

NOVÝ KONCEPT PŘÍPRAVY REFERENČNÍCH MATERIÁLŮ S HETEROZYGOTNÍM GENOTYPEM PRO MOLEKULÁRNĚ DIAGNOSTICKÉ ÚČELY

MARTIN BERÁNEK, MONIKA DRASTÍKOVÁ
a VLADIMÍR PALIČKA

Ústav klinické biochemie a diagnostiky, Lékařská fakulta
Univerzity Karlovy a Fakultní nemocnice Hradec Králové,
Sokolská 581, 500 05 Hradec Králové
beranek@lfhk.cuni.cz

Došlo 24.3.14, přijato 10.4.14.

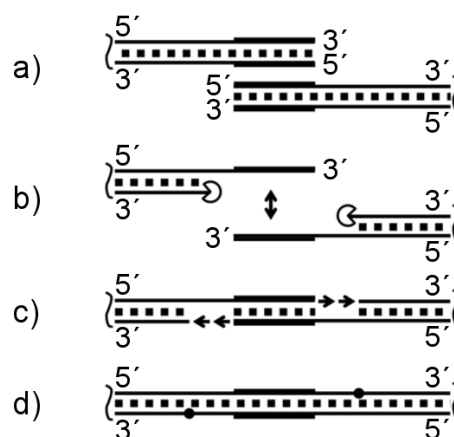
Klíčová slova: syntéza genů, referenční materiál,
klonování, trombofilní mutace, komutabilita

Úvod

Příprava rekombinantních molekul DNA je technologický proces používaný přes čtyřicet let v experimentální i klinické medicíně. Jednu z klinických aplikací těchto molekul představují referenční materiály určené k zajištění spolehlivé diagnostiky genetických onemocnění. Rekombinantní postupy jsou založeny na fragmentaci molekul DNA, cílené mutagenezi, ligaci, amplifikaci a celé řadě dalších technik prováděných v podmínkách *in vitro*¹. Tyto tradiční přístupy doplnila v roce 2009 technika syntézy genů, která je schopna laboratorně vytvořit genové úseky dlouhé až desítky kilobází².

Podstatou zmíněné syntézy genů je jednak příprava polynukleotidových vláken, využívaná pro tvorbu oligonukleotidů a dlouhých ultramerů^{3,4}, a jednak metoda spojování dlouhých fragmentů DNA (enzymatic assembly)⁵. Na obr. 1 je znázorněn princip tohoto analytického postupu. Genetická informace obsažená ve spojovaných fragmentech DNA může pocházet z jednoho nebo více přírodních druhů. Pro jejich spojení je nezbytná přítomnost překrývajících se koncových úseků o délce několika desítek nukleotidů (obr. 1a, zmíněné úseky jsou znázorněny tučně). Katalytickým účinkem 5'-exonukleas nejprve na 3'-konicích vláken dojde ke vzniku jednořetězcových přesahů, které jsou v rámci dvou sousedících fragmentů vzájemně komplementární (obr. 1b). Po hybridizaci těchto fragmentů umožní DNA polymerasa dostavění chybějících nukleotidů (obr. 1c) a poté následuje zkompletování řetězců DNA ligasou (obr. 1d). Celý proces je izotermický, trvá několik minut a lze během něj spojit až několik desítek různých genových segmentů.

Po zaklonování syntetických genů do vektoru je lze použít např. jako kalibrační standardy pro kvantitativní PCR³. Kromě fragmentů vyšetřovaného genu může vytvořený klon obsahovat také sekvence některého z genů pro-



Obr. 1. Schéma spojení dlouhých fragmentů DNA pomocí jednořetězcových přesahů

vozních (house-keeping genes), jež jsou nezbytné pro normalizaci množství transkriptů RNA^{6,7}.

Další možnou diagnostickou aplikací syntetických genů představují referenční materiály používané při analýze genotypů, které zahrnují sekvence homologní se zdravou (wild-type) či mutovanou alelou⁸. Tyto materiály se stávají součástí diagnostických postupů pro vyšetření vzácných alelických variant, jejichž populační výskyt je zpravidla hluboko pod 1 %.

Referenční materiály s heterozygotním genotypem se v současnosti připravují smícháním klonů obsahujících sekvence různých alel téhož genu. Analytickým problémem těchto heterozygotních materiálů je však jejich nižší komutabilita v porovnání s reálnými vzorky DNA. Rozdíly jsou způsobeny nepřesnostmi při odhadu množství molekul příslušných vektorů. Projevují se zejména u metod pracujících na bázi alelické diskriminace prováděné pomocí real-time PCR a hydrolyzačních sond.

Cílem naší studie bylo vytvořit pomocí genové syntézy heterozygotní referenční materiál, který by obsahoval fragmenty jak zdravé, tak mutované alely v jediném klonu. Pro studii byly jako modelové vybrány dvě mutace související s etiologií žilní trombózy, a to mutace G20210A v genu pro faktor II (protrombin) a mutace G1691A v genu pro faktor V krevního srážení⁹. Výsledný genotyp získaného klonu by tak měl mít charakter složeného (dvojitého) heterozygota.

Experimentální část

Syntéza genů

Navržené fragmenty DNA obsahovaly část genu *FII* s rozsahem 321 nukleotidů (přístupový kód v databázi NCBI: NG_008953.1, nukleotidy 20153–20473; pozice nukleotidu pro mutaci G20210A: 20313) a část genu *FV* s rozsahem 429 nukleotidů (přístupový kód: NG_011806.1, nukleotidy 36507–36935; pozice nukleotidu pro mutaci

G1691A: 36721). Podle řazení jednotlivých alel byly navrženy tři typy inzertů, resp. klonů: K0, K1 a K2 (obr. 2a–c). Při přípravě klonu K0 bylo použito řazení fragmentů podle genové příslušnosti, u klonu K1 bylo použito střídavé řazení – nejprve byly syntetizovány sekvence obou alel typu wild-type, a poté alely mutované. Také u klonu K2 bylo použito střídavé řazení, ovšem s inverzně orientovanými fragmenty obou mutovaných alel. Oblasti kódujících sekvencí jsou na obr. 2 znázorněny šipkami směřujícími zleva doprava, u reverzních vláken je orientace šipek opačná. Syntéza inzertů byla podle našeho návrhu provedena společností IDT, USA.

Příprava klonů a referenčního materiálu

Inzerty byly vloženy do klonovacího vektoru pIDT SMART. Po genovém přenosu a kultivaci v kompetentních buňkách *E. coli* byla izolována plazmidová DNA, provedena spektrofotometrická analýza a lyofilizace (IDT, USA). Pro rozpuštění lyofilizované DNA a sériové ředění (10^9 – 10^1 kopií/ μ l) byl použit AE pufr (Qiagen, SRN). Získaný referenční materiál alikvotovaný po 20 μ l do plastových zkumavek byl uchováván při -80 °C.

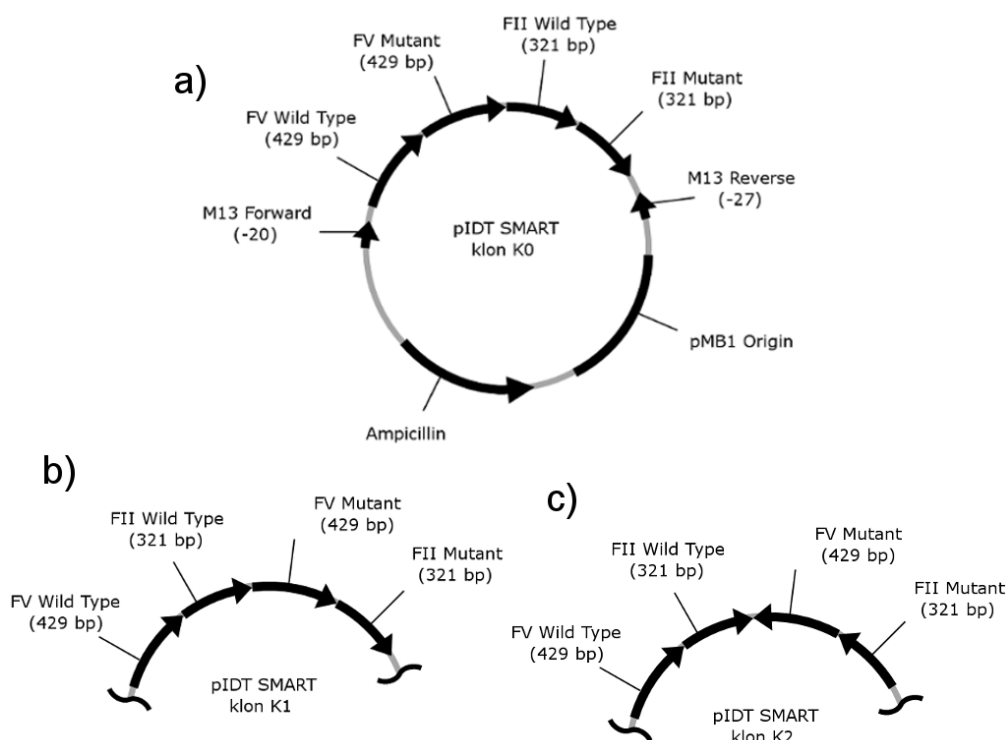
PCR amplifikace

Amplifikační směsi objemu 25 μ l obsahovaly 2,5 μ l $10\times$ koncentrovaného PCR pufru s chloridem hořečnatým

o koncentraci 15 mmol l^{-1} (Takara, Japonsko), deoxynukleotidy (finální koncentrace ve směsi $200 \mu\text{mol l}^{-1}$, Takara, Japonsko), amplifikační primery ($0,4 \mu\text{mol l}^{-1}$, Generi Biotech, ČR), $1,5 \text{ U Taq HS DNA polymerasy}$ (Takara, Japonsko) a 2 μ l plazmidové DNA. Sekvence primerů byly následující: přímý (forward) primer pro *FII* 5'- CAG TTT GGA GAG TAG GGG GC -3', reverzní primer pro *FII* 5'- GGT GGT GGA TTC TTA AGT CTT CT -3', přímý primer pro *FV* 5'- CAG TTC AAC CAG GGG AAA CCT A -3', reverzní primer pro *FV* 5'- TCA CAC TGG TGC TAA AAA GGA -3'. Teplotní podmínky pro PCR v termocyklu T3000 (Biometra, SRN) zahrnovaly počáteční denaturaci 5 min při 95 °C a 30 amplifikačních cyklů (denaturace 30 s při 95 °C, hybridizace 30 s při 59 °C, elongace 30 s při 72 °C). Amplikony měly při analýze mutace G20210A délku 208 bp, u vyšetření mutace G1691 délku 332 bp. Detekce probíhala na 2% horizontálním agarosovém gelu s ethidiumbromidem.

Sekvenování DNA

Amplikony získané za podmínek PCR popsaných výše byly purifikovány soupravou QIAquick Gel Purification System (Qiagen, SRN). Sekvenační reakce obsahovala 8 μ l komerční reakční směsi „BigDye X Terminator Kit“ verze 3.1 (Applied Biosystems, USA), 5 μ l purifikovaného amplikonu, 5 μ l demineralizované vody a 1 μ l přímého nebo reverzního primeru ředěného na koncentraci $10 \mu\text{mol l}^{-1}$.



