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

Genetic analysis of STR markers on chromosome 21 in a Han population from southeast China

Y.N. Zhu, S.M. Lu, M. Wang, F.X. Shen, Y. Chen and J.J. Hu

Key Laboratory of Reproductive Genetics (Zhejiang University), Ministry of Education, Women's Hospital, Zhejiang University, Hangzhou, Zhejiang, China

Corresponding author: S.M. Lu
E-mail: lusm@zju.edu.cn

Genet. Mol. Res. 14 (1): 1718-1725 (2015)
Received February 10, 2014
Accepted August 4, 2014
Published March 6, 2015
DOI http://dx.doi.org/10.4238/2015.March.6.18

ABSTRACT. Short tandem repeats (STRs) are highly polymorphic sequences and have been extensively used as genetic markers in mapping studies, disease diagnosis, and human identity testing. In this study, 11 STR markers on chromosome 21, including D21S1432, D21S11, D21S1246, D21S1412, D21S1437, D21S1442, D21S2039, D21S1270, D21S1435, D21S1409, and D21S1446, were analyzed in 740 unrelated Han individuals from southeast China. A total of 132 alleles, ranging from 7-21 for each locus, were named according to the guidelines of the International Society for Forensic Haemogenetics. The distributions of allelic frequencies for the 11 STRs and population genetic parameters were determined. All 11 STR markers showed high polymorphism and heterogeneity in the southeast Han population, with polymorphism information content of 0.61-0.87, heterogeneity of 64.5-86.1%, and power of discrimination of 0.835-0.973. Among the 11 STR markers, D21S1412, D21S1270, D21S11, and D21S1442 showed relatively higher heterogeneity. Their combination was relatively informative and was used in a quantitative fluorescence-polymerase chain reaction

assay to diagnose Down syndrome (trisomy 21) in a southeast Chinese Han population. The genetic information and population data for these 11 STRs may be used not only in quantitative fluorescence-polymerase chain reaction assays but also in forensic studies and other genetic tests.

Key words: Southeast China; Down syndrome; Han population; Quantitative fluorescence polymerase chain reaction; Chromosome 21; Short tandem repeats

INTRODUCTION

Short tandem repeats (STRs), known as microsatellites, are highly polymorphic sequences of nucleotides that are abundant in the eukaryotic genome. Because of their high level of polymorphism and low mutation rate, STRs are widely used as genetic markers in mapping studies, disease diagnosis, and human identity testing. Each STR acts as a marker for a particular chromosome, and thus has recently been utilized in quantitative fluorescence-polymerase chain reaction (QF-PCR) assays for the prenatal detection of common aneuploidies, such as Down syndrome, trisomy 18, and trisomy 13 (Cirigliano et al., 2009; Hills et al., 2010). In a QF-PCR assay, at least 2 markers are required to be concurrently heterozygous in order to confirm the copy number of a particular chromosome (Association for Clinical Cytogenetics and Clinical Molecular Genetics Society, 2012). A marker must be heterozygous for the ratio of its allele peak areas to show a disomic (1:1) or trisomic (2:1, 1:2 or 1:1:1) chromosome complement. A marker is uninformative if only a single peak is observed. Using highly polymorphic STR markers, QF-PCR has been shown to be a fast and efficient method for detecting common aneuploidies in Western countries (Cirigliano et al., 2001; Mann and Ogilvie, 2012). However, these QF-PCR assays are typically designed for the Caucasian population. STR markers vary among population groups. Therefore, STR markers must be evaluated for their polymorphisms and heterozygosity before using in QF-PCR assays for a particular population (Lee et al., 2010; Jain et al., 2012).

