiler® (Life Technologies, USA), which had 15 overlapped autosomal STR loci with GoldenEye™ 20A, including the loci D8S1179 and FGA. Polymerase chain reaction (PCR) was performed according to the manufacturer's recommendations. The amplified PCR products were separated by capillary electrophoresis on a genetic analyzer (ABI PRISM 3130XL, Applied Biosystems, USA) and analyzed with GeneMapper ID-X (Applied Biosystems, USA).

Sequencing reactions were performed to analyze the alleles of the loci D8S1179, FGA and D22GATA198B05. The primer sequences of loci D8S1179 and FGA were obtained from the STRbase website (http://www.cstl.nist.gov/biotech/strbase/) and the loci D22GATA198B05 was designed using Primer5 software according to the databases UCSC Genome Browser (http://genome.ucsc.edu/), and were as follows: D8S1179 Forward – ATTGCAACTTATATGTATTTTTGTATTTCATG, D8S1179 Reverse – ACCAAATTGTGTTCATGAGTATAGTTTC, FGA Forward – GGCTGCAGGGCATAACATTA, FGA Reverse – ATTCTATGACTTTGCGCTTCAGGA, D22GATA198B05 Forward – GGTCTCCAGGCGGCCTGTCGT, D22GATA198B05 Reverse – TTGATAAGATTTAGGATTGAT. PCR products were cloned and sequenced using the TOPO TA cloning kit (Invitrogen, Carlsbad, CA, USA) and BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) using M13 forward and reverse primers according to the manufacturer's recommendations. Capillary electrophoresis was performed on a genetic analyzer (ABI PRISM 3130XL, Applied Biosystems, Foster City, CA) with Data Collection Software V3.0. The sequence data was analyzed using DNA sequencing analysis software V5.2 (Applied Biosystems).

2.2. Karyotype analysis

Karyotype analysis was performed on conventional GTG-banded metaphases prepared from peripheral blood samples obtained from the child and his parents using standard protocols [4].

2.3. CNV analysis with whole genome SNPs Chip

250 ng genomic DNA for each sample was used to performe the CNV analysis with Illumina HumnaOmni ZhongHua-8 BeadChip (Illumina, USA) according to the manufacturer's protocols. The 900,015 markers on this chip captured 81% variation with minor allele frequency >5% in East Asians. The genotyping data, B allele frequency (BAF) and log R ratio (LRR) of each marker were collected, and the copy number of each genome loci was calculated using GISTIC 2.0 software package [5].

3. Results

The 19 autosomal STR genotypes of the putative father, the mother, and the child were showed in Table 1. At the D8S1179 locus, the putative father, the mother, and the child had genotypes "15", "13", and "13, 15, 16", respectively. The same results were obtained from both Identifiler® and GoldenEye™ 20A (Fig. 1). In addition, there was a mutation at the STR locus FGA, in which the putative father was a "22, 24", the mother was a "22, 25", and the child was a "22, 23". The Combined Paternity Index (CPI) for trios paternity excluding D8S1179 was 3,226,464.

Table 1
Nineteen autosomal STR genotyping of the father, mother and the child. A Tri-allele was found in D8S1179 (underlines) and one-step mutation was found in FGA (italic).

Loci	Father	Child	Mother	Paternity index
D19S433	13,15.2	13,14	14	1.871
D5S818	10,12	10,11	11	2.203
D21S11	31.2,32.2	28,32.2	28,32.2	1.779
D18S51	13,15	13,14	12,14	4.878
D6S1043	14,20	16,20	16	9.259
D3S1358	15,17	15	15,17	1.219
D13S317	11,12	11,12	8,12	2.1
D7S820	11,12	10,11	10	1.233
D16S539	12	9,12	9	5.195
CSF1PO	10,11	11	11	2.433
Penta D	9,14	9,14	8,9	12.755
vWA	16,18	16,19	17,19	3.253
TPOX	8	8,9	8,9	3.516
Penta E	9,10	10,20	20	12.5
TH01	7,9	9	9	0.973
D12S391	17,19	17,19	17,20	2.033
D2S1338	20,22	22,23	23,24	9.671
FGA	*22,24*	*22,23*	*22,25*	0.007
D8S1179	15	13,15,16	13	–
CPI				32,26,464

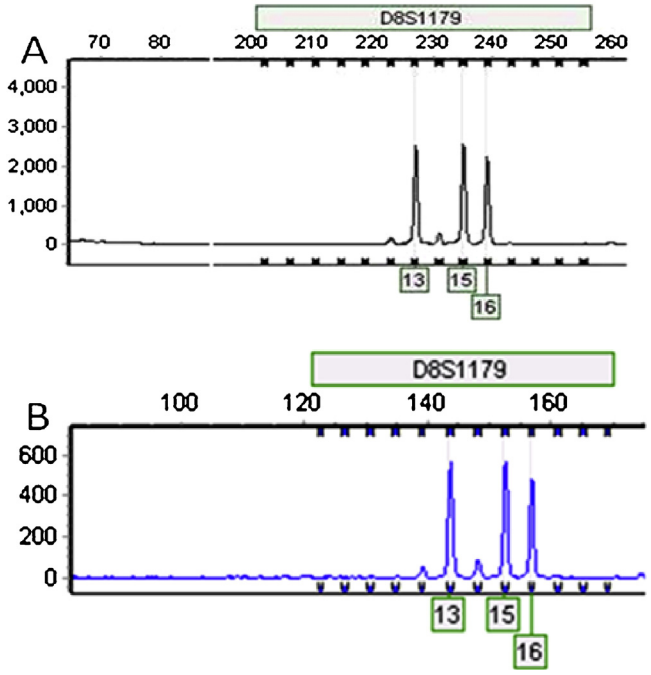

Fig. 1. Tri-allelic patterns observed at locus D8S1179. (A) Electropherogram profile amplified with AmpFlSTR Identifiler®. (B) Electropherogram profile amplified with GoldenEye™ 20A.

The clone-based sequencing results of the D8S1179 locus showed that sequence heterogeneity was observed between the two 15 alleles of the putative father, with one repeat structure was $[TCTA]_2[TCTG]\,[TCTA]_{12}$ and the other was $[TCTA]_1[TCTG]\,[TCTA]_{13}$. Sequence heterogeneity was also observed between the two 13 alleles of the mother, in which the repeat structures were $[TCTA]_1[TCTG]\,[TCTA]_{11}$ and $[TCTA]_{13}$, respectively. The repeat structures of the alleles 13, 15 and 16 of the child were $[TCTA]_1[TCTG]\,[TCTA]_{11}$, $[TCTA]_1[TCTG]\,[TCTA]_{13}$ and $[TCTA]_2[TCTG]\,[TCTA]_{13}$, respectively. The sequencing result confirmed that the allele 13 and allele 15 of the child was maternal and paternal, respectively. But the source of the allele 16 of child was not clear.

The genotypes at FGA locus of the putative father, the mother, and the child are 22/24, 22/25, and 22/23, respectively, with the corresponding sequences as follows: $[TTTC]_3TTTTTTCT[CTTT]_{14}CTCC[TTCC]_2/[TTTC]_3TTTTTTCT[CTTT]_{16}CTCC[TTCC]_2$, $[TTTC]_3TTTT\backslash TTCT[CTTT]_{14}CTCC[TTCC]_2/[TTTC]_3TTTTTTCT[CTTT]_{17}CTCC[TTCC]_2$, $[TTTC]_3TTTTTTCT[CTTT]_{14}CTCC[TTCC]_2/[TTTC]_3TTTT\ TTCT[CTTT]_{15}CTCC[TTCC]_2$. No SNP was found at the franking regions of the allele 22 for all three individuals.

According to the sequences of the alleles and franking regions of the alleles, the allele 22 of the child could come from either the putative father (a one-step mutation) or the mother (a two-step mutation). Since the higher of the mutation steps, the less likely the mutation would happen, it is more likely that the FGA allele 23 of the child was inherited from the putative Father's 24 with a one-step mutation. The PI of FGA can be calculated accordingly [6].

The result of Karyotype analysis showed that the G-banding revealed the abnormal non-mosaic 47, XY + mar Karyotype [7] in metaphase cells in the child (Fig. 2). Both parents had normal karyotype (data not shown). The size of an additional marker chromosome was similar to G group chromosomes.

By analyzing the genotyping data of whole genome SNPs Chip, we identified a 46 Mb copy number gain on the q arm of chromosome 8 (chr8: 126,033,450-146,364,022; UCSC Genome Browser hg19) and a 10 Mb copy number gain on p arm of chromosome 22(chr22: 1-20,189,390; UCSC Genome Browser hg19) on the genome of the child (Fig. 3). BAF calculates the ratio of hybridization intensity between alternative allele to the total intensity of one SNP locus. LRR compares the observed total hybridization intensity in the experiment to the expected total hybridization intensity which obtained from large pilot study of Illumina bead chips in one SNP locus. In this case, the copy numbers of these 2 loci were 3 according to LRR of SNPs in these regions. After calculating BAF of heterozygote SNPs in these 2CNV regions, we found the paternal alleles show an average frequency of 0.66, and maternal alleles show 0.34. This result indicated the copy number gains were present in the paternal genome. Considering the blood sample from the putative father showed normal copy number in these 2 regions, the amplifications might be caused by copy number alteration during spermatogenesis.

As the result of SNP genotyping and Karyotype both showed that there were copy number variants not only at chromosome 8q24.3 but also at chromosome 22q11.21, a third STR typing kit Microreader™ 23sp ID system (Microread Genetics Incorporation, China) which had 22 new autosomal STR loci was used to verify this. According to the product manual, one of the 22 loci, the locus D22GATA198B05 was located within the chromosome 22q11.21 region. No inconsistencies were observed between the putative father, mother and the child on all these 22 loci (Table 2). The electropherogram of D22GATA198B05 locus showed obviously peak height imbalance (Fig. 4). This result suggested the presence of the copy number variants at this region. This was also confirmed by clone-based sequencing. The sequencing result of the putative father, mother and the child for the D22GATA198B05 locus showed that the repeat structures were $[TCTC]_1[TATC]_{11}[ACCT]_5/[TCTC]_2[TATC]_{16}[ACCT]_5$, $[TCTC]_1[TATC]_{10}[ACCT]_6/[TCTC]_2[TATC]_{15}[ACCT]_5$ and $[TCTC]_1[TATC]_{11}[ACCT]_5/[TCTC]_1[TATC]_{10}[ACCT]_6/[TCTC]_2[TATC]_{16}[ACCT]_5$, respectively. Two sequences with different repeat structure for allele 17 were also observed and both of them could be detected from the sequencing result of the child. This clearly indicated that the putative father provide not only the allele 23 but also the allele 17 to the child.

