Preparing compound heterozygous reference material using gene synthesis technology: a model of thrombophilic mutations

Martin Beranek[a], Monika Drastikova[a], Petr Dulicek[b], Vladimir Palicka[a]

Aims. The aim of our study is to present a novel approach for preparing a compound heterozygous reference material (*het*RM) using gene synthesis technology with inverted insertion of wild-type and mutant fragments into a single cloning vector. *Factor II* (G20210A) and *Factor V* (G1691A Leiden) gene mutations were used as an experimental model.
Methods. During the gene synthesis, DNA fragments were aligned in the following order: G1691 *FV* wild-type forward strain, G20210 *FII* wild-type forward strain, 1691A *FV* mutant reverse strain, 20210A *FII* mutant reverse strain. The complete chain was inserted into a pIDT SMART cloning vector and amplified in an *E. coli* competent strain. For assessing *het*RM characteristics and commutability, we used real-time PCR with subsequent melting curve analysis, real-time PCR with hydrolysis probes, allele-specific amplification, reverse hybridization, and dideoxynucleotide DNA sequencing.
Result. All five methods yielded concordant results of DNA analysis of the *het*RM. Differences in real-time PCR cycle threshold values after six-months of storage at -80 °C were not statistically significant from those obtained from freshly prepared *het*RM aliquots, which is a good indication of their stability.
Conclusion. By applying the procedures of gene synthesis and cloning technology, we prepared and verified a model genetic reference material for *FII* G20210A and *FV* G1691A testing with a compound heterozygous genotype. The *het*RM was stable, commutable, and available in large quantities and in a wide concentration range.

Key words: thrombophilic mutation, gene synthesis, reference material, commutability

Received: April 1, 2014; Accepted with revision: July 8, 2014; Available online: July 18, 2014
http://dx.doi.org/10.5507/bp.2014.041

[a]Institute of Clinical Biochemistry and Diagnostics, Faculty of Medicine in Hradec Kralove, Charles University in Prague and University Hospital Hradec Kralove, Czech Republic
[b]4th Department of Internal Medicine - Hematology, Faculty of Medicine in Hradec Kralove, Charles University in Prague and University Hospital Hradec Kralove
Corresponding author: Martin Beranek, e-mail: beranek@lfhk.cuni.cz

INTRODUCTION

Reference materials (RMs) ensure the analytical accuracy and reliability of molecular genetics investigation. They are used for quality assurance, interlaboratory proficiency testing, and preparation of calibrators or materials to convert quantitative PCR data to the international scale[1]. RMs are usually based on fragments of genomic DNA, cDNA molecules or amplicons of patients suffering from a genetic disease, immortalized cell lines with specific genetic alterations[2,3], recombinant molecules obtained by targeted mutagenesis[4], and synthetically prepared genes[5].

In genotyping assays for rare allelic variants, heterozygous control materials are formed by mixing equal numbers of cells or DNA clones containing wild-type and mutant sequences. This approach enabled the development of certified RMs for G20210A mutation in the *Factor II* (*FII*) gene[6] and G1691A Leiden mutation in the *Factor V* (*FV*) gene[7]. Both these mutations are associated with a higher risk of venous thrombosis and appear in healthy people of European ancestry at frequencies of 1-2% (G20210A) and 5-6% (G1691A), respectively[8,9].

The mixing ratio between different DNA clones is determined by measuring the optical density or fluorescence of intercalating dyes. Imprecision of these methods, however, can create a disproportionate amount of wild-type and mutant DNA molecules in the mixture, thus impairing RM commutability. These discrepancies must be resolved through a series of real-time PCR experiments leading to the normalization of the ratio. Unfortunately, results of the normalization process are only valid for the tested batches of wild-type and mutant clones.

The aim of our study is to present a novel approach for preparing a compound heterozygous RM (*het*RM) using the technology of hybrid gene synthesis with inverted insertion of wild-type and mutant fragments into a single cloning vector. *Factor II* (G20210A) and *Factor V* (G1691A Leiden) gene mutations were used as an experimental model. For assessing *het*RM characteristics and commutability, five different methods were applied: real-time PCR with subsequent melting curve analysis, real-time PCR with hydrolysis probes, allele-specific amplification, reverse hybridization, and dideoxynucleotide DNA sequencing.

