Analysis of NGS-Ministr as A New Detection Strategy Chromosomal Trisomy 16 of Fetal Materials in Miscarriages

Lilan Yao

Affiliated Hospital of Zunyi Medical University

Yafei Tian

Fudan University

Yijian Zhu

chongqing population and family planing science and technology research institute

Daru Lu

Chongqing population and family planing science and technology research institute

Linbo Chen

Chongqing population and family planing science and technology research institute

Li Mou

Chongqing population and family planing science and technology research institute

Rui Yin

chongqing population and family planing science and technology research institute

Li Luo

Zunyi Medical College Affiliated Hospital: Affiliated Hospital of Zunyi Medical College

Siyu Chai

Affiliated Hospital of Zunyi Medical University

Jingmin Yang

Chongqing population and family planing science and technology research institute

Pengyu Chen (✉ pychenfs@163.com)

Affiliated Hospital of Zunyi Medical University <https://orcid.org/0000-0002-6519-9079>

Research

Keywords: miniSTR, NGS, trisomy 16, miscarriage, spontaneous abortion

Posted Date: July 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-658285/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Although many technologies can identify chromosomal abnormalities, they can't completely replace the advantages of short tandem repeat (STR) genotyping technology. The miniSTR typing based on next generation sequencing (NGS-miniSTR) with high throughput, accuracy and sensitivity has the potential to determine the trisomy and overcome the limitations of conventional STR techniques, which never be tried. The study aimed to explore the value of the NGS-miniSTR in trisomy testing, which is an extension and improvement of the methodology of chromosomal detection technology.

Methods: A total of 140 fetal materials in miscarriages (70 trisomy 16 and 70 normal karyotypes) screened by copy-number variation sequencing (CNV-seq) were genotyped using the customized panel containing 11 miniSTRs based on NGS. Direct counting method, chi-square (χ^2) test, *k*-means clustering analysis and Mahalanobis distance were operated to compare the difference of miniSTR between trisomy 16 and normal group, and then analyze whether trisomy 16 could be identified.

Results: All miniSTR results based on NGS were successfully obtained, except for D16S771 in 21 samples. There were significant differences in the average depth of coverage at each locus in each sample. Direct counting method and chi-square (χ^2) test showed significant differences in allelic pattern and read ratio between trisomy and normal group, respectively. Almost all samples correctly divided into 2 clusters based on diallelic STR reads ratios according to *k*-means clustering analysis. In addition, the Mahalanobis distance showed that D16S771 had multiple outliers.

Conclusion: A new strategy of miniSTR-NGS was firstly introduced to successfully detect all samples of trisomy 16 in the fetal material in miscarriages, which revealed that the allelic pattern and allelic read ratio could be effective indexes to identify the number of chromosomes.

Introduction

Chromosome examination of spontaneous abortions has shown that a significantly risk is chromosomally abnormal, especially in trisomy. Trisomy 16 is the most common, accounting for more than 30% of autosomal trisomy[1].

Short tandem repeats (STRs), or microsatellites, are stretches of simple nucleotide repetitions in the genome and are often polymorphic as a result of the varied number of repeat units forming alleles of different lengths. Because of highly heterozygosity and large amount of genetic information, STRs have been widely used in forensic medicine[2, 3], human evolution[4], anthropology[5, 6], and genetic diseases research[7].

During the past nearly 30 years, trisomy and other aneuploid tests based on STR typing widely adopted the capillary electrophoresis (CE) technology[8]. Since 1993, the quantitative fluorescent polymerase chain reaction (QF-PCR) and other technologies relied on CE have been repeatedly demonstrated that the prenatal diagnosis of chromosome trisomy can be rapidly and accurately achieved by short tandem repeats typing[9, 10]. Additionally, the STR-based techniques for the trisomy have more advantages in detecting maternal contamination, determining the parental source of variation and inferring the period of chromosomal inseparability[8], which cannot be achieved by other non-STR trisomy detection technology, such as traditional karyotyping, fluorescent in situ hybridization (FISH), comparative genomic hybridization (CGH), CNV-seq, et al. The throughput limitation, highly maternal cell contamination and degraded DNA are the main obstacles for traditional STR-based techniques in the detection of embryonic karyotypes in miscarriage.

