Short Tandem Repeats Used in Preimplantation Genetic Testing of B-Thalassemia: Genetic Polymorphisms For 15 Linked Loci in the Vietnamese Population

Dang Tien Truong, Ngo Van Nhat Minh, Dinh Phuong Nhung, Hoang Van Luong, Do Quyet, Tran Ngoc Anh, Trinh The Son, Nguyen Thanh Tung, Nguyen Thi Thu Ha, Duong Thi Phuong Anh, Le Hoang, Nguyen Le Thuy, Nguyen Thi Hoa, Nguyen Duy Bac, Vu Thi Ngoc, Toi Chu Dinh

1Vietnam Military Medical University, Hanoi, Vietnam; 2National Institute of Hematology and Blood Transfusion, Hanoi, Vietnam; 3Tam Anh General Hospital, Hanoi, Vietnam; 4Institute for Research and Development, Duy Tan University, Danang, Vietnam; 5Department of Human and Animal Physiology, Faculty of Biology, Hanoi National University of Education, Hanoi, Vietnam

Abstract

BACKGROUND: β-thalassemia is one of the most common monogenic diseases worldwide. Preimplantation genetic testing (PGT) of β-thalassemia is performed to avoid affected pregnancies has become increasingly popular worldwide. In which, the indirect analysis using short tandem repeat (STRs) linking with HBB gene to detect different β-globin (HBB) gene mutation is a simple, accurate, economical and also provides additional control of contamination and allele-drop-out ADO.

AIM: This study established microsatellite markers for PGT of Vietnamese β-thalassemia patient.

METHODS: Fifteen (15) STRs gathered from 5 populations were identified by in silico tools within 1 Mb flanking the HBB gene. The multiplex PCR reaction was optimized and performed on 106 DNA samples from at-risk families.

RESULTS: After estimating, PIC values were ≥ 0.7 for all markers, with expected heterozygosity and observed heterozygosity values ranged from 0.81 to 0.92 and 0.53 to 0.86, respectively. One hundred percent of individuals had at least seven heterozygous markers and were found to be heterozygous for at least two markers on either side of the HBB gene.

CONCLUSION: In general, a pentadecaplex marker (all ≤ 1 Mb from the HBB gene) assay was constituted for β-thalassemia PGT on Vietnamese population.

Introduction

β-Thalassemia is one of the most common monogenic diseases, accounting for 1.5% of the population [1], concentrated in Central and Southern Asia, the Middle East, Northern Africa and the Mediterranean including Vietnam with the carrier. HBB mutations have more than 200 different types which have been knowns [2] lead to insufficient β-globin synthesis. The frequency of Vietnamese carriers ranged from 1.5 to 25.0% depending on ethnic group [3], [4], [5]. PGT-M (Preimplantation genetic testing for monogenic disease) promises to prevent monogenic disease in children born to at-risk couples by avoiding transferring mutation embryos to women in vitro fertilization. In which, PGT-M for β-Thalassemia has become the most common application among monogenic disorders.

There have been various studies that established direct or indirect PGT procedure for β-Thalassemia. Among numerous techniques, microsatellite markers such as STRs have provided many advantages in the indirect analysis. STRs are the repetitive DNA fragments of 2-6 bp which structure is highly conservative, inherited through generations and characteristic for each individual. The STRs is also highly diverse and can be amplified by PCR.
Thus, microsatellite markers linking with HBB gene has played an essential role in linkage analysis for β-Thalassemia.

Currently, there have been many typical studies on indirect linkage-based PGT for β-thalassemia published in the world. Wen Wang (2009) combined the Nested-PCR method amplifying STR markers and minisequencing method on nine embryos and concluded five unaffected embryos for transferring [6]. Li Fan (2017) used STR markers to perform PGT on WGA products from 147 day-5 embryos and identified 24 non-mutations, 38 carriers, and 18 mutation embryos [7].