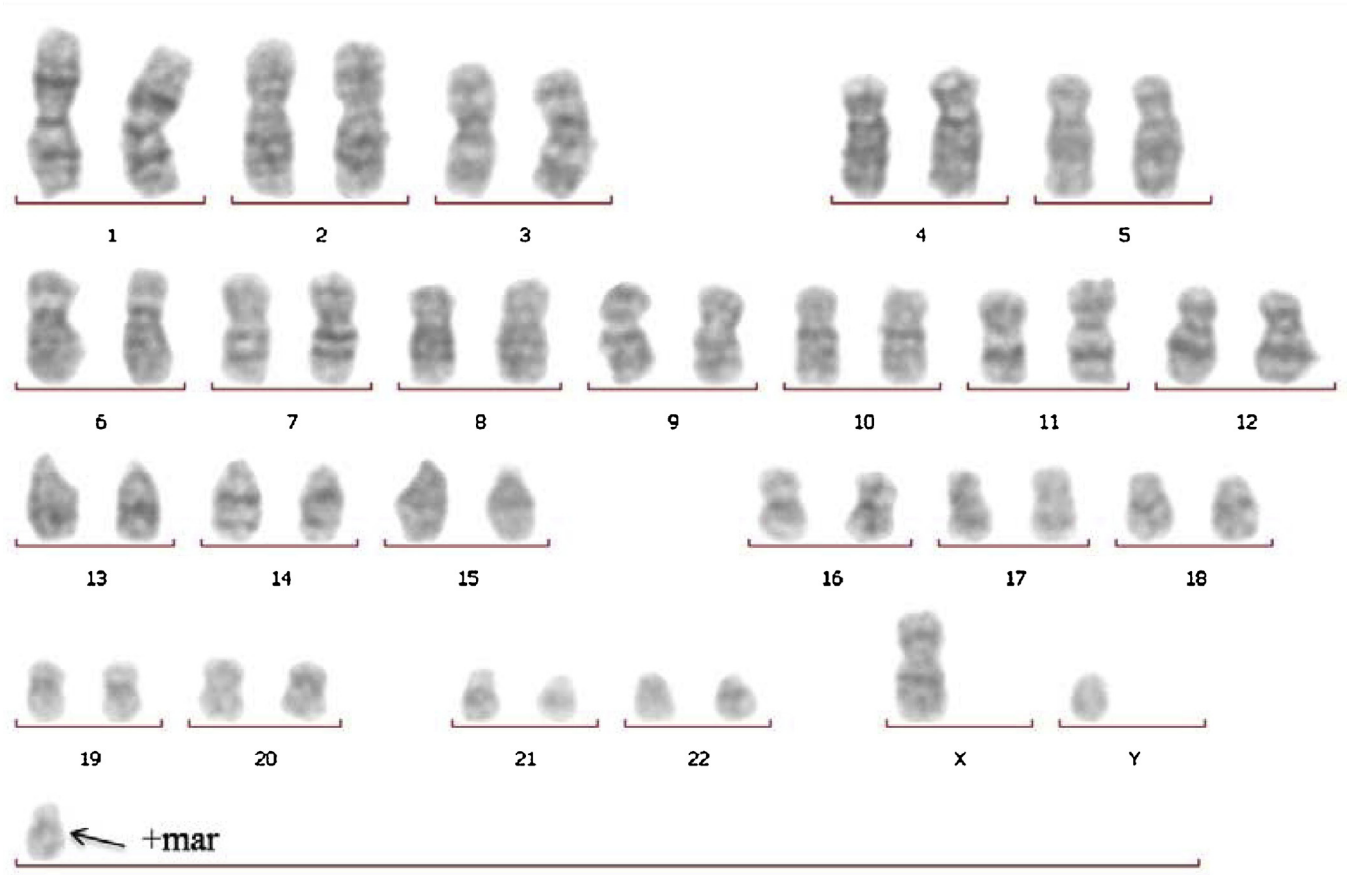

Fig. 2. An additional marker chromosome was similar to G group chromosomes revealed the abnormal karyotype of the child.

