Rapid screening for sickle cell disease by polymerase chain reaction-high resolution melting analysis

LIANG YUE¹,²*, MIN LIN¹,²*, JIANG-TAO CHEN³,⁴, XIAO-FEN ZHAN¹, DE-SHANG ZHONG⁴,⁵, SANTIAGO-M. MONTE-NGUBA⁶, PEI-FEN LIU¹,²,⁵, XUE-FEN PAN³, JIANG-HUA HUANG¹, XI WANG⁴, JUAN CARLOS SALAS EHAPOB, URBANO MONSUY EYI⁶, HUI-TIAN YANG² and LI-YE YANG¹,²

¹Medical College of Shantou University, Shantou, Guangdong 515041; ²Laboratory Medical Center, Chaozhou Central Hospital, Chaozhou, Guangdong 521021; ³Laboratory Medical Center, Huizhou Central People’s Hospital, Huizhou, Guangdong 516001; ⁴The Twenty-Fifth Bath of Chinese Medical Aid Team to the Republic of Equatorial Guinea, Guangzhou, Guangdong 511000; ⁵Huizhou Chinese Traditional Hospital, Huizhou, Guangdong 516001, P.R. China; ⁶Hospital Regional De Malabo, Malabo 999115, The Republic of Equatorial Guinea

Received August 20, 2013; Accepted March 5, 2014

DOI: 10.3892/mmr.2014.2130

Abstract. Each year, ~300,000 individuals with sickle cell disease (SCD), a hemoglobinopathy caused by β-globin gene mutation, are born, and >75% of those are in Africa. The present study examined 511 individuals on the island of Bioko (Equatorial Guinea) and attempted to establish a method for rapid sickle cell disease screening. Following DNA extraction and polymerase chain reaction (PCR) amplification, high resolution melting (HRM) analysis was used to assess the specificity of fluorescence signals of the PCR products and to differentiate various genotypes of these products. The analytical results of HRM were validated using DNA sequencing. By HRM analysis, 80 out of 511 samples were classified as hemoglobin S (Hb S) heterozygotes, while 431 out of 511 samples were classified as wild-type. No mutant homozygote was identified. DNA sequencing indicated that within the 431 wild-type samples as indicated by HRM analysis, one case was actually a Hb S heterozygote and another case was a rare hemoglobin S-C genotype (sickle-hemoglobin C disease). One out of 80 suspected Hb S heterozygotes as indicated by HRM was confirmed as wild-type by DNA sequencing and the results of residual 508 cases were consistent for HRM analysis and sequencing. In conclusion, HRM analysis is a simple, high-efficiency approach for Hb S screening and is useful for early diagnosis of SCD and particularly suitable for application in the African area.

Introduction

Sickle cell disease (SCD) or drepanocytosis, one of the most common monogenetic inherited diseases worldwide, was first discovered in 1949 by Pauling et al (1), and this led to the hypothesis of hemoglobinopathies. This disease results from point mutations (HBB: c.20 A>T, rs334) that cause structural changes (glutamic acid to valine) in β-globin protein and leads to a functionally defective hemoglobin S. Cases of SCD of the homozygous genotype often result in stillbirth or the subject succumbing to the disease prior to adulthood, while heterozygous genotypes tend to present with numerous complications, including severe anemia, infectious complications, acute organ failure and chronic end-organ damage (2,3).

SCD is mainly prevalent in Sub-Saharan Africa where malaria is endemic. A recent study confirmed that patients with SCD of the heterozygotic type have increased survival rates in malaria-endemic regions (4). The heterozygote of the sickle gene provides protection against infection with Plasmodium falciparum and shows less severe symptoms when infected. Natural selection factors are the main reason for the high frequencies of the disease in tropical countries and in countries to which large populations have migrated from these regions. The World Health Organization reported that ~2% of newborns in Nigeria and Ghana were affected by sickle cell anemia, with a total of 150,000 affected children born each year in Nigeria alone (5,6). The frequency of SCD ranges from 10 to 40% across equatorial Africa, decreasing to 1-2% at the North African coast and <1% in South Africa (5).

