Molecular assay of $-\alpha^{3.7}$ and $-\alpha^{4.2}$ deletions causing α -thalassemia by denaturing high-performance liquid chromatography

Chia-Cheng Hung^{a,1}, Chien-Nan Lee^{b,1}, Chih-Ping Chen^{c,d,e,f}, Yuh-Jyh Jong^{g,h},
Wu-Shiun Hsiehⁱ, Win-Li Lin^a, Yi-Ning Su^{j,k,*}, Su-Ming Hsu^l

^a Institute of Biomedical Engineering, College of Medicine and College of Engineering, National Taiwan University, Taipei, Taiwan

^b Department of Obstetrics and Gynaecology, National Taiwan University Hospital, Taipei, Taiwan

^c Department of Obstetrics and Gynaecology, Mackay Memorial Hospital, Taipei, Taiwan

^d Department of Medical Research, Mackay Memorial Hospital, Taipei, Taiwan

^e Department of Biotechnology and Bioinformatics, Asia University, Taichung, Taiwan

^f College of Chinese Medicine, China Medical University, Taichung, Taiwan

^g Department of Paediatrics, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan

^h Graduate Institute of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

ⁱ Departments of Paediatrics, National Taiwan University Hospital, Taipei, Taiwan

^j Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan

^k Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan

^l Department of Pathology, National Taiwan University Hospital, Taipei, Taiwan

Received 12 January 2007; received in revised form 7 March 2007; accepted 9 March 2007

Available online 6 April 2007

Abstract

Objectives: α -Thalassemia, the most common single gene disorder in humans, is due to the absence of one ($-\alpha/\alpha$) or both ($--/\alpha$) of the two functional α -globin genes ($\alpha 1$ and $\alpha 2$). The $-\alpha^{3.7}$ and $-\alpha^{4.2}$ single gene deletions are common in Southeast Asian populations. Southern blotting analysis and gap PCR assay are commonly used for the detection of such α -thalassemia genotypes. The two genes are located on chromosome 16, with high homology (>96%).

Design and methods: Based on the sequence variation within the two Z boxes, a denaturing high-performance liquid chromatography (DHPLC)-based assay was developed for rapid genotyping of the $-\alpha^{3.7}$ and $-\alpha^{4.2}$ alleles. To demonstrate the utility of this approach, 40 DNA samples with known genotypes were analyzed, including $-\alpha^{3.7}/\alpha\alpha$ (7 cases), $-\alpha^{4.2}/\alpha\alpha$ (4 cases), $-\alpha^{3.7}/--^{SEA}$ (6 cases), $-\alpha^{4.2}/--^{SEA}$ (3 cases), and 20 unaffected subjects ($\alpha\alpha/\alpha\alpha$).

Results: We successfully distinguished all of the α -thalassemia genotypes through their characteristic chromatograms of $\alpha 1$ and $\alpha 2$ genes. The accuracy of this technique for our sample was 100% sensitivity and specificity.

Conclusion: This novel and alternative DHPLC-based α -thalassemia genotype assay is easy, rapid, and highly accurate. This technique enables the diagnosis of silent α^+ thalassemia and hemoglobin H disease for large scale population screening.

© 2007 The Canadian Society of Clinical Chemists. Published by Elsevier Inc. All rights reserved.

Keywords: Denaturing high-performance liquid chromatography; α -Thalassemia; $-\alpha^{3.7}$; $-\alpha^{4.2}$

Introduction

α -Thalassemia is the most common recessively inherited hemoglobin disorder [1]. Unlike β -thalassemia, the most common causes of α -thalassemia are deletions of one ($-\alpha/\alpha$) or both ($--/\alpha$) of the two duplicated α -globin genes ($\alpha 1$ and $\alpha 2$) located on chromosome 16 (HBA1-OMIM #141800, GenBank #NM_000558.3; HBA2-OMIM #141850, GenBank

* Corresponding author. Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, No. 7, Chung-Shan South Road, Taipei, Taiwan. Fax: +886 2 2392 0470.

E-mail address: ynsu@ntu.edu.tw (Y.-N. Su).

¹ Chia-Cheng Hung and Chien-Nan Lee contributed equally to this study.

#NM_000517.3) [1, 2]. The two α -globin genes are highly homologous with sequence identity greater than 96% between the two Z1 and Z2 boxes with high GC content [3]. In Southeast Asia, the Southeast Asian ($-\alpha^{SEA}$), Filipino ($-\alpha^{FIL}$) and Thai ($-\alpha^{THAI}$) deletions are the most prevalent two-gene *cis* deletions in the population [4,5]. The other common deletions are the $-\alpha^{3.7}$ and $-\alpha^{4.2}$ single gene deletions [6]. The loss of three of the four functional α -globin genes results in compound heterozygotes of the ($---$) haplotype and the ($-\alpha$) haplotype, which comprise hemoglobin H (HbH) disease [7]. In Southeast Asia, different populations exhibit various single gene or two-gene deletion frequencies of α -globin locus. The population-based study in Thailand showed that 4.3% of populations were carriers of $---/\alpha\alpha$ genotype, and 12.0% of $-\alpha/\alpha\alpha$ genotype [8]. The carrier rates for Malays were 2.5% with $---/\alpha\alpha$ genotype, and 11.7% with $-\alpha/\alpha\alpha$ genotype [9]. The Chinese in Hong Kong, 4.5% were carriers with $---/\alpha\alpha$ genotype, and 0.5% of populations were carriers with $-\alpha/\alpha\alpha$ genotype [10]. Similar findings were reported in the Taiwanese population, 6.1% were at risk for $---/\alpha\alpha$ genotype, and 4.5% at risk for $-\alpha/\alpha\alpha$ genotype [11].

Traditionally, DNA analyses are carried out in α -thalassemias using Southern blot analysis [12], but this method is time-consuming, labor-intensive and is incompatible with high-throughput diagnostic strategies. For the majority of α -thalassemias in the world, gap PCR-based assays are more sensitive and more rapid and suitable for routine use. Therefore, a universal PCR approach was developed, based on the principles of multiplex PCR assays, which allows rapid detection of common deletions [13–15]. The analysis approach by denaturing high-performance liquid chromatography (DHPLC) is a relatively new method optimized for the screening of single base mutations, small deletions, and insertions that separate heteroduplex and homoduplex molecules on a stationary phase under partially denaturing conditions [16], and it is suitable for high homology genes [17,18].

The $\alpha 1$ and $\alpha 2$ genes contain two duplicated boxes (Z1 and Z2) [19]. Unequal crossing over between the homology generates $-\alpha^{3.7}$ single gene deletions. In the Z boxes, there are 26 nucleotide differences. Moreover, the 7-bp insertion/deletion within the second intron of $\alpha 1/\alpha 2$ branch migration inhibition [20]. Based on the three single nucleotide variations within the two Z boxes, a DHPLC-based approach was developed for rapid genotyping of α -thalassemia. We describe the relatively simple and semi-automatable technique of DHPLC for the molecular identification of α -globin gene with $-\alpha^{3.7}/\alpha\alpha$, $-\alpha^{4.2}/\alpha\alpha$, $-\alpha^{3.7}/---^{SEA}$, $-\alpha^{4.2}/---^{SEA}$, and $\alpha\alpha/\alpha\alpha$ genotypes in a single tube assay.

Material and methods

Subjects and DNA extraction

All study subjects were patients at the National Taiwan University Hospital. A total of 40 DNA samples with known genotype were analyzed in the study, including $-\alpha^{3.7}/\alpha\alpha$ (7

cases), $-\alpha^{4.2}/\alpha\alpha$ (4 cases), $-\alpha^{3.7}/---^{SEA}$ (6 cases), $-\alpha^{4.2}/---^{SEA}$ (3 cases), and 20 subjects unaffected by α -thalassemia ($\alpha\alpha/\alpha\alpha$). DNA was extracted from peripheral whole blood using the Chemagic DNA Blood Kit (Chemagen), according to the manufacturer's instructions.

Primer design and PCR conditions

PCR primers were designed using the GenBank reference sequence (NG_000006s) for the α -globin gene to amplify an amplicon (353 bp). A single PCR reaction was used to amplify both $\alpha 1$ and $\alpha 2$ Z boxes that different in three nucleotide changes. Primer sequences were 5'-GGT AGA GGA GTC TGA ATC TGG-3' (forward) and 5'-CTA GCC AGC GAG AAA GTC A-3' (reverse). PCR was performed in a final volume of 25 μ L containing 100 ng of genomic DNA, 100 μ M of each dNTP, 0.12 μ M of each primers, 0.5 U of AmpliTaq Gold enzyme (PE Applied Biosystems), 2 mM $MgCl_2$ and 2.5 μ L of 10 $\times$ buffer II as provided by the manufacturer. PCR was performed in a multiblock system (MBS) Satellite thermal cycler (ThermoHybaid) with an initial denaturation step at 95 $^{\circ}$ C for 10 min, followed by 35 cycles consisting of denaturation at 94 $^{\circ}$ C for 30 s, annealing at 59 $^{\circ}$ C for 45 s, elongation at 72 $^{\circ}$ C for 45 s, and then a final extension step at 72 $^{\circ}$ C for 10 min.

Gap PCR

The gap PCR protocol was modified as reported in a previous study [21]. Multiplex gap PCR was performed in a final volume of 25 μ L containing 100 ng of genomic DNA, 400 μ M of each dNTP, 0.6 μ M of each primers (A, B, D and E), 8% DMSO, 1.5 U of Expand Long Template DNA polymerase (Roche), and 2.5 μ L of 10 $\times$ buffer-1 with 1.75 mM $MgCl_2$ as provided by the manufacturer. Gap PCR was performed in a MBS Satellite thermal cycler (Thermo) with an initial denaturation step at 95 $^{\circ}$ C for 2 min, followed by 10 cycles consisting of denaturation at 94 $^{\circ}$ C for 40 s, annealing at 59 $^{\circ}$ C for 60 s, elongation at 68 $^{\circ}$ C for 6 min, followed by another 25 cycles where the extension time was increased by 20 s after each cycle, and then a final extension step at 72 $^{\circ}$ C for 10 min.

DHPLC assay

Heteroduplex analyses were performed according to the protocol of the manufacturer and previous studies [16,18]. PCR amplicons which amplified both $\alpha 1$ and $\alpha 2$ Z boxes were run under partially denaturing conditions in the Transgenomic Wave Nucleic Acid Fragment Analysis System (Transgenomic Inc.). PCR products were analyzed in a linear acetonitrile gradient with triethylammonium acetate as the mobile phase, using buffer A (0.1 M TEAA) and buffer B (0.1 M TEAA with 25% acetonitrile) (Transgenomic). Buffer concentrations at the start of the program were 55% B, with a gradient over the 6.5-min run time to 68% B, and the flow rate was 0.9 mL/min. Run temperature was constant at 64 $^{\circ}$ C,

and eluted DNA fragments were detected at a wavelength of 260 nm.

Results

Based on the sequence information, we used three nucleotide variants located in the Z boxes, which allowed the $\alpha 1$ gene to be distinguished from the $\alpha 2$ gene in this assay (Fig. 1). The boldface red letters shows that $\alpha 1$ gene has ATC-pair sequence and $\alpha 2$ has GCT-pair sequence. It would be helpful for distinguish two alleles by DHPLC.

The partial denaturing conditions of DHPLC were used to detect heteroduplex formation by a change in peak retention pattern. The four species shown in Fig. 2a are two heteroduplex and two homoduplex. The heteroduplex formation of the PCR amplicons is because $\alpha 1$ and $\alpha 2$ are highly similar which can be expected [16]. These stutter peaks (heteroduplex) were eluted much earlier in the chromatographic process than the main peaks (homoduplex) so that the main peaks were eluted unambiguously. According to the ranking of stabilities as described previous [16], we would expect the homoduplex with the ATC-pair ($\alpha 1$) had less retention time than the GCT-pair ($\alpha 2$). Using DHPLC analysis, we unambiguously identified the $\alpha 1$ and $\alpha 2$ peaks at 64 °C, as shown in Fig. 2. The peak patterns were related to the formation of heteroduplexes and homoduplexes. Samples were eluted from the column with peaks at 5.22 min for the $\alpha 1$ -allele and at 5.37 min for the $\alpha 2$ -allele samples with clear distinction. Our methodology is not only a qualitative but also a quantitative comparison between the two homoduplexes in order to differentiate genotypes.