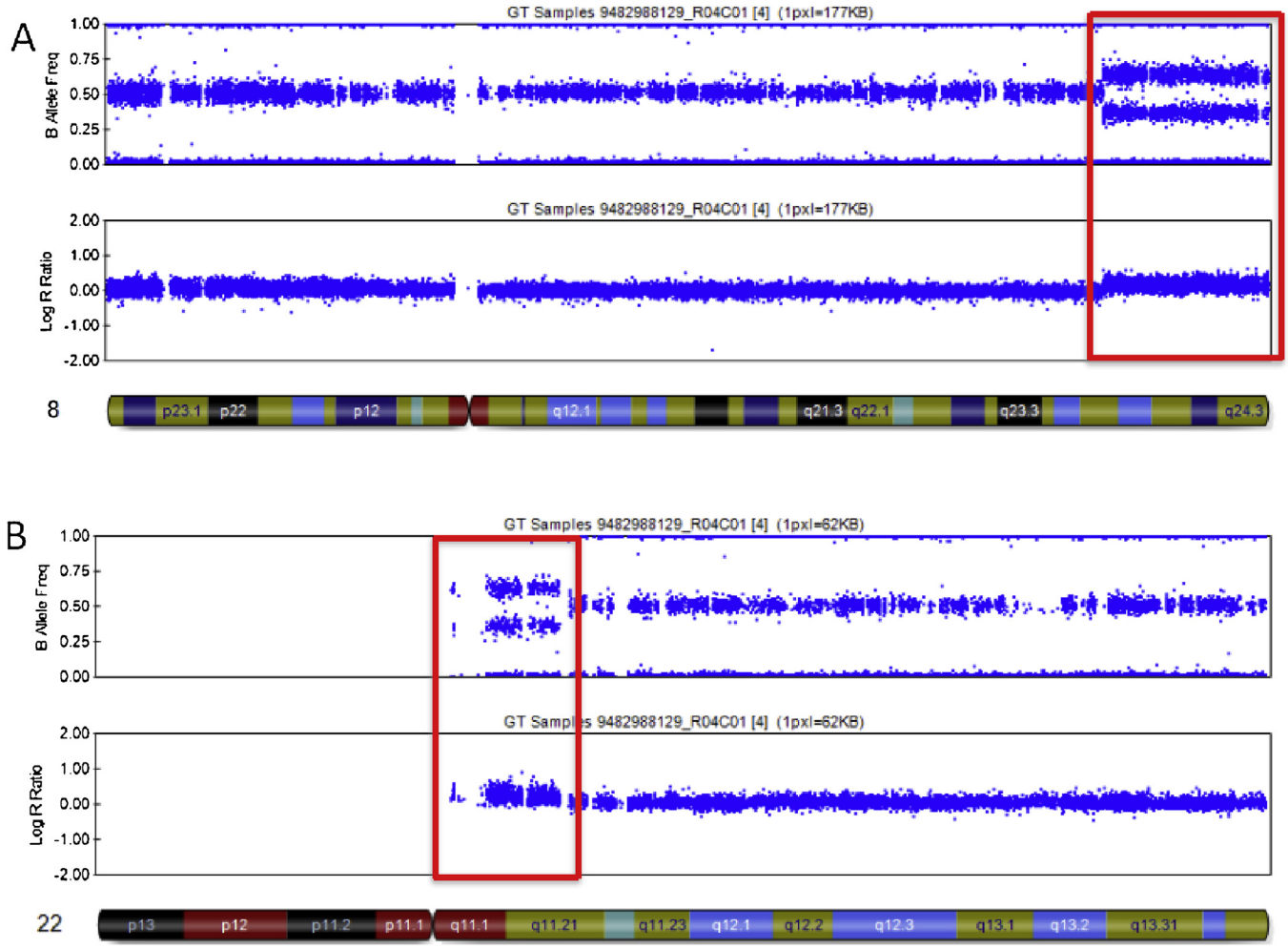

Fig. 3. CNV regions identified by whole genome SNPs genotyping. B allele frequency (upper dot plot of each figure) and log R ratio (lower dot plot) of each genotyped locus on chromosome 8 (A) and 22 (B) were showed. CNV region were indicated by red box with deviated B allele frequency and higher log R ratio.

Table 2
Twenty-two autosomal STR genotyping of the father, mother and the child.

Loci	Father	Child	Mother	Paternity index
D6S477	15	13,15	13,14	3.249
D18S535	9,13	9,14	13,14	2.638
D19S253	13	12,13	12	5.203
D15S659	15,16	12,16	12,16	1.31
D11S2368	19,20	19,20	19,21	2.62
D20S470	15,16	13,15	9,13	2.756
D1S1656	16,17	13,16	13,17.3	2.282
D22-GATA198B05	17,23	17,23	17,22	16.174
D16S539	12	9,12	9	4.482
D7S3048	19	19,25	24,25	10.941
D8S1132	19	19,21	19,21	3.166
D4S2366	11,12	10,12	10,12	2.09
D21S1270	12.3,13	10,12.3	10,14	4.895
D13S325	20,24	19,24	19,22	186
D9S925	15,17	15,16	14,16	2.514
D3S3045	12,15	12,15	12,13	6.764
D14S608	9,12	9,11	9,11	1.329
D10S1435	10,14	12,14	12	3.413
D12S391	17,19	17,19	17,20	2.447
D2S1338	20,22	22,23	23,24	10.054
D17S1290	16	16	15,16	3
D5S2500	11,15	12,15	11,12	1.73
CPI				4,52,98,65,05,28,438.90

4. Discussion

CNVs were structurally variant regions in which copy number differences had been observed between genomes [8]. This variation might range from about one kilobase to several megabases in size. Deletions, insertions, duplications and complex multi-site variants were found in all humans and other mammals [9]. Current annotated CNVs cover about 68.7% of the genome, and, to date, over 15,963 non-overlapping human CNV loci have been identified (http://projects.tcag.ca/variation). CNVs are a major source of human genetic diversity, and have been shown to influence rare genomic disorders [10] as well as complex traits and diseases [11].

In forensic science, paternity testing usually follows simple Mendelian inheritance where a child inherits one allele from the mother and another from the father at every locus. In some rare cases, abnormal patterns of inheritance can be encountered. Since CNVs correspond to relatively large regions of the genome that have been duplicated or deleted on certain chromosomes, the genotyping result of the genetic marker located in these regions might be affected, such as null allele, tri-allele and imbalanced in allele peak height in STR markers routinely used in parentage testing. In general, the presence of two STR locus inconsistencies between a child and his biological father was a rare event and could lead to a false exclusion. When encountered with such case, extended analyses need to be conducted to avoid the false