Obr. 2. Složení inzertů syntetických genů v klonovacích vektorech

Profil sekvenační reakce se skládal z denaturace 20 s při 96 °C, hybridizace 20 s při 50 °C a elongace 4 min při 60 °C (35 reakčních cyklů). Produkty reakce byly přečištěny soupravou „BigDye X Terminator Purification Kit“ (Applied Biosystems, USA). Pro kapilární elektroforézu byl použit elektrokinetický nástřík při 1,6 kV po dobu 10 s; separace probíhala v 50cm kapiláře při 8,5 kV a teplotě 60 °C.

Real-time PCR

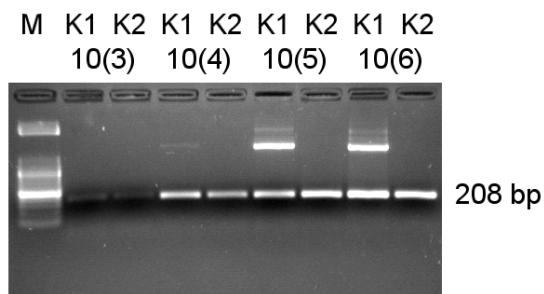
Pro vyšetření mutací G20210A a G1691A byly použity soupravy určené pro *in vitro* diagnostiku (CE-IVD) od společnosti Geneti Biotech, ČR: gb HEMO FII Kit a gb HEMO FV Kit. Analýza byla provedena na analyzátoru Rotorgene 6000 (Corbett Research, Australia). Délky ampliconů byly 91 bp (*FII*) a 110 bp (*FV*).

Výsledky

Pomocí molekulového klonování byly vytvořeny modelové klony K1 a K2, které se lišily složením použitého inzeru. Úspěšnost vložení fragmentů DNA do těchto vektorů potvrdil výrobce sekvenační analýzou s primery komplementárními s úseky M13 (obr. 2a). Inzert navržený pro klon K0 nesplňoval strukturální požadavky pro přípravu syntetického genu (příliš vysoká homologie sousedících fragmentů s toutéž orientací), a k jeho přípravě proto nedošlo.

Kontrola amplifikace syntetických fragmentů obsažených v klonech K1 a K2 zahrnovala standardní PCR a detekci produktů na agarosovém gelu. Z obr. 3 je patrné, že klon 1 při analýze mutace G20210A provedené na čtyřech koncentračních hladinách (10^3 , 10^4 , 10^5 a 10^6 kopií/ μ l) vykazoval známky nespecifické amplifikace. Stejná situace nastala i u vyšetření mutace G1691A. Naopak, u klonu 2 k tvorbě nespecifických produktů nedošlo, a byl proto použit pro další testování. Písmeno M označuje na elektroforetickém gelu velikostní marker XIII (Roche, SRN).

Oboustranným sekvenováním ampliconů byl u klonu 2 potvrzen genotyp složeného heterozygota pro mutace G20210A a G1691A. Výšky pík odpovídající zdravé



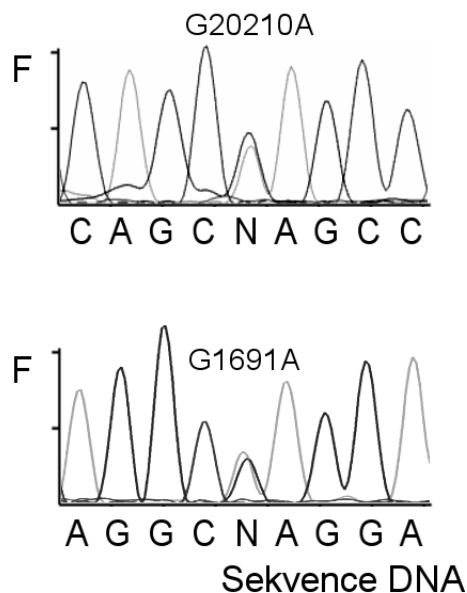
Obr. 3. PCR amplifikace části genu *FII* provedená u klonů K1 a K2 o různé koncentraci DNA (10^3 – 10^6 kopií/ μ l); délka ampliconů byla 208 bp

a mutované alele byly v obou sledovaných místech přibližně stejně vysoké (obr. 4). Na obr. 5 je ukázán záznam analýzy mutace G1691A v genu *FV* provedené metodou alelické diskriminace. Plnými čarami jsou zde znázorněny amplifikační křivky pro zdravou alelu G1691 a čarami přerušovanými křivky pro mutovanou alelu 1691A. Křivky *a* a *c* byly získány testováním referenčního materiálu připraveného tradičním způsobem, tedy smícháním plazmidové DNA z klonů obsahujících zdravou nebo mutovanou alelu^{10,11}. Křivky *b* a *d* pocházející z analýzy klonu 2 (použitá koncentrace 10^5 kopií/ μ l) vykazovaly menší vzájemné rozdíly ve srovnání s křivkami tradičně připravených heterozygotních kontrol. Křivka *e* představuje mez detekce (threshold) nastavenou pro obě vyšetření na hodnotu 0,05.

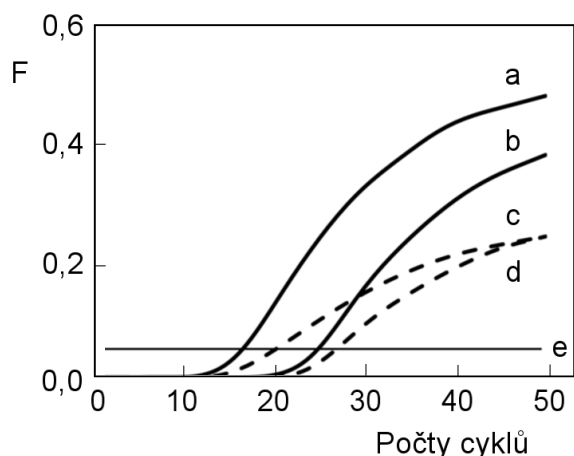
Diskuse

Analýza mutací je na pracovištích molekulární genetiky prováděna validovanými metodami schválenými Českým institutem pro akreditaci (ČIA). Součástí každé analytické série jsou referenční kontrolní materiály, jejichž výsledky potvrzují spolehlivost použitého vyšetřovacího postupu. V řadě případů však certifikované referenční materiály nejsou komerčně dostupné a pro jejich přípravu je nutno použít klonovací techniky nebo o ně požádat jinou akreditovanou laboratoř.

V naší studii jsme vyvinuli a ověřili vlastnosti modelového referenčního materiálu s genotypem složeného heterozygota pro mutace G20210A v genu *FII* a G1691A v genu *FV*. Jedinci s tímto genotypem se v naší populaci vyskytují s frekvencí nižší než 1 %. Struktura referenčního



Obr. 4. Analýza sekvence DNA v genech *FII* (horní část) a *FV* (dolní část) potvrzující složenou heterozygotitu klonu 2 pro mutace G20210A a G1691A



Obr. 5. Alelická diskriminace pro mutaci G1691A metodou real-time PCR

materiálu byla vytvořena pomocí technologie syntézy genů a molekulového klonování. K přípravě inzertu byly použity fragmenty obou genů obsahující sekvence zdravých i mutovaných alel. Genotyp referenčního materiálu založeném na klonu 2 byl potvrzen sekvenační analýzou.

Pro funkčnost tohoto referenčního materiálu byla zcela zásadní orientace fragmentů DNA, zejména invertovaná pozice mutovaných fragmentů vůči fragmentům se sekvencí wild-type. Střídavé řazení fragmentů obou genů s délkou 300–500 bp a s 40–60% zastoupením nukleotidů s guaninem a cytosinem snížily pravděpodobnost vzniku sekundárních struktur, jež negativně ovlivňují účinnost PCR. Námi provedené experimenty ukázaly uspokojivou amplifikační účinnost metody real-time PCR aplikované pro vyšetření obou testovaných mutací. Přetrvávající rozdíly v amplifikačních účinnostech patrné u křivek *b* a *d* na obr. 5 nebyly způsobeny vlastnostmi použitého kontrolního materiálu, nýbrž jinými faktory – strukturou použitých sond, teplotním profilem PCR reakce, sekvencemi použitých primerů, stabilitou použitých reagensů, apod. Tyto faktory lze eliminovat další optimalizací vyšetřovací soupravy.

Příprava syntetických genů různého biologického původu v délce jednotek až desítek kilobází je v současnosti komerčně dostupná. Z tohoto pohledu je námi navržený postup vytvoření heterozygotního referenčního materiálu jednou z možností, jak získat kontrolní vzorky pro vyšetření jakéhokoliv bialelického polymorfního místa s řídce se vyskytujícími minoritními variantami.

Závěr

Použitím technologie genové syntézy a molekulového klonování jsme připravili a ověřili vlastnosti modelového referenčního materiálu s genotypem složeného heterozygota pro mutace G20210A v genu *FII* a G1691A v genu *FV* určeného pro molekulární diagnostiku žilních trombóz.

Tato studie byla finančně podpořena prostředky MZ ČR – RVO (FNHK, 00179906) a grantem SVV 260057/2014.

LITERATURA

1. Cederbaum S. D., Fareed G. C., Lovett M. A., Shapiro L. J.: *West. J. Med.* 141, 210 (1984).
2. Hughes R. A., Miklos A. E., Ellington A. D.: *Methods Enzymol.* 498, 277 (2011).
3. Viljoen C. D., Thompson G. G., Sreenivasan S.: *Gene* 516, 143 (2013).
4. Beránek M., Drastíková M., Sirák I., Paulíková S., Vošmik M., Petera J.: *Chem. Listy* 107, 880 (2013).
5. Gibson D. G., Young L., Chuang R. Y., Venter J. C., Hutchison C. A., Smith H. O.: *Nat. Methods* 6, 343 (2009).
6. Kruttgen A., Razavi S., Imohl M., Ritter K.: *Med. Microbiol. Immunol.* 200, 137 (2011).
7. Drastíková M., Gabalec F., Čáp J., Beránek M.: *Klin. Biochem. Metab.* 21, 129 (2013).
8. Huang M. C., Cheong W. C., Lim L. S., Li M. H.: *Electrophoresis* 33, 78 (2012).
9. Ender G., Mannhalter C.: *Clin. Chim. Acta* 330, 31 (2003).
10. Klein C. L., Marki-Zay J., Corbisier P., Gancberg D., Cooper S., Gemmati D., Halbmayr W. M., Kitchen S., Melegh B., Neumaier M., Oldenburg J., Leibundgut E. O., Reitsma P. H., Rieger S., Schimmel H. G., Spannagl M., Tordai A., Tosetto A., Visvikis S., Zadro R., Mannhalter C.: *Clin. Chem. Lab. Med.* 43, 862 (2005).
11. Gray E., Hawkins J. R., Morrison M., Hawkins M., Byrne E., Kitchen S., Jennings I., Makris M., Preston F. E., Metcalfe P.: *Thromb. Haemostasis* 96, 215 (2006).

M. Beránek, M. Drastíková, and V. Palička
(*Institute of Clinical Biochemistry and Diagnostics, Faculty of Medicine and Faculty Hospital, Hradec Králové*):
A New Concept in Preparation of Reference Materials with Heterozygous Genotype for Molecular Diagnostic Purposes

Using the technique of enzymatic gene assembly we developed a compound heterozygous reference material for analysis of G20210A mutation in the Factor II and G1691A in the Factor V genes. The prepared clone contained both wild-type and mutant fragments of the genes. The DNA fragments carrying the mutations were inserted into the cloning vector in the inverse orientation relative to the wild-type ones. The clone performance was verified by DNA sequencing and by the allele discrimination real-time PCR methods. Our novel approach makes it possible to prepare a compound heterozygous reference material for any rare variants at biallelic polymorphic sites if analyzed by real-time PCR or other methods.