Noninvasive prenatal diagnosis for Hemophilia A by haplotype-based approach using cell-free fetal DNA in maternal plasma

Running title: Noninvasive prenatal diagnosis for Hemophilia A

Zhihui He[1], Chao Chen[2], Fengyu Guo[2], Lu Jiang[2],Yaping Zhu[2], Dan Li[1], Renhua Wu[2], Rong Lu[1], Mei Zhao[3], Fang Chen[4], Peixiang Ni[2], Jun Sun[2*], Zhiyu Peng[4*]

[1]Department of Obstetrics and Gynecology, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China;

[2]Tianjin Medical Laboratory, BGI-Tianjin, BGI-Shenzhen, Tianjin 300308,China;

[3]BGI-Guangzhou Medical Laboratory, BGI-Shenzhen, Guangzhou 510006, China;

[4]BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China

The first two authors contributed equally to this work.

Word count 1677, table 2 and figure 3

***Correspondence author:**

Zhiyu Peng (Bei Shan Industrial Zone, Yantian District, Shenzhen, China, +8618665899278, pengzhiyu@genomics.cn)

Jun Sun (Central Avenue 55, Airport Business Park East Building E3, Tianjin, China, +8613622027365, sunjun@genomics.cn)

Funding sources: The Natural Science Foundation of Guangdong Province China (2015A030313472)

Conflicts of interest:

There is no potential conflicts about this manuscript.

Abstract

Objectives: This study was used to demonstrate the feasibility of haplotype-based noninvasive prenatal diagnosis for Hemophilia A.

Methods: Two families affected by Hemophilia A participated in our study. Maternal haplotypes associated with pathogenic mutation was built using targeted region genotypes of mother and the proband. Combined with maternal pathogenic haplotype, a Hidden Markov Model was constructed to deduce fetal haplotype using high-coverage targeted sequencing of maternal plasma. The presence of pathogenic haplotype in male indicated affected fetus, in female indicated carrier. Prenatal diagnosis was confirmed with amniocentesis by long distance PCR.

Results: The haplotype-based noninvasive prenatal diagnosis was successfully performed in Hemophilia A. One fetus was identified to be normal, another fetus was identified as carrier, and the results were confirmed by amniocentesis .

Conclusion: Our research demonstrated the feasibility of noninvasive prenatal diagnosis for Hemophilia A by haplotype-based approach.

Introduction

Hemophilia A (HA) is one of the most common X-linked recessive inherited bleeding disorder with a incidence of approximately 1/5,000 in males [1]. HA was caused by genetic mutation of *F8* gene, resulting in deficiency of coagulation factor factor VIII in plasma levels. The clinically mainifesting was spontaneous or traumatic bleeding in several organs. Prenatal diagnosis is important for HA families because of lacking in effective curing method. But the invasive prenatal diagnosis may led a small risk of miscarriage and infection [2]. The discovery of cell free fetal DNA (cff-DNA) in maternal plasma makes noninvasive prenatal diagnosis (NIPD) possible [3]. The intron 22 inversion in *F8* gene accouts for approximately 45% of severe HA patients. Because of the vast background of maternal DNA [4] and technical obstacles in determining the inversion breakpoint in plasma, the direct detection is impracticable. The haplotype-based method could potentially solve the problem through analyzing the mutation-linked polymorphism sites [5]. This was the few use of NIPD for HA by haplotype-based approach to demonstrate our approach to be an affordable strategy for clinical feasibility of NIPD of X-linked disorders.

Materials and Methods

HA cases recruitment and clinical information

Two families F01-F02 diagnosed with HA were recruited including two pregnant women, their spouses, and their brother affected by HA. The causative mutation in the brothers and the pregnant women were previously identified to be intron 22 inversion in *F8* gene by long

distance PCR (LD-PCR). Genetic counseling were given to the families and the parents decided to take our noninvasive prenatal test. All participants signed the informed consent. The ethics committee of The First Affiliated Hospital of Guangzhou Medical University approved this study. The 5 mL blood samples obtained from each of family members including mother, father and proband were transported to BGI-Tianjin. Maternal plasma was separated using double-centrifugation method and stored frozen in time. Amnioticfluid was obtained through amniocentesis for routine prenatal diagnosis. Invasive prenatal diagnosis and NIPD were performed in a double-blind manner by two independent investigators.