Nevertheless, these direct analysis methods have no probability of controlling contamination and ADO phenomenons, which are considered as the main reasons leading to misdiagnosis in PGT. Thus, β-Thalassemia with wide range of gene and variety of mutations is recommended applying with indirect analysis method. These methods were proved to be sensitive, accurate, reliable and rapid to control the pitfalls of PCR-based PGT, including PCR failure, contamination, and ADO.

Despite the preeminence of STRs, the limited number of researches and available markers for the Vietnamese population, this hampers their utility in linkage-based β-Thalassemia PGT. In this study, we developed a multi-marker panel consisting of 15 STRs for Vietnamese β-Thalassemia patients. The data suggested that the STRs set was qualified to perform PGT-M with high heterozygous values, number of heterozygous markers on each individual and the equal distribution of markers on either side of HBB gene.

Materials and methods

**Control Samples for Method Optimization**

One hundred six genomic DNA samples were extracted from blood and amniotic fluid of at-risk families at the Vietnam National Institute of Hematology and Blood Transfusion (NIHBT), Vietnam from 07/2016 to 6/2018. DNA was used either to pre-screen of microsatellite markers or to determine heterozygosity values of them.

**Short tandem repeat**

Initial selected STRs were identified based on the STR database and Tandem Repeat Finder provided by Gary Benson (http://tandem.bu.edu/trf/trf.html). DNA sequence within 1 Mb upstream and downstream of the HBB gene (11p15.4) (genome assembly GRCh37/hg19. Feb 2009, annotation) was extracted from the UCSC Genome Browser. Initial selection criteria for the STRs followed Machado 2009 [8]. The first microsatellite markers were subsequently compared and selected for Vietnamese population based on report from populations of Malaysia, China, and India [9]. Primers were designed by Primer3 Tool. UCSC In-silico PCR with downloaded reference DNA sequence (genome assembly GRCh37/hg19. Feb 2009, annotation) and BLAST from NCBI were used to determine and exclude the primer complementing with Alu and non-specific sequences.

**DNA Extraction**

DNA was extracted from blood or amniotic fluid by Blood DNA Extraction QIAamp® DNA Mini Kit (Cat No./ID: 51304, QIAGEN) following the optimal instruction from QIAGEN Producer. Purified DNA was qualified and quantified by NanoDrop One Spectrophotometer with criteria: OD A260/A280 between 1.7-2.0 and concentration above 10 ng/µl.

**PCR amplification**

Single PCR: Followed T_m of 15 primers on http://www.operon.com/tools/oligo-analysis-tool.aspx, determined the average theoretical annealing temperature of these 15 primer pairs which is 60°C. Thus, PCR single primer was conducted according to the temperature range set at 55°C-60°C-65°C, and products were analyzed on the agarose gel. Single PCR amplification was performed in a 50 µL reaction volume consisting of 2 µL genomic DNA (concentration: 10-20 ng/µl), 25 µL 2X QIAGEN Multiplex Master Mix, and 0.2 µM of each primer (Table 1). Thermal cycling involved an initial 15 minutes enzyme activation at 95°C, 30 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 1 minute 30 seconds, and extension at 72°C for one minute, and a final extension at 60°C for 30 minutes.

**Table 1: Singleplex PCR components (Total volume: 50 µL)**

| Component                  | Concentration | Volume (µL/tube) |
|----------------------------|---------------|-----------------|
| PCR inhibitors             |               | 21              |
| 2X QIAGEN Multiplex MasterMix | 1 X           | 25              |
| Forward primer             | 0.2 µM        | 1               |
| Reverse primer             | 0.2 µM        | 1               |
| DNA template               |               | 2               |

**Multiplex PCR optimization**

Multiplex PCR was performed at optimal annealing temperature and primer concentrations initially keeping at the same concentration of 0.2 µM. Then, each primer concentration was adjusted based on product signal strength by increasing or decreasing 0.05 µM. The multiplex PCR amplification was performed in 50 µl reaction consisted of 1X QIAGEN Multiplex PCR Master Mix, 0.05-0.4 µM of each primer and 100 ng DNA template; followed the protocol: an
initial 15 min enzyme activation at 95°C, 30 cycles of denaturation at 95°C for 30 seconds, annealing at 60°C for 1 min 30 s at the first cycle and 6 additional seconds for every next cycle, and extension at 72°C for 1 min.