MATERIAL AND METHODS

Preparation of *het*RM

The synthesized DNA chain (IDT, USA) contained wild-type and mutant *FV* fragments spanning 429 nucleo-

tides (chromosome 1q23, NCBI reference sequence code NG_011806.1, positions 36507–36935) and 321 nucleotides of wild-type and mutant *FII* fragments (chromosome 11p11, NG_008953.1, 20153–20473). The fragments included parts of the genes commonly used for G1691A and G20210A laboratory testing. During the synthesis, DNA fragments were aligned in the following order: G1691 *FV* wild-type forward strain, G20210 *FII* wild-type forward strain, 1691A *FV* mutant reverse strain, 20210A *FII* mutant reverse strain, as shown in Fig. 1.

The complete chain was inserted into a pIDT SMART cloning vector; the constructs were transferred into an *E. coli* competent strain and amplified (IDT, USA). Then, purification, spectrophotometric determination of plasmid DNA concentration, and freeze-drying followed. After reconstituting the freeze-dried DNA molecules in Tris-EDTA buffer, serial dilution (10^9–10^1 copies/μL) in plastic tubes was performed. 20 mL aliquots of the diluted DNA were stored at -80 °C.

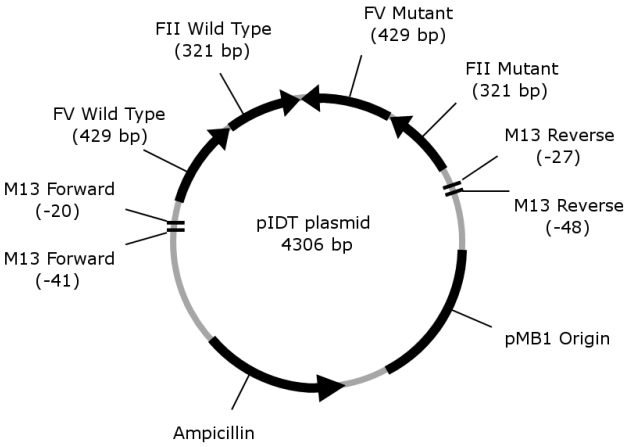

Fig. 1. Structure of pIDT SMART cloning vector containing synthetically prepared fragments of *Factor V* and *Factor II* genes. The fragments carrying mutations *FV* G1691A and *FII* G20210A were synthesized in inverse orientation to the wild-type fragments of the genes (marked by arrows). M13 are flanking sequences of M13 bacteriophage inside the cloning vector.

Real-time PCR assays

For *het*RM testing we used commercial kits for *in vitro* diagnostics (CE-IVD certification). Factor II (Prothrombin) G20210A Kit and Factor V Leiden Kit were based on the principle of real-time PCR with fluorescent hybridization probes followed by melting curve analysis[10] in the LightCycler 1.5 (Roche Diagnostics, Germany). 165 bp products for *FII* and 222 bp products for *FV* were amplified.

Allelic discrimination assays using real-time PCR (ref.[11]) and fluorescent hydrolysis probes (gb HEMO FII Kit and gb HEMO FV Kit, Generi Biotech, Czech Republic) were performed on the Rotorgene 6000 (Corbett Research, Australia). The sizes of the amplicons were 91 bp (*FII*) and 110 bp (*FV*).

Allele-specific amplification

PCR mixtures (25 μL) for *FII* and *FV* contained 10× concentrated PCR buffer, 200 μM each of deoxynucleotides, 400 nM primers (Generi Biotech, Czech Republic), 2.5 mM magnesium chloride, 50 ng of DNA, and one unit of *Taq* polymerase HS (Takara, Japan). PCR for *FII* included forward primer 5´- CCG CCT GAA GAA GTG GAT AC -3´ and allele-specific reverse primers 5´- CAC TGG GAG CAT TGA GGA TC -3´ for the wild-type allele or 5´- CAC TGG GAG CAT TGA GGA TT -3´ for the mutant one. After initial denaturation (5 min at 95 °C), PCRs were run for 30 cycles consisting of 30 s denaturation at 95 °C, 30 s annealing at 60 °C, and 60 s extension at 72 °C in the ABI 2720 thermal cycler (Applied Biosystems, USA).