Bioko, also known as Fernando Póo, formerly (1973-79) Macias Nguema Biyogo, is an island in the Gulf of Guinea and is located ~100 km off the coast of Southern Nigeria and 160 km northwest of continental Equatorial Guinea, West Africa. The island has a population of 260,000 (2001) (7). The mean annual rainfall is ~2,000 mm/yr. Findings of previous epidemiological

Correspondence to: Dr Li-Ye Yang, Medical College of Shantou University, 22 Xinling Road, Shantou, Guangdong 515041, P.R. China
E-mail: yangleeyee@sina.com

*Contributed equally

Key words: sickle cell anemia, gene mutation, high-resolution melting, newborn screening, Bioko island
optimal sample size \( n \) was estimated. The overall incidence rate of SCD \( \pi \) was roughly estimated as 50% (as this results in the largest sample size) and the allowed error extent \( \delta \) was estimated as ±10%, while the statistic value of \( u_{\alpha/2} \) is 1.96, >95% confidence limits. For simple random sampling, the sample size formula used was:

\[
n \geq \frac{u^2}{\alpha/2} \times \pi(1 - \pi) \delta^2 = 96
\]

Thus, >96 patients carrying the Hb S gene were required. In regard to the SCD prevalence of 20% in Bioko in previous screening, ≥480 subjects (100/20x96) were required. In total, 511 samples were selected for the present study.

**Subjects and DNA preparation.** Of all 2,000 dried peripheral blood spots from residents in Bioko, 511 subjects (aged 1-80 years), consisting of 232 males and 279 females were randomly selected for the present study. The present study was approved by the Ethics Committee of the De Malabo Regional Hospital (Malabo, Equatorial Guinea) and Chaozhou Central Hospital (Chanzhou, China). Blood samples were obtained using the dried-blood spot method and extraction was conducted by Chelex-100, a chelating resin (10). Extracted DNA samples were frozen at -20°C until required.

**Primer design.** The nucleotide sequences of the β-globin gene were retrieved from the PubMed database (www.ncbi.nlm.nih.gov/pubmed). The HRM and sequencing primers were designed with the Primer 3 software (http://bioinfo.ut.ee/primer3-0.4.0/) and additionally checked for structures including primer dimers and hairpins using the oligo 7.37 primer design software Molecular Biology Insights Inc. (Cascade, CO, USA; Table I). In the HRM assay, polymerase chain reaction (PCR)-amplified fragments were 80 bp long (located in exon 1). Sequencing primers were designed to amplify the sequence of the human β-globin gene (GenBank: NC_000007.3) ranged from exon 1 to intron 2. The length of the amplicon was 957 bp (Fig. 1).

**HRM conditions and melting curve analysis.** Fluorescence quantitative PCR (qPCR) in 96-well plates was implemented with the thermal cycler LightCycler®480II (Roche Diagnostics), amplified in duplicate in a total volume of 20 µl containing 50 ng of genomic DNA, 1 µM of each primer (10 µM/l), 100 µM of each deoxyribonucleotide triphosphate (dNTP), 4 µl of 5X PCR buffer, 0.2 units hotstart polymerase (Takara, Dalian, China) and 1 µl LC Greens PLUS DNA binding dye (Idaho Technology, Salt Lake City, UT, USA). DNA with one known heterozygous control, one known wild-type control and one blank control were included in each assay. The following cycle program was used: Denaturation at 94°C for 3 min, followed by 35 cycles of 98°C for 10 sec, 68°C for 20 sec, a final cycle at 95°C for 1 min, and 40°C for 1 min to achieve heteroduplex formation. The fluorescence data were collected from 65 to 95°C with temperature increases of 0.05°C/sec for amplicon scanning. HRM analysis was performed using the software LightCycler 480 SW 1.5 (Roche Diagnostics GmbH, Mannheim, Germany). The ‘auto-group’ function was applied to generate automatic groups for the amplicons. Normalized melting curves showed the fluorescence signal against the temperature, and derivative plots showed the melting temperature (Tm) peaks. The samples with known genotypes, which were validated by DNA sequencing, were used as standard genotype references. When the plots of samples were classified in correlation with the standard genotype references, they were identified as the identical genotype of the standard. Furthermore, each DNA sample was amplified twice to validate the coherence of curve shapes and melting peaks.

**DNA fragment amplification and sequencing.** To validate the accuracy of HRM analysis, sequencing was performed in all the samples. The PCR reaction was performed in a reaction volume of 50 µl containing 100 ng of genomic DNA, 0.2 µM each of the primer, 200 µM of dNTP, 25 µl of 10X buffer I and 0.5 units polymerase enzyme (Takara). The reaction was performed in a MJ Mini Personal Thermal Cycler. The cycle program was as follows: Denaturation at 94°C for 1 min, 60°C for 45 sec, and extension at 72°C for 1 min with 35 cycles, followed by 10 min extension at 72°C. Subsequently, 10 µl of the PCR products was fractionated on a 2% agarose gel following staining with ethidium bromide to confirm product integrity. PCR products were then bi-dimensionally sequenced by an ABI 3730xL DNA Sequencer (PE Biosystems, Norwalk, CT, USA).

**Statistical analysis.** Screening abilities of HRM and the consistency of the analysis were assessed by classical metrics, including sensitivity, specificity, likelihood ratio and the Youden index. Additionally, statistical analysis was performed using SPSS 16.0 software (SPSS, Inc., Chicago, IL). For all statistical tests, a difference between values with a P-value <0.05 was considered statistically significant.
Results

Hb S genotype scanning and SCD incidence in Bioko. In a blind study, the graphs of 511 selected samples were divided into two groups by HRM analysis (Fig. 2), curve shapes of 80 samples were identified as heterozygotes based on reference DNA, while the remaining 431 samples were regarded as wild-type. Following sequencing of all samples, it was confirmed that, among 431 samples with indication of being wild-type by HRM analysis, one case was a Hb S heterozygote and another case was a hemoglobin S-C genotype (two false-negative results), and one out of 80 Hb S heterozygotes as indicated by HRM analysis was identified as wild-type (one false-positive result). The results of the residual 508 cases were consistent for HRM and sequencing (Fig. 3).

In the analysis, no mutant homozygotes were identified. Two potential factors were taken into account: a) The principle of natural selection (stillbirth and early death resulting from severe clinical symptoms) restricted the number of homozygote SS carriers in the population. b) For the reason of finite samples and convenience sampling, surviving patients with hemoglobin genotype SS were not covered.

The results of the HRM detection were accurate except for the two false-negative results and one false-positive result. According to the actual sequencing, 15.9% (81/511) of all specimens were Hb S carriers (including one case of sickle-hemoglobin C disease), and the occurrence rate of SCD obtained in the present study was lower than that of a previous study in Equatorial Guinea (8).

Assessment of methodology. The validity and consistency of the methodology were assessed by the method of likelihood ratio and the $\kappa$ test. Considering DNA sequencing as the accepted golden standard, the performance of the HRM method was compared to the standard reference using the paired 4-fold table (Table II).

In contrast to DNA sequencing, the accuracy of the HRM method was 99.6%, the test sensitivity (Se) was 97.5%
Figure 2. Genotypical differentiation by HRM profiling. Based on the distinction of relative fluorescence signals and melting temperatures, melting curves of heterozygous and wild-type samples were able to be grouped (A, B and C). The melting curve of Hb S-C was not able to not be differentiated from the wild-type (A). HRM, high resolution melting; Hb, hemoglobin.

Figure 3. DNA sequence of PCR products. The arrows show wild-types (A), heterozygous mutation of Hb S (B), SNP (rs713040) and hemoglobin S-C mutational site (C), respectively. PCR, polymerase chain reaction; HB, hemoglobin; SNP, single-nucleotide polymorphism.
[95% confidence interval (CI), 0.914-0.993], the test specificity (Sp) was 99.7% (95% CI, 0.987-0.999), the positive and negative predictive value were 98.8 and 99.5%, respectively, the positive likelihood ratio LR+ was 419.4 (95% CI; 59.192-2971.4), while the negative likelihood ratio LR- was 0.025 (95% CI, 0.006-0.097). Furthermore, the Youden index was 0.973 and the Kappa index was 0.978 (p<0.01). As a rule of thumb, LR+ >10 or LR- <0.1 generated significant conclusions (11), and in regard to the selected method, the greater the LR+ and the lower the LR-, the higher was the power of detection. In a consistency check, the Kappa index ranging from 0.75 to 1 indicated a high degree of correspondence between the two diagnostic tests. Similarly, a higher value of Se or Sp (which is close to 1) tends to produce more objective results. In general, HRM analysis generated results consistent with direct sequencing with good discrimination and reliability for the aim of genotyping.

Analysis of sequencing results. The single-nucleotide polymorphism (SNP) HBB: c.9 C>T, rs713040 is a bi-allele SNP, which is located 11 bp in the 5′ end direction from the HB S mutational site. In the fluorescence qPCR, an impurity peak originating from an unexpected SNP interfered with the interpretation of the mutational curve. Therefore, primer design was optimized, and the disturbance of the SNP (rs713040) was eliminated. In addition, the allele frequency of rs713040 was assessed by DNA sequencing and the frequency for the three genotypes (CC, CT, TT) in 511 samples is 0.73 (372/511), 0.24 (125/511) and 0.03 (14/511), respectively. Furthermore, a rare sample of hemoglobin S mixed with hemoglobin C (HBB: c.19 G>A, rs33930165) was identified by sequencing.

Discussion

The Chelex-100 method for DNA isolation appears well-suited for fast DNA preparation and long-term storage. In the assay of the present study, genomic DNA was isolated using the Chelex-100 method and filter paper due to three reasons: a) Chelex-100 has already been successfully used to recover DNA from the following: Forensic specimens, formalin-fixed and paraffin wax-embedded tissue, whole blood, dried blood spots on filter paper and clinical samples. The method has a wide scope of applications and meets the requirements of a prenatal screening test to extract DNA from multifarious samples, including peripheral whole blood, chorion cells and amniotic fluid; b) although conventional commercial kits are superior to the Chelex-100 method in regard to obtaining highly purified DNA, HRM analysis with high sensitivity and DNA extraction by Chelex-100 is still sufficient for PCR amplification, while this method is more cost- and time-efficient ($0.25 per sample and 3 min per sample); c) the entire procedure of the Chelex-100 method is easily performed in a single container, reducing the chances of sample contamination and loss of sample; d) recovery of DNA from filter paper impregnated with whole blood is a useful method of preparing DNA for PCR-based genetic testing, and dried blood spots on filter paper are able to be easily transported and stored; DNA on filter paper remains relatively stable at room temperature for ≥4.5 months (10).

Primer design is essential for a sensitive, specific and stable HRM analysis. The basic principle of HRM analysis is that samples with mutations or polymorphisms are able to form heteroduplexes due to sequence variations and generate abnormal melting curves compared with normal controls. Thus, the following items should be considered: a) Relatively short PCR amplicons (50-150 bp) are preferred, since the effect of single-base mutation on the melting behavior of short amplicons in fluorescence qPCR is more significant than that of long amplicons; b) melting curves of unintended SNPs in a target fragment are able to be simultaneously detected by HRM; however, for the purpose of single-base mutation screening, multiple melting peaks may become too complex to be distinguished and should be avoided; c) the presence of unexpected SNPs located in the primers may block the proceeding of the PCR amplification. In the presented assay, a primer sequence was designed that overlaid this SNP by locating it near the 5′ end of the primer in order to prevent the interference of the SNP (rs713040) near the targeted mutation. As base alterations in the 5′ region of the primer hardly influence the amplification efficiency, the accuracy and specificity of the trial were assured.