**Capillary electrophoresis**

The PCR products were fluorescently labeled in a 20 µL mixture consisted of 1X QIAGEN Multiplex MasterMix, 0.2 µm each fluorescent primer with M13 sequences as in Table 2, following cycling condition: an initial denaturation step at 95°C for 15 min, followed by 5 cycles of denaturation at 95°C for 45 s, annealing at 60°C for 1 min, and extension at 72°C for 1 min, and a final extension at 72°C for 5 min.

One µL aliquot of fluorescent PCR product was mixed with 8.5 µL of Hi-Di Formamide (Applied Biosystems, Foster City, CA, USA) and 0.5 µL of GeneScan 500 LIZ size standard, denatured at 95°C for 5 minutes, cooled to 4°C, and resolved in ABI 3130XL Fragment Analyzer (Applied Biosystems). Post-electrophoresis analysis was performed using GeneMapper 5.0 software (Applied Biosystems).

**Statistical analysis**

Allele frequency, PIC, expected heterozygosity (He) and Observed Heterozygosity (Ho) of the 15 microsatellite markers were calculated using Microsoft Excel.

**Results**

**Identification of STR**

Fifteen (15) STRs were identified within 1 Mb around the HBB gene (6 upstream STRs and nine downstream STRs) (Table 2). HBB4506 located farthest to HBB gene (0.74 Mb) and all other markers were comparatively closer. Thus, all STRs had high linkage to the HBB gene.

| No. | STR   | Repeat | Size (bp) | Location with HBB gene |
|-----|-------|--------|-----------|------------------------|
| 1   | HBB4506 | (AC) n | 366-398   | Downstream             |
| 2   | D11S988 | (TG) n | 103-147   | Downstream             |
| 3   | HBB4677 | (AC) n | 172-214   | Downstream             |
| 4   | D11S2362 | (AT) n | 87-123    | Downstream             |
| 5   | HBB5089 | (AC) n | 241-265   | Downstream             |
| 6   | D11S1243 | (TG) n | 220-256   | Downstream             |
| 7   | HBB5138 | (AC) n | 404-428   | Downstream             |
| 8   | HBB5178 | (AC) n | 158-192   | Downstream             |
| 9   | HBB5205 | (AGAT) n | 401-449   | Downstream             |
| 10  | D11S769 | (CA) n | 195-241   | Upstream               |
| 11  | HBB5576 | (AAGG) n | 327-369   | Upstream               |
| 12  | HBB5655 | (AC) n | 272-320   | Upstream               |
| 13  | HBB5620 | (AG) n | 311-331   | Upstream               |
| 14  | HBB5659 | (ATCT) n | 375-417   | Upstream               |
| 15  | D11S1338 | (AC) n | 137-157   | Upstream               |

Forward primer had an additional sequence for fluorescent primer M13 were shown in Table 3.