The sequences of *FV* allele-specific forward primers were: 5´- CAG ATC CCT GGA CAG ACG -3´ for wild-type allele and 5´- CAG ATC CCT GGA CAG ACA -3´ for 1691A allele; the consensual *FV* reverse primer was 5´- TGT TAT CAC ACT GGT GCT TAA -3´ (ref.[12]). The PCR conditions were the same as above except the annealing temperature (56 °C). Amplicons were electrophoresed on a 3% agarose gel with ethidium bromide (100 V, 90 min). The sizes of the amplicons were 180 bp for *FII* and 174 bp for *FV*.

Reverse hybridization assay

Plasmid DNA molecules at a concentration of 10^5 copies/μL were used for multiplex PCR with biotinyled primers followed by reverse hybridization on nitrocellulose strips according to manufacturer's instructions (FV-PTH StripAssay Kit, ViennaLab, Austria). The lengths of the PCR products were 173 bp (*FII*) and 223 bp (*FV*).

DNA sequencing

FII amplicons (208 bp) were generated by PCR with forward primer 5´- CAG TTT GGA GAG TAG GGG GC -3´ and reverse primer 5´- GGT GGT GGA TTC TTA AGT CTT CT -3´. For *FV* amplicons (332 bp) we used forward primer 5´- CAG TTC AAC CAG GGG AAA CCT A -3´ and reverse primer 5´- TCA CAC TGG TGC TAA AAA GGA -3´ (Generi Biotech, Czech Republic). The same primers were used for bidirectional sequencing with a BigDye Terminator v3.1 Cycle Sequencing Kit and an ABI 3130 Genetic Analyzer (Life Technologies, USA). Unincorporated dye terminators, salts, and unused primers were removed from the mixtures by a BigDye XTerminator Purification Kit (Life Technologies, USA).

Testing *het*RM stability

Storage stability of DNA molecules in the *het*RM aliquots with 10^4, 10^5 and 10^6 copies/μL was examined using real-time PCR (gb HEMO FV Kit and gb HEMO FII Kit) under the same conditions as above. Differences in cycle thresholds (CT) were analyzed after six months of storage. The storage period was interrupted six times by thawing the tubes to room temperature for 30 min.

Statistical analysis

All experiments were done in triplicate. Differences between CT means were tested using t-tests. *P* values < 0.05 were considered statistically significant. All calculations were performed using the Statistica software (version 11, StatSoft, Tulsa, USA) for Windows.

RESULTS

All five methods yielded concordant results of DNA analysis of the *het*RM. Bands appearing on reverse hybridization test strips at positions for wild-type and mutant *FII* and *FV* alleles were of the same intensity and imitated the results of real heterozygous specimens. Allelic bands visible in the agarose gel after allele-specific PCR amplification of the *het*RM were comparable for wild-type and mutant fragments (Fig. 2). Sequencing analysis revealed

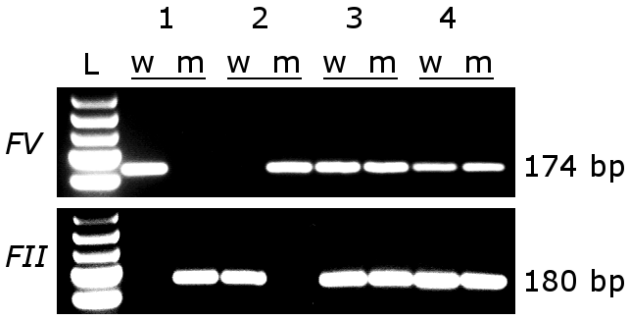

Fig. 2. Allele-specific amplification of *Factor V* (upper part) and *Factor II* (lower part) genes in mixtures specific for wild-type (w) and mutant (m) alleles. Line L: Low Molecular Weight DNA Ladder (New England Biolabs, UK); sample 1: *FII* G20210A homozygote; sample 2: *FV* G1691A homozygote; sample 3: compound heterozygote for *FII* G20210A and *FV* G1691A mutations; sample 4: *het*RM with compound heterozygosity for both mutations.

two very similar peaks apparent at *FV* 1691 and *FII* 20210 polymorphic sites, thus confirming the compound heterozygous character of the *het*RM (Fig. 3).

Fig. 4 shows the results of a melting curve analysis. A double heterozygous character of the *het*RM for *FII* G20210A and *FV* G1691A was found in the range of five concentration orders (from 10^2–10^6 copies/μL). Similarly, no significant influence of various *het*RM copy numbers in the mixture was found when allelic discrimination real-time PCR was carried out (Fig. 5).

Evaluating the stability of the *het*RM via real-time PCR, we achieved highly consistent triplicate data which gave coefficients of variability lower than 5%. The differences in CT values after six-months of storage at -80 °C were not statistically significant from those obtained from freshly prepared *het*RM aliquots, which is a good indication of their stability (Table 1). Six freezing-thawing cycles did not impair amplification efficiencies of *FII* and *FV* PCRs performed with the *het*RM.

DISCUSSION

One of the crucial steps in accurate and reliable genetic analysis is receiving the correct result of RMs investigated simultaneously with clinical specimens. The effort of the World Health Organization, European Society for Human Genetics, external quality providers, and RMs manufacturers is to prepare certified reference materials

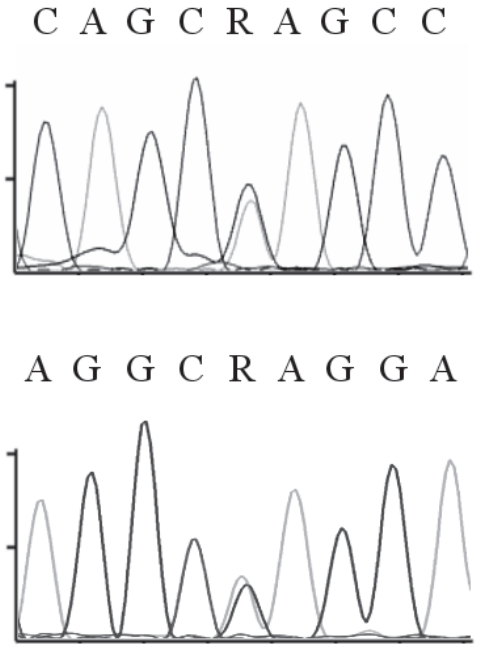

Fig. 3. Sequencing analysis of the *het*RM. Upper part: *FII* sequenced fragments; lower part: *FV* fragments. Letter R: site of heterozygosity (*FII* G20210A and *FV* G1691A).

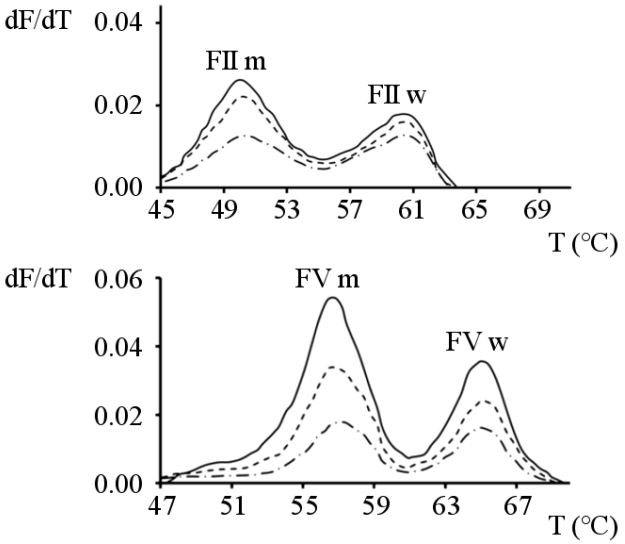

Fig. 4. Melting curve analysis. Upper part: *FII* analysis of the *het*RM heterozygosity, *FII m* is the melting temperature peak for 20210A mutant allele, *FII w* is the wild-type peak (G20210). Lower part: *FV m* is the melting temperature peak for 1691A mutant allele, *FV w* is the wild-type peak (G1691). The concentration range used: 10^2 (dash-and-dotted curves), 10^4 (dotted curves), 10^6 (solid curves) copies/μL.

Table 1. Stability of the *het*RM analyzed by real-time PCR cycle thresholds

Day	Allele	DNA concentration		
		10^4 copies/μL	10^5 copies/μL	10^6 copies/μL
Day 0	*FII* w	29.2 (0.6)	24.8 (0.5)	20.7 (0.4)
	FII m	29.3 (0.5)	24.8 (0.5)	20.5 (0.4)
	FV w	26.9 (0.5)	23.5 (0.3)	20.5 (0.3)
	FV m	28.2 (0.5)	25.3 (0.5)	21.8 (0.4)
Day +180	*FII* w	28.9 (0.9)	25.8 (0.9)	21.6 (0.4)
	FII m	28.6 (0.9)	25.4 (0.9)	21.3 (0.4)
	FV w	26.9 (0.2)	23.5 (0.2)	20.4 (0.2)
	FV m	29.0 (0.2)	25.9 (0.2)	22.5 (0.2)

Values are expressed as means (standard deviations), w is wild-type allele, m is mutant allele

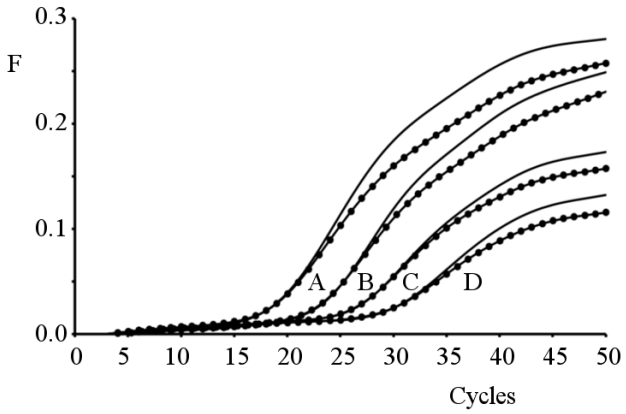

Fig. 5. Method of allelic discrimination for *FII* G20210A mutation. The *het*RM was used at concentrations of 10^3 (part D), 10^4 (part C), 10^5 (part B), and 10^6 (part A) copies/μL manifested by different cycle thresholds. The threshold value was 0.05. *FII* wild-type alleles are indicated by solid lines, *FII* 21210A mutant alleles by dotted lines.

(CRMs) fully commutable with real biological samples[13]. Unfortunately, CRMs for most genetic alterations associated with inheritable diseases are not available yet. A list of currently offered genetic CRMs can be found at the following site: http://www.nibsc.org.

We chose *FII* G20210A and *FV* G1691A Leiden thrombophilic mutations as model alterations since a number of methods were validated for analysis[3,7]. We developed a unique reference material with a nucleotide sequence including wild-type and mutant alleles, and forming a *FII*/*FV* compound heterozygous genotype.

Both mutations are considered independent risk factors for venous thrombosis[14]. The G20210A mutation appearing in the 3´ untranslated region of the *FII* gene elevates plasma prothrombin concentrations[15]. The Leiden mutation prevents the cleavage of activated factor V by activated protein C, and thus contributes to the development of thrombosis in its carriers[16]. We have genotyped 2095 patients at risk of inherited thrombophilia. *FII* G20210A and *FV* G1691A mutations were found at frequencies of 4.7% and 14.4%, respectively. The distribution of genotypes was as follows: 1369 double wild-type subjects (65.5%), 510 *FV* G1691A heterozygotes (24.3%), 144 *FII* G20210A heterozygotes (6.9%), 24 *FV* G1691A homozygotes (1.1%), and four *FII* G20210A homozygotes (0.2%). Forty-four subjects (2.1%) were compound heterozygotes for *FII* G20210A and *FV* G1691A mutations with a described accelerated risk of recurrent thrombosis in comparison to subjects heterozygous for G1691A alone[17].

