GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Polymorphisms of three new microsatellite sites of the dystrophin gene

R.F. Sun[1,2], Y.S. Zhu[1,2,4], J.L. Feng[1,2], Z. Tian[1,2], W.J. Kuang[1,2], Y. Liu[1,3], H.B. Zhang[1,2] and S.B. Li[1,2]

[1]Key Laboratory of the National Ministry of Health for Forensic Sciences, College of Medicine, Xi'an Jiaotong University, P.R. China
[2]Key Laboratory of Environment and Gene Related Diseases, Xi'an Jiaotong University, Ministry of Education, P.R. China
[3]Material Evidence Identification Center, Ministry of Public Security, Building, Nanli Muxudi, Xicheng District, Beijing, P.R. China
[4]Key Laboratory of Reproduction and Genetics of Ningxia Hui Autonomous Region, Yinchuan, P.R. China

Corresponding author: S.B. Li and H.B. Zhang
E-mail: shbinlee@mail.xjtu.edu.cn / zhanghb@mail.xjtu.edu.cn

Genet. Mol. Res. 10 (2): 744-751 (2011)
Received June 26, 2010
Accepted February 23, 2011
Published April 26, 2011
DOI 10.4238/vol10-2gmr962

ABSTRACT. To look for novel microsatellites in the dystrophin gene for the diagnosis of Duchenne muscular dystrophy, candidate microsatellite sites in the dystrophin gene were analyzed with the SSRHunter software and were also genotyped. Among the 15 candidate microsatellite sites, three novel microsatellite sites in the 60th, 30th, and 2nd intron were found to have a high degree of polymorphism. We submitted these three new loci to the European Molecular Biology Laboratory, under accession Nos. FN547040, FN547041 and FN557526, which were called DXSDMD-in60, DXSDMD-in30 and DXSDMD-in2, respectively. In these three loci, we found 9, 6 and 11 alleles, respectively, in the 205 individuals. In addition, we also detected 20, 19 and 20 genotypes for the three loci in female samples, with a polymorphism information content of more than 0.600. In conclusion, the three microsatellite sites in the intron region of the

dystrophin gene have a high degree of polymorphism, and they can be used in population genetics, as well as to provide a theoretical basis for genetic diagnosis and elucidation of molecular mechanisms in Duchenne muscular dystrophy.

Key words: Duchenne muscular dystrophy; Dystrophin gene; Microsatellite; Polymorphism; Han population genetics

INTRODUCTION

Duchenne muscular dystrophy (DMD) is a fatal, X-linked recessive genetic disease, which is characterized clinically by progressive denaturation and necrosis of skeletal muscles. The pathogenesis of DMD is a mutation in the dystrophin gene, which leads to structural and functional changes of the dystrophin protein on the cell membrane (Koenig et al., 1987; Sifringer et al., 2004). Duchenne muscular dystrophy and Becker muscular dystrophy (BMD) are allelic disorders that cause progressive muscle atrophy, where BMD is a milder variant of DMD. BMD is less severe but displays variable phenotypes, ranging from a slightly less severe DMD-like condition to a very mild condition, and the majority of patients can be diagnosed according to their phenotype (Sewry, 2010). However, a recent clinical casebook described a boy without any signs of DMD but with confirmation of molecular findings (Dubowitz, 2006).

The clinical phenotypes caused by dystrophin mutation are complex and variable. For example, deletion of exon 44, which is a small-sized exon, can lead to a typical DMD phenotype, while some patients who have the 50% deletion in the dystrophin gene show the BMD phenotype. There are intact reading frames with deletion in exons 32-44, 48-51 or 48-53 in some patients; however, the concentration of dystrophin protein is normal or essentially normal (Melis et al., 1998). Many thousands of mutations have been recorded in the dystrophin gene, but 65% of DMD patients harbor dystrophin gene deletions in a mutation-rich area or ‘‘hot-spot’’ in the central genomic region (exons 41-54). About 6-10% of DMD cases are caused by duplication, and most of them can lead to a new stop codon, resulting in premature truncation. The hypothesis of “reading frame” considering the severity of the phenotype depends on whether the mutations affect the reading frame, rather than depending on the size of deletion or duplication fragment (Walmsley et al., 2010). Recently, animal studies provided similar findings (Banks and Chamberlain, 2008; Ambrosio et al., 2009; Walmsley et al., 2010).

Although there is no specific therapy in humans for this disease, prenatal diagnosis is a useful measure to prevent the birth of a baby with this disease. Linkage analysis can improve diagnostic yield, where several highly polymorphic microsatellite markers have been reported, but not enough. We therefore aimed to analyze these microsatellite sites in order to identify new polymorphic DNA markers for linkage analysis. We performed whole gene scanning and DNA sequencing to screen and characterize the polymorphic microsatellite sites. Three new microsatellite sites of dystrophin showing high polymorphism were discovered for the first time. Based on this, the genetic distributions of these loci in the Han population in Xi’an were characterized to provide the basis for population genetics and the early diagnosis and prevention of DMD disease.

MATERIAL AND METHODS

Sampling

Based on the principle of informed consent, 205 individuals (100 males and 105 females, with no genetic connections between each other) (Gao and Li, 2008) were recruited into this study. Peripheral blood, 3~5 mL, was collected in an EDTA anticoagulant vacuum tube. The quantity of DNA was determined using the Blood Genomic DNA Extraction System kit (Tiangen Biotech, Foster City, Beijing) in accordance with the manufacturer protocol. All samples were stored at -20°C until needed.

Selection of microsatellite genetic markers

We extracted the dystrophin sequence from NCBI database and then performed whole gene scanning to find out the microsatellite sites of dystrophin using the SSRHunter software (http://www.biosoft.net) (Li and Wan, 2005). Microsatellite sequences were selected according to the following criteria (Yan and Hou, 2004): 1) the core sequence is 3~4 bp long; 2) the core sequence is strictly regular repeats; 3) there is only one repeat unit contained in each fragment (i.e., only one motif); 4) the repeat number of the core sequence is ≥5, and 5) the locus has not been reported. Based on the search results of the ReapeatMask software, we confirmed that the microsatellite sites screened according to the above criteria did not belong to the Alu family or other repetitive sequences. The 15 candidate microsatellite sequences of the dystrophin gene selected using the SSRHunter software were designated M1, M2, M15, according to their chromosome position (see Figure 1).

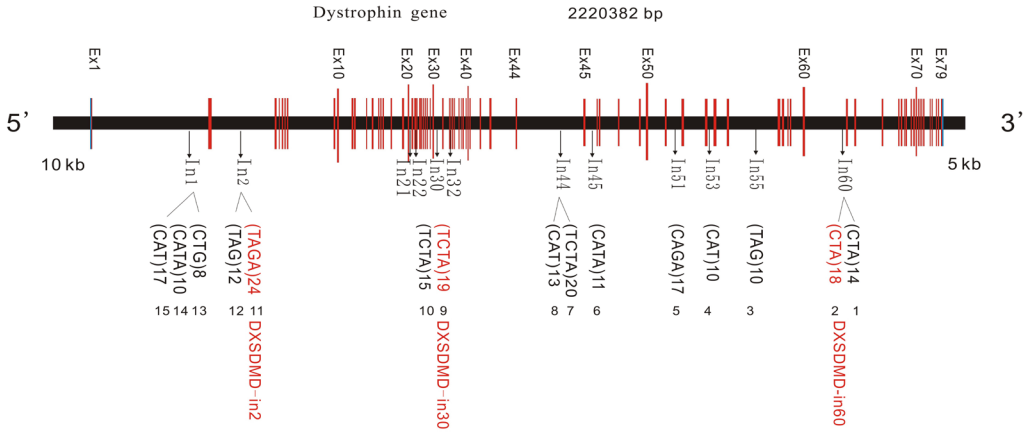

Figure 1. Distribution of 15 microsatellite sites in dystrophin gene. Ex = exon; In = intron.