Library preparation and targeted sequencing

In this study we used haplotype-based noninvasive prenatal testing of HA using a 110.08 kb SeqCap kit (Roche) including *F8* gene coding region and 445 surrounding highly heterozygous SNPs (Minor Allele Frequency, MAF=0.3-0.5) distributed within 1Mb chromosome X region（Figure 1）.Inclusion of 10.37kb chromosome Y-specific region facilitates the simultaneous determination of fetal gender.

Genomic DNA (gDNA) and cell free DNA was extracted from blood or umbilical cord blood and maternal plasma using QIAamp DNA Mini kit (Qiagen) and QIAamp Circulating Nucleic Acid Kit (Qiagen), respectively. The library was prepared according to our mature protocal. The libraries of gDNA and cell free DNA were pooled into one and then captured using our customized probe, respectively. Using BWA software, the sequencing reads were aligned to the human reference genome (hg19). SNP calling was performed using GATK software.

Determination of fetal gender and DNA fraction

The mean depth of target region in Y chromosome was utilized to determine the fetal sex. If the fetus is female, the depth of Y chromosome region would be extremely low, resulting from short non-specific region reads could be mapped to Y chromosome specific region.

The fetal DNA concentration in maternal plasma fraction was calculated by taking advantage of SNP sets with different parental homozygosity according to the formula: $\varepsilon=2dF/(dF+dM)$, where ε represents for fetal DNA fraction in plasma, dF is the depth of reads supporting fetal specific alleles and dM is the depth of reads supporting maternal shared alleles in plasma. Assuming maternal genotype is AA and paternal genotype is TT, the fetal genotype should be AT. So fetal specific allele was T and maternal shared allele was A.

Haplotype-based noninvasive prenatal diagnosis of *F8* gene

A noninvasive haplotype-based linkage analysis method as previous described [6] was used to determine fetal haplotype using the SNP information flanking the 1M and coding region of *F8* gene of proband and parents . We defined haplotype 0 (Hap0) as the pathogenic haplotype and haplotype 1 (Hap1) as the wildtype haplotype. Because the candidate Hap0 was inherited from maternal chromosome X and we directly phased the maternal haplotypes of *F8* gene. . Based on maternal haplotype and SNP representation in the plasma, an HMM was constructed. We chose the SNP sets which were heterozygous in mother but homozygous in father to acquire the maternal inheritance. The fetal DNA fraction and sequencing error rate were utilized to determine the maternal haplotypes inheritance in the fetus. The probability of each

state of maternal inheritance was calculated for each SNP marker in the maternal plasma. The fetal haplotype was deduced using Viterbi algorithm used to decode the most hidden states.

Prenatal diagnosis of HA using amniotic fluid

LD-PCR using primer P, Q and B [7] of amnioticfluid was utilized to assess the accuracy of haplotype-based NIPT of HA. Electrophoresis showed 12 kb band (Fragment PQ) for wild-type, 11 kb band (Fragment PB) for male patient and PQ and PB bands for female carriers .

Results

Sequencing data

For maternal plasma , we obtained a mean 41.35 million data mapped to target region with a mean depth of 375.61x and 99.48% coverage more than 20 reads. A mean depth of 151.08x was obtained in genomic DNA of pregnant woman, her brother and her spouse, respectively and the 20x coverage was 97.93% (Table 1). The data satisfied subsequent analysis.

Estimation of fetal gender and DNA fraction

The mean depth and 20x coverage mapped to chromosome Y in F01 plasma was 26.5x and 42.59%, respectively. The fetus of F01 was inferred to be male. However, no reads was mapped to target region of chromosome Y, the fetus of F02 was determined to be female. The fetal DNA fraction of F01 and F02 was 11.09% and 9.37%, respectively.

Haplotype-based noninvasive prenatal diagnosis

The maternal haplotype could be constructed based on these SNPs. The number of informative phased SNPs used for NIPD for F01 and F02 was 144 and 49, respectively. In F01, the number of SNPs (Table 2, Figure 2) supported fetus inherited maternal wild haplotype (Hap1) was 144, while none supported the pathogenic haplotype (Hap0). The fetus could be subsequently predicted to be unaffected with HA considering the fetal male gender. In F02, the number of SNPs (Table 2, Figure 2) supported fetus inherited maternal pathogenic haploype (Hap0) was 49, while none supported the wild haplotype (Hap1). The fetus could be subsequently predicted to be carrier with HA considering the fetal female gender.