| STR | Primer | Sequence (5’-3’) | T<sub>m</sub> (°C) |
|-----|--------|------------------|------------------|
| HBB | B1F<sup>+</sup> | GTAAAAACGGCGCAGCTGTTGACCATATCTGTGGAA | 71.9 |
| 4506 | B1R | GTTTTCGAAGGAGAAGGCTTGACATACGCTGCTGAA | 62.9 |
| D11S988 | B2S<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 72.3 |
| B2R | CATACTG | 63.2 |
| HBB | B3F<sup>+</sup> | GTAAAAACGGCGCAGCTGTTGACCATATCTGTGGAA | 72.1 |
| 4677 | B3R | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 61.7 |
| D11S2362 | B4F<sup>+</sup> | GTAAAAACGGCGCAGCTGTTGACCATATCTGTGGAA | 73.2 |
| 2 | B4R | CATACTG | 60.4 |
| HBB | B5F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 71.4 |
| 5089 | B5R | CATACTG | 60.8 |
| D11S1243 | B6F<sup>+</sup> | GTAAAAACGGCGCAGCTGTTGACCATATCTGTGGAA | 73.3 |
| 3 | B6R | CATACTG | 59.9 |
| HBB | B7F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 71.1 |
| 5138 | B7R | CATACTG | 59.6 |
| HBB | B8F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 71.2 |
| 5178 | B8R | CATACTG | 59.7 |
| HBB | B9F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 73.3 |
| 5205 | B9R | CATACTG | 61.3 |
| D11S176 | B10F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 72.9 |
| 0 | B10R | CATACTG | 63.8 |
| HBB | B11F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 73.3 |
| 5576 | B11R | CATACTG | 59.9 |
| HBB | B12F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 71.4 |
| 5655 | B12R | CATACTG | 59.4 |
| HBB | B13F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 71.1 |
| 5820 | B13R | CATACTG | 63.3 |
| HBB | B14F<sup>+</sup> | GTTTTTCGACTGACCGGAGCAAGGAAGTTGAA | 72.5 |
| 5859 | B14R | CATACTG | 60.6 |
| D11S133 | B15F<sup>+</sup> | GTAAAAACGGCGCAGCTGTTGACCATATCTGTGGAA | 71.5 |
| 8 | B15R | CATACTG | 66.6 |

* Primer fluorescently labeled with HEX, ** Primer fluorescently labeled with FAM.

**Multiplex PCR Optimization**

Annealing temperature (T<sub>a</sub>) and primer concentration optimization were performed. The results of T<sub>a</sub> optimization of 15 primer were shown in Figure 1.

![Figure 1: Annealing temperature optimization of 15 STRs. Annealing temperature at 60°C gave clear bands and most uniform signal. The products were electrophoresed on 1% agarose gel, at 120V for 30 minutes. The 10 µl sample was each well](image-url)
The optimal annealing temperature showed clear bands and fair signals for all 15 primers (Figure 1). We finally concluded that 60°C was the optimal annealing temperature.

After adjusting primer concentrations (0.05-0.4 µM), optimal Multiplex PCR gave clear peaks. The PCR products were fluorescently labeled for capillary electrophoresis (Figure 2).

### Polymorphism evaluation

To determine the polymorphism and informativeness of the 15 markers, the allele frequencies, PIC, He, and Ho values of each marker were calculated. Most STRs sequences were successfully amplified in all individuals except D11S1760 (84/106) and HBB5655 (81/106) caused by PCR Failure.

In summary, 270 alleles were observed with 11-28 alleles per marker. Allele frequency ranged from 0.0047 to 0.3255 (Table 4). All markers had PIC values of ≥ 0.7 and among these markers, HBB5178 showed the lowest polymorphism (0.79), and D11S1760 showed the highest polymorphism (0.91).

The data showed that He values ranged from 0.81 (HBB5138) to 0.92 (D11S1760), and all markers had Ho values of ≥ 0.5, guaranteeing high polymorphic information. The Ho values of markers ranged from 0.53 (HBB5655)-0.86 (HBB5205). Thus, HBB5205 was the most informative marker, while HBB5655 was the least informative marker. Furthermore, the number of heterozygous markers of each individual was also counted.

Data showed 100% of individuals had at least seven heterozygous markers (Figure 4). Also, all were observed to be heterozygous for at least two markers on either side of the HBB gene (Figure 5). Based on the results, all 15 markers are high in polymorphism and informativeness for the Vietnamese population.