Recently, the next-generation sequencing (NGS) has provided high-throughput, higher accuracy and sensitivity; meanwhile, it has proved to be available and efficient for miniSTR length and sequence polymorphism in previous researches[11, 12]. The mini-STR, or smaller STR, refers to the shorter STR products produced by designing primers as close to the core repeat sequence as possible[13]. Compared with the routine STRs, the fragments of miniSTR loci are

shorter and the number of repeats of the core sequence remains the same, which improves the sensitivity and detectable rate in mixture/degraded DNA, and is more suitable for NGS[11, 14–16]. Therefore, the miniSTR typing technology based on NGS platform (“NGS-miniSTR”) has the potential to cover all chromosomes based on one sequencing and adapt to the different samples such as fetal material in miscarriages. However, no attempt was done to detect the chromosomal trisomy by NGS-miniSTR.

Herein, we studied trisomy 16, the most common trisomy in spontaneous abortion, and then analyzed the NGS-miniSTR results of 140 aborted samples confirmed by CNV-seq to explore whether trisomy 16 could be identified. The research aimed to provide a basis for NGS-miniSTR as a new detection strategy in trisomy 16 and other chromosomal abnormalities.

Materials And Methods

Sample collection and preparation

A total of 140 pregnancy products of spontaneous abortion experienced fetal chromosomal examination in the third affiliated Hospital of Zunyi Medical University (Guizhou, China) during January 2021 to June 2021 were selected and collected, including 70 cases of trisomy 16 and 70 cases of normal karyotype. All participants have signed the written informed consent, and the research was in accordance with the Helsinki Declaration and approved by the Ethics Committee of the third hospital affiliated to the Zunyi Medical University.

The tissue fragments of fetal materials were confirmed to be placental chorionic villi by gross observation and microscopic examination. Carefully, regions of chorionic villi were dissected from the tissue and thoroughly washed at least three times in normal saline to reduce any maternal cell contamination. Then, genomic DNA was extracted and quantified according to Tissue DNA Kit (OMEGA) and dsDNA HS Assay Kit for Qubit® (YEASEN) protocols, respectively.

CNV-Seq

CNV-Seq was performed as previously described[17], using > 50 ng of genomic DNA as the starting template. Then, library construction and amplification adopted the DNA Hieff NGS OnePot DNA Library Prep Kit for Illumina (YEASEN, Shanghai) according to the manufacturer's instruction. Next, the purified library was sequenced on Illumina Nextseq500 platform (Illumina, USA). Uniquely mapped reads per sample were aligned to the human reference genome GRCH37/hg19. Furthermore, the copy number variation was determined public databases such as Database of Genomic Variants (<http://dgv.tcag.ca/dgv/app/home>), Decipher (<https://www.deciphergenomics.org/>), ClinGen (<https://dosage.clinicalgenome.org/>), OMIM (<http://omim.org/>), UCSC (<http://www.genome.ucsc.edu/>) and PubMed (<https://pubmed.ncbi.nlm.nih.gov/>), etc.

Selection of miniSTR markers

In order to improve test performance, as many miniSTRs as possible were screened into this study through various literatures and biological databases. The miniSTR selection and primer design of the customized STR panel (including D16S2619, D16S2620, D16S2621, D16S2622, D16S2624, D16S2625, D16S3253, D16S3393, D16S752, D16S769 and D16S771) were completed by Shanghai We-health Biomedical Technology Co., Ltd. Accordingly, all primer sequences were listed in Table 1 and these loci were mapped on chromosome 16 using the UCSC Genome Browser (<http://genome.ucsc.edu/cgi-bin/hgBlat>). The inclusion criteria of these markers were as following: (i) STR flank sequences were stable and suitable for the design of miniSTR; (ii) the amplicon size of the miniSTR markers about 150 bp because of the reads limits on Illumina Nextseq500 system; (iii) there are plenty of miniSTRs available to set up composite systems.