Invasive prenatal diagnosis

Prenatal diagnosis performed by LD-PCR on the F01 was proved to be unaffected fetus with HA (Figure 3A).The fetus of F02 was diagnosed as carrier with HA (Figure 3B).

Discussion

As a simple procedure, the NIPD method is easier to access comparing to invasive prenatal diagnosis. NIPD for HA would allow testing in the early weeks of gestation, give more rooms and time for making decision of terminating the pregnancy. Moreover, it gives a choice for women who refuse invasive methods to have more information about the possibility of an affected fetus. The discovery of cell free fetal DNA (cff-DNA) in maternal plasma bring a new start for noninvasive prenatal diagnosis. Heretofore, several studies have reported that noninvasive prenatal detection of fetal chromosome aneuploidy is high accuracy [8, 9]. In addition, whole genome sequencing (WGS) of maternal plasma DNA can restore the

fetal specific allele with accuracy more than 95%, thus providing the possibility of single gene detection in maternal plasma [10]. Nevertheless, the extremely high cost of whole genome deep sequencing limited its clinical application. Deep sequencing of target region is a non-invasive alternative to single- gend disease detection, which could realize the deep sequencing of interest region at a lower cost [11]. Lam et al. developed a noninvasive prenatal algorithm for diagnosing fetal single-gene diseases using target region capture sequencing for single base variations and indels [12]. Using this method, two fetuses were correctly diagnosed as β-thalassemia carriers. Cell-free fetal DNA analysis in maternal plasma provides a noninvasive method for assessing fetal sex in HA pregnancies. Whereas, without mutation-specific confirmatory analysis, the disease status of male fetuses remains unknown. Tsui et al. uesed microfluidics digital PCR method to do noninvasive prenatal diagnosis for HA [13]. But it can only detect the point mutation. The high incidence of *F8* inversion 22 (Inv 22) mutation cannot be detected by this method. Microfluidics digital PCR is expensive, because a microfluid digital PCR chip need thousands of dollars. Our team developed a haplotype-based strategy for one step restoration of fetal haplotype and genotype on both parentally heterozygous sites in the maternal plasma [5].Using this method, we successfully predicted the fetal mutation status of many autosomal recessive genetic diseases based on the maternal plasma, including congenital adrenal hyperplasia, maple syrup urine disease, congenital deafness and Duchenne Muscular Dystrophy [14-16]. This is the few application of the haplotype-based strategy to HA prenatal detection. New technologies (for example, microfluidics digital PCR and next-generation sequencing) greatly accelerates the clinical

application of NIPT, including fetal gender, RhD status, and aneuploidy screening. Our haplotype-assisted strategy provides an opportunity for the diagnosis of single-gene diseases by NIPD.

At present, fetal gene sequencing analysis method based on capture technologies is expensive because the capture range is large. All libraries were captured with a target region of 110.08kb including *F8* gene coding area, 445 highly heterozygous SNPs distributed within 1 Mb flanking regions of *F8*, and 10.37 kb chromosome Y-specific region. So, we can reduce the cost of sequencing significantly.

Gender diagnosis is very important of NIPD of X-linked inherited diseases. We use the specific Y chromosome probe to do gender diagnosis easily and accurately. The capture range and depth of the Y chromosome is significantly difference between male and female fetuses according to cff-DNA sequencing data analysis. It is approximately 0 for female fetus. NIPD may predict fetal gender accurately.

Even though NIPD on HA has a lot of advantages compared with the invasive methods, but the accuracy needs to be further improved before clinical application. As accuracy increases and cost reduce, we hypothesized that NIPD will be used as an alternative for pregnant women who are likely to have fetuses with HA and other x-linked diseases in the future, faster and easier.

Acknowledgments

This work was funded by The Natural Science Foundation of Guangdong Province China (2015A030313472). We are grateful to the participants and their families. We also thank our coinvestigators for their assistance in research design, data analysis..

References

1. Dai J, Lu Y, Ding Q, et al. The status of carrier and prenatal diagnosis of haemophilia in China. Haemophilia, 2012, 18 (2): 235-240.

2. Kozlowski P, Knippel A, Stressig R. Individual risk of fetal loss following routine second trimester amniocentesis: a controlled study of 20,460 cases. Ultraschall Med. 2008; 29: 165-172.