In southeast China, little information has been reported regarding STR polymorphism markers on chromosome 21, other than D21S11 (Zhu et al., 2009). To identify informative STR markers for Down syndrome and detection assays for the Han population from Southeast China, we developed a multiplex typing system and analyzed 11 STR markers on chromosome 21, including D21S1432, D21S11, D21S1246, D21S1412, D21S1437, D21S1442, D21S2039, D21S1270, D21S1435, D21S1409, and D21S1446. All candidate markers were selected from either commercial kits (Aneufast, Wollerau, Switzerland; TrueScience, Santa, UT, USA; Elucigene, Manchester, UK) or population reports (Mann et al., 2001; Lee et al., 2010). In this study, we describe the genetic information and population data for these 11 STRs, which may be used not only in QF-PCR assays but also for forensic studies and other genetic tests.

MATERIAL AND METHODS

Sample preparation and DNA extraction

We analyzed 740 blood samples (0.2 mL) from unrelated individuals belonging to the Han population in southeast China. Informed consent was obtained from all subjects prior to their participation in this study. Genomic DNA was extracted using the Chelex-100 protocol

as described by Walsh et al. (1991).

PCR amplification

Amplification was carried out in a 25-μL PCR volume containing 0.2-2 ng DNA, 200 μM of each dNTP, 1X GeneAmp PCR Buffer II, 1.5 mM $MgCl_2$, and 1.0 U AmpliTaq Gold™ DNA polymerase (Applied Biosystems, Foster City, CA, USA). Primers for 4 of the 11 loci (D21S1432, D21S11, D21S1437, and D21S2039) were obtained from NCBI (http://www.ncbi.nlm.nih.gov/sites/entrez). However, we modified all previously reported primer sets in order to amplify the fragments in a single PCR. For the other 7 loci (D21S1246, D21S1412, D21S1442, D21S1270, D21S1435, D21S1409, and D21S1409), new primers were designed using the PRIMER3 software (http://frodo.wi.mit.edu/primer3). Primer hybridization specificity was checked against the human genome using BLASTn (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Secondary structures and primer dimerization were screened using the Autodimer software (Vallone and Butler, 2004). STR primer sequences and other details are presented in Table 1.

Table 1. Details about the primers for the 11STR loci on chromosome 21.

Marker	UniSTSid[a]	Chromosomal location[b]	Physical location[a] (Mb)	Primer sequences[a] (5'→3')
D21S1432	39747	q21.1	17343446	CTTAGAGGGACAGAACTAATAGGC
				GTGTCTTAGCCTATTGTGGGTTTGTGA
D21S11	57855	q21.1	20554263	ATCAATTCCCCAAGTGAATTGC
				GTTGTATTAGTCAATGTTCTCCAG
D21S1246	45212	q22.2	40871104	GATAAAGTAGACAGGTAAACA
				GGATTATAATTCAAGATGAGAT
D21S1412	34813	q22.2	40750288	AAGGCTATGGAGGAGAGCCAGACT
				GAGTTGAGATCGCACCATTG
D21S1437	20052	q21.1	21646839	ATGTACATGTGTCTGGGAAGG
				TTCTCTACATATTTACTGCCAACA
D21S1442	37142	q21.3	28818479	GCCTTTATACTTGGCTGTGATAG
				TTAAGACTTCTCGATCTCCAGAATCAC
D21S2039	473678	q22.11	34717022	TTACGTTCTTCATTTGATCTTAGCC
				GTGTCTTCCAGGCATGATGGCACAC
D21S1270	31268	q22.11	31706761	CCCACTGTATTATTCAGGGC
				ACACACACACACACACATGC
D21S1435	32036	q21.3	27848874	TTAGGA AGAGCCCCATTC
				CAAGAAAAATGAGATTTGAC
D21S1409	52053	q21.2	24348727	TTCTGAATTTTCTAAGAATGCACC
				GGAACATACGCTCTCTCCCTTATTC
D21S1446	20055	q22.3	48037585	ATGTACGATACGTAATACTTGACAA
				GTCCCAAAGGACCTGCTC

[a]Data obtained from NCBI (http://www.ncbi.nlm.nih.gov/sites/entrez). [b]Data obtained from Human Genome Browser (http://www.genome.ucsc.edu/cgi-bin/hgGateway).