miniSTR sequencing and typing

All samples were blindly coded, and then tested by NGS-miniSTR. Initially, DNA template amplification and library construction of targeted miniSTRs for NGS were performed using the customized STR panel. Above progress contained 2 PCR reactions: (i) The first PCR reaction was performed in a 20 µL reaction volume containing 20–200 ng of template DNA, 2.5 µL of Panel mix, 7 µL of 3×T enzyme, and nuclease-free water. PCR condition consisted of pre-degeneration of 3 min at 95°C; 18–20 cycles of 30 s at 96°C, 4 min at 60°C, 4 min at 72°C; then extension of 4 min at 72°C, and hold at 10°C; (ii) A total of 10 µL of the second PCR reaction volume were employed, including 5 µL of 2xKAPA Multiplex PCR mix, 0.5 µL of primers, and 4.5 µL of Nuclease-free water. PCR condition consisted of pre-degeneration of 3 min at 96°C; 6–8 cycles of 15 s at 96°C, 15 s at 58°C, 30 s at 72°C; then extension of 5 min at 72°C, and hold at 10°C. According to the manufacturer's protocol, all library products were sequenced using the NextSeq 500/550 MID Output Reagent Cartridge V2 (150cycle) kit on the Illumina NextSeq500 sequencing platform (Illumina, USA). Then, sequence alignment was performed between the original Fastq file and the human reference genome GRCH37/hg19. Reads were filtered through primer sequences, and specific sequences with length of 10-18bp at both ends of the miniSTR were selected for matching and screening sequence information in the target region. Finally, the accurate allelic typing was obtained by the SSR.pl software and We-health's own self-script.

Statistical analysis

Read coverage distribution and allelic pattern frequencies of 11 loci were estimated by direct counting method. The single allele was treated as monoallelic pattern, two alleles was considered as diallelic pattern, and three alleles was deemed as tri-allelic pattern in each locus. Then, the chi-square (χ^2) test was used to statistically analyze the difference among different allelic pattern (significance level = 0.05). To further describe allele discrepancy between the trisomy 16 and normal samples, the *k*-means clustering analysis (*k* = 2) was conducted with R software[18] based on the biallelic read-ratio of 11 miniSTR loci. Furthermore, the Mahalanobis distance was obtained using SPSS software to exclude outliers, and corresponding outlier was greater than the critical value.

Result

The characteristics of miniSTR

Genomic coordinates for the 11 miniSTRs in Fig. 1 covered the long and short arms of chromosome 16, with eight markers on the long arm and three markers on the short arm. All primer sequence and locus information of 11 miniSTRs on chromosome 16 were listed in Table 1. The sequence length of miniSTR ranged from 107 bp to 159 bp, which can be directly detected by NGS. The loci' GC contents spanned 0.27 (D16S3393) to 0.46 (D16S2619 and D16S771). Notably, the marker D16S771's forward primer contained the repeat region (GAAG), which were different with other loci. Among 140 samples at 11 loci (a total of 1540 loci) in this study, except D16S771 in 21 samples, the genotypes of 1520 (98.7%) loci can be successfully obtained by multiple PCR combine with NGS technology. The coverage of average read of each locus in each sample was 95 (D16S771) to 7054 (D16S2622). The read coverage distribution of each locus was showed in Fig. 2, indicating that different marker presented biased reads. In addition, the lowest reads were observed at the locus D16S771, but highest reads were found at locus D16S2622.

Classification analysis of the STR profiles

Raw miniSTR data on chromosome 16 for 70 trisomy and 70 normal karyotypes from the fetal materials in miscarriages were obtained by NGS (S1 Table). There were three distinct allelic patterns for DNA sample detection in this blind study, including monoallelic, diallelic, and tri-allelic patterns. As expected, there was clear difference in allelic pattern between the trisomy 16 and the normal karyotype. Two allelic patterns were observed in normal karyotype samples, including monoallelic and diallelic patterns. However, three distinct allelic patterns have been found in almost all trisomy samples except T29 and T36 samples, suggesting that the trisomy detection rate was about 98.6%(138/140) based on the allelic