3. Lo YM, Corbetta N, Chamberlain PF, et al. Presence of fetal DNA in maternal plasma and serum. Lancet. 1997; 350: 485-487.

4. Lo YM, Tein MS, Lau TK, et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. Am J Hum Genet. 1998; 62: 768-775.

5. Chen SP, Ge HJ, Wang XB, et al. Haplotype-assisted accurate non-invasive fetal whole genome recovery through maternal plasma sequencing. Genome Med. 2013; 5: 18.

6. Xu Y, Li X, Ge HJ, et al. Haplotype-based approach for noninvasive prenatal tests of Duchenne muscular dystrophy using cell-free fetal DNA in maternal plasma. Genet Med. 2015; 17(11): 889-896.

7. He ZH, Chen SF, Chen J, Jiang WY. A Modified I-PCR to detect the factor VIII Inv22 for genetic diagnosis and prenatal diagnosis in haemophilia A. Haemophilia, 2012, 18(3): 452-456.

8. Norton ME, Brar H, Weiss J, et al. Non-Invasive Chromosomal Evaluation (NICE) Study: results of a multicenter prospective cohort study for detection of fetal trisomy 21 and trisomy 18. Am J Obstet Gynecol. 2012; 207: 137 e131-138.

9. Palomaki GE, Deciu C, Kloza EM, et al. DNA sequencing of maternal plasma reliably identifies trisomy 18 and trisomy 13 as well as Down syndrome: an international collaborative study.Genet Med. 2012; 14: 296-305.

10. Lo YM, Chan KC, Sun H, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. Sci Transl Med. 2010; 2: 61ra91.

11. Liao GJ, Chan KC, Jiang P, et al. Noninvasive prenatal diagnosis of fetal trisomy 21 by allelic ratio analysis using targeted massively parallel sequencing of maternal plasma DNA. PLoS One. 2012; 7: e38154.

12. Lam KW, Jiang P, Liao GJ, et al. Noninvasive prenatal diagnosis of monogenic diseases by targeted massively parallel sequencing of maternal plasma: application to beta-thalassemia. Clin Chem. 2012; 58: 1467-1475.

13. Tsui NB, Kadir RA, Chan KC, et al. Noninvasive prenatal diagnosis of hemophilia by microfluidics digital PCR analysis of maternal plasma DNA Blood. 2011; 117(13): 3684-3691.

14. Ma DY, Ge HJ, Li XC, et al. Haplotype-based approach for noninvasive prenatal diagnosis of congenital adrenal hyperplasia by maternal plasma DNA sequencing. Gene. 2014; 2: 252-258.

15. You Y, Sun Y, Li XC, et al. Integration of targeted sequencing and NIPT into clinical practice in a Chinese family with maple syrup urine disease. Genet Med. 2014; 16: 594-600.

16. Meng M, Li XC, Chen F, et al. Noninvasive prenatal testing for autosomal recessive conditions by maternal plasma sequencing in a case of congenital deafness. Genet Med. 2014; 16: 972-976.

Figures Legends

Figure 1. Target region of F8 gene and SNPs used for haplotyping. Custom NimbleGen probes within the 110.08 kb region was designed including the F8 gene coding region and 445 surrounding highly heterozygous SNPs distributed within 1 Mb region.

Figure 2. Results of haplotype-based NIPD for Hemophilia A. X-axis represents the target region of our customized probe and blue vertical lines represent interval of F8 gene; Y-axis represents the logarithm of the ratios of fetal-inherited maternal haplotype. Hap0 represents the pathogenic haplotype and Hap1 represents the wild haplotype. The black lines above zero indicate that the fetus inherited the pathogenic haplotype (Hap0), and the black lines below zero indicate that the fetus inherited the normal haplotype (Hap1).

Figure 3. Results of invasive diagnosis of studied families. A Electrophoresis result LD-PCR of amniotic fluid (F01). *M: Marker, Lane 1and 2: Amniotic fluid, Lane 3: Mother, Lane 4: Carrier control of the kit, Lane 5: Normal control of the kit, Lane 6: Blank control. Finished by the researchers of The First Affiliated Hospital of Guangzhou Medical University. B Electrophoresis result LD-PCR of amniotic fluid (F02). *PB was amplified by primer P and B, PQ was amplified by primer P and Q, N: Normal Control, P: Positive Control(Male patient), 714R2: amniotic fluid. Finished by the researchers of BGI Genomics, BGI-Shenzhen.