Allelic detection and genotyping

Primers specific for the 15 candidate microsatellite sites were designed with Primer 5.0. The reaction system for polymerase chain reaction (PCR) amplification was 12 μL and consisted of the following reagents: 6 μL 2X PCR Master Mixture (Tianwei-shidai Company, China), 1~1.5 μL of each primer (5 μM), 1~2 μL sterile deionized water and 2~3 μL template DNA (30 ng/μL). The primers and reaction conditions for PCR are listed in Table 1. Allele typing and DNA sequencing were carried out using 6% polyacrylamide gel electrophoresis and silver staining, and the alleles were named according to the number of repeats (Kimmel and Chakraborty, 1996).

Table 1. Sequence and polymerase chain reaction (PCR) information for three polymorphic microsatellite sites.

Name	Primer sequence	Annealing	PCR product (bp)	Chr start	Chr stop
DXSDMD-in60	5' CGAGGGGATCAGGGTAATA 3' 5' CTGTTCTCTTCTCTGGTCATCA 3'	58.4°C, 40 s	138-162	31402427	31402480
DXSDMD-in30	5' GTTAGTCCCTATTCTATTCCTTTC 3' 5' AAGAATGCCACCAAAATGAC 3'	56.6°C, 30 s	152-168	32426983	32427061
DXSDMD-in2	5' TAGATTAGATAGGTAGGTAGGTAGG 3' 5' CTTGCCAGTTTTTTTCAGTG 3'	62.0°C, 30 s	110-150	32967440	32967538

Preparation of allelic genotyping standard material

We recovered and purified the allele fragments from the electrophoretic gels, and each fragment and pGEM-T Easy vector (Applied Promega, Foster City, Wisconsin) were connected and then transferred to *Escherichia coli* DH5α competent cells (Tiangen Biotech). At the end, the selection of positive clones by the blue-white screening method was confirmed by sequencing. This was the standard material for genotyping.

Statistical analysis

The modified Powerstats software was employed for the Hardy-Weinberg equilibrium (HWE) test, calculating allele frequency, genotype frequency, heterozygosity, polymorphism information content, and haplotype diversity [$HD = (n/n - 1)(1 - \sum Pi^2)$], where *Pi* is the haplotype frequency and *n* is the number of samples. Linkage disequilibrium (LD) between the three loci was determined by the Arlequin software.

RESULTS

Three of 15 loci were found to be polymorphic, namely DXSDMD-in60, DXSDMD-in30 and DXSDMD-in2. No significant differences were observed between allele distributions in male and female groups, and therefore, the two groups were pooled into a single group to analyze the frequency distributions for the respective loci. All loci analyzed were in HWE. DXSDMD-in60 was located in the 60th intron of dystrophin, whose repetitive unit was CTA. The repeat number was found to be 13~21 in this study. DXSDMD-in30 was located in the 30th

intron of dystrophin, whose repetitive unit was TCTA. The repeat number was found to be 9~13 in this study. DXSDMD-in2 was located in the 2nd intron of dystrophin, whose repetitive unit was TAGA. The repeat number was found to be 6~16 in this study. The other 12 loci did not display polymorphism in the Han population in Xi'an, China. Our study showed that in the Han population in Xi'an, there were 9, 6 and 11 alleles for the three microsatellites, while 20, 19 and 20 genotypes were detected in female samples, all of which with polymorphic information content more than 0.600. The frequency distribution is shown in Tables 2 and 3. Allele frequencies of DXSDMD-in60, DXSDMD-in30 and DXSDMD-in2 loci were 0.0032~0.5226, 0.0097~0.4258 and 0.0032~0.3903, respectively. In female samples, genotype 14,14, genotype 11,12 and genotype 12,13 were the most frequent in the DXSDMD-in60, DXSDMD-in30 and DXSDMD-in2 loci, respectively. The analysis for LD was performed for all pairs of the three microsatellite sites in female samples, no evidence of LD was detected. In the 100 male samples, we detected 56 haplotypes consisting of these three microsatellite sites, with a haplotype diversity of 0.9784 (see Table 4).

Table 2. Population genetics index of the three microsatellite sites (males: N = 100; females: N = 105).

	DXSDMD-in60	DXSDMD-in30	DXSDMD-in2
6			0.0032
7			0.0000
8			0.0258
9		0.0806	0.0032
10		0.1774	0.0065
11		0.4258	0.1452
12		0.2258	0.3903
13	0.0065	0.0806	0.2419
14	0.5226	0.0097	0.1548
15	0.0484		0.0258
16	0.1613		0.0032
17	0.1065		
18	0.1194		
19	0.0258		
20	0.0032		
21	0.0065		
H	0.61	0.724	0.676
PIC	0.65	0.69	0.7

H = heterozygosity; PIC = polymorphism information content.

DISCUSSION

The dystrophin gene is located in Xp21.2, spanning 2.4 Mb in genomic DNA, which is about 1.5% of the whole length of the X-chromosome. It contains 79 exons and 78 introns (Davies, 1997; Suminaga et al., 2002). There are two kinds of mutation of dystrophin: about 60% (Roberts et al., 1994; Mendell et al., 2001) of the mutations are deletions, which usually occur in the region of the 5' end and exon 1-11 and 41-54 regions (Den Dunnen et al., 1989; Basak et al., 2006; Lai et al., 2006); the other 40% are non-deletions, which include point mutations, small deletions and insertions of the dystrophin gene. Microsatellite-PCR technique has been the preferred technique for linkage analysis of non-deletion DMD families. Several

Table 3. Frequencies of the genotype at three microsatellite sites (females, N = 105).

No.	DXSDMD-in60			DXSDMD-in30			DXSDMD-in2		
	Genotype	N	Frequency	Genotype	N	Frequency	Genotype	N	Frequency
1	13,16	2	0.02	9,9	1	0.01	6,13	1	0.01
2	14,14	31	0.30	9,10	3	0.029	8,8	1	0.01
3	14,15	5	0.05	9,11	7	0.067	8,13	1	0.01
4	14,16	17	0.16	9,12	4	0.038	8,14	1	0.01
5	14,17	8	0.08	9,13	1	0.01	9,12	1	0.01
6	14,18	12	0.11	9,14	1	0.01	10,12	1	0.01
7	14,19	4	0.04	10,10	3	0.029	11,11	5	0.05
8	14,20	1	0.01	10,11	15	0.143	11,12	9	0.09
9	15,15	1	0.01	10,12	8	0.076	11,13	8	0.08
10	15,16	3	0.03	10,13	3	0.029	11,14	5	0.05
11	15,18	1	0.01	10,14	1	0.01	11,15	1	0.01
12	15,19	1	0.01	11,11	19	0.181	12,12	18	0.17
13	16,16	3	0.03	11,12	20	0.19	12,13	21	0.20
14	16,17	2	0.02	11,13	7	0.067	12,14	10	0.10
15	16,18	5	0.05	11,14	1	0.01	12,15	1	0.01
16	17,17	3	0.03	12,12	5	0.048	13,13	7	0.07
17	17,18	1	0.01	12,13	4	0.038	13,14	8	0.08
18	18,18	2	0.02	12,14	1	0.01	13,15	1	0.01
19	18,19	2	0.02	13,13	1	0.01	14,14	3	0.03
20	21,21	1	0.01				14,15	2	0.02
Total	20	105		19	105		20	105	

microsatellite sites in the dystrophin gene have been reported to be ideal loci for family linkage analysis, including microsatellite sites located in the 5' promoter region, introns 44, 45, 49 and 50, and 3' non-translational region (Giliberto et al., 2003). However, linkage analysis is a kind of probabilistic diagnosis (Lai et al., 2002). It is very important to increase the polymorphic loci for linkage analysis to avoid the wrong analysis results caused by gene recombination and to increase the reliability of the analysis results. Our study uncovered three novel microsatellite sites in the intron region of dystrophin, and the results indicated that heterozygosity of the three microsatellite sites DXSDMD-in60, DXSDMD-in30 and DXSDMD-in2 was 0.610, 0.724 and 0.676, respectively. Polymorphism information content was 0.650, 0.690 and 0.700, respectively, which will provide more genetic markers for linkage analysis in the future.

In our study, the three loci screened were located in different introns of dystrophin, and the distance between DXSDMD-in60 and DXSDMD-in2 is 1.67 Mbp. After linkage study, no evidence of LD was detected. There are in total 56 haplotypes consisting of DXSDMD-in60-DXSDMD-in30-DXSDMD-in2 among 100 male individuals, with frequencies varying from 0.010 to 0.080. The most common haplotype was 14-12-12, and the haplotype diversity was 0.9784.

In conclusion, these results indicate that these three microsatellite sites show high polymorphism, which can be used in linkage analysis, gene location and diagnosis of dystrophin-related diseases, as well as population genetics studies (Ribeiro Rodrigues et al., 2008). Further study of the polymorphism of these three microsatellite sites of the dystrophin gene still needs a larger sample size, as well as other regional and national population data, and this polymorphism also needs further investigation and verification in different ethnic, regional populations, as well as in disease studies.

Table 4. Haplotype distribution of the three microsatellite sites (males, N = 100).

No.	DXSDMD-in60	DXSDMD-in30	DXSDMD-in2	N	Fre	No.	DXSDMD-in60	DXSDMD-in30	DXSDMD-in2	N	Fre
1	14	9	13	2	0.02	29	16	12	8	1	0.01
2	14	9	14	1	0.01	30	16	12	11	1	0.01
3	14	10	11	1	0.01	31	16	12	13	1	0.01
4	14	10	12	4	0.04	32	16	13	12	2	0.02
5	14	10	13	2	0.02	33	17	9	12	1	0.01
6	14	10	14	1	0.01	34	17	9	13	1	0.01
7	14	11	8	1	0.01	35	17	9	14	1	0.01
8	14	11	10	1	0.01	36	17	10	12	1	0.01
9	14	11	11	4	0.04	37	17	10	14	1	0.01
10	14	11	12	5	0.05	38	17	10	15	1	0.01
11	14	11	13	6	0.06	39	17	11	12	5	0.05
12	14	11	14	5	0.05	40	17	11	13	1	0.01
13	14	11	15	1	0.01	41	17	11	14	1	0.01
14	14	12	8	1	0.01	42	17	12	11	1	0.01
15	14	12	11	1	0.01	43	17	12	12	1	0.01
16	14	12	12	8	0.08	44	17	12	13	1	0.01
17	14	12	13	2	0.02	45	18	9	12	1	0.01
18	14	12	14	3	0.03	46	18	10	13	1	0.01
19	14	13	8	1	0.01	47	18	10	14	1	0.01
20	14	13	12	3	0.03	48	18	11	11	1	0.01
21	15	10	11	1	0.01	49	18	11	12	2	0.02
22	15	10	12	1	0.01	50	18	11	14	1	0.01
23	15	11	12	1	0.01	51	18	12	11	1	0.01
24	16	10	13	2	0.02	52	18	12	12	1	0.01
25	16	10	15	1	0.01	53	18	13	11	1	0.01
26	16	10	16	1	0.01	54	18	13	12	1	0.01
27	16	11	12	5	0.05	55	18	13	14	1	0.01
28	16	11	13	1	0.01	56	19	11	13	1	0.01
HD = 0.9784											

HD = haplotype diversity.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation. We thank Zhaoqin Chai, MD, for their help.

REFERENCES

Ambrosio CE, Fadel L, Gaiad TP, Martins DS, et al. (2009). Identification of three distinguishable phenotypes in Golden Retriever muscular dystrophy. *Genet. Mol. Res.* 8: 389-396.

Banks GB and Chamberlain JS (2008). The value of mammalian models for Duchenne muscular dystrophy in developing therapeutic strategies. *Curr. Top. Dev. Biol.* 84: 431-453.

Basak J, Dasgupta UB, Banerjee TK, Senapati AK, et al. (2006). Analysis of dystrophin gene deletions by multiplex PCR in eastern India. *Neurol. India* 54: 310-311.

Davies KE (1997). Challenges in Duchenne muscular dystrophy. *Neuromuscul. Disord.* 7: 482-486.

Den Dunnen JT, Grootscholten PM, Bakker E, Blonden LA, et al. (1989). Topography of the Duchenne muscular dystrophy (DMD) gene: FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Am. J. Hum. Genet.* 45: 835-847.

Dubowitz V (2006). Enigmatic conflict of clinical and molecular diagnosis in Duchenne/Becker muscular dystrophy. *Neuromuscul. Disord.* 16: 865-866.

Gao Y and Li SB (2008). Effects of sample size on the observed number of allele of 9 STR loci with various genetic data. *Yi Chuan* 30: 313-320.

Giliberto F, Ferreiro V, Dalamon V, Surace E, et al. (2003). Direct deletion analysis in two Duchenne muscular dystrophy symptomatic females using polymorphic dinucleotide (CA)n loci within the dystrophin gene. *J. Biochem. Mol. Biol.* 36: 179-184.

Kimmel M and Chakraborty R (1996). Measures of variation at DNA repeat loci under a general stepwise mutation model. *Theor. Popul. Biol.* 50: 345-367.

Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, et al. (1987). Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* 50: 509-517.

Lai KK, Lo IF, Tong TM, Cheng LY, et al. (2006). Detecting exon deletions and duplications of the DMD gene using multiplex ligation-dependent probe amplification (MLPA). *Clin. Biochem.* 39: 367-372.

Lai PS, Takeshima Y, Adachi K, Van Tran K, et al. (2002). Comparative study on deletions of the dystrophin gene in three Asian populations. *J. Hum. Genet.* 47: 552-555.

Li Q and Wan JM (2005). SSRHunter: development of a local searching software for SSR sites. *Yi Chuan* 27: 808-810.

Melis MA, Cau M, Muntoni F, Mateddu A, et al. (1998). Elevation of serum creatine kinase as the only manifestation of an intragenic deletion of the dystrophin gene in three unrelated families. *Eur. J. Paediatr. Neurol.* 2: 255-261.

Mendell JR, Buzin CH, Feng J, Yan J, et al. (2001). Diagnosis of Duchenne dystrophy by enhanced detection of small mutations. *Neurology* 57: 645-650.

Ribeiro Rodrigues EM, Leite FP, Hutz MH, Palha TJ, et al. (2008). A multiplex PCR for 11 X chromosome STR markers and population data from a Brazilian Amazon Region. *Forensic Sci. Int. Genet.* 2: 154-158.

Roberts RG, Gardner RJ and Bobrow M (1994). Searching for the 1 in 2,400,000: a review of dystrophin gene point mutations. *Hum. Mutat.* 4: 1-11.

Sewry CA (2010). Muscular dystrophies: an update on pathology and diagnosis. *Acta Neuropathol.* 120: 343-358.

Sifringer M, Uhlenberg B, Lammel S, Hanke R, et al. (2004). Identification of transcripts from a subtraction library which might be responsible for the mild phenotype in an intrafamilially variable course of Duchenne muscular dystrophy. *Hum. Genet.* 114: 149-156.

Suminaga R, Takeshima Y, Adachi K, Yagi M, et al. (2002). A novel cryptic exon in intron 3 of the dystrophin gene was incorporated into dystrophin mRNA with a single nucleotide deletion in exon 5. *J. Hum. Genet.* 47: 196-201.

Walmsley GL, Arechavala-Gomeza V, Fernandez-Fuente M, Burke MM, et al. (2010). A Duchenne muscular dystrophy gene hot spot mutation in dystrophin-deficient Cavalier King Charles Spaniels is amenable to exon 51 skipping. *PLoS One* 5: e8647.

Yan J and Hou YP (2004). Exploring Novel STR Loci on Human Chromosome 21 for Forensic and Medical Genetics. Doctoral thesis, Sichuan University, Chengdu.