pattern difference between chromosomal trisomy and normal diploidy. The sample T29 include 1 monoallelic locus and 10 diallelic loci, while the sample T36 contained 2 monoallelic loci and 9 diallelic loci. To illustrate the STR profiles in the trisomy 16 and normal samples, the allelic pattern frequency of different miniSTRs were described in Fig. 3. All markers had the tri-allelic pattern, but just in trisomy samples. Among 11 miniSTRs on chromosome 16, the most diallelic pattern was displayed at D16S769, which amounted 94 diallelic combinations ($D_t = 51$, $D_n = 43$), and the tri-allelic pattern was most common at D16S3393 ($T_t = 36$). Then, the least diallelic pattern was observed at D16S771 ($D_t = 33$, $D_n = 29$) and the minimum tri-allelic pattern was found at D16S2621. Meanwhile, the monoallelic pattern was most frequently discovered at D16S771 ($M_t = 20$, $M_n = 41$). Also, the genetic data of locus D16S771 was not available for 21 samples, including 18 trisomy samples and 3 normal samples.

Quantitative analysis of the STR profiles

There was significant deviation in read ratio among different allelic patterns ($p < 0.05$). Although the allelic read-ratios were variable, it usually varied within a certain range. No significant difference was found in the read ratio distribution of monoallelic miniSTR pattern between normal and trisomy samples ($p > 0.05$). The diallelic patterns from normal DNA had two alleles with similar read ratio. In contrast, the trisomy STR profiles presented the different allele read distribution. For trisomy samples, the tri-allelic pattern had three alleles, each of which had similar read ratio, while diallelic patterns had two alleles with a read ratio of nearly 2:1. According to the *k*-means clustering analysis by R software, the cluster result presented in S1 Table and Fig. 4. The read ratios of diallelic pattern in each miniSTR were divided into 2 clusters with an accurate recognition rate of 96.2% (932/938). Cluster 1 mainly covered the trisomy 16, while cluster 2 contained the majority of normal samples, indicating that trisomy-16 and normal diploidy can be distinguished by read ratio differentiation. Under the diallelic pattern, the number of miniSTRs per sample changed from 3 to 11, which can be used to determine the karyotypes. Meanwhile, the sample T29 and T36 can be correctly identified as cluster 1 (namely trisomy group). However, the clustering results showed that there were certain errors in some loci, such as D16S771, which were not consistent with the known CNV-seq results. Moreover, the total allelic read ratio of the locus D16S771 for each sample was lower than other loci. Among 36 misclustered loci, D16S771 accounted for 25 samples, including 11 trisomy samples and 14 normal samples. To further illustrate the bias of these loci, the Mahalanobis distance of outliers for diallelic and tri-allelic patterns was provided in Table 2. There were 9, 17 and 11 outliers in the tri-allelic pattern of trisomy samples, diallelic pattern of trisomy samples, diallelic pattern of normal samples, respectively. Here outliers of the D16S771 were observed 31 times from different allelic patterns, which implicated that D16S771 had a high error rate. Compare with D16S771, other outlier loci (D16S2624, D16S3393, D16S2620 and D16S769) were discovered less frequently, no more than three times.

Discussion

In this study, the miniSTR typing based on NGS was performed on the detection of abnormal products of conception and almost all loci could be successfully typed within the sequence length allowed by NGS. Compared with different allelic pattern, the tri-allelic pattern only existed in the trisomy sample, which means that the tri-allelic combination can reflect the chromosome number. Whereas a few samples without tri-allelic pattern could not be identified by relying only on allelic patterns, such as T29 and T36. Previous researches have proposed that allelic read numbers were associated with peak height in CE, or rather they are equivalent[15]. Although the allele read number (allele coverage) varied tremendously from the various loci or different samples, the read ratio for each locus showed a certain regularity. The allele read ratio of each locus described a quantitative relationship to measure chromosome changes, which could be reflected by *k*-means clustering analysis. It was further suggested that there was sufficient quantitative difference between normal and trisomy in diallelic STR profiles to allow for analysis of chromosomal copies.