Samples from wild type, $-\alpha^{3.7}$ and $-\alpha^{4.2}$ individuals can all be detected and identified with 100% reproducibility. Furthermore, the difference in DHPLC pattern between combined heterozygous $-\alpha^{3.7}/-SEA$ and $-\alpha^{4.2}/-SEA$ individuals could also be determined (Fig. 2). These results were proved by

Fig. 1. The α -globin gene structure and the sequence variants in the Z boxes. The primers are indicated in boldface blue letters, the sequence variants are marked in boldface red letters. The star means the nucleotide sequence in the $\alpha 2$ gene. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. Molecular identification of specific α -globin gene genotypes detected by DHPLC-based assay of (a) $\alpha\alpha/\alpha\alpha$, (b) $-\alpha^{3.7}/\alpha\alpha$, (c) $-\alpha^{4.2}/\alpha\alpha$, (d) $-\alpha^{3.7}/-SEA$, and (e) $-\alpha^{4.2}/-SEA$.

multiplex gap PCR (Fig. 3). The results demonstrated that all the major genotypes ($-\alpha^{3.7}/\alpha\alpha$, $-\alpha^{4.2}/\alpha\alpha$, $-\alpha^{3.7}/-SEA$, $-\alpha^{4.2}/-SEA$, and $\alpha\alpha/\alpha\alpha$) were easily distinguishable based on their respective DHPLC profiles. Nonetheless, because of the possibility of variation of overall elution times with different columns and buffers, internal controls of known genotype must be run with each batch of samples.

Discussion

Currently, there are numerous reports describing the use of gap PCR as the gold standard methodology for detection of the α -thalassemia genotypes [13–15,21]. Previously, DHPLC was reported as a rapid, accurate genotyping platform for the detection of α -thalassemia; however, it can only distinguish the three major genotypes ($-\alpha^{4.2}/\alpha\alpha$, $-\alpha^{4.2}/-SEA$ and $\alpha\alpha/\alpha\alpha$)

Fig. 3. Molecular identification of specific α -globin gene genotypes detected by multiplex gap PCR: (a) $\alpha\alpha/\alpha\alpha$, (b) $-\alpha^{3.7}/\alpha\alpha$, (c) $-\alpha^{4.2}/\alpha\alpha$, (d) $-\alpha^{3.7}/-SEA$, and (e) $-\alpha^{4.2}/-SEA$. The DNA fragments were separated by 1.0% agarose gel electrophoresis followed by staining with ethidium bromide for 45 min.

[22]. Our assay uses DHPLC to separate PCR amplicons relying on the same primer pair, enabling each sample to be genotyped within a single reaction more easily. The results demonstrated that the technique can identify all the major genotypes ($-\alpha^{3.7}/\alpha\alpha$, $-\alpha^{4.2}/\alpha\alpha$, $-\alpha^{3.7}/\text{---SEA}$, $-\alpha^{4.2}/\text{---SEA}$, and $\alpha\alpha/\alpha\alpha$).

Our method had 100% sensitivity and selectivity when compared with the gap PCR method for the 40 samples that we analyzed. The method was proved to be suitable for the genotype of $-\alpha^{3.7}/\text{---SEA}$ and $-\alpha^{4.2}/\text{---SEA}$ carriers; at a certain threshold, the gap PCR-based method might be more efficient. With our method, only a single PCR reaction with one primer pair was needed and the further analysis by DHPLC takes only about 10 min. Over one hundred samples can be processed overnight automatically. This throughput enables this technique to be adapted for large-scale population screening. Our rapid and low-cost assay requires no labeling for genotyping of the α -globin allele. In addition, the PCR amplicon region in our design is less GC-rich, thus, we did not need to use dimethyl sulphoxide (DMSO) in our PCR assay, making this approach highly robust, reproducible and likely to be suitable for DHPLC assay in routine diagnostic and screening use.