Table 1 Sequencing data of studied families.

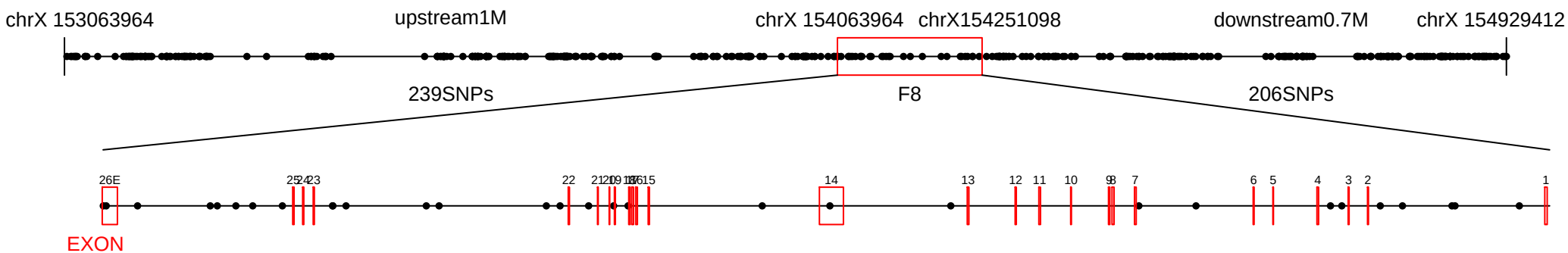
Family	Sample	Target region data (M)	Mean depth (x)	20x Coverage fraction	Mean depth of chr Y	20x Coverage fraction of chr Y	SNPs	Fetal DNA fraction
F01	mother	21.29	193.36	98.70%	0	0	248	11.09%
	father	15.50	140.84	98.31%	45.28	72.2%		
	proband	15.14	137.51	98.11%	43.05	71.1%		
	plasma	29.93	271.87	99.37%	26.50	42.59%		
	AF	11.00	99.94	96.80%	113.67	96.74%		
F02	mother	22.73	206.49	98.79%	0	0	106	9.37%
	father	10.22	92.82	95.81%	102.40	95.4%		
	proband	14.91	135.47	97.83%	153.40	97.4%		
	plasma	52.77	479.35	99.59%	0	0		
	AF	21.06	191.30	98.63%	0	0		

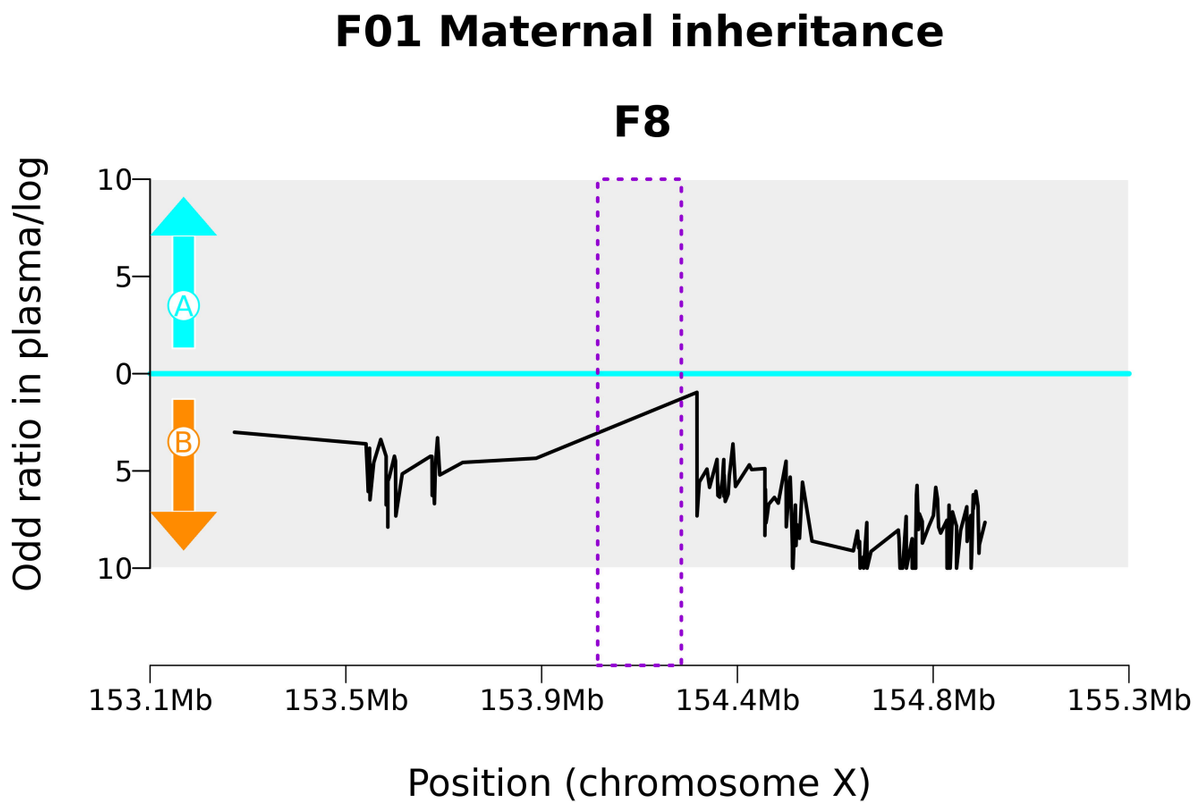
Notes: AF, Amniotic fluid; the proband was brother of pregnant woman; SNPs represents the number of SNPs detected in studied families.

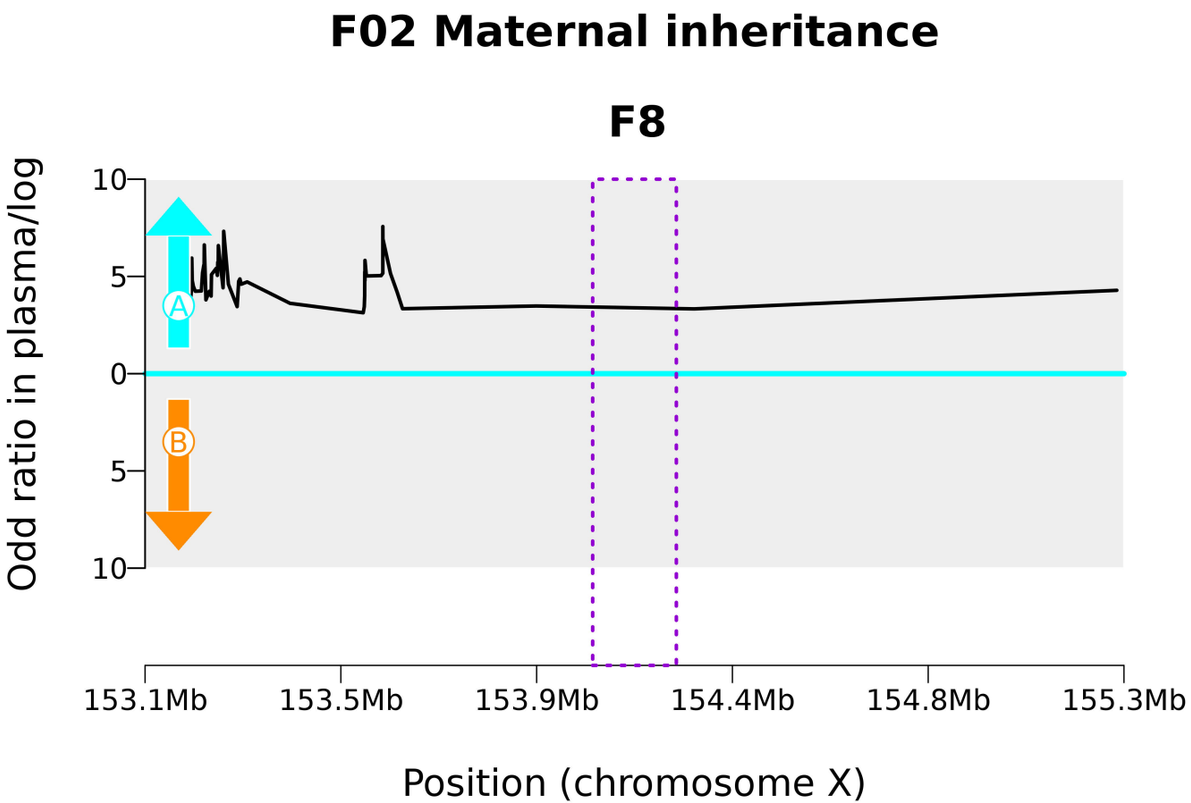
Table 2 The NIPD results of *F8* gene

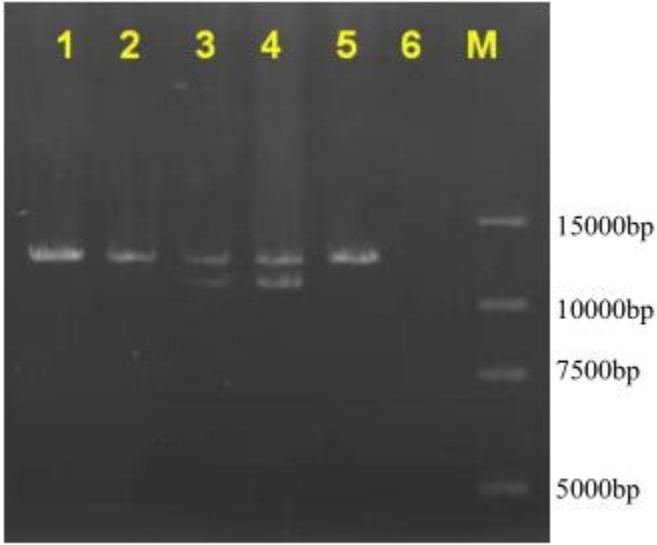
Family	Fetal Gender	Phased SNPs[a]	SNPs For Hap0	SNPs For Hap1	NIPD results	Predicted results
F01	male	144	0	144	Hap1	normal
F02	female	49	49	0	Hap0	carrier

Notes: SNPs[a] represents SNPs that were used to predict maternal haplotypes. Hap0, fetal-inherited maternal mutant haplotype; Hap1, fetal-inherited maternal wild-type haplotype.

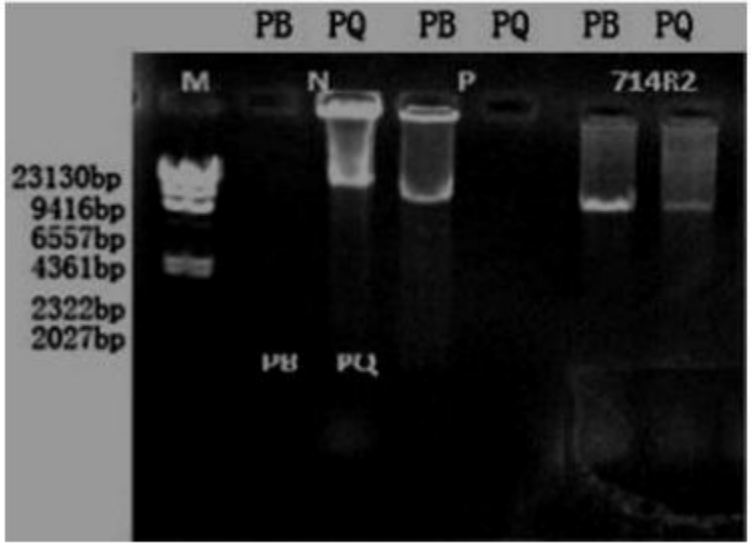

chrX 153063964
upstream1M
chrX 154063964
chrX154251098
downstream0.7M
chrX 154929412
239SNPs
F8
206SNPs
26E
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
EXON


F01 Maternal inheritance
F8
Odd ratio in plasma/log
10
5
0
5
10
A
B
153.1Mb
153.5Mb
153.9Mb
154.4Mb
154.8Mb
155.3Mb
Position (chromosome X)


F02 Maternal inheritance
F8
Odd ratio in plasma/log
10
5
0
5
10
A
B
153.1Mb
153.5Mb
153.9Mb
154.4Mb
154.8Mb
155.3Mb
Position (chromosome X)


1
2
3
4
5
6
M
15000bp
10000bp
7500bp
5000bp

A


PB
PQ
PB
PQ
PB
PQ
M
N
P
714R2
23130bp
9416bp
6557bp
4361bp
2322bp
2027bp
PB
PQ

B