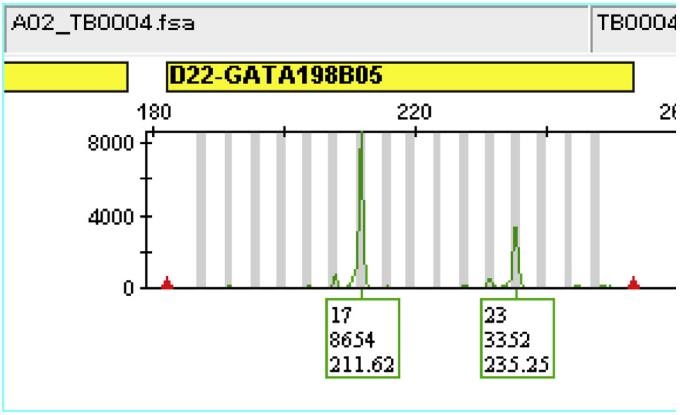

Fig. 4. Peak height imbalance of heterozygote allele at locus D22GATA198B05 in Microreader™ 23sp ID electropherogram profiles.

exclusion. In the present rare case, tri-allele and one-step mutation was concurrently observed at STR locus D8S1179 and FGA. After further validated by more STR loci were investigated, cloning sequencing, karyotype and whole-genome SNP analysis, copy number variants at chromosome 8q24.3 and 22q11.21 was discovered. Since the putative father had contributed both the allele 15 and 16, Paternity Index for tri-allele of the D8S1179 locus was calculated using the formula: $PI = X/Y = AM13 \times AF15 \times AF15 \rightarrow 16/AM13 \times RF15 \times RF16 = 1 \times 1 \times \mu \times 1/2 \times (1/10)^{S-1}/P15 \times P16$ in which μ is the mutation rate and S is the mutation step [6,12]. Although the final CPI was reduced to 292,317 when considering the D8S1179 locus, it still could affirm the paternity.

In conclusion, we suggested that for parentage testing cases with tri-allele, more investigations (e.g., cloning sequencing, karyotyping and whole genome analysis, as well as appropriate statistical estimations) could be conducted to further confirm or exclude the relationship.

Conflict of interest

The authors have no conflicts of interest.

Acknowledgment

This project was supported by the National Natural Science Foundation of China (NSFC, No. 81172909, 81330073).

References

[1] P. Stankiewicz, J.R. Lupski, Structural variation in the human genome and its role in disease, Annu. Rev. Med. 61 (2010) 437–455.
[2] P.S. Walsh, D.A. Metzger, R. Higuchi, Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material, Biotechniques 10 (1991) 506–513.
[3] Y.M. Huang, J. Wang, Z. Jiao, L. Yang, X. Zhang, H. Tang, et al., Assessment of application value of 19 autosomal short tandem repeat loci of GoldenEye 20A kit in forensic paternity testing, Int. J. Legal Med. 127 (2013) 587–590.
[4] M. Seabright, A rapid banding technique for human chromosomes, Lancet (1971) 971–972.
[5] C.H. Mermel, S.E. Schumacher, B. Hill, M.L. Meyerson, R. Beroukhim, G. Getz, GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, Genome Biol. 12 (2011) R41.
[6] H. Lu, D. Lv, Paternity index calculation in mutation autosomal STR locus, Chin. J. Forensic Sci. (2009) 32–36.
[7] L.G. Shaffer, M.L. Slovak, L.J. Campbell, An international system for human cytogenetic nomenclature, Hum. Genet. 126 (2009) 603–604.
[8] L. Feuk, A.R. Carson, S.W. Scherer, Structural variation in the human genome, Nat. Rev. Genet. 7 (2006) 85–97.
[9] R. Redon, S. Ishikawa, K.R. Fitch, L. Feuk, G.H. Perry, T.D. Andrews, et al., Global variation in copy number in the human genome, Nature (2006) 444–454.
[10] J.L. Freeman, G.H. Perry, L. Feuk, R. Redon, S.A. McCarroll, D.M. Altshuler, et al., Copy number variation: new insights in genome diversity, Genome Res. 16 (2006) 949–961.
[11] L. Feuk, C.R. Marshall, R.F. Wintle, S.W. Scherer, Structural variants: changing the landscape of chromosomes and design of disease studies, Hum. Mol. Genet. 15 (2006) R57–R66 Spec no. 1.
[12] B. Brinkmann, M. Klintschar, F. Neuhuber, J. Huhne, B. Rolf, Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat, Am. J. Hum. Genet. 62 (1998) 1408–1415.