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

Martin Beranek, Monika Drastikova, Petr Dulicek, Vladimir Palicka

Aims. The aim of our study is to present a novel approach for preparing a compound heterozygous reference material (hetRM) 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 hetRM 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 hetRM. 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 hetRM 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 hetRM was stable, commutable, and available in large quantities and in a wide concentration range.

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

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. 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, recombinant molecules obtained by targeted mutagenesis, and synthetically prepared genes.

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 and G1691A Leiden mutation in the Factor V (FV) gene. 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.

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 (hetRM) 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 hetRM 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 hetRM

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 (10⁹–10¹ copies/µL) in plastic tubes was performed. 20 mL aliquots of the diluted DNA were stored at -80 °C.

The sequences of FV allele-specific forward primers were: 5′- CAG ATC CCT GGA CAG ACA -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.

**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.

### DNA sequencing

Allele 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 10⁻⁶ concentrated PCR buffer, 200 nM 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 ACA -3′ for wild-type allele and 5′- CAG ATC CCT GGA CAG ACA -3′ for 1691A allele; the consensual 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 10⁹ copies/µL were used for multiplex PCR with biotinylated 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).

**Real-time PCR assays**

For hetRM 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 analysis² 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).
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 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. 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.

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: 10² (dash-and-dotted curves), 10⁴ (dotted curves), 10⁶ (solid curves) copies/µL.

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.
Table 1. Stability of the hetRM analyzed by real-time PCR cycle thresholds

<table>
<thead>
<tr>
<th>Day</th>
<th>Allele</th>
<th>DNA concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$10^4$ copies/µL</td>
</tr>
<tr>
<td>Day 0</td>
<td>FII w</td>
<td>29.2 (0.6)</td>
</tr>
<tr>
<td></td>
<td>FII m</td>
<td>29.3 (0.5)</td>
</tr>
<tr>
<td></td>
<td>FV w</td>
<td>26.9 (0.5)</td>
</tr>
<tr>
<td></td>
<td>FV m</td>
<td>28.2 (0.5)</td>
</tr>
<tr>
<td>Day +180</td>
<td>FII w</td>
<td>28.9 (0.9)</td>
</tr>
<tr>
<td></td>
<td>FII m</td>
<td>28.6 (0.9)</td>
</tr>
<tr>
<td></td>
<td>FV w</td>
<td>26.9 (0.2)</td>
</tr>
<tr>
<td></td>
<td>FV m</td>
<td>29.0 (0.2)</td>
</tr>
</tbody>
</table>

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 hetRM was used at concentrations of $10^4$ (part D), $10^5$ (part C), $10^6$ (part B), and $10^7$ (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 samples13. 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 analysis3,7. 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 thrombosis14. The G20210A mutation appearing in the 3´ untranslated region of the FII gene elevates plasma prothrombin concentrations15. The Leiden mutation prevents the cleavage of activated factor X by activated protein C, and thus contributes to the development of thrombosis in its carriers16. 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 alone17.

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 hetRM 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 hetRM 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 hetRM compound heterozygosity.

The stability of the hetRM 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 hetRM 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 hetRM 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, figures.

Conflict of interest statement: None declared.

REFERENCES