Commercially available RMs for thrombophilic mutations contain *FII* and *FV* DNA fragments distributed in separated vials. In genetic testing, however, parallel analyses of both mutations in the same analytical run are often required. Thus, a compound heterozygosity of the *het*RM with an equal ratio of wild-type and mutant fragments for G1691A and G20210A seems to be a useful tool for routine laboratory work.

The construction of the *het*RM was based on short, synthetic single-strand DNA fragments hybridized together via their cohesive ends. Due to the almost complete homogeneity of wild-type and mutant fragments differing in one nucleotide only, combining them into two direct tandems was impossible. For that, we used alternating *FV* and *FII* fragments: G1691 *FV* wild-type forward strain, G20210 *FII* wild-type forward strain, 1691A *FV* mutant reverse strain, and 20210A *FII* mutant reverse strain. Inversely oriented mutant fragments were used to reduce the possibility of non-specific amplification. All testing methods provided results confirming *het*RM compound heterozygosity.

The stability of the *het*RM after six months of storage at -80 °C, interrupted by repeated freezing-thawing cycles, was acceptable. No signs of plasmid DNA degradation during the storage period were recorded. The *het*RM concentrations from 10^4 to 10^6 copies/μL provided real-time amplification curves comparable with real DNA extracts.

The combination of synthetic fragments presented on *FII* and *FV* gene models is a universal approach for obtaining a compound heterozygous material for two or more biallelic polymorphic sites. Since the maximal length of synthetic genes is currently several kb, the size of each fragment should be no longer than 200-400 bp. Such lengths are suitable for most manufactured diagnostic assays.

As DNA specimens or RMs containing rare allelic variants are not widely available, using a synthetic heterozygous RM as a template enables easier validations of in-house real-time PCR methods. Taking into account the perfect concentration match between wild-type and

mutant fragments, possible discrepancies in amplification curves for different alleles could not be caused by the unsatisfactory quality or poor performance of DNA in the mixture. Other reasons should be considered: suboptimal temperature profile, varying quality of hydrolysis probes or primers, lower PCR effectiveness for one of the tested allelic variants, stabilities of reagents, etc.

In conclusion, by applying the procedures of gene synthesis and cloning technology, we prepared and confirmed a model genetic reference material for *FII* G20210A and *FV* G1691A testing with a compound heterozygous genotype. The *het*RM was stable, commutable, and available in large quantities and in a wide concentration range. It eliminated the necessity for preparing four individual clones as well as obviating the need to normalize their concentrations ratio.

ACKNOWLEDGEMENTS

The study was supported by the project MZ ČR – RVO (FNHK, 00179906) of the Ministry of Health, Czech Republic, by the grant SVV 260057/2014, and by the project PRVOUK P37 of the Charles University in Prague, Czech Republic. We would like to thank Dr. Jiri Kruml, KRD s.r.o., Prague, Czech Republic for technical assistance and useful discussions.

Authorship contributions: MB, PD: literature search; MB, MD, PD: manuscript writing; VP: study design; MD: data collection; MB, MD: data analysis; MB, PD, VP: data interpretation; MB: statistical analysis, fugures.

Conflict of interest statement: None declared.

REFERENCES

1. White HE, Hedges J, Bendit I, Branford S, Colomer D, Hochhaus A, Hughes T, Kamel-Reid S, Kim DW, Modur V, Müller MC, Pagnano KB, Pane F, Radich J, Cross NC, Labourier E. Establishment and validation of analytical reference panels for the standardization of quantitative BCR-ABL1 measurements on the international scale. Clin Chem 2013;59:938-48.
2. Hawkins M, Boyle J, Wright KE, Elles R, Ramsden SC, O'Grady A, Sweeney M, Barton DE, Burgess T, Moore M, Burns C, Stacey G, Gray E, Metcalfe P, Hawkins JR. Preparation and validation of the first WHO international genetic reference panel for Fragile X syndrome. Eur J Hum Genet 2011;19:10-7.
3. Boyle J, Hawkins M, Barton DE, Meaney K, Guitart M, O'Grady A, Tobi S, Ramsden SC, Elles R, Gray E, Metcalfe P, Hawkins JR. Establishment of the first WHO international genetic reference panel for Prader Willi and Angelman syndromes. Eur J Hum Genet 2011;19:857-64.
4. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 1989;77:51-9.
5. Hughes RA, Miklos AE, Ellington AD. Gene synthesis: methods and applications. Methods Enzymol 2011;498:277-309.
6. Klein CL, Marki-Zay J, Corbisier P, Gancberg D, Cooper S, Gemmati D, Halbmayer WM, Kitchen S, Melegh B, Neumaier M, Oldenburg J, Leibundgut EO, Reitsma PH, Rieger S, Schimmel HG, Spannagl M, Tordai A, Tosetto A, Visvikis S, Zadro R, Mannhalter C. Reference materials (RMs) for analysis of the human factor II (prothrombin) gene G20210A mutation. Clin Chem Lab Med 2005;43:862-8.
7. Gray E, Hawkins JR, Morrison M, Hawkins M, Byrne E, Kitchen S, Jennings I, Makris M, Preston FE, Metcalfe P. Establishment of the 1st International Genetic Reference Panel for Factor V Leiden, human gDNA. Thromb Haemost 2006;96:215-9.
8. Crowther MA, Kelton JG. Congenital thrombophilic states associated with venous thrombosis: a qualitative overview and proposed classification system. Ann Intern Med 2003;138:128-34.
9. Bauer KA. The thrombophilias: well-defined risk factors with uncertain therapeutic implications. Ann Intern Med 2001;135:367-73.
10. Crockett AO, Wittwer CT. Fluorescein-labeled oligonucleotides for real-time PCR: using the inherent quenching of deoxyguanosine nucleotides. Anal Biochem 2001;290:89-97.
11. Louis M, Dekairelle AF, Gala JL. Rapid combined genotyping of factor V, prothrombin and methylenetetrahydrofolate reductase single nucleotide polymorphisms using minor groove binding DNA oligonucleotides (MGB probes) and real-time polymerase chain reaction. Clin Chem Lab Med 2004;42:1364-9.
12. Stirling D. Multiplex amplification refractory mutation system for the detection of prothrombotic polymorphisms. In: Bartlett JMS, Stirling D, editors. PCR protocols, second edition. New Jersey: Humana Press; 2003. p.323-327.
13. Ramsden SC, Deans Z, Robinson DO, Mountford R, Sistermans EA, Grody WW, McQuaid S, Patton SJ, Stenhouse SA. Monitoring standards for molecular genetic testing in the United Kingdom, the Netherlands, and Ireland. Genet Test 2006;10:147-56.
14. Endler G, Mannhalter C. Polymorphisms in coagulation factor genes and their impact on arterial an venous thrombosis. Clin Chim Acta 2003;330:31-55.
15. Poort SR, Rosendaal FR, Reitsma PH, Bertina RM. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. Blood 1996;88:3698-703.
16. Rosendaal FR, Siscovick DS, Schwartz SM, Beverly RK, Psaty BM, Longstreth WT Jr, Raghunathan TE, Koepsell TD, Reitsma PH. Factor V Leiden (resistance to activated protein C) increases the risk of myocardial infarction in young women. Blood 1997;89:2817-21.
17. De Stefano V, Martinelli I, Mannucci PM, Paciaroni K, Chiusolo P, Casorelli I, Rossi E, Leone G. The risk of recurrent deep venous thrombosis among heterozygous carriers of both factor V Leiden and the G20210A prothrombin mutation. N Engl J Med 1999;341:801-6.