Samples were amplified using a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems) under the following conditions: initial denaturation at 95°C for 11 min, followed by 30 cycles of 94°C for 60 s, 58°C for 60 s, and 72°C for 60 s, and an additional cycle at 60°C for 60 min.

Sample electrophoresis and data analysis

PCR products were detected by capillary electrophoresis using the ABI PRISM 310

Genetic Analyzer with denaturing polymer 310 POP-4™ (Applied Biosystems). Fragment sizing was conducted Genescan™-500 LIZ™ size standards. The genotypes of the STRs were determined using the GeneMapper ID ver. 3.2 software (Applied Biosystems). Allele typing was based on homemade allelic ladders, and DNA from the 9947A cell line (Promega, Madison, WI, USA) was typed to calibrate the allelic ladder.

Sequence analysis

PCR products were purified and sequenced on an ABI 3130 Genetic Analyzer using a BigDye® Terminator Cycle Sequencing Kit (Applied Biosystems) according to manufacturer instructions. All different alleles identified in the 11 loci were named based on variable tandem repeat motifs according to the guidelines of the International Society for Forensic Haemogenetics (Bär et al., 1997).

Statistical analysis

Allele frequencies, expected heterozygosity, polymorphism information content, and probability of paternity exclusion were calculated using Powerstats (http://www.promega.com/geneticidtools). The ARLEQUIN ver. 3.1 software (Excoffier et al., 2005) was used to perform the following statistical analyses: the exact test for the Hardy-Weinberg equilibrium and the linkage disequilibrium test between all pairs of markers. Test size α was equal to 0.05; P values were corrected by Bonferroni's adjustment. $P < 0.0009$ (0.05/55, 55 = number of comparisons performed) was considered to be statistically significant in the exact test for linkage disequilibrium.

RESULTS

The 11 STR markers on chromosome 21 were amplified. For the 9947A cell line, the STR markers D21S1432, D21S11, and D21S1446 showed a homozygous/non-informative pattern (single peak), while D21S1246, D21S1412, D21S1437, D21S1442, D21S2039, D21S1270, D21S1435, and D21S1409 were heterozygous/informative (2 peaks with 1:1 ratio; Figure 1).

The sequencing data for different alleles of the 11 STR loci are listed in Table 2. Of the 11 STR loci, 8 contained variations in a single repeat region (D21S1432, D21S1437, D21S1442, D21S2039, D21S1270, D21S1435, D21S1409, and D21S1446), 2 contained simple repeats with motif sequence variants (D21S1246 and D21S1412), and 1 showed a complex structure (D21S11).

Hardy-Weinberg equilibrium tests demonstrated no significant deviation from the expected values for all 11 STR loci in the Han population in southeast China. Linkage disequilibrium for all pairs of loci was tested in the populations studied. One significant result from 55 pairwise comparisons was found between D21S1246 and D21S1412 ($P = 0.0000$); no clear evidence for linkage disequilibrium among the other 54 pairs was found using the significant threshold P value of 0.0009 (by Bonferroni adjustment). A total of 132 alleles, ranging from 7-21 for each locus, were examined in southeast Han individuals. Allele frequencies and statistical analysis are presented in Table 3. All 11 STR markers showed a high degree of polymorphism and heterogeneity in the southeast Han population: polymorphism information content, 0.61-0.87; heterogeneity, 64.5-86.1%; and power of discrimination, 0.835-0.973. Among the 11 STR markers, D21S1412, D21S11, D21S1442, and D21S1270 showed rela-

tively higher heterogeneity, with informativeness of 86.1, 81.4, 80.8, and 78.2%, respectively. Based on these data, if these 4 STRs were combined in a QF-PCR assay, the probability that at least 2 STR markers will concurrently be heterozygous is 100%.

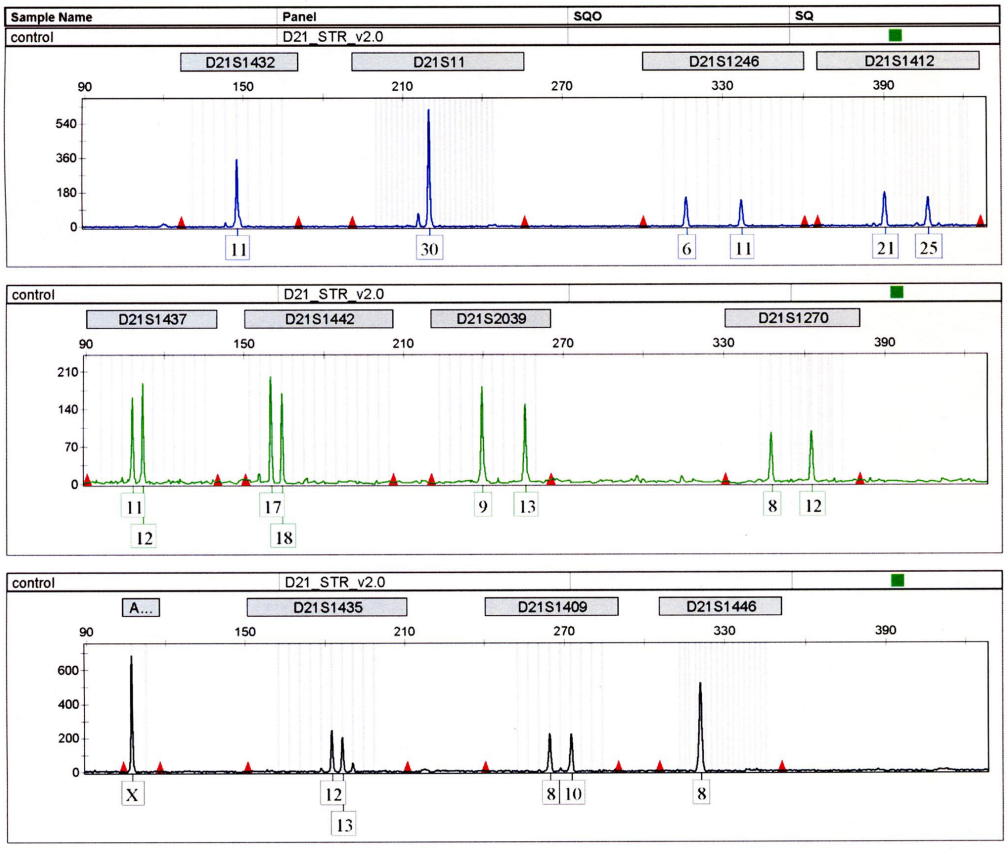

Figure 1. Genotype profile of the 9947A cell line.

Table 2. Sequencing data for different alleles of the 11 STR loci.

Locus	Alleles	Repeat structure (5'→3')
D21S1432	7	$(GATA)_n$
D21S11	17	$(TCTA)_{4-11}(TCTG)_{3-6}(TCTA)_3TA(TCTA)_{2-3}TCA(TCTA)_2TCCATA(TCTA)_{8-15}$
D21S1246	8	$(ATAG)_3(ACAG)(ATAG)_{1-12}(ACAG)_{0-1}$
D21S1412	21	$(CTTT)_{10-19}(CTGT)(CTCT)_{1-3}(CTTT)_{3-6}$
D21S1437	9	$(GGAA)_n$
D21S1442	13	$(GATA)_n$
D21S2039	7	$(TTAT)_n$
D21S1270	15	$(GATA)_n(GACA)$
D21S1435	9	$(ATCT)_n$
D21S1409	7	$(TAGA)_n$
D21S1446	12	$(TCTA)_n$

Table 3. Allele frequencies and statistical parameters of the 11 STRs on chromosome 21 in the southeast Han Chinese population.

Alleles	D21S1432	D21S11	D21S1246	D21S1412	D21S1437	D21S1442	D21S2039	D21S1270	D21S1435	D21S1409	D21S1446
5	-	-	0.002	-	-	-	-	-	-	-	-
6	-	-	0.304	-	-	-	-	-	-	0.002	0.002
7	-	-	-	-	-	-	0.003	0.006	-	0.375	0.003
7.2	-	-	-	-	-	-	-	-	-	-	0.001
7.3	-	-	-	-	-	-	-	0.001	-	-	-
8	0.042	-	-	-	-	-	0.306	0.290	0.001	0.012	0.540
9	0.039	-	-	-	0.007	-	0.385	0.055	0.004	0.023	0.172
9.2	-	-	-	-	-	-	-	-	-	-	0.030
9.3	-	-	-	-	-	-	-	0.002	-	-	-
10	0.314	-	-	-	0.095	-	0.026	0.045	0.104	0.230	0.028
10.2	-	-	-	-	-	-	-	-	-	-	0.018
10.3	-	-	-	-	-	-	-	0.101	-	-	-
11	0.335	-	0.004	-	0.149	-	0.201	0.108	0.364	0.352	0.011
11.2	-	-	-	-	-	-	-	-	-	-	0.176
11.3	-	-	-	-	-	-	-	0.036	-	-	-
12	0.211	-	0.022	-	0.011	-	0.073	0.228	0.237	0.005	-
12.2	-	-	-	-	-	-	-	-	-	-	0.019
12.3	-	-	-	-	-	-	-	0.007	-	-	-
13	0.056	-	0.349	-	0.043	-	0.006	0.109	0.220	-	-
13.2	-	-	-	-	-	-	-	-	-	-	0.002
13.3	-	-	-	-	-	-	-	0.001	-	-	-
14	0.004	-	0.247	-	0.489	-	-	0.010	0.059	-	-
14.3	-	-	-	-	-	-	-	0.001	-	-	-
15	-	-	0.067	-	0.146	-	-	-	0.011	-	-
16	-	-	0.005	-	0.058	0.002	-	-	0.001	-	-
16.2	-	-	-	0.003	-	-	-	-	-	-	-
17	-	-	-	0.132	0.004	0.005	-	-	-	-	-
18	-	-	-	0.005	-	0.015	-	-	-	-	-
18.3	-	-	-	-	-	0.002	-	-	-	-	-
19	-	-	-	0.004	-	0.084	-	-	-	-	-
20	-	-	-	0.017	-	0.123	-	-	-	-	-
21	-	-	-	0.035	-	0.310	-	-	-	-	-
21.1	-	-	-	0.006	-	-	-	-	-	-	-
22	-	-	-	0.121	-	0.259	-	-	-	-	-
22.1	-	-	-	0.009	-	-	-	-	-	-	-
23	-	-	-	0.163	-	0.106	-	-	-	-	-
23.1	-	-	-	0.006	-	-	-	-	-	-	-
23.2	-	-	-	0.021	-	-	-	-	-	-	-
24	-	-	-	0.159	-	0.052	-	-	-	-	-
24.1	-	-	-	0.009	-	-	-	-	-	-	-
24.2	-	-	-	0.007	-	-	-	-	-	-	-
25	-	-	-	0.169	-	0.027	-	-	-	-	-
25.1	-	-	-	0.006	-	-	-	-	-	-	-
26	-	-	-	0.089	-	0.012	-	-	-	-	-
26.1	-	-	-	0.003	-	-	-	-	-	-	-
27	-	0.001	-	0.023	-	0.002	-	-	-	-	-
28	-	0.054	-	0.013	-	-	-	-	-	-	-
28.2	-	0.003	-	-	-	-	-	-	-	-	-
29	-	0.296	-	-	-	-	-	-	-	-	-
29.2	-	0.004	-	-	-	-	-	-	-	-	-
30	-	0.268	-	-	-	-	-	-	-	-	-
30.2	-	0.007	-	-	-	-	-	-	-	-	-
30.3	-	0.003	-	-	-	-	-	-	-	-	-
31	-	0.092	-	-	-	-	-	-	-	-	-
31.2	-	0.076	-	-	-	-	-	-	-	-	-
32	-	0.024	-	-	-	-	-	-	-	-	-
32.2	-	0.125	-	-	-	-	-	-	-	-	-
32.3	-	0.001	-	-	-	-	-	-	-	-	-
33	-	0.006	-	-	-	-	-	-	-	-	-
33.2	-	0.037	-	-	-	-	-	-	-	-	-
34.2	-	0.003	-	-	-	-	-	-	-	-	-
36	-	0.001	-	-	-	-	-	-	-	-	-
H_E	73.00	81.40	72.20	86.10	70.40	80.80	69.80	78.20	76.50	68.40	64.50
PIC	0.691	0.783	0.672	0.873	0.672	0.774	0.663	0.807	0.712	0.620	0.615
PD	0.887	0.933	0.870	0.973	0.881	0.928	0.865	0.947	0.893	0.832	0.835
PE	0.477	0.624	0.462	0.717	0.435	0.614	0.425	0.565	0.535	0.404	0.348
HW-P	0.864	0.663	0.334	0.152	0.710	0.537	0.051	0.002	0.128	0.021	0.791

H_E = expected heterozygosity; PIC = polymorphism information content; PD = power of discrimination; PE = power of paternity exclusion; HW-P = probability values of exact tests for Hardy-Weinberg disequilibrium.

DISCUSSION

Over the past few years, rapid diagnosis of chromosome aneuploidies has been successfully achieved using QF-PCR amplification of chromosome-specific STR markers (Mann and Ogilvie, 2012). STR markers should be highly polymorphic and heterogeneous, and thus more likely to provide informative results. The Association for Clinical Cytogenetics and the Clinical Molecular Genetics Society recommend that a minimum of 4 markers be tested for each chromosome to reduce uninformative results (Association for Clinical Cytogenetics and Clinical Molecular Genetics Society, 2012). Therefore, the most critical step in QF-PCR is selecting sufficient chromosome-specific STR markers that are informative for a given population.

The heterogeneity of several STR markers in this study in the Han population of southeast China differed from that in other populations. Mann et al. (2001) reported that the heterogeneities of D21S11 and D21S1270 were 0.90 and 0.86, respectively, in a UK population, which is much higher than that observed in the southeast Han population (0.81 and 0.782, respectively); the heterogeneity of D21S1435 was 0.701 in an Indian population, which was much lower than that in the southeast Han population (0.765) (Jain et al., 2012). However, the heterogeneity of D21S1412 and D21S1435 in the southeast Han population was similar to that in a Korean population (D21S1412: 0.861 *vs* 0.876 and D21S1435: 0.765 *vs* 0.761, respectively) (Lee et al., 2010). STR heterogeneity varied among populations, and the most appropriate chromosome-specific STR markers may be different for different populations.

Among the 11 candidate STR markers, all of which showed high heterogeneity (>0.60) in the southeast Han population, D21S1412, D21S11, D21S1442 and D21S1270, showed relatively higher heterogeneities (86.1, 81.4, 80.8, and 78.2%, respectively), and were more suitable for the QF-PCR assay. In addition, the combination of these 4 STRs was associated with a 100% probability that 2 or more of the STRs would be concurrently heterozygous, further confirming their reliability for diagnosing Down syndrome in the southeast Han Chinese population.