However, the $\alpha\alpha/\alpha\alpha$ would be expected to be identical DHPLC pattern to that compound heterozygous $-\alpha^{3.7}/-\alpha^{4.2}$ genotype. The hematological indices of mean cell hemoglobin (MCH) and mean cell volume (MCV) values are helpful in the molecular diagnosis. When $\text{MCV} > 80$ fL or $\text{MCH} > 25$ pg, we clarify as normal individual. It is noted that the $-\alpha^{3.7}/-\alpha^{3.7}$ genotype will have the same DHPLC pattern as the $-\alpha^{3.7}/\text{---}$ genotype. Likewise, the $-\alpha^{4.2}/-\alpha^{4.2}$ and $-\alpha^{4.2}/\text{---}$ genotypes cannot be distinguished from each other. We combined use of parent's genotype if available (the parental genotypes are already known determined by another more definitive method) and patient's phenotype to resolve the problem in such situation. Other genotypes such as the anti-3.7/ $\alpha\alpha$ or anti-4.2/ $\alpha\alpha$ genotypes might produce $\alpha 1$ and $\alpha 2$ peaks with ratios that are very similar to the $-\alpha^{4.2}/\alpha\alpha$ and $-\alpha^{3.7}/\alpha\alpha$ genotypes, respectively (3:2 vs. 2:1). By our method, it is difficult to distinguish one genotype from another. To resolve the problem, we will further combine use of the multiplex PCR and capillary electrophoresis system for quantitative the deletion or duplication genotype of α -thalassemia [23,24].

Our methodology offers a practical alternative for rapid detection of $-\alpha^{3.7}/\alpha\alpha$, $-\alpha^{4.2}/\alpha\alpha$, $-\alpha^{3.7}/\text{---SEA}$ and $-\alpha^{4.2}/\text{---SEA}$ genotypes causing α -thalassemia. The new techniques are highly accurate, rapid, inexpensive, automatable, and easily interpretable. This method could be useful in predicting the clinical phenotype of thalassemia intermedia and feasible for large-scale population screening.

Acknowledgments

The authors thank Dr. Fon-Jou Hsieh for his expert assistance. This work was supported by grants from the National Science Council of Taiwan (NSC 94-3114-P-002-Y5 and NSC 95-2314-B-002-034).

References

- [1] Higgs DR, Vickers MA, Wilkie AO, Pretorius IM, Jarman AP, Weatherall DJ. A review of the molecular genetics of the human alpha-globin gene cluster. *Blood* 1989;73(5):1081–104.
- [2] Kazazian Jr HH. The thalassemia syndromes: molecular basis and prenatal diagnosis in 1990. *Semin Hematol* 1990;27(3):209–28.
- [3] Lacerra G, Fiorito M, Musollin G, Di Noce F, Esposito M, Nigro V, et al. Sequence variations of the alpha-globin genes: scanning of high CG content genes with DHPLC and DG-DGGE. *Hum Mutat* 2004;24(4):338–49.
- [4] Eng B, Patterson M, Borys S, Chui DH, Wayne JS. PCR-based diagnosis of the Filipino (-(FIL)) and Thai (-(THAI)) alpha-thalassemia-1 deletions. *Am J Hematol* 2000;63(1):54–6.
- [5] Chui DH, Wayne JS. Hydrops fetalis caused by alpha-thalassemia: an emerging health care problem. *Blood* 1998;91(7):2213–22.
- [6] Chen FE, Ooi C, Ha SY, Cheung BM, Todd D, Liang R, et al. Genetic and clinical features of hemoglobin H disease in Chinese patients. *N Engl J Med* 2000;343(8):544–50.
- [7] Chui DH, Fucharoen S, Chan V. Hemoglobin H disease: not necessarily a benign disorder. *Blood* 2003;101(3):791–800.
- [8] Sriroongrueng W, Pornpatkul M, Panich V, Fucharoen S. Alpha-thalassemia incidence in southern Thailand by restriction endonuclease analysis of globin DNA from placental blood at Songklanagarind Hospital. *Southeast Asian J Trop Med Public Health* 1997;28(Suppl 3):93–6.
- [9] Wee YC, Tan KL, Chow TW, Yap SF, Tan JA. Heterogeneity in alpha-thalassemia interactions in Malays, Chinese and Indians in Malaysia. *J Obstet Gynaecol Res* 2005;31(6):540–6.
- [10] Lau YL, Chan LC, Chan YY, Ha SY, Yeung CY, Wayne JS, et al. Prevalence and genotypes of alpha- and beta-thalassemia carriers in Hong Kong—implications for population screening. *N Engl J Med* 1997;336(18):1298–301.
- [11] Ko TM, Chen TA, Hsieh MI, Tseng LH, Hsieh FJ, Chuang SM, et al. Alpha-thalassemia in the four major aboriginal groups in Taiwan. *Hum Genet* 1993;92(1):79–80.
- [12] Olivieri NF, Chang LS, Poon AO, Michelson AM, Orkin SH. An alpha-globin gene initiation codon mutation in a black family with HbH disease. *Blood* 1987;70(3):729–32.
- [13] Chong SS, Boehm CD, Higgs DR, Cutting GR. Single-tube multiplex-PCR screen for common deletion determinants of alpha-thalassemia. *Blood* 2000;95(1):360–2.
- [14] Chong SS, Boehm CD, Cutting GR, Higgs DR. Simplified multiplex-PCR diagnosis of common southeast asian deletion determinants of alpha-thalassemia. *Clin Chem* 2000;46(10):1692–5.
- [15] Liu YT, Old JM, Miles K, Fisher CA, Weatherall DJ, Clegg JB. Rapid detection of alpha-thalassaemia deletions and alpha-globin gene triplication by multiplex polymerase chain reactions. *Br J Haematol* 2000;108(2):295–9.
- [16] Xiao W, Oefner PJ. Denaturing high-performance liquid chromatography: a review. *Hum Mutat* 2001;17(6):439–74.
- [17] Hung CC, Lee CN, Chen CP, Jong YJ, Chen CA, Cheng WF, et al. Quantification of relative gene dosage by single-base extension and high-performance liquid chromatography: application to the *SMN1/SMN2* gene. *Anal Chem* 2005;77(21):6960–8.
- [18] Su YN, Hung CC, Li H, Lee CN, Cheng WF, Tsao PN, et al. Quantitative analysis of *SMN1* and *SMN2* genes based on DHPLC: a highly efficient and reliable carrier-screening test. *Hum Mutat* 2005;25(5):460–7.
- [19] Higgs DR, Hill AV, Bowden DK, Weatherall DJ, Clegg JB. Independent recombination events between the duplicated human alpha globin genes; implications for their concerted evolution. *Nucleic Acids Res* 1984;12(18):6965–77.
- [20] Michelson AM, Orkin SH. Boundaries of gene conversion within the duplicated human alpha-globin genes. Concerted evolution by segmental recombination. *J Biol Chem* 1983;258(24):15245–54.
- [21] Shaji RV, Srivastava A, Chandy M, Krishnamoorthy R. A single tube multiplex PCR method to detect the common alpha+ thalassemia alleles. *Blood* 2000;95(5):1879–80.

- [22] Ou-Yang H, Hua L, Mo QH, Xu XM. Rapid, accurate genotyping of the common $\alpha(4.2)$ thalassaemia deletion based on the use of denaturing HPLC. *J Clin Pathol* 2004;57(2):159–63.
- [23] Hung CC, Chen CP, Lin SP, Chien SC, Lee CN, Cheng WF, et al. Quantitative assay of deletion or duplication genotype by capillary electrophoresis system: application in Prader–Willi syndrome and Duchenne muscular dystrophy. *Clin Chem* 2006;52(12):2203–10.
- [24] Hung CC, Su YN, Tsao PN, Chen PC, Lin SJ, Lin CH, et al. Unequal crossover recombination-population screening for *PHOX2B* gene polymorphism using CE. *Electrophoresis* 2007;28(6):894–9.