The melting curves for the wild-type and mutant hemoglobins were almost indistinguishable. Theoretically, adding one equivalent of reference DNA (wild-type) into homozygous samples produces heteroduplex DNA (heterozygotes), so that the curve profiles of the former and latter are able to be distinguished. This approach has been successfully verified in a previous study by our group (12).

At present, except for HRM analysis, various techniques, including the amplification restriction mutation system (ARMS) and restriction fragment length polymorphisms (RFLP) PCR as well as direct DNA sequencing, are widely applied for screening and diagnosing Hb variants. Conventional methods for large-scale detection of Hb S gene mutations have disadvantages in terms of simplicity, performance, sensitivity, turnaround time and cost. For instance, RFLP assays are time-consuming and are able to lead to misdiagnosis in the case of a supplementary SNP on the restriction sequence (13). Moreover, RFLP and ARMS assays are required for confirmation testing (14). Although the DNA sequencing technique is widely recognized as the gold standard for screening genomic variations, its widespread use is hampered by its cost- and time-consuming nature (15).

HRM is an inexpensive, high-throughput and closed-tube genetic scanning technique. On average, it is completed within 1 h (including PCR amplification and HRM analysis) and the cost is only $0.25 per sample in the present study. It is also amendable to high-throughput mass screening in 96- or 384-well plates. For accurate genotypical analysis, direct sequencing, reverse dot-blot or allele-specific PCR methods are options, and unlabeled probe melting analysis was also developed on the same platform (16). Furthermore, close-tube operation eliminates the occurrence of cross-contamination. DNA quality and complex genotypes (including hemoglobin S mixed with hemoglobin C) may interfere with HRM analysis. As for certain samples, results of HRM and sequencing were inconsistent and were repeated under non-modified HRM conditions. The curves for these samples were still incorrectly grouped and DNA template purity
coupled with concentration may have accounted for the incorrect grouping. Additionally, hemoglobin C disease, resulting from substitution at the first position of HBB codon six, is widely distributed in West Africa (17). HRM analysis showed that the curve shape of the hemoglobin S-C genotype was broadly similar to that of the wild-type and was difficult to distinguish. Therefore, this technology was not able to differentiate hemoglobin S-C genotype from the wild-type.

In routine clinical practice, the demand for rapid and reliable detection and identification of Hb variants is increasing in order to analyse clinical samples and apply targeted therapies or counseling. In most genetic conditions, prenatal diagnosis with SCD is possible in the first trimester through chorionic villus sampling or in the second trimester through amniocentesis. Following PCR amplification and HRM analysis, the DNA samples of the heterozygous or homozygous genotype Hb S are detected. Based on the results, genetic counseling and prenatal diagnosis would then be offered to the couples. Additionally, the HRM method may also be a suitable tool for mass SCD screening at the molecular level, particularly in the African areas currently lacking appropriate means for neonatal screening and early diagnosis (18). Besides, the HRM method, which is used for Hb S screening, may also be applied in malaria-associated studies.

For the purpose of large-scale SCD screening and early diagnosis, there may be a broad prospect in the application of the HRM method developed in the present study.

Acknowledgements

The present study was supported by the National Natural Science Foundation of China (no. 81101329 to Min Lin), Social Development Program of Guangdong Province (no. 2011B031800329 to Min Lin). The authors thank the Department of Health of Guangdong Province and Department of Aid to Foreign Countries of Ministry of Commerce of People's Republic of China for their help.

References