Notably, the locus D16S771 deviated from expectations. Not only were the reads at D16S771 significantly lower than other loci, but genotypes could not be obtained in some samples. Besides, the locus D16S771 had multiple outliers and clustering errors. It turned out that D16S771 in this customized panel existed the relatively high noise proportion, which might be affected by the primer sequence or library construction process. Overall, the miniSTR sequencing results in our study demonstrated that the trisomy 16 and normal diploids could be successfully distinguished by NGS-miniSTR strategy. Meanwhile, the election of effective marker is crucial, which should meet the length limitation of NGS and be balanced.

Currently, the study tried to analyze the miniSTR sequencing data which is based on allelic pattern and read ratio, and explore the reliability of the detection of trisomy 16 by NGS-miniSTR strategy. As expected, the allele pattern and allelic ratio in each marker can be effective indexes to quantify the number of chromosomes. Thus, NGS-miniSTR may be a new method for chromosomal trisomy detection. Compare with other methods, NGS-miniSTR with the characteristic of high throughput, sensitivity and accuracy, has potential to performed directly on mixture/degraded DNA samples such as abnormal products of conception, amniotic fluid, blood, etc. Meanwhile, the new strategy can cover all chromosomes base on one sequencing, which cannot be achieved by capillary electrophoresis. Moreover, the NGS-miniSTR strategy have retained the advantages of traditional detection methods in the detection of maternal contamination[19], determining the parental source of variation and inferring the period of chromosomal inseparability. However, above specific advantages cannot be achieved with non-STR detection techniques, such as traditional karyotyping[17], FISH[21], CGH[21], CNVseq[22], etc.

The limitation of our study is that the sample type is relatively small, and only trisomy 16 was tested, which is due to the difficulty of sample collection. So far, the new strategy based on NGS-miniSTR was the first introduced to detect the chromosomal trisomy. This study about the exploration of the new trisomy detection method based on the NGS-miniSTR, is an attempt to expand and improve the methodology of chromosomal abnormality detection technology. Further, it provided a basis for the new strategy in the determination of other aneuploidies from different samples, which was helpful for clinicians to carry out effective treatment and genetic counseling, and further provided scientific basis for digging the genetic etiology of chromosomal abnormalities and for prenatal care.

Conclusion

In the present research, the performance efficiency of the allele pattern and allelic read rate based on the NGS-miniSTR method was 98.6% and 96.2%, respectively, which combined to distinguish almost all trisomy samples 16 from normal samples. Herein, a new detection strategy for chromosome trisomy of fetal material in miscarriages was introduced, which can be rapid, High throughput and highly-sensitive.

Abbreviations

STR: short tandem repeat; NGS: next generation sequencing; NGS-miniSTR: miniSTR typing based on next generation sequencing; CNV-seq: copy-number variation sequencing; CE: capillary electrophoresis; QF-PCR: quantitative fluorescent polymerase chain reaction; FISH: fluorescent in situ hybridization; CGH: comparative genomic hybridization; mini-STR: namely smaller STR.

Declarations

Ethics approval and consent to participate

This work was approved by Committee on Ethics in Institute of Mother and Child, Warsaw, Poland.

Consent for publication

Consent to participate in this study was signed by the patient.

Availability of data and materials

All relevant data are within the manuscript and its additional files.

Competing interests

The authors have declared that no competing interests exist.

Funding

This work was founded by Chongqing Population and Family Planning Science and Technology Research Institute (Grant No: 2020sys-kfkt-0005 and Grant No: 2020cstc-jxj1-01703) and Chongqing Natural Science Foundation (Grant No: cstc2020jcyj-msxm3584).

Author Contributions

PYC and JMY conceived the idea for the study. LLY, YFT, YJZ, DRL, LBC, LM, RY, LL and SYC performed or supervised laboratory work. LLY analyzed the data and wrote the manuscript.

Acknowledgements

We would like to appreciate all members who participated in this study.

References

1. Ogasawara M, Aoki K, Okada S, Suzumori K. Embryonic karyotype of abortuses in relation to the number of previous miscarriages. *Fertil Steril*. 2000;73(2):300–4.
2. Diegoli Toni M. Forensic typing of short tandem repeat markers on the X and Y chromosomes. *Forensic Sci Int Genet*. 2015;18:140–51.
3. Thompson Robyn Z, Silvia, McCord B. An overview of DNA typing methods for human identification: past, present, and future. *Methods Mol Biol*. 2012; 830: 3–16.
4. De Knijff P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, et al. Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med*. 1997;110(3):134–49.
5. Rosenberg NA, Pritchard K, Weber JL, Cann HM, Kidd KK, L.A.Zhivotovsky, et al. Genetic Structure of Human Populations. *Science*. 2002; 298: 2381–2385.
6. Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, Seielstad MT, et al. The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *Am J Hum Genet*. 2000;66(3):979–88.
7. Tang Haibao, Kirkness Ewen F, Lippert Christoph, Biggs William H, Fabani Martin G, Ernesto, et al. Profiling of Short-Tandem-Repeat Disease Alleles in 12,632 Human Whole Genomes. *Am J Hum Genet*. 2017;101(5):700–15.
8. Donaghue C, Mann K, Diack JSW, Dunlop N, Hardy CA, Levett LJ, et al. ACC/CMGS best practice guidelines for QF-PCR-based diagnosis of aneuploidy. *British Human Genetics Conference*. 2004.
9. Mansfield ES. Diagnosis of Down syndrome and other aneuploidies using quantitative polymerase chain reaction and small tandem repeat polymorphisms. *Hum Mol Genet*. 1993;2(1):43–50.

10. Prenatal Diagnosis Cooperative Group on Quantitative Fluorescent Polymerase Chain Reaction. Expert consensus on quantitative fluorescent polymerase chain reaction in the prenatal diagnosis of aneuploidies. *Zhonghua Fu Chan Ke Za Zhi*. 2016;51(5):321–4.
11. Børsting, Claus. Morling Niels. Next generation sequencing and its applications in forensic genetics. *Forensic Sci Int Genet*. 2015;18:78–89.
12. Scheible Melissa L, Odile J, Rebecca I, Jodi. Short tandem repeat typing on the 454 platform: strategies and considerations for targeted sequencing of common forensic markers. *Forensic Sci Int Genet*. 2014;12:107–19.
13. Van N Filip, Van HD, Van N Christophe, Deforce Dieter. Retrospective study of the impact of miniSTRs on forensic DNA profiling of touch DNA samples. *Sci Justice*. 2014;54(5):369–72.
14. Butler John M, Shen Yin, McCord Bruce R. The development of reduced size STR amplicons as tools for analysis of degraded DNA. *J Forensic Sci*. 2003;48(5):1054–64.
15. Wang Z, Zhou D, Wang H, Jia Z, Liu J, Qian X, et al. Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler NGS STR Panel and the Ion PGM System. *Forensic Sci Int Genet*. 2017; 126–134.
16. Lin X, Li H, Liu L, Kong L, Zhang Y. Next generation sequencing as a new detection strategy for maternal cell contamination in clinical prenatal samples. *Ginekologia polska*. 2018;89(6):326–34.
17. Lui S, Song L, Cram DS, Xiong L, Wang K, Wu R, et al. Traditional karyotyping vs copy number variation sequencing for detection of chromosomal abnormalities associated with spontaneous miscarriage. *Ultrasound in Obstetrics & Gynecology*. 2015; 472–477.
18. Kassambara A, Fabian Mundt. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. 2020; <https://CRAN.R-project.org/package=factoextra>.
19. Noveski Predrag T, Marija V, Marija K, Maja SS, Emilija. Plaseska-Karanfilska Dijana. Multilevel regression modeling for aneuploidy classification and physical separation of maternal cell contamination facilitates the QF-PCR based analysis of common fetal aneuploidies. *PLoS One*. 2019;14(8):e0221227.
20. Fejgin MD, Pomeranz M, Liberman M, Fishman A, Amiel A. Fluorescent in situ hybridization: an effective and less costly technique for genetic evaluation of products of conception in pregnancy losses. *Acta Obstetrica et Gynecologica Scandinavica*. 2005; 84(9).
21. Repnikova EA, Rosenfeld JA, Bailes A, Weber C, Pyatt RE. Characterization of copy number variation in genomic regions containing STR loci using array comparative genomic hybridization. *Forensic Science International: Genetics*. 2013;7(5):475–81.
22. Zhao X, Fu L. Efficacy of copy-number variation sequencing technology in prenatal diagnosis. *J Perinat Med*. 2019;47(6):651–5.