### Table 4: Distribution of observed allele frequencies of 15 microsatellite markers

| Markers | Allele Frequency | Allele Frequency | Allele Frequency | Allele Frequency | Allele Frequency | Allele Frequency | Allele Frequency |
|---------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| HBB5450 | 0.0094           | 0.0233           | 0.0197           | 0.0118           | 0.0223           | 0.0085           | 0.0238           |
| D11S995 | 0.047            | 0.0097           | 0.0197           | 0.0118           | 0.0223           | 0.0085           | 0.0238           |
| HBB477  | 0.0364           | 0.0263           | 0.0197           | 0.0118           | 0.0223           | 0.0085           | 0.0238           |
| HBB5859 | 0.047            | 0.0097           | 0.0197           | 0.0118           | 0.0223           | 0.0085           | 0.0238           |
| D11S1243| 0.047            | 0.0097           | 0.0197           | 0.0118           | 0.0223           | 0.0085           | 0.0238           |

### Table 5: Observed Heterozygosity, Expected Heterozygosity and Polymorphic Information Content values of 15 STRs

| Markers | Heterozygosity | Total | HET | PIC |
|---------|----------------|-------|-----|-----|
| HBB506  | 0.92           | 0.88  | 0.84 |
| HBB5186 | 0.86           | 0.84  | 0.82 |
| HBB477  | 0.86           | 0.84  | 0.82 |
| D11S262 | 0.86           | 0.84  | 0.82 |
| HBB509  | 0.70           | 0.68  | 0.64 |
| D11S1243| 0.70           | 0.68  | 0.64 |
| HBB5178 | 0.68           | 0.66  | 0.62 |
| D11S1760| 0.68           | 0.66  | 0.62 |

In general, the study had established a high polymorphic STR panel and relatively close with HBB gene. The panel should be potential to perform PGT for β-thalassemia.

Figure 2: Representative electropherogram of multiplex PCR product after optimization

![Figure 2](https://www.idpress.eu/mjms/index)
Discussion

For many years, despite medical advances, the treatment of a β-Thalassemia patient still confronts many obstacles because it is impossible to completely treat the disease. Along with the treatment of individuals with β-Thalassemia, the application of PGT-M to help at-risk couples of having healthy babies is an extensively studied field. STRs contain different polymorphism and informativeness, but all showed the high values of PIC, Ho, He proving their clinical application. Our method had high sensitivity and specificity along with sufficient contamination and ADO monitoring. Microsatellite markers became a reliable tool when using alone in indirect mutation detection or conjunction with direct mutation detection for more precise diagnosis. Besides, this method required highly trained staffs, accurate instrument, and optimal procedures. Therefore, a particular training program needs carrying out before performing this method.

In conclusion, All 15 STR markers were polymorphic and informative with high Ho, He and PIC. At least 7 of 15 markers were informative for 106 studied individuals, and all were observed to be heterozygous for at least two markers on either side of the HBB gene. Thus, these STRs marker have significant meaning when applied in PGT-M widely.

Authors' contributions

DTT, NVNM, DPN, HVL, DQ, TNA, TTS, NT, NTTH, DTPA, NLT, HTH, LH and NDB designed and performed experiments, and collected data and

---

Figure 4: Percentage of individuals for number of heterozygous markers

Figure 5: Percentage of individuals heterozygous for different numbers of upstream and downstream flanking microsatellite markers (n = 106)
informed consents. DTT, NVNM, DPN, HVL, DQ, TNA, TTS, NTT, NTTH, DTPA, NLT, HTH, LH, NDB, VTN and DTC analyzed and interpreted the results, and edited and corrected the manuscript. DTT, NVNM, VTN and DTC wrote the manuscript. All authors approved the final manuscript.

Ethical approval

This study is approved by the ethics committee of the Tam Anh General Hospital on 21 April 2018 following the Decision No 59/QĐ/BĐKTA by the director of the Tam Anh General Hospital about the establishment of the ethical committee.

Informed consent

The informed consents were signed by patients and their male partners.

References

1. Higgs DR, Engel JD, Stamatoyannopoulos G. Thalassaemia. The lancet. 2012; 379(9813):373-383. https://doi.org/10.1016/S0140-6736(11)60283-3
2. Giardine B, et al. Updates of the HbVar database of human hemoglobin variants and thalassemia mutations. Nucleic Acids Res. 2013; 42(D1):D1063-D1069. https://doi.org/10.1093/nar/gkt911 PMid:24137000 PMcid:PMC3964999
3. Khanh N. The frequency of hemoglobinopathies in Vietnam. Vietnamese J Med. 1993; 8(174):11-16.
4. Fucharoen S, Winichagoon P. Hemoglobinopathies in Southeast Asia: molecular biology and clinical medicine. Hemoglobin. 1997; 21(4):299-319. https://doi.org/10.3109/03630269709000664 PMid:9255610
5. Bowman JE, editor. Distribution and Evolution of Hemoglobin and Globin Loci: Proceedings of the Fourth Annual Comprehensive Sickle Cell Center Symposium on the Distribution and Evolution of Hemoglobin and Globin Loci at the University of Chicago, Chicago, Illinois, USA, October 10-12. 1982. Elsevier Publishing Company; 1983.
6. Yap C, et al. First successful preimplantation genetic diagnosis in Singapore-avoidance of beta-thalassaemia major. Ann Acad Med Singapore. 2009; 38(8):720-3.
7. Fan L, et al. Genetic diagnosis of beta-thalassemia preimplantation using short tandem repeats in human cryopreserved blastocytes. International Journal of Clinical and Experimental Pathology. 2017; 10(7):7586-7595.
8. Machado F, Medina-Acosta E. High-resolution combined linkage physical map of short tandem repeat loci on human chromosome band Xq28 for indirect haemophilia A carrier detection. Haemophilia. 2009; 15(1):297-308. https://doi.org/10.1111/j.1365-2516.2008.01866.x PMid:18752533
9. Chen M, et al. Identification of novel microsatellite markers< 1 Mb from the HBB gene and development of a single-tube pentadecaplex PCR panel of highly polymorphic markers for preimplantation genetic diagnosis of beta-thalassemia. Electrophoresis. 2015; 36(23):2914-2924. https://doi.org/10.1002/elps.201500148 PMid:26331357
10. Rechitsky S, et al. Reliability of preimplantation diagnosis for single gene disorders. Molecular Cellular Endocrinology. 2001; 183:S65-S68. https://doi.org/10.1016/S0303-7207(01)00576-7
11. Rechitsky S, et al. Preimplantation genetic diagnosis with HLA matching. Reproductive BioMedicine Online. 2004; 9(2):210-221. https://doi.org/10.1016/S1472-6483(10)62132-3
12. Fiorentino F, et al. Short tandem repeats haplotyping of the HLA region in preimplantation HLA matching. European Journal of Human Genetics. 2005; 13(8):953. https://doi.org/10.1038/sj.ejhg.5201435 PMid:15886713
13. Kuliev A, et al. Preimplantation diagnosis and HLA typing for haemoglobin disorders. Reproductive BioMedicine Online. 2005; 11(3):362-370. https://doi.org/10.1016/S1472-6483(10)60845-0
14. Fiorentino F, et al. Strategies and clinical outcome of 250 cycles of preimplantation genetic diagnosis for single gene disorders. Human Reproduction, . 2005; 21(3):670-684. https://doi.org/10.1093/humrep/dei382 PMid:16311287
15. Kuliev A, et al. Preimplantation genetic diagnosis for hemoglobinopathies. Hemoglobin. 2011; 35(5-6):547-555. https://doi.org/10.3109/03630269.2011.60845 PMid:21910603
16. Zachaki S, et al. Novel and known microsatellite markers within the β-globin cluster to support robust preimplantation genetic diagnosis of β-thalassemia and sickle cell syndromes. Hemoglobin. 2011; 35(1):56-66. https://doi.org/10.3109/03630269.2010.544620 PMid:21250882
17. Durmaz B, et al. Genotyping of β-Globin Gene Mutations in Single Lymphocytes: A Preliminary Study for Preimplantation Genetic Diagnosis of Monogenic Disorders. Hemoglobin. 2012; 36(3):230-243. https://doi.org/10.3109/03630269.2012.675881 PMid:22524255.