In most QF-PCR assays, STR alleles are differentiated from each other by length rather than by name (Mann et al., 2001; Jain et al., 2012; Mann and Ogilvie, 2012). In the present study, we explored the genetic nature of the 11 STR markers and named the alleles according to International Society for Forensic Haemogenetics guidelines (Bär et al., 1997), which may enable the use of these STR markers for other genetic tests such as forensic identification. Of the 11 STR markers included in this study, D21S11 is a CODIS autosomal STR locus, which is consistent with the allele sequencing data in our study (Möller et al., 1994). Except for D21S11, the other 10 markers have rarely been used in forensic study. However, all 11 STR markers showed a high power of discrimination (>0.83), which is useful for forensic applications. In particular, D21S1412 had the highest polymorphism information content (0.87), the highest power of discrimination of the loci (0.973) and the highest probability of paternity exclusion (0.717) among the loci studied. Notably, D21S1246 and D21S1412 existed in linkage disequilibrium, with their loci too close to each other on chromosome 21. As alleles of linked loci form haplotypes that recombine during meiosis, the haplotype frequencies of D21S1246-D21S1412 allowed further analysis for forensic applications.

In conclusion, we found that all 11 STR markers included in this study were informative for diagnosing Down syndrome in the southeast Han Chinese population, and that 4 of these loci (D21S1412, D21S1270, D21S11, and D21S1442) were preferable. This is the first report regarding the nomenclature and allele frequency data for 10 STR markers (apart from

D21S11) in the southeast Han Chinese population. Therefore, all 11 STRs on chromosome 21 can be used in not only QF-PCR but also forensic and other genetic tests.

ACKNOWLEDGMENTS

Research supported by the Program for Zhejiang Leading Team of S&T Innovation, China (#2011R50013-15).

REFERENCES

Association for Clinical Cytogenetics and Clinical Molecular Genetics Society (2012). QF-PCR for the Diagnosis of Aneuploidy Best Practice Guidelines (2012) v3.01. Available at [http://www.devyser.com/Download.ashx?download=40&file=QF-PCR_Best_practice_guidelines.pdf]. Accessed January 2012.

Bär W, Brinkmann B, Budowle B, Carracedo A, et al. (1997). DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. International Society for Forensic Haemogenetics. *Int. J. Legal Med.* 110: 175-176.

Cirigliano V, Lewin P, Szpiro-Tapies S, Fuster C, et al. (2001). Assessment of new markers for the rapid detection of aneuploidies by quantitative fluorescent PCR (QF-PCR). *Ann. Hum. Genet.* 65: 421-427.

Cirigliano V, Voglino G, Ordoñez E, Marongiu A, et al. (2009). Rapid prenatal diagnosis of common chromosome aneuploidies by QF-PCR, results of 9 years of clinical experience. *Prenat. Diagn.* 29: 40-49.

Excoffier L, Laval G and Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47-50.

Hills A, Donaghue C, Waters J, Waters K, et al. (2010). QF-PCR as a stand-alone test for prenatal samples: the first 2 years' experience in the London region. *Prenat. Diagn.* 30: 509-517.

Jain S, Panigrahi I, Gupta R, Phadke SR, et al. (2012). Multiplex quantitative fluorescent polymerase chain reaction for detection of aneuploidies. *Genet. Test. Mol. Biomarkers* 16: 624-627.

Lee MH, Park SY, Kim DJ, Kim MJ, et al. (2010). Genetic variation of three autosomal STR loci D21S1435, D21S1411, and D21S1412 in Korean population. *Mol. Biol. Rep.* 37: 99-104.

Mann K and Ogilvie CM (2012). QF-PCR: application, overview and review of the literature. *Prenat. Diagn.* 32: 309-314.

Mann K, Fox SP, Abbs SJ, Yau SC, et al. (2001). Development and implementation of a new rapid aneuploidy diagnostic service within the UK National Health Service and implications for the future of prenatal diagnosis. *Lancet* 358: 1057-1061.

Möller A, Meyer E and Brinkmann B (1994). Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. *Int. J. Legal Med.* 106: 319-323.

Vallone PM and Butler JM (2004). AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques* 37: 226-231.

Walsh PS, Metzger DA and Higuchi R (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10: 506-513.

Zhu Y, Lu S, Xie Z, Chen Y, et al. (2009). Genetic analysis of 15 STR loci in the population of Zhejiang Province (Southeast China). *Forensic Sci. Int. Genet.* 3: e139-e140.