Tables

Table1. The primer sequences of 11 miniSTRs on chromosome 16.

LOCUS	Primer sequence[upstream sequence _ downstream sequence]
D16S2619	AAGGGTAATTGGACTCTG_CCTATCTCTATCCATGTA
D16S2620	CCCGGTTGTCAAAAGAGG_AAAAGAGAAAGTGAGAGA
D16S2621	CCCGCCTCAGGTAGGTGA_TCCATAGTCATATGGGCC
D16S2622	CCATTCCATTTCTAACT_ATTTCACTCTAGCTTGGG
D16S2624	GAGGCAATTTGTTACAGA_TGAAAAGTAATAATGTAC
D16S2625	GGGGACACATGAGACCCT_TCAGATCATTCTTCCCTT
D16S3253	TGTGTTCAAGCATGGGTG_TTACAATAGCACCACTCC
D16S3393	GGGGAGATCCTGCCTCTA_GGAGTGCCTGAAGGAGTT
D16S752	GAGGATTGGAGGAGGGTG_CAACCCTGTCCCTAAGGA
D16S769	GCTCAGAGATTCTCAGAA_TTCCTCTTAGCCTTGGTC
D16S771	TGCTTTAAGACAGAAGAA_ACAACAAGTGTGCCTTCT

Table 2. The Mahalanobis distance of outliers for diallelic and tri-allelic patterns.

ID	Loci	MAH_Tt(12.84)	ID	Loci	MAH_Dt(10.60)	ID	Loci	MAH_Dn(10.60)
T58	D16S2624	40.89	T11	D16S771	49.03	N59	D16S771	119.64
T20	D16S771	38.59	T17	D16S2624	46.02	N52	D16S771	73.07
T43	D16S771	27.09	T19	D16S769	44.04	N54	D16S771	47.95
T32	D16S771	18.13	T55	D16S771	40.23	N50	D16S771	41.77
T11	D16S3393	17.69	T66	D16S771	27.31	N37	D16S771	40.02
T46	D16S771	16.49	T8	D16S771	24.13	N18	D16S771	26.10
T69	D16S771	16.22	T21	D16S771	20.31	N40	D16S771	20.44
T62	D16S771	16.15	T18	D16S771	19.91	N51	D16S771	20.10
T53	D16S771	16.15	T8	D16S771	18.59	N45	D16S771	18.21
			T1	D16S771	18.26	N54	D16S3393	11.82
			T47	D16S3393	17.91	N56	D16S771	11.76
			T65	D16S771	14.85	N17	D16S2620	10.91
			T25	D16S771	13.17			
			T50	D16S771	12.99			
			T57	D16S771	12.82			
			T29	D16S771	11.16			
			T47	D16S771	10.75			

Note: MAH_Tt(12.84), MAH_Dt(10.60) and MAH_Dn(10.60) presented the Mahalanobis distances of the tri-allelic locus for trisomy samples, the diallelic locus for trisomy samples, diallelic locus for normal samples, respectively, and the corresponding critical values were 12.84, 10.60 and 10.60.

Figures

Figure 1

Genomic coordinates for the 11 miniSTR on chromosome 16.

Figure 2

The read distribution of each locus for all samples.

Figure 3

The allelic pattern frequency of different miniSTRs in the trisomy 16 and normal samples. Note: Mt: monoallelic pattern of the trisomy 16; Mn: monoallelic pattern of the normal sample; Dt: diallelic pattern of the trisomy 16; Dn: diallelic pattern of the normal sample; Tt: tri-allelic pattern of the trisomy.

Figure 4

The k-means clustering result for the diallelic read ratios.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [S1Table.xls](#)