



## Restrikční endonukleázy a genetické testy

Bakalářská práce

*Vedoucí práce:*  
prof. RNDr. Aleš Knoll, Ph.D.

*Vypracovala:*  
Veronika Vilímová

zadání

### Čestné prohlášení

Prohlašuji, že jsem práci **Restrikční endonukleázy a genetické testy** vypracovala samostatně a veškeré použité prameny a informace uvádím v seznamu použité literatury. Souhlasím, aby moje práce byla zveřejněna v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách ve znění pozdějších předpisů a v souladu s platnou *Směrnici o zveřejňování vysokoškolských závěrečných prací*.

Jsem si vědoma, že se na moji práci vztahuje zákon č. 121/2000 Sb., autorský zákon, a že Mendelova univerzita v Brně má právo na uzavření licenční smlouvy a užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

Dále se zavazuji, že před sepsáním licenční smlouvy o využití díla jinou osobou (subjektem) si vyžádám písemné stanovisko univerzity, že předmětná licenční smlouva není v rozporu s oprávněnými zájmy univerzity, a zavazuji se uhradit případný příspěvek na úhradu nákladů spojených se vznikem díla, a to až do jejich skutečné výše.

V Brně dne .....

.....

podpis

## **PODĚKOVÁNÍ**

Děkuji prof. RNDr. Aleši Knollovi, Ph.D. za metodické vedení mé bakalářské práce. Děkuji také Ing. Elišce Horecké za poskytnutí dat z databáze školní laboratoře a odbornou pomoc. Rovněž děkuji svým přátelům za pomoc s formální úpravou textu, za podporu a užitečné rady. Za podporu při studiu a za vloženou důvěru velmi děkuji také příteli, své a jeho rodině.

## **ABSTRAKT**

Bakalářská práce Restrikční endonukleázy a genetické testy se zabývá problematikou restrikčních enzymů. V literárním přehledu je popsán charakter restrikčních endonukleáz, jejich rozdělení a mechanismus působení, dále se práce dotýká tématu vlivů podmínek restrikční reakce a pojednává o metodách využití těchto enzymů, jako jsou například klonování DNA, analýza variability nebo mapování genomu. V praktické části je zahrnuta práce se softwarem na automatické vyhledávání restrikčních míst a restrikční analýzu DNA.

Klíčová slova: restrikční endonukleázy, genetické testy, restrikční analýza, RFLP, PCR-RFLP

## **ABSTRACT**

This bachelor thesis Restriction endonucleases and genetic tests engages in the issue of restriction enzymes. In the literary review I describe the character of restriction endonucleases, their classification and mechanism of operation. The thesis furthermore touches a topic concerning the effect of the restriction reaction conditions and discusses techniques of utilization of these enzymes, where for example DNA cloning, analysis of variability or genome mapping belong. The practical part of the thesis includes some work with a software for an automatic search for restriction sites and for restriction DNA evaluation.

Key words: restriction endonucleases, genetic tests, restriction analysis, RFLP, PCR-RFLP

## OBSAH

|          |  |    |
|----------|--|----|
| 1        | ÚVOD.....                                      | 7  |
| 2        | CÍL PRÁCE .....                                | 8  |
| 3        | LITERÁRNÍ PŘEHLED .....                        | 9  |
| 3.1      | Charakteristika restrikčních endonukleáz ..... | 9  |
| 3.1.1    | Nomenklatura.....                              | 9  |
| 3.1.2    | Stavy enzymů.....                              | 10 |
| 3.1.3    | Význam.....                                    | 10 |
| 3.1.4    | Efektivita restrikčních endonukleáz.....       | 10 |
| 3.2      | Typy restrikčních endonukleáz .....            | 11 |
| 3.2.1    | Typ I.....                                     | 11 |
| 3.2.1.1  | Rodiny .....                                   | 12 |
| 3.2.1.2  | Enzymatická aktivita .....                     | 12 |
| 3.2.2    | Typ II .....                                   | 13 |
| 3.2.2.1  | Subtyp IIA .....                               | 14 |
| 3.2.2.2  | Subtyp IIB .....                               | 15 |
| 3.2.2.3  | Subtyp IIC .....                               | 16 |
| 3.2.2.4  | Subtyp IIE .....                               | 17 |
| 3.2.2.5  | Subtyp IIF.....                                | 18 |
| 3.2.2.6  | Subtyp IIG .....                               | 19 |
| 3.2.2.7  | Subtyp IIH .....                               | 20 |
| 3.2.2.8  | Subtyp IIM .....                               | 20 |
| 3.2.2.9  | Subtyp IIP.....                                | 21 |
| 3.2.2.10 | Subtyp IIS .....                               | 22 |
| 3.2.2.11 | Subtyp IIT.....                                | 23 |
| 3.2.3    | Typ III .....                                  | 24 |
| 3.2.4    | Typ IV .....                                   | 24 |
| 3.3      | Specifita restrikčních endonukleáz .....       | 25 |
| 3.3.1    | „Star“ aktivita .....                          | 26 |
| 3.3.2    | Vlivy podmínek reakce .....                    | 28 |
| 3.3.2.1  | Teplota inkubace .....                         | 28 |
| 3.3.2.2  | pH.....  | 28 |
| 3.3.2.3  | Kofaktory .....                                | 28 |
| 3.3.2.4  | Složení soli a iontová síla.....               | 28 |

|         |   |    |
|---------|---|----|
| 3.3.2.5 | Stabilizátory .....                                     | 29 |
| 3.3.2.6 | Pufry .....   | 29 |
| 3.3.2.7 | Objem .....   | 30 |
| 3.3.2.8 | Doba trvání reakce .....                                | 30 |
| 3.4     | Využití v molekulární genetice a diagnostice .....      | 31 |
| 3.4.1   | RFLP .....  | 31 |
| 3.4.2   | PCR-RFLP .....  | 32 |
| 3.4.3   | Restrikční mapy .....                                   | 33 |
| 3.4.4   | Genetické testy .....                                   | 33 |
| 3.4.5   | Klonování DNA .....                                     | 35 |
| 3.4.6   | Tvorba genomových knihoven .....                        | 38 |
| 3.4.7   | Tvorba genomových knihoven pro sekvenování genomů ..... | 38 |
| 4       | MATERIÁL A METODIKA .....                               | 40 |
| 4.1     | Geny .....  | 40 |
| 4.1.1   | <i>MC1R</i> .....                                       | 40 |
| 4.1.2   | <i>MATP</i> .....                                       | 40 |
| 4.2     | Počítačový software a webové aplikace .....             | 40 |
| 4.2.1   | Webcutter 2.0 .....                                     | 41 |
| 4.2.2   | WatCut .....  | 41 |
| 4.2.3   | NEBcutter V2.0 .....                                    | 42 |
| 5       | VÝSLEDKY A DISKUZE .....                                | 43 |
| 5.1     | Sequence Scanner v1.0 .....                             | 43 |
| 5.2     | Webcutter 2.0 .....                                     | 43 |
| 5.3     | WatCut .....  | 44 |
| 5.4     | NEBcutter V2.0 .....                                    | 45 |
| 5.5     | Shrnutí a diskuze .....                                 | 47 |
| 6       | ZÁVĚR .....   | 49 |
| 7       | POUŽITÁ LITERATURA .....                                | 50 |
| 8       | SEZNAM ZKRATEK .....                                    | 56 |
| 9       | SEZNAM OBRÁZKŮ A PŘÍLOH .....                           | 57 |
| 10      | PŘÍLOHY .....   | 58 |
| A.      | Doplňující tabulky .....                                | 58 |
| B.      | Doplňující obrázky .....                                | 60 |

# 1 ÚVOD

Restrikční endonukleázy jsou enzymy, které se vážou na krátké specifické sekvence párů bází v DNA a štěpí molekulu DNA v blízkosti této sekvence, a to v tzv. restrikčním místě. Restrikční enzymy jsou přirozeným produktem mikroorganismů, zejména však bakteriálních kmenů a archeí, jež tyto enzymy tvoří pro ochranu proti začlenění virové či jiné parazitické DNA do svého vlastního genomu (Wilson et al. 2012).

Aby se organizmy vyhnuly štěpení vlastní DNA, musí endonukleázu doprovázet ochranné prvky, které zabrání této degradaci – methyltransferázy. Tyto proteiny rozpoznávají stejnou specifickou sekvenci jako odpovídající restrikční enzymy, ale místo štěpení připojují k bázím na obou řetězcích DNA methylovou skupinu. Methylovaná sekvence DNA již tedy není náchylná k hydrolýze, jelikož ji endonukleáza nemůže rozpoznat (Wilson et al. 2012).

Restrikční endonukleázy (RES) a DNA-methyltransferázy (MT) jsou tedy součástí restrikčně – modifikačních systémů, které se dělí dle molekulární struktury, rozpoznávání sekvence, pozice štěpení a požadavků na kofaktory na čtyři skupiny (Tab. P1) – Typ I, II, III a IV (Rao, Dryden a Bheemanaik 2013).

Již více než čtyřicet let jsou tyto enzymy známy a používány, přesto však ještě není jejich potenciál plně prozkoumán. Díky nim se vědcům otevřela cesta k významným, dříve nemyslitelným metodám. Fragmenty DNA, které vznikly štěpením restrikčními endonukleázami, mohou být dále využity různými způsoby. Lze je mimo jiné klonovat, sekvenovat, rekombinovat nebo porovnávat mezi sebou a hledat tak polymorfizmy. Restrikční enzymy významně ovlivnily vývoj mnoha biologických oborů, ať už se jedná o lékařství či forenzní vědy (Snustad a Simmons 2009).



## **2 CÍL PRÁCE**

Bakalářská práce Restrikční endonukleázy a genetické testy je založena na rešerši literatury na téma restrikční endonukleázy, ale také na praktické práci s aktuálně dostupným softwarem pro analýzu restrikčních míst v DNA. Cílem teoretické části práce tedy je shrnout problematiku restrikčních endonukleáz a především prozkoumat nejdůležitější oblasti využití těchto enzymů. Metu praktické části práce představuje nalezení a návrh vhodné restrikční endonukleázy pro testování polymorfizmu metodou PCR-RFLP, ale také výsledná komparace účelnosti, přehlednosti a funkčnosti programů pro automatické vyhledávání restrikčních míst.

## 3 LITERÁRNÍ PŘEHLED

### 3.1 Charakteristika restrikčních endonukleáz

#### 3.1.1 Nomenklatura

Dle původního návrhu Smithe a Nathanse (1973) jsou restrikční enzymy, doprovodné DNA-methyltransferázy a proteiny pojmenovány na základě taxonomie organismů, v nichž byly objeveny. První písmeno názvu enzymu odpovídá prvnímu písmenu rodového jména organismu, zatímco druhá dvě písmena názvu enzymu jsou prvními dvěma písmeny druhového jména. Mohou být přidána další písmena či čísla pro indikaci individuálního kmene nebo sérotypu daného organismu. Na konec názvu je připojena římská číslice odlišující jednotlivé enzymy, pokud jich bylo v organismu objeveno více. Například enzym *HindIII* byl izolován z bakterie *Haemophilus influenzae* (Lehmann and Neumann, 1896) Winslow et al., 1917 sérotypu d a je odlišný od enzymů *HindI* a *HindII*, které jsou v bakterii také přítomny.

Roberts et al. (2003) ve své publikaci upravuje původní názvosloví. Navrhuje několik změn obecných pravidel, mezi nimiž vystupuje záměr zrušit užívání kurzívy pro třípísmenný akronym v názvu restrikčních endonukleáz i methyltransferáz. Tento krok obhájí třemi argumenty: velké množství časopisů kurzívu již vůbec nepoužívá; zachování úmluvy o užívání kurzívy se špatně převádí do počítačů; kurzíva neslouží k žádnému zásadnímu účelu. Toto pravidlo se od roku 2003 v České republice neuchytilo a tradice psaní kurzívou stále pokračuje. Ostatní změny již nejsou tak kontroverzní. Kolektiv autorů se například shoduje na užívání zkratk, ruší psaní znaku střední tečka „·“ před předponami a další.

Není-li z kontextu jasné, o jaký typ proteinu se jedná, přidává se k názvu předpona. „R“ značí restrikční endonukleázy, „M“ methyltransferázy – pokud je přítomno více jednotlivých methyltransferáz, značí se „M1“, „M2“, pro spojené methyltransferázy užíváme značku „M1~M2“. Předpona „RM“ označuje enzymy, v nichž se restrikční a modifikační aktivity objevují ve stejném polypeptidovém řetězci. Předponou můžeme označit také doplňkové proteiny: „N“ pro klasické vrubovací enzymy (štěpí pouze jeden řetězec DNA), „V“ pro vrubovací enzymy štěpící nesouhlasné vazby G/T (vzniklé deaminací m5C – methylovaný cytosin – uvnitř rozpoznávané sekvence methyltransferázy) a „C“ pro kontrolní proteiny (Roberts et al. 2003).

### **3.1.2 Stavy enzymů**

Postupné objevování jednotlivých restričních enzymů je zaznamenáno na webu REBASE – Databáze restričních enzymů a datuje se od roku 1968 do současnosti. První výrazný nárůst objevů nastal v roce 1985 (272 enzymů), dále v roce 1990 došlo k objevu nejvyššího počtu potvrzených enzymů vůbec (512 enzymů). S rokem 1997 začalo přibývat enzymů, které byly identifikovány z analýzy mikrobiálních genomických sekvencí (tedy nebyly ještě potvrzeny), a od roku 2002 již objevy těchto předpokládaných enzymů zcela převažují nad potvrzenými a popsányými enzymy (Roberts 2016a). Ke 28. 1. 2016 bylo zaevidováno již 4 205 regulérních a 38 917 experimentálně nepotvrzených restričních enzymů (Roberts 2016b).

### **3.1.3 Význam**

Restriční endonukleázy v průběhu let změnilы tvář molekulární biologie tak, jak si to před několika desetiletími nedovedl nikdo představit. Díky několika schopným vědcům se restriční enzymy staly nástrojem k analýze DNA, klonování a mapování genů a tvorbě rekombinantní DNA. Technologie využívající endonukleázy přetvořily přírodní vědy a medicínu a daly vzniknout velkému množství společností, což mělo významný dopad na rozvoj ekonomiky (Pingoud, Wilson a Wende 2014).

Z konkrétních významných událostí spojených s restričními endonukleázami je možné jmenovat například produkci inzulínu z rekombinantních bakterií a kvasinek, čímž se výrazně zvětšily zásoby inzulínu pro diabetiky. Důležitá je také produkce rekombinantní vakcíny na hepatitidu B nebo překonstruování RES pro výrobu umělých nukleáz, které mají potenciál pro využití v genové terapii a metodě cílení genů (Loenen et al. 2014a).

### **3.1.4 Efektivita restričních endonukleáz**

Jakmile cizí DNA vstoupí do organismu (např. virová DNA do bakteriální buňky), velmi záleží na tom, který prvek R–M systému se k ní dostane dříve. Jestliže je cizí DNA nalezena restriční endonukleázou, bude rozštěpena. Bude-li však cizí DNA nejdříve methylována, restriktce již nemůže proběhnout a bakterie bude napadena. Typická virová DNA obsahuje 10–20 rozpoznatelných míst pro RES i MT. Aby nedocházelo často k virovým infekcím, stačí endonukleáze rozštěpit cizí DNA pouze jednou či dvakrát, kdežto MT by musela modifikovat všechna tato místa pro zabránění restriktce. RES tedy mají jasnou výhodu.

R–M systémy všech typů byly zkoumány měřením efektivity tvorby plak fágů (EOP = efektivita výsevu) na různých bakteriálních hostitelích. EOP se dá vyložit také jako míra pravděpodobnosti, že MT modifikuje všechny rozpoznané sekvence před tím, než RES rozštěpí pouze jedno toto místo.

Stejně jako plně methylovaná DNA je i hemi-methylovaná DNA, vzniklá replikací plně methylované DNA, imunní k rozštěpení. Proběhla-li by však další replikace, stala by se jedna z výsledných molekul náchylnou ke štěpení. Proto je třeba, aby byla methyltransferáza vždy přítomna při replikaci hostitelské DNA (Wilson et al. 2012).

## 3.2 Typy restričních endonukleáz

### 3.2.1 Typ I

Restričně – modifikační systémy Typu I (konkrétně *Escherichia coli* K-12 (Migula, 1895) Castellani a Chalmers, 1919 a *E. coli* B) byly objeveny jako první ze všech R–M systémů, nicméně byly zastíněny systémy Typu II, jelikož ty, na rozdíl od Typu I, štěpí na samostatné fragmenty a jsou tedy lépe použitelné (Murray 2000).

R–M systémy Typu I jsou kódovány třemi geny, a tudíž obsahují tři podjednotky. Gen *hsdR* kóduje tvorbu restriční podjednotky HsdR. Gen *hsdM* kóduje modifikační podjednotku HsdM a gen *hsdS* zase rozpoznávací podjednotku HsdS (S = specifita). Jedná se tedy o komplexní systémy schopné rozpoznání, modifikace i štěpení DNA. Enzymy Typu I jsou také schopny translokovat DNA do nahodilého spojení s výraznou změnou konformace molekuly, přičemž jsou stále napojeny na rozpoznávací místo.

Výhodou R–M systémů Typu I oproti Typu II z biologického hlediska je zabránění bakteriofágům získat rezistenci na činnost restričních endonukleáz. Viry a jiné mobilní genetické elementy dokáží získat požadovanou modifikaci nebo syntetizovat své vlastní methyltransferázy, umí také modifikovat svou DNA včleněním nezvyklých bází, případně syntetizovat antirestriční proteiny, které napodobují DNA, a zvládnou se také vyvarovat specifických rozpoznávacích sekvencí. Řešením pro hostitele je periodická změna specifity rozpoznávací sekvence, a to jak pro restriční enzymy, tak pro methyltransferázy. R–M systémy Typu I vyvinuly účinný způsob, jak tuto změnu provést synchronizovaně, díky rozpoznávací podjednotce S, která sestává ze dvou individuálních domén TRD (target recognition domain). Tyto domény mohou díky genetickému přeskupení vytvořit nové kombinace a tím účinně mnohokrát změnit specifitu. Množství různých specifit potenciálně vytvořených těmito mikroby daleko přesahuje množství *hsdS* genů, které vlastní. Z 22 genů *hsdS* může

vzejít až 500 rozličných specifit, což představuje mimořádný obranný repertoár (Loenen et al. 2014b).

### 3.2.1.1 Rodiny

Restrikční enzymy Typu I jsou rozděleny do pěti rodin označených písmeny A–E (Tab. P2). Rozdělení proběhlo na základě testů komplementarity kódujících genů, cross-reaktivity protilátek a aminokyselinové sekvence (Loenen et al. 2014b).

Základní enzymy skupiny IA *EcoKI* a *EcoBI* jsou kódovány komplementárními geny, které se v genetických testech spolu chovají jako dvě alely, nicméně mají velmi rozdílné sekvence nukleotidů. Protilátky proti enzymům skupiny IA fungují pouze v rámci skupiny, neovlivňují však enzymy ze skupiny jiné, tzn., že nejsou cross-reaktivní. Skupinu IB reprezentuje enzym *EcoAI* kódovaný geny bakterie *E. coli* 15T<sup>-</sup>. Třetí skupina IC představuje enzymy kódované plazmidy a hlavním prototypem je zde endonukleáza *EcoR124I*. Do ID patří například enzym *StySBLI*, který se nachází v bakterii *Salmonella enterica* (ex. Kauffmann a Edwards, 1952) Le Minor a Popoff, 1987 sérovar *blegdam* (Murray 2000). Prozatím poslední skupinu IE charakterizuje enzym *KpnBI* z *Klebsiella pneumoniae* Trevisan, 1885. Nepochybně existují další skupiny Typu I, nicméně dosud nebyly prozkoumány, jelikož ještě donedávna byly tyto enzymy těžko charakterizovatelné. S příchodem sekvenování genomu, bioinformatiky a hlavně nejnovější analýzy methylomů (soubor methylových modifikací nukleových kyselin v genomu organismu) realizované Single-Molecule Real-Time sekvenováním (SMRT, sekvenování jednotlivých molekul v reálném čase) se zpřístupnily možnosti identifikace enzymů a určení jejich rozpoznávaných sekvencí (Loenen et al. 2014b).

### 3.2.1.2 Enzymatická aktivita

Enzymy R–M systémů Typu I jsou pentamerické proteiny skládající se ze dvou restrikčních, dvou modifikačních podjednotek a jedné rozpoznávací podjednotky (2R+2M+S). Pro své fungování požadují přítomnost ATP, hořčnatých iontů a S-adenosylmethioninu (SAM nebo AdoMet). Kromě velkého R–M komplexu mají bakterie také trimer 2M+S, který vystupuje pouze jako methyltransferáza.

Enzymy rozpoznávají dvojstranné sekvence DNA skládající se ze dvou polosekvencí rozdělených řadou nespecifických nukleotidů (N). Každá polosekvence je rozpoznána jednou TRD. Příklad:

**AACNNNNNGTGC (AAC N<sub>6</sub> GTGC)**

Podjednotka R kontaktuje jádro trimeru 2M+S prostřednictvím podjednotky M. Není-li ani jedna polosekvence rozpoznávaného místa methylována, restriční endonukleáza translokuje DNA a štěpí ji v různé vzdálenosti, přibližně v polovině mezi sousedícími rozpoznávanými místy, díky čemuž nevytváří žádné typické vzorce fragmentů. V důsledku toho nemohou být v buněčných extraktech jednoduše identifikovány a jejich rozpoznávaná místa se velmi těžko mapují (Loenen et al. 2014b).

### 3.2.2 Typ II

Skupina restričních endonukleáz Typu II je konglomerace mnoha různých proteinů, které mají společnou schopnost štěpit dvoušroubovici DNA ve fixní pozici uvnitř rozpoznávané sekvence, nebo v její blízkosti. Toto štěpení vytváří reprodukovatelné fragmenty DNA a předvídatelné vzorce při provedení gelové elektroforézy. Pro tyto své vlastnosti se enzymy Typu II staly neocenitelnými laboratorními prvky pro manipulaci s DNA a pro její zkoumání.

Geny kódující produkci R–M systémů Typu II se nacházejí na chromozomech, příležitostně však i na přenosných elementech, jako třeba na plazmidech, transpozonech a inzertovaných sekvencích (Pingoud, Wilson a Wende 2014).

První objevenou restriční endonukleázou Typu II se stala *HindII* izolovaná z bakterie *Haemophilus influenzae* sérotypu d v laboratoři Hamiltona Smithe na Johns Hopkins Medical School (Kelly a Smith 1970). Zajímavostí je, že Smith tehdy netušil o přítomnosti druhé RES *HindIII* v prvním preparátu *HindII*. Přítomnost *HindIII* tehdy mohla těžce narušit analýzu rozpoznávané sekvence *HindII*. K velkému štěstí Kellyho a Smithe byla jako substrát použita DNA fága T7, která neobsahuje restriční místo pro *HindIII* (Pingoud, Wilson a Wende 2014).

Jak postupem času následovaly další a další objevy nových enzymů, vyšlo najevo, že se od sebe na úrovni aminokyselinového složení všechny velmi liší. Stále nedošlo ke všeobecné dohodě, co tato diverzita z evolučního hlediska znamená. S přihlédnutím k rychlosti metabolismu prokaryot, jejich velkému množství nik, rychlému množení a nekonečné délce života je nadmíru pravděpodobné, že při formování jejich nynější podoby došlo ke všem možným evolučním scénářům, a to nejspíše mnohokrát, na mnoha místech, nezávisle na sobě (Pingoud, Wilson a Wende 2014).

Vyvstává tedy otázka, jak klasifikovat enzymy Typu II, když vykazují takovou míru diference. Dnešní přístup, navržený Robertsem (Roberts et al. 2003) a přijatý na konsenzu před deseti lety, sdružuje skupiny enzymů dle fenotypu, jež je založen

na jejich chování a na vlastnostech štěpení. Tento způsob však má své nedostatky a zřejmě bude v nedaleké budoucnosti přepracován. Další metody klasifikace zahrnují dělení enzymů dle podobnosti struktury nebo organizace podjednotek a domén, případně dle podobnosti restričních míst (Pingoud, Wilson a Wende 2014).

Dělení enzymů Typu II dle chování je založeno na jedenácti subtypech (A, B, C, E, F, G, H, M, P, S, T), které jsou definovány specifickým znakem. Ten však nemusí být nutně jedinečný. Díky tomuto přístupu se může stát, že jeden enzym bude patřit do více než jednoho subtypu (Pingoud, Wilson a Wende 2014).

### 3.2.2.1 *Subtyp IIA*

Restriční endonukleázy zařazené do skupiny IIA rozpoznávají asymetrické sekvence a štěpí buď uvnitř dané sekvence, nebo v definované vzdálenosti od ní. Mnoho enzymů je doprovázeno jednou či dvěma methyltransferázami, z nichž každá modifikuje jeden řetězec rozpoznávané sekvence. Ostatní enzymy mají kombinované R–M vlastnosti. K některým z nich se ještě přidává separovaná methyltransferáza (Pingoud, Wilson a Wende 2014).

Rozpoznávaná sekvence a štěpná místa jsou reprezentována na následujícím příkladu: *HphI*: GGTGA (8/7) – první číslo v závorce udává pozici štěpení na pozitivním vlákně a druhé číslo udává štěpnou pozici na komplementárním vlákně (Roberts et al. 2003).

5' **GGTG**ANNNNNNNN ↓ 3'

3' **CCACT**NNNNNNN ↑ 5'

Dalo by se předpokládat, že vzhledem k častějšímu výskytu asymetrických rozpoznávaných sekvencí v DNA se budou endonukleázy IIA uplatňovat hojněji než IIP, které štěpí symetrické sekvence. Není tomu tak, jelikož IIP enzymy na sebe berou podobu homodimerů, čímž nejefektivněji využívají genetické zdroje. Většina IIA enzymů zřejmě také kvůli štěpení dimerizuje, nicméně pouze krátce (Bitinaite et al. 1998).

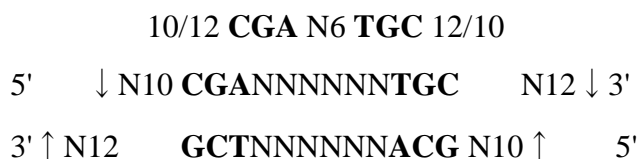
V případě separovaných R a M komponentů se rozpoznávací doména naváže na DNA individuálně a asymetricky, přičemž restriční doména tvoří dočasný dimer s identickou doménou z jiné molekuly. Toto chování naznačuje, že štěpení je kooperativní. Míra štěpení často nepoměrně roste se zvyšující se koncentrací enzymů a je obvykle vyšší na substrátech s větším množstvím rozpoznávaných sekvencí.

Ne všechny IIA enzymy využívají dimerizace. Některé pracují se dvěma rozdílnými katalytickými místy pocházejícími z různých podjednotek (např. enzym *BbvCI*) nebo z různých domén uvnitř stejného proteinového řetězce (např. *Mva1269I*). Tyto enzymy mají sklon ke štěpení uvnitř sekvence či velmi blízko ní, což jsou pozice nepřístupné pro oddělené dimerové katalytické domény (Pingoud, Wilson a Wende 2014).

### 3.2.2.2 *Subtyp IIB*

Enzymy ze skupiny IIB jsou typické štěpením DNA na obou stranách rozpoznávané sekvence. Vytváří tak menší fragment o 20–34 bp, který obsahuje rozpoznávanou sekvenci. Jsou to velké komplexní R–M enzymy, jež nejsou doprovázeny žádnou další methyltransferázou. IIB enzymy mají mnoho společných rysů s endonukleázami Typu I a jsou s nimi příbuzné.

Specifická rozpoznávací sekvence je dvojstranná – obsahuje dvě půlsekvence rozdělené od sebe krátkou řadou nespecifických bází. Zde je příklad enzymu *BcgI*:



Výsledkem štěpení jsou fragmenty s kohezními konci, které na 3' konci přesahují dvěma až pěti bázemi. Modifikační jednotka enzymů IIB methyloje rezidua adeninu. Pro správný průběh methylace a u některých enzymů i pro průběh restrikce je požadován AdoMet.

IIB enzymy vykazují obecně tři typy chování, pokud se setkají s rozpoznávanou sekvencí. V neutrálním módu sekvenci ignorují, v restrikčním módu štěpí a v modifikačním módu methylují. Vliv na chování enzymů má zřejmě stav rozpoznávané sekvence (Pingoud, Wilson a Wende 2014). Pokud jsou obě polosekvence methylované, jsou enzymem ignorovány, jelikož indikují plně modifikovanou hostitelskou DNA. Je-li sekvence methylována pouze na jednom z komplementárních řetězců, jedná se o nově replikovanou hostitelskou DNA, jejíž přítomnost by měla indukovat modifikační činnost enzymu. Narazí-li enzym na nemethylované rozpoznávané místo na DNA, je zde velká pravděpodobnost, že jde o cizí DNA. Oproti enzymům subtypu IIP, které ihned takovouto sekvenci štěpí, enzymy subtypu IIB nejdříve hodnotí situaci. Jestliže naleznou více nemodifikovaných sekvencí na stejné molekule DNA, potvrdí se předpoklad a dojde ke štěpení. Je zde však



možnost, že nedopatřením nedošlo k methylaci pouze této jediné sekvence, což enzym vyhodnotí jako nemodifikovanou hostitelskou DNA a zahájí methylaci.

Enzymy *BcgI*, *CspCI* a *BsaXI* vlastní samostatné podjednotky S, které sestávají ze dvou protilehlých rozpoznávacích domén, z nichž každá se váže na jednu polosekvenci. Podjednotky těchto enzymů jsou 2RM+1S formátu, ale existují pouze jako heterohexamery – tedy dimery složené z trimerů 2×(2RM+1S), a to jak v roztoku, tak ve vazbě s DNA. Pro proběhnutí štěpení je nutné, aby byl enzym navázán na dvě rozpoznávané sekvence, přičemž k restrikci dojde na všech čtyřech řetězcích zároveň. Co se týče modifikace, váže se *BcgI* a jemu podobní na jednotlivé sekvence také jako hexamer. Hemimethylované sekvence modifikuje rychleji než nemethylované, takže představuje spíše udržovací MT (Pingoud, Wilson a Wende 2014).

Ostatní IIB enzymy, jako například *AloI*, *PpiI* a *CjeI*, jsou jednořetězcové proteiny, které také obsahují dvě rozpoznávací domény (Pingoud, Wilson a Wende 2014). Není zcela jasné, jak se dvě rozpoznávací domény dělí o jednu katalytickou, nicméně Smithová, Pernstich a Halford (Smith, Pernstich a Halford 2014) při zkoumání enzymu *TstI* došli k závěru, že takovéto proteiny zřejmě vystupují jako homotetramery, které se vážou na dvě dvojstranné rozpoznávané sekvence spíše než na čtyři. Většina IIB enzymů je tedy aktivní jako endonukleáza, až když je navázána na dvě rozpoznávaná místa, a to buď v konfiguraci *cis* na stejné molekule DNA, nebo v konfiguraci *trans* na dvou oddělených molekulách DNA.

Stejně jako u endonukleáz Typu I fungují u skupiny IIB rozpoznávací domény individuálně, takže se mohou mezi sebou vyměňovat, aby došlo k vytvoření nových kombinací rozpoznávaných sekvencí (Pingoud, Wilson a Wende 2014).

Aby mohly enzymy IIB štěpit četné DNA duplexy naráz, potřebují se shlukovat do velkých oligomerů s molekulovou hmotností přesahující 500 kDa, což je činí největšími známými restrikčními endonukleázami (Pingoud, Wilson a Wende 2014).

### 3.2.2.3 *Subtyp IIC*

Enzymy ze skupiny IIC mají uvnitř jednoho polypeptidového řetězce restrikční i modifikační domény. Zahrnují všechny známé IIB restrikční endonukleázy, ale i další typy enzymů. Mohou a nemusí být doprovázeny doplňkovými methyltransferázami.

Enzymy IIC štěpí vždy mimo svou rozpoznávanou sekvenci, a to v různé vzdálenosti s odchylkou  $\pm 1$  nebo  $\pm 2$  bp. Rozpětí mezi rozpoznávanou sekvencí a restrikčním místem je spíše dáno fyzickou vzdáleností, než počtem párů bází,

kteří leží mezi nimi. Tím pádem má vliv na vzdálenost štěpení hlavně topologie DNA, ionické podmínky a sekvence párů bází. U enzymů provádějících restrikci na obou stranách dvojstranné sekvence se vzdálenost měří spíše od modifikovaného adeninu nacházejícího se v každé polosekvenci, než od hranic rozpoznávané sekvence. Často je to stejný adenin jak napravo, tak nalevo, jelikož stejný protein katalyzuje reakci na obou stranách.

Většina IIC endonukleáz se váže na cílové sekvence ve formě monomeru. Některé enzymy rozpoznávají kontinuální asymetrické sekvence, které štěpí na jednom místě přibližně ve vzdálenosti jedné otočky šroubovice DNA – *Tth111II* (CAARCA 11/9, R = A nebo G), případně ve vzdálenosti jedné a půl otočky – *Eco57I* (CTGAAG 16/14) anebo až na místě vzdáleném dvě otočky – *MmeI* (TCCRAC 20/18). Jiné enzymy, včetně subtypu IIB, mají dvě rozpoznávací domény, které rozpoznávají dvojstranné asymetrické sekvence a štěpí je na obou jejich stranách. Několik enzymů, např. *HaeIV* (7/13 GAY N5 RTC 14/9, Y = T nebo C), se váže ve formě homodimerů, přičemž používají dvě kopie stejné rozpoznávací domény. Jejich rozpoznávané sekvence jsou také dvojstranné, ovšem jsou symetrické.

R–M proteiny subtypu IIC vlastní pouze jedno katalyzační místo pro štěpení, přesto však štěpí oba řetězce DNA. Štěpení totiž zahrnuje přechodnou dimerizaci mezi štěpnými doménami sousedících molekul. Obsahuje-li substrát více rozpoznávaných míst nebo jsou-li přidány oligonukleotidy obsahující danou sekvenci, zlepšuje se efektivita štěpení. Větší množství míst má zřejmě za následek zvýšení lokální koncentrace enzymů, což podporuje krátkodobou dimerizaci (Pingoud, Wilson a Wende 2014).

#### **3.2.2.4 Subtyp IIE**

Restrikční endonukleázy subtypu IIE jsou enzymy skupiny IIP s alosterickým efektem, který stimuluje katalýzu, pokud se naváže na další rozpoznávanou sekvenci. Biologická funkce alosterické aktivace je nejasná, ale je možné, že by měla zabránit štěpení buněčné DNA na neúmyslně nemodifikovaných místech v průběhu doby, kdy methylace DNA není schopna udržet tempo se syntézou DNA. Tato schopnost může být také vlastní enzymům subtypu IIF.

Příkladem může být enzym *EcoRII* (|CCWGG), který se jako homodimer váže na dvě nebo tři kopie rozpoznávané sekvence najednou. Jedna z těchto sekvencí je koordinovaně vázána dvěma podjednotkami v DNA-rozlišovacím žlábků mezi nimi a bude štěpena. Báze A nebo T (=W) jsou při štěpení nečekaně vyjmuty ven ze

struktury šroubovice, takže z rozpoznávané sekvence zůstane pouze CC-GG. Toto chování vykazují i podobné enzymy *PspGI* (|CCWGG) a *Ecl18kI/SsoII* (|CCNGG). Ostatní sekvence se vážou jiným způsobem než štěpená sekvence, a to na N-koncové domény, které působí jako alosterické aktivátory. Dalším příkladem endonukleázy IIE je *NaeI* (GCC|CCG), která má funkční spojitost s DNA topoizomerázami (Pingoud, Wilson a Wende 2014).

### 3.2.2.5 Subtyp IIF

Enzymy patřící do skupiny IIF se vážou na dvě rozpoznávané sekvence a koordinovaně štěpí všechny čtyři DNA řetězce naráz. Je zde určitá podobnost s enzymy subtypu IIE, jelikož ty se také vážou na více rozpoznávaných sekvencí, nicméně pouze jedna z těchto sekvencí je opravdu štěpena a ostatní jsou alosterické. Enzymy subtypu IIF naproti tomu vystupují jako homotetramery. Jejich vazná místa jsou identická, mají samozřejmě katalytickou funkci, ale také alosterické či kooperativní vlastnosti.

Některé IIF enzymy rozpoznávají asymetrické sekvence (IIB *TstI*, IIS *BspMI*), jejichž strukturální organizace zůstává nejasná. Jiné IIF enzymy rozpoznávají symetrické sekvence jako páry tzv. „back-to-back“ homodimerů, které nejsou schopny štěpit, dokud obě vazná místa nejsou obsazena. Pro lepší stabilitu je preferována vazba daných sekvencí v konformaci *cis*. Není jasné, jak si katalytická místa mezi sebou předávají signál k vazbě na DNA, zřejmě to však zahrnuje konformační změny, které se šíří napříč tetramerovým rozhraním. U enzymů *Bse634I* a *SfiI* se vyskytla mutace v tomto rozhraní, která jim dovolila vázat se na jednotlivou sekvenci a efektivně ji štěpit, přičemž enzymy zůstaly zachovány v tetramerické formě.

Zvláštností enzymu *SgrAI* (CR|CCGGYG) je, že po sestavení do homotetrameru pokračuje formováním průběžných oligomerů tvořených řadou helikálně uspořádaných homodimerů navázaných na DNA. Sousední homodimery jsou od sebe odchýleny přibližně o 90 ° tak, že čtyři homodimery společně tvoří téměř jednu otočku levotočivé šroubovice, která se může skládat až z osmnácti homodimerů. Oligomerická forma *SgrAI* je vysoce aktivní jak na vlastní základní sekvenci, tak na „star“ sekvenci (CR|CCGGYN, N = libovolná báze), která má jen o 4 % nižší efektivitu štěpení než základní sekvence. Z toho vyplývá, že homodimer enzymu *SgrAI* je asymetrický v tom smyslu, že jedna podjednotka důsledně rozpoznává vnější pár bází dané sekvence, kdežto druhá podjednotka není vždy tak důsledná, což může být způsobeno výraznými konformačními změnami při sestavování oligomerů.

Dalším netypickým úkazem je skutečnost, že příbuzné RES se mohou chovat jako IIE i jako IIF enzymy – např. *EcoRII* (|CCWGG, TypIIE) a *Ecl18kI/SsoII* (|CCNGG, Typ IIF). A evolučně odlišné verze stejného enzymu se mohou chovat odlišně. Příkladem je *Cfr42I* a *Eco29kI*, které rozpoznávají CCGC|GG. *Cfr42I* má v roztoku formu tetrameru, váže se na dvě rozpoznávané sekvence a štěpí současně. *Eco29kI* naopak formuje v roztoku monomer, váže se na DNA jako homodimer a dvojici sekvencí štěpí vždy postupně (Pingoud, Wilson a Wende 2014).

### 3.2.2.6 Subtyp IIG

Pro enzymy subtypu IIG je důležitou prerekvizitou přítomnost sloučeniny AdoMet. Jsou jí buď stimulovány, nebo ji přímo vyžadují. Do této skupiny patří většina IIB a IIC enzymů. Typickým znakem těchto kombinovaných R–M enzymů je přítomnost restriční domény a methylační domény třídy gama v jediném polypeptidovém řetězci. K R a M doméně náleží jediná S doména, která se vyskytuje v podobě samostatné podjednotky, nebo může být součástí C-terminální části proteinu a rozpoznává jak kontinuální, tak dvojstranné (nespojité) sekvence, z nichž každá může být symetrická či asymetrická. Enzymy subtypu IIG se vyskytují v celé škále oligomerických forem bez doprovodných methyltransferáz nebo s nimi.

AdoMet je pro IIG enzymy důležitý z hlediska restrikce, ale i z hlediska methylace, protože poskytuje methylovou skupinu pro methylační reakci. Výhodou závislosti na AdoMetu může být sebezáchova – při nedostatku AdoMetu, z něhož vyplývá nedostatečná modifikace DNA, se snižuje pravděpodobnost štěpení genetické informace vlastní hostitelské buňky. Na nevýhodu lze však narazit u štěpení *in vitro*, jež často nelze dokončit, protože část sekvencí bude modifikována již při inkubaci, což je činí rezistentní k následnému štěpení.

Stane-li se, že se změní sekvence rozpoznávaná IIG enzymy, a to například mutací S podjednotky nebo výměnou domén, stane se tak pro restrikci i modifikaci zároveň. Tato funkční synchronizace umožnila specifickým sekvencím určitých IIG enzymů (např. rodina *MmeI*) široce se od sebe rozlišit. V kostře proteinové struktury rozpoznávací domény se vyvinula struktura dovolující různým aminokyselinovým párům specifikovat alternativní páry bází v rozpoznávané sekvenci (Tab. P3). Toto chování je pro RES netypické, jelikož se jako celek vyvíjely v naprosto opačném směru, a to k neměnnosti rozpoznávaných sekvencí.

Většina prokaryot vlastní pouze jednu nebo dvě endonukleázy subtypu IIG spolu s velkým množstvím enzymů Typu I, IIP, IIS a s o něco menším množstvím enzymů Typu III. Lze však nalézt i výjimky, jako například *Helicobacter pylori* (Marshall et al., 1985) Goodwin et al., 1989, který má často až 20 systémů subtypu IIP a IIS, ale jen čtyři až šest IIG systémů. *Borrelia burgdorferi* Johnson et al., 1984 naopak obsahuje ke dvaceti systémům subtypu IIG a ostatní systémy nemusí obsahovat vůbec. O výhodách a nevýhodách těchto variací je známo poměrně málo (Pingoud, Wilson a Wende 2014).

### 3.2.2.7 Subtyp IIH

Skupina IIH obsahuje enzymy, které jsou označovány jako hybridní, jelikož se skládají z restričních endonukleáz náležících Typu II, ale doprovází je adeninové methyltransferázy třídy gama, o kterých se dříve myslelo, že náleží pouze systémům Typu I. Dnes je však již známo, že tyto MT jsou obecně rozšířené i mezi endonukleázami Typu II – např. tvoří jádro většiny IIG endonukleáz.

Prvním objeveným zástupcem této skupiny je enzym *AdhI* (GACNNN|NNGTC), jehož methyltransferáza modifikuje adenin (GAC N GTC) na N6-methyladenin a funguje v podobě tetrameru 2M+2S. Po objevu byl zařazen do neformální skupiny Typu I ½, jelikož se zdálo, že je to chybějící článek evolučního vývoje mezi enzymy Typu I a Typu II. Nyní se tomuto typu endonukleáz již nepřikládá takový význam a není již příliš využíván (Pingoud, Wilson a Wende 2014).

### 3.2.2.8 Subtyp IIM

Enzymy subtypu IIM rozpoznávají pouze methylované sekvence, které štěpí na fixních místech. Co do IIM však neřadíme, jsou enzymy ze skupiny IV (např. enzym *McrBC*), protože neštěpí na fixních místech a mohou být k methylaci méně či více tolerantní, ačkoliv pro RES subtypu IIM je methylace naprosto esenciální (Pingoud et al. 2004).

Nejznámějším zástupcem této skupiny je enzym *DpnI* (Gm6A|TC), který ve formě monomeru štěpí jeden řetězec po druhém. *DpnII* je alelickou alternativou k *DpnI*, takže *in vivo* štěpí stejnou sekvenci, která však musí být nemethylovaná (GATC).

Není jasné, co stojí za plnou závislostí na methylovaném adeninu v rozpoznávané sekvenci *DpnI*. Je však možné, že methylová skupina indukuje strukturální změny (např. pozměnění struktury postranního řetězce aminokyselin s dlouhým řetězcem, jako je lysin nebo arginin), jimiž upraví konformaci proteinu, která narušuje vazbu nebo jí brání, na konformaci, která je kompatibilní a umožňuje navázání (Pingoud, Wilson a Wende 2014).

### 3.2.2.9 Subtyp IIP

Subtyp IIP obsahuje ty nejrozšířenější a nejrozmanitější enzymy z endonukleáz Typu II. Členové IIP skupiny rozeznávají palindromické sekvence a štěpí buď uvnitř těchto sekvencí (*EcoRI* G|AATTC), nebo, méně často, na jejich okraji (*EcoRII* |CCWGG). Palindromy jsou symetrické sekvence, které mají na obou řetězcích stejné pořadí bází, a to ve stejném směru (5'→3' nebo 3'→5').

5' G|AATTC 3'

3' CTTAA|G 5'

Endonukleázy subtypu IIP jsou téměř vždy doprovázeny jednou methyltransferázou rozpoznávající stejnou sekvenci. V ojedinělých případech jsou methyltransferázy dvě. Některé RES fungují ve formě monomerů, nejčastěji na sebe však berou podobu homodimerů a homotetramerů. Multimery většinou štěpí oba řetězce v jedné vazebné události. Monomery štěpí jeden řetězec DNA po druhém, nicméně po štěpení prvního řetězce zůstávají stále ve vazbě, což nasvědčuje tomu, že oba řetězce štěpí jediná molekula. Z důvodu polarizace katalytické reakce a opačné orientace řetězců se však monomer musí po prvním štěpení oddělit od rozpoznávané sekvence, aniž by přitom zrušil vazbu s DNA, poté se otočí o 180° a znovu se připojí k rozpoznávané sekvenci, aby zde mohl štěpit.

Například monomerní enzym *BcnI* (CC|SGG, kdy S = G nebo C) v prvním kroku rychle provede hydrolýzu jednoho z řetězců. Zde byla zjištěna nepatrná preference řetězce obsahujícího v místě S guanin místo cytosinu. Poté se pomalu odpojí, otočí a vrátí ke druhému řetězci. V posledním kroku stejným způsobem hydrolyzuje zbylý řetězec.

Rozpoznávané sekvence tohoto typu endonukleáz jsou obvykle 4–8 bp dlouhé. Mohou být kontinuální (*HindIII* A|AGCTT) či přerušené nespecifickými bazickými páry, a to jedním (*HinfI* G|ANTC), dvěma (*Hpy188III* TCN|NGA), třemi (*DraIII* CACNNN|GTG), čtyřmi (*XmnI* GAANN|NNTTC), pěti (*BglI* GCCNNNN|NGGC) a více, až do počtu devíti nespecifických párů bází (*XcmI* CCANNNNN|NNNTGG). Tyto rozdíly se odvíjí od geometrického vztahu mezi dvěma podjednotkami homodimeru. Rozpoznávané sekvence mohou být tvořeny pouze jedním párem bází (*XmaI* C|CCGGG, *DraI* TTT|AAA) či oběma páry. Mnoha enzymům však mohou vyhovovat alternativní páry bází na určitých pozicích, jako například R:Y (purin:pyrimidin = A:T nebo G:C), W:W (slabé párování bází = A:T nebo T:A), S:S (silné párování bází = G:C nebo C:G) a M:K (methylatelná báze = A:T nebo

C:G). Štěpení může produkovat fragmenty s tupými konci (bez přesahů), nebo s konci kohezními, které mají přesahy 1–4 bází na 5' nebo 3' konci.

Dnes jsou již známy stovky různých specifit pro enzymy subtypu IIP. Pro každou z nich existuje alespoň několik izoschizomerů – enzymů s identickou specifitou a podobným aminokyselinovým složením, které často představují odlišné varianty jednoho původního enzymu, jehož gen se horizontálně přenášel mezi prokaryoty a v průběhu času akumuloval neutrální změny. Skupiny endonukleáz s úzce příbuznými specifitami mají často podobné aminokyselinové složení, což ukazuje na nedávnou adaptivní radiaci ze společného předka. Enzymy s nepříbuznými specifitami nevykazují podobnost v AMK složení, což může být znakem individuálního vývoje, nebo vymizení stop po společném předkovi následkem plynutí času (Pingoud, Wilson a Wende 2014).

### **3.2.2.10 Subtyp IIS**

Enzymy subtypu IIS štěpí DNA ve fixní pozici mimo jejich rozpoznávanou sekvenci. Štěpení je posunuto v jednom směru od sekvence v rámci jedné či dvou otoček dvojité šroubovice. Do této skupiny lze technicky zařadit i enzymy skupin IIB, IIC a IIG, které se nicméně dále liší od zbytku enzymů IIS.

Rozpoznávané sekvence enzymů subtypu IIS jsou většinou asymetrické, a to zřejmě z toho důvodu, že existuje mnohem více asymetrických sekvencí, než těch symetrických. Štěpení proto probíhá pouze na jedné straně sekvence. Nejznámější a jedna z prvních objevených IIS endonukleáz *FokI* štěpí v místě vzdáleném devět bází od sekvence na jednom řetězci a třináct bází daleko od sekvence na druhém řetězci, čímž produkuje fragmenty s přesahem čtyř bází na 5' konci. Na základě dohody se rozpoznávaná sekvence enzymů subtypu IIS uvádí ve směru, ve kterém ke štěpení dochází napravo od sekvence, takže specifická sekvence enzymu *FokI* je psána stylem GGATG 9/13, nikoliv 13/9 CATCC, ačkoliv je tento způsob stejně přesný.

Aminokyseliny zodpovídající za rozpoznávání a za katalýzu jsou umístěny v odlišných doménách, které mohou být rozděleny na samostatné proteinové řetězce. Enzymy IIS se tedy skládají z N-koncové rozpoznávací domény, spojovací části neboli „paže“ a C-koncové štěpící domény bez sekvenční specifity. Obvykle jsou enzymy subtypu IIS doprovázeny dvěma samostatnými methyltransferázami, z nichž každá modifikuje jeden řetězec rozpoznávané sekvence tak, že methyluje jeden adenin a jeden cytosin. V některých případech (např. u *FokI*) jsou methyltransferázy spojeny do jednoho proteinového řetězce. Pouze několik R–M systémů subtypu IIS vlastní jen

jednu doprovodnou methyltransferázu. Tyto systémy rozpoznávají tzv. kvazipalindromické sekvence, které jsou asymetrické z pohledu restriční endonukleázy, nicméně methyltransferázy je vnímají jako symetrické (a dvojznačné). V důsledku toho jsou jedinou methyltransferázou modifikovány oba řetězce dané sekvence. Příkladem těchto systémů jsou *BbvI* (RES: GCAGC 8/12, MT: GCWGC) a *MlyI* (RES: GAGTC 4/4, MT: GASTC, S = G nebo C).

Studiem struktury enzymu *FokI* bylo zjištěno, že pokud je enzym krystalizován, katalytická doména je blokována vazebnou doménou v pozici nevýhodné pro štěpení. Tato skutečnost naznačuje možnost kontroly katalytické domény pro zabránění štěpení nescifické DNA. Uvolnění domény nastává zřejmě díky konformačním změnám ve vazebné doméně, jakmile je dosaženo rozpoznávaného místa. Nejsou sice známa všechna fakta o štěpení enzymy subtypu IIS, ale nynější představa je taková, že je třeba, aby proběhla dimerizace katalytických domén blízkých molekul, z nichž alespoň jedna je navázána na rozpoznávanou sekvenci. Druhá se může nacházet volně v roztoku nebo může být nescificky vázána na DNA, ale nejstabilnější bude komplex tehdy, naváže-li se také specificky na rozpoznávanou sekvenci. Jsou-li obě domény specificky navázány, není třeba, aby k sobě byly určitým způsobem orientovány, nebo aby byly blízko sebe, jelikož mezi nimi dojde ke stočení DNA (Pingoud, Wilson a Wende 2014).

### 3.2.2.11 Subtyp IIT

Enzymy subtypu IIT obsahují dvě odlišná katalytická místa a jednájí buď ve formě heterodimeru (*BbvCI* CC|TCAGC, *Bpu10I* CC|TNAGC, *BtsI* CCNNNNN|NNGG, aj.), nebo jako jednořetězcový protein (*Mva1269I*, *BtsCI*, *AciI*, *BsrI*, atd.).

Všechny tyto proteiny rozpoznávají asymetrické sekvence, které štěpí uvnitř, nebo velmi blízko jedné strany sekvence. Některé enzymy mají své dvě podjednotky podobně velké. Ty se pak obě podílí na rozpoznávání sekvence i na štěpení (př. *BbvCI*, *BmgBI*). Ostatní enzymy obsahují podjednotky rozdílné velikosti. Větší podjednotka potom rozpoznává sekvenci a štěpí jeden řetězec, zatímco menší podjednotka pouze štěpí druhý řetězec (př. *BsrDI*, *BtsI*, *BspD6I*).

Endonukleázy ze skupiny IIT jsou doprovázeny dvěma samostatnými methyltransferázami – každá slouží k modifikaci jednoho řetězce. MT mohou být sloučeny do jednoho bílkovinného řetězce, nebo obsaženy ve dvou oddělených proteinech.



Jelikož mají enzymy subtypu IIT dvě odlišná katalytická místa, mohou být pozměněny na vrubovací endonukleázy, zaměřené svou specifitou na řetězce, a to buď změnou jednoho či druhého místa, nebo odstraněním menší podjednotky (Pingoud, Wilson a Wende 2014).

### 3.2.3 Typ III

Přestože bylo charakterizováno pouze malé množství (Tab. P4) endonukleáz Typu III, jejich přítomnost ve většině sekvenovaných bakteriálních genomů naznačuje jejich důležitost a nutnost dalšího zkoumání.

Vznik R–M systémů Typu III je dán dvěma geny – gen *res* kóduje restriční podjednotku R a gen *mod* je zodpovědný za modifikační podjednotku M. Enzymy s katalytickou aktivitou jsou komplexy 2R+2M, nicméně se ukázalo, že *EcoP15I* RES se vyskytuje i v heterotrimerické formě 1R+2M. Enzymy s funkcí modifikační existují ve formě dimeru 2M, který methyluje adenin do formy m6A (Rao, Dryden a Bheemanaik 2013).

Enzymy Typu III štěpí DNA mimo asymetrické 5 až 6 bp dlouhé rozpoznávané sekvence, a to ve fixní pozici 25–27 párů bází od oné sekvence ve směru řetězce (Loenen et al. 2014a). K účelnému a výkonnému štěpení je zapotřebí, aby endonukleázy Typu III interagovaly se dvěma samostatnými nemethylovanými sekvencemi v inverzní orientaci „head-to-head“. Aby k tomu mohlo dojít, musí mezi sebou restriční místa komunikovat, a to na velkou vzdálenost. Pro tento účel mají enzymy Typu III, podobně jako enzymy Typu I, vyvinutou sekvenčně specifickou aktivitu adenosintrifosfatázy, která umožňuje komunikaci díky hydrolýze ATP (Rao, Dryden a Bheemanaik 2013). Je-li však hydrolýza ATP nedostatečná, musí dojít ke komunikaci jiným způsobem. Dříve se zdálo, že translokace DNA se děje pomocí smyčkování (DNA looping), nicméně vyvstala nová teorie enzymatické difúze podél DNA, způsobené konformační změnou závisící na ATP (Loenen et al. 2014a).

### 3.2.4 Typ IV

Restriční systémy Typu IV se skládají z jednoho či dvou genů a štěpí pouze modifikovanou DNA, jsou však málo selektivní co do výběru specifické sekvence. Výjimkou je nejlépe prozkoumaná endonukleáza *EcoKMcrBC* pocházející z *E. coli*, která štěpí přibližně 30 bp od jednoho ze dvou rozpoznávaných dinukleotidů obecného vzorce RmC (purin, methylovaný cytosin m4C, či m5C), jež jsou rozděleny 40 až 3 000

páry bází (Roberts et al. 2003). Mezi další relativně dobře charakterizované enzymy patří *Sau*USI (Sm5CNGS) a *Pvu*RtsII (mC[N11–13/N9–10]G).

Revel (1967) objevil geny *rglA* a *rglB* pro restrikci methylované DNA. V 80. letech 20. století byly obdobné geny zjištěny u *E. coli* a byly přejmenovány na *mcrA* a *mcrBC*. Později byl v *E. coli* objeven třetí gen *mrr*. Enzymy vzniklé translací těchto genů pak jsou senzitivní k modifikacím cytosinu (*mcrA*, *mcrBC*), případně cytosinu a adeninu (*mrr*).

Nejznámějšími modifikacemi (Tab. P5) bází DNA jsou právě methylace cytosinu a adeninu, konkrétně C5-methylcytosin (m5C), N4-methylcytosin (m4C) a N6-methyladenin (m6A), které jsou také nejrozšířenější. Další modifikace typické pro bakteriofágy jsou 5-hydroxymethyluracil (hm5U), 5-hydroxymethylcytosin (hm5C) a jeho deriváty s cukernatými rezidui (ghm5C). Samozřejmě existuje také spousta dalších, méně častých a exotických modifikací bází, nicméně změněna může být i páteř DNA, a to konkrétně výměnou –SH skupiny za nevazebný kyslík na fosfodiesterové vazbě, čímž vzniká fosfothioesterová vazba.

Dosud známé enzymy Typu IV mají oddělené domény pro restrikci a pro vazbu na DNA uloženy buď na stejném řetězci, nebo dokonce na různých polypeptidových řetězcích. Enzymy lze zatím rozdělit do šesti rodin – *McrA*, *McrBC*, *Sau*USI, *Mrr*, *Pvu*RtsII a *GmrSD* –, z nichž každá má svá specifika (Loenen a Raleigh 2014c).

### 3.3 Specifita restrikčních endonukleáz

Přesnost, se kterou restrikční endonukleázy identifikují své cílové sekvence, je významnou charakteristickou vlastností těchto enzymů a mechanismy jejího fungování jsou dnes stále nejasné (Wilson et al. 2012).

První hypotéza (Seeman, Rosenberg a Rich 1976) předpokládá sekvenční specifitu založenou na vodíkových vazbách DNA. Na krajích bazických párů, které jsou odkryty v menším a větším žlábků dvojité šroubovice, se nachází atomy vodíku a dusíku, které mohou přijmout či poskytnout vodíkové vazby. Jejich pozice a polarita jsou pro každý bazický pár odlišné. Seeman, Rosenber a Rich navrhli, že by dvě či tři vodíkové vazby mezi těmito atomy a aminokyselinami mohly sloužit k jedinečné identifikaci každého bazického páru. Tento předpoklad vedl k předvídaní vazeb mezi konkrétními aminokyselinami a bázemi, např. ve větším žlábků se bude na adenin vázat asparagin a glutamin, zatímco na guanin se naváže arginin, v menším žlábků se naopak asparagin s glutaminem budou vázat na guanin. O deset let později (McClarin et al. 1986) byla

rozluštěna krystalická struktura *EcoRI*, která po revizi (Kim et al. 1990) změnila pohled na tuto hypotézu. Aby mohla specifita fungovat na základě vodíkových vazeb, musela by existovat klíčová vodíková vazba zprostředkovaná molekulou vody. Voda ale formuje vodíkové vazby jak s donory, tak s akceptory a nemusí mezi nimi nutně rozlišovat. Toto rozlišování je ovšem základním předpokladem pro rozpoznávání podle této hypotézy (Wilson et al. 2012). V dnešní době jsou již rozluštěny krystalické struktury více než 30 restrikčních enzymů vázaných na svou specifickou sekvenci DNA. U některých endonukleáz, jako *HindIII*, *MvaI* a *Eco29kI*, je přítomno příliš málo vodíkových vazeb, které by mohly zodpovídat za specifitu, ale ve většině ostatních enzymů se zdá být hypotéza funkční. I když je tento fakt brán jako důkaz hypotézy, jedná se pouze o důkaz nepřímý a bylo by nutno experimentálně potvrdit její pravdivost (Wilson et al. 2012).

Výzkum Wilsona et al. (2012) vyvrátil hypotézu vodíkových vazeb, jelikož bylo zjištěno, že vodíkové vazby mohou být odebrány a specifita je nadále zachována. Wilson et al. (2012) také věří, že základním předpokladem rozpoznávání sekvencí je precizní uspořádání atomů ve vazebném místě bazického páru. Každé takové místo může řádně přijmout pouze jeden „správný“ bazický pár, jelikož by při navázání jiného bazického páru nastaly prostorové a elektrostatické konflikty, které vylučují správné fungování systému. Sekvenční specifita je tedy určena tvarem a elektrostatikou povrchu vazebného místa DNA.

Mechanismus specifity začíná nespecifickým volným navázáním proteinu na DNA. Enzym klouže tam a zpět, přeskakuje z molekuly na molekulu, dokud nenajde rozpoznávanou sekvenci. Poté se protein pevně naváže k této sekvenci a provede katalytickou reakci. V procesu hledání restrikčního místa se enzym setká s tisíci odlišnými sekvencemi, které nemají s rozpoznávanou sekvencí příliš společného, a tak dojde k mnoha konfliktům na vazebném místě. Efekty těchto konfliktů jsou aditivní, takže se všechny malé individuální konflikty kombinují do jedné velké nekompatibility (Wilson et al. 2012).

### **3.3.1 „Star“ aktivita**

Ačkoliv mají restrikční endonukleázy většinou velmi přesnou schopnost rozpoznávat svou specifickou sekvenci, často mohou také štěpit DNA nespecificky díky vlivu organických rozpouštědel, iontů, malých koncentrací molekul a enzymů a podobně. Mnoho běžně užívaných restrikčních endonukleáz vykazuje jistou „star“ aktivitu i při

standardních podmínkách reakce. Totiž i substrát DNA může pozměnit místo štěpení, protože nukleotidy obklopující rozpoznávanou sekvenci se z velké části podílí na jeho energetické bilanci (Kamps-Hughes et al. 2013).

Promiskuita je u velkého množství endonukleáz potlačena při standardních reakčních podmínkách za přítomnosti kationtů hořečnatých (Vasu et al. 2013), ale také kobaltnatých a zinečnatých (Pingoud, Wilson a Wende 2014). Při změněných podmínkách, jako je malá iontová síla, zvýšené pH nebo nahrazení hořečnatých iontů za ionty manganaté, endonukleázy projevují výraznou „star“ aktivitu. Existují však výjimky – například enzym *KpnI* se stává promiskuitním za působení hořečnatých iontů, nicméně je-li místo kovového iontu použit ion vápenatý, „star“ aktivita mizí (Vasu et al. 2013). Také při optimálních pufracích podmínkách a vysoké koncentraci některých endonukleáz se ukazuje omezená přesnost štěpení a je indukována „star“ aktivita, která však může být do jisté míry potlačena spermidinem, hydrostatickým tlakem a mutacemi (Pingoud, Wilson a Wende 2014).

Promiskuita zřejmě poskytuje endonukleáze schopnost adaptability na hostitelskou bakterii a je tak bránou k vývoji nových enzymatických aktivit (Vasu et al. 2013). Avšak pro většinu použití v molekulární genetice a diagnostice je „star“ aktivita nevýhodná, jelikož jsou její produkty málo předvídatelné a výsledek pak může obsahovat velmi málo požadovaných fragmentů. Naopak prospěšná byla „star“ aktivita v letech, kdy ještě nebylo objeveno velké množství enzymů, takže některé RES byly využívány jako zástupci endonukleáz, které ještě nebyly nalezeny (Wei et al. 2008).

Příkladem „star“ aktivity může být enzym *EcoRI*, který za malé iontové síly a alkalického pH kromě své typické sekvence G|AATTC štěpí i N|AATTN (N = A, C, G, T), přičemž preferované sekvence jsou GGATTT, AAATTT, GAATTT a GAATTA, ale CAATTG štěpena není (Pingoud, Wilson a Wende 2014).

Wei et al. (2008) navrhli způsob, jak kvantitativně definovat „star“ aktivitu. Tzv. Fidelity Index (FI, „index věrnosti“) se vypočítá jako poměr nejvyššího množství enzymu, které nevykazuje při štěpení „star“ aktivitu, a nejnižšího množství enzymu potřebného pro kompletní štěpení na analogických rozpoznávaných místech. FI tedy indikuje, jak specifický je daný enzym. Čím vyšší je hodnota indexu, tím je vyšší specifita enzymu a tím vyšší je rozsah koncentrací enzymu, které mohou být použity, aniž by se vyskytla „star“ aktivita. FI konkrétní endonukleázy se samozřejmě bude lišit na různých substrátech a při použití různých reakčních pufrů, teploty a dalších faktorů.

Proto tento index poskytuje pouze obecná vodítka pro výběr optimálního enzymu a vhodných reakčních podmínek pro jeho použití.

### **3.3.2 Vlivy podmínek reakce**

Každý restriční enzym má individuální optimální podmínky, za kterých reakce probíhá. Nejen tyto podmínky, ale také způsob skladování, jsou určeny výrobcem a jejich dodržením uživatel dosáhne nejvyšší aktivity, nejlepší přesnosti a největší stability enzymu. Faktory, které je nutno sledovat při reakci restričních enzymů, zahrnují teplotu, pH, kofaktory, složení soli, iontovou sílu a stabilizátory. Důležitá je však i doba trvání inkubace či množství reagentů (Promega Corporation 2016).

#### **3.3.2.1 Teplota inkubace**

Převážná většina enzymů vykazuje nejvyšší aktivitu při 37 °C, nicméně jisté endonukleázy potřebují vyšší či nižší teploty pro optimální aktivitu. Inkubační teplota představuje zvýšení teploty bakteriálního kmene, z něhož byla endonukleáza izolována. Při inkubacích vysokými teplotami delších než jedna hodina se doporučuje pokrýt roztok s reakcí kapkou minerálního oleje pro prevenci vypařování (Promega Corporation 2016).

#### **3.3.2.2 pH**

Typická hodnota pH pro velké množství enzymů se pohybuje mezi 7,2 a 8,5. Tato hodnota se sleduje hlavně při inkubaci (Promega Corporation 2016).

#### **3.3.2.3 Kofaktory**

Všechny komerčně dostupné restriční enzymy vyžadují především hořčičnaté ionty. Jejich koncentrace se může pohybovat v relativně velkém rozmezí 5–30 mM v reakci (Promega Corporation 2016). Enzymy Typu I však potřebují také ATP a AdoMet (Loenen et al. 2014b), enzymy subtypu IIG (zahrnuje většinu IIB a IIC enzymů) vyžadují AdoMet (Pingoud, Wilson a Wende 2014) a endonukleázy Typu III zase potřebují ATP a některé z nich i AdoMet (Loenen et al. 2014b).

#### **3.3.2.4 Složení soli a iontová síla**

Pro stimulaci reakce se užívá nejčastěji chlorid sodný, některé enzymy však preferují chlorid draselný nebo acetáty. Jednotlivé restriční endonukleázy reagují různě na iontovou sílu roztoku, ale největší množství je stimulováno 50–150 mM soli. Některé jsou však inhibovány již při koncentracích vyšších než 20 mM (Promega Corporation 2016).

### 3.3.2.5 *Stabilizátory*

Při delších reakcích či při skladování je doporučeno užívat stabilizátory, jako hovězí sérový albumin (BSA), merkaptoetanol, dithioerytriol (DTE) nebo dithiotreitol (DTT). Tyto látky chrání restriční enzymy před oxidací, proteázami, nespecifickou adsorpcí a škodlivými vlivy prostředí, jako je teplo, povrchové pnutí či rušivé substance. Většinou se přidává 0,1 mg/ml i pro zvýšení enzymatické aktivity v běžné reakci (Promega Corporation 2016).

### 3.3.2.6 *Pufry*

Pufry jsou částečně neutralizované roztoky slabých kyselin či slabých zásad a jejich konjugovaných protějšků (pro kyselinu zásada a naopak). Koncept pufrování, tedy udržování stálého pH, je založen na částečné disociaci slabých kyselin či zásad ve vodném roztoku. Vstoupí-li do systému látka o rozdílném pH, dochází k disociaci nedisociované formy kyseliny či zásady. Schopnost pufru udržovat pH je závislá na dostatečné koncentraci obou složek konjugovaného páru (Stoll a Blanchard 1990).

Při výběru vhodného pufru je třeba mít na paměti hlavně optimální pH pro činnost vybraného enzymu, nespecifické efekty pufru na enzym a interakce pufru se substrátem a kovy. Důležité je také zvolit vhodnou koncentraci pufru, jelikož velká iontová síla může snížit aktivitu enzymu (Stoll a Blanchard 1990). Pufry jsou běžně dodávány 10krát koncentrované, a tudíž je třeba použít pouze množství o 1/10 celkového objemu reakce, aby bylo dosaženo normální koncentrace (New England Biolabs 2016).

Společnosti vyrábějící a prodávající restriční enzymy mají vyvinuty své pufry, které jsou vhodné pro dostatečnou aktivitu jimi nabízených endonukleáz. Těchto pufrů je vždy několik a liší se jak složením, tak koncentracemi složek. Nejčastěji se používá Tris-HCl, MgCl<sub>2</sub> a NaCl. Některé enzymy však lépe fungují v acetátovém prostředí, jiné zase preferují draslík místo sodíku anebo nepotřebují k aktivitě sůl vůbec. Nejčastěji používanými stabilizátory jsou DTT a BSA (New England Biolabs 2016, Promega Corporation 2016, Takara Bio Company 2016, Thermo Fisher Scientific Inc. 2010).

Je-li prováděno štěpení jednoho substrátu dvěma restričními enzymy, je třeba použít pufr vhodný pro oba enzymy. Jestliže jeden z enzymů má aktivitu menší než 75 %, je třeba upravit reakční čas, nebo množství enzymu, nicméně se také zvýší riziko „star“ aktivity. Další možností je použít izoschizomer, či neoschizomer, který se více hodí pro vybraný pufr. V případě, že nelze použít výše uvedená doporučení, lze provést

štěpení prvním enzymem, inaktivovat jej a následně upravit podmínky pro štěpení druhým enzymem a provést druhou reakci (Promega Corporation 2016).

Pro ředění restričních endonukleáz se používají ředící pufrů. Ředění by mělo proběhnout těsně před použitím enzymu a doporučená konečná koncentrace enzymu by měla být nejméně tisíc jednotek na mililitr (New England Biolabs 2016).

### **3.3.2.7 Objem**

Doporučený objem reakce se pohybuje okolo 50  $\mu$ l na  $\mu$ g substrátu. Enzym by se měl v celkovém objemu reakce vyskytovat do 10 %, jelikož přebytek glycerolu, který se přidává do skladovacích pufrů, může způsobovat „star“ aktivitu. Aditiva ve skladovacím pufru enzymu, ale také kontaminanty substrátu mohou představovat problém při malých reakčních objemech, a tudíž je nutné změnit objem složek reakce, nebo případně rozředit enzym ředícím pufr pro zvětšení jeho objemu (New England Biolabs 2016). Také příliš viskózní a příliš zředěný substrát může neblaze ovlivnit aktivitu enzymu (Promega Corporation 2016).

### **3.3.2.8 Doba trvání reakce**

Ve většině případů reakce probíhá jednu hodinu. Tento čas může být zkrácen buď přebytkem enzymu, nebo speciálními enzymy, které nabízí například firma New England Biolabs – 1  $\mu$ l jejich Time-Saver<sup>TM</sup> enzymu rozštěpí 1  $\mu$ g substrátu za 5 až 15 minut. Reakční čas lze také prodloužit až na 16 hodin, a to pouze použitím menšího množství enzymu. Potřebné množství je samozřejmě pro každý enzym jiné. Některé endonukleázy, jako například *AccI*, zvládnou za 16 hodin rozštěpit substrát při objemu 0,13 jednotek, jiné RES (např. *BsaXI*) potřebují 0,25 jednotek pro úplné proběhnutí reakce, další enzymy (např. *Acc65I*) musí mít alespoň 0,5 jednotek a ty nejméně aktivní zvládnou štěpení až při objemu 1 jednotky, což je stejné množství, jaké je potřeba k reakci trvající jednu hodinu (New England Biolabs 2016). Při výběru enzymu pro dlouhodobé štěpení je tedy nutné myslet i na ekonomiku laboratoře.

Reakci lze ukončit několika způsoby. Není-li již třeba dále manipulovat s DNA, užívá se roztok s obsahem kyseliny ethylendiamintetraoctové (EDTA), který vyváže dvojmocné ionty. Bude-li DNA dále používána, přeruší se reakce fenolovou či chloroformovou extrakcí, kolonou, anebo tepelnou inaktivací. Při ní je reakce inkubována po dobu 20 minut při teplotě potřebné k vyřazení enzymu z provozu. Většině enzymů, které mají optimální inkubační teplotu 37 °C, stačí k inaktivaci působení teploty 65 °C. Enzymy s vyšší optimální inkubační teplotou lze většinou

zastavit při 80 °C, nicméně na některé endonukleázy tento způsob není možné použít a je třeba se uchýlit k výše uvedeným metodám (New England Biolabs 2016).

### **3.4 Využití v molekulární genetice a diagnostice**

Restrikční endonukleázy mají mnoho praktických použití a jsou ceněny zejména pro svou dostupnost, jednoduchou a rychlou manipulaci i přesnost. Od 70. let 20. století se staly nedílnou součástí molekulární biologie díky využití při klonování, sekvenování DNA, detekci a nadprodukcii enzymů a hormonů, objevování mutací, identifikaci genů chorob nebo také při určování paternity (Loenen et al. 2014a).

#### **3.4.1 RFLP**

V sekvenci DNA homologních chromozomů jedince či v sekvenci DNA stejných genů různých jedinců mohou být rozdíly. Podstoupí-li tyto sekvence restrikční štěpení správně vybranou endonukleázou, pak se rozdíly budou nacházet právě v místě rozpoznávaném enzymem a u jedné sekvence tak nedojde ke štěpení. Délky fragmentů z restrikce obou sekvencí budou vůči sobě rozdílné. Tento jev se nazývá polymorfismus délků restrikčních fragmentů neboli RFLP (Snustad a Simmons 2009).

Substrát je vystaven působení restrikčních enzymů, které jej rozštěpí. Při gelové elektroforéze na fragmenty substrátu uložené v agarózovém či polyakrylamidovém gelu působí stejnosměrný elektrický proud. Záporně nabitá DNA putuje od katody k anodě, čímž se oddělují a seřazují dle délky jednotlivé fragmenty. Výsledky mohou být analyzovány Southernovým přenosem (Obr. P1), kdy je DNA z elektroforetického gelu přenesena na nylonovou membránu, na kterou je fixována osvětlením UV světlem. Poté je inkubována v roztoku obsahujícím denaturované radioaktivně označené sondy DNA, jež se naváží na komplementární fragmenty na membráně. Výsledky jsou pak zobrazeny autoradiograficky (Pritchard a Korf 2013).

Genotypování pomocí RFLP má několik výhod oproti ostatním technikám. Lokusy vhodné pro RFLP jsou rozprostřeny po celém genomu a markery jsou vysoce polymorfní. Existuje široký výběr enzymů pro provedení reakce. Markery pro RFLP mají kodominantní dědičnost a jsou velmi dobře reprodukovatelné. Polymorfní lokusy identifikované při RFLP jsou stabilně detekovatelné nehledě na vlivy prostředí a interakci genů (Dai a Long 2015).



### 3.4.2 PCR-RFLP

Při této technice předchází samotnému štěpení amplifikace substrátové DNA metodou polymerázové řetězcové reakce (tzv. PCR). Ta obsahuje tři opakující se kroky. Při denaturaci jsou za teploty 94 °C od sebe odděleny komplementární řetězce DNA. Následuje fáze připojení primerů, kdy při 60 °C nasedají speciálně navržené krátké oligonukleotidové řetězce na substrátovou DNA. Při elongaci enzym polymeráza za teploty 72 °C syntetizuje komplementární vlákna DNA od místa napojení primerů. Takovýto cyklus při PCR proběhne celkem třicet dva a v každém cyklu je množství DNA zdvojnásobeno (Pritchard a Korf 2013).

Amplifikovaná DNA projde restriční reakcí a výsledný štěpný produkt je pak analyzován gelovou elektroforézou. Prozářením etidiumbromidem obarveného gelu UV světlem lze rovnou vidět rozložení fragmentů (National Center for Biotechnology Information 2014b).

Metoda PCR-RFLP slouží k detekci krátkých indel polymorfizmů a SNP polymorfizmů. Indel, neboli inzerce–delece, představují mutace vzniklé přidáním či odebráním jedné a více bází v sekvenci DNA. SNP jsou změny jediné báze v analogických sekvencích DNA jednotlivců. V lidském genomu se takovýto polymorfizmů nachází několik stovek tisíc uvnitř genů a ještě mnohem více v nekódujících oblastech genomu, celkově tedy zhruba 3 miliony SNP, které jsou rozprostřeny náhodně s frekvencí každých 300 až 1000 bází.

Pro provedení detekce SNP je třeba, aby jeden polymorfizmus byl součástí rozpoznávané sekvence restričního enzymu a druhý ne. Endonukleáza rozštěpí DNA pouze v prvním případě a vzniknou tak různě dlouhé fragmenty, které je možno detekovat na elektroforetickém gelu (Griffiths et al. 2012).

Detekce SNP slouží hlavně k identifikaci genů chorob a k testování jednotlivců na fenotypové variace v reakci na farmaceutika. Zde hraje roli především detekce alel genu cytochromu P450, který řídí degradaci mnoha cizorodých látek, a tedy i různých léků. Každá z alelických variant může vykazovat jinou aktivitu enzymů, takže lidé s alelou pro sníženou aktivitu metabolizují daný lék pomaleji a s větší pravděpodobností se u nich projeví vedlejší účinky. Na druhou stranu pacienti s duplikacemi alely mají rychlý metabolismus, který nechává medikamentu velmi krátkou dobu působnosti. Individuální analýza SNP odhalí, která alela je přítomna a pacientovi může být předepsána dávka léku přizpůsobená na míru (Clark a Pazdernik 2013).

### 3.4.3 Restrikční mapy

Restrikční mapy lze chápat jako lineární mapy pozic restrikčních míst, které ukazují jejich pořadí a vzdálenosti v segmentu DNA (Griffiths et al. 2005). Jedná se o mapy fyzické, které vyjadřují skutečnou molekulovou vzdálenost v párech bází (bp), kilobázích (kb) či megabázích (Mb) (Snustad a Simmons 2009). Hlavně dříve byly cenným nástrojem v genomové analýze převážně při mapování virů, ale uplatnily se také při klonování DNA, jelikož rozložení restrikčních míst určuje, kde může být vytvořen klonovatelný fragment DNA s lepivými konci. Nejčastější jsou restrikční mapy lokalizovaných oblastí na chromozomech nebo mapy menších úseků DNA, například těch v plazmidech a organelách (Griffiths et al. 2005).

První způsob tvorby takovéto mapy zahrnuje porovnání štěpení jedním a dvěma enzymy. Dva restrikční enzymy jsou použity v oddělených reakcích a poté jsou oba použity v reakci společně. Po skončení štěpení jsou velikosti fragmentů DNA analyzovány elektroforeticky. Štěpení dvěma enzymy ukazuje, zda fragment stvořený jedním enzymem obsahuje restrikční místo pro enzym druhý. Pokud ano, je daný fragment nahrazen dvěma či více menšími fragmenty. Porovnání velikostí fragmentů z různých štěpení umožňuje přibližné určení restrikčních míst (Obr. P2).

Další ze způsobů tvorby restrikční mapy počíná označením 5' konce vzorku DNA fosforem  $^{32}\text{P}$ . V následujícím štěpení prvním restrikčním enzymem jsou produkovány fragmenty s fosforovou značkou na jednom konci. Nejdelší fragment je izolován a štěpen druhým restrikčním enzymem, ale reakce se předčasně ukončí, aby nedošlo k úplné restrikci. Tímto se vytvoří skupina různě velkých označených fragmentů, z nichž každý má jedno z restrikčních míst na neoznačeném konci. Po elektroforéze může být odhadnuta vzdálenost každého restrikčního místa od označeného konce.

Po vzniku restrikční mapy se provede Southernův přenos, čímž se do mapy vloží geny. Pokud gen hybridizuje s fragmentem A od restrikčního enzymu 1 a zároveň také s fragmentem B od restrikčního enzymu 2, musí být lokalizován v oblasti, kde se fragmenty A a B překrývají (Griffiths et al. 2000).

### 3.4.4 Genetické testy

Pod tímto tématem se skrývá využití restrikčních enzymů k detekci polymorfizmů v jednoduchých, rychlých a levných testech, které mají uplatnění v různých odvětvích lidské činnosti, od forezních věd, diagnostiky zdravotního stavu lidí i zvířat až po určování vlastností vztahujících se k užitečnosti zvířat (Dai a Long 2015).

Pro určení původu biologických materiálů a příbuznosti jsou důležité jednoduché repetitivní sekvence DNA obsahující krátké, tandemově se opakující motivy, tzv. genetické otisky prstů. Tyto tandemové repetece o variabilním počtu (VNTR, z angl. variable number tandem repeats) jsou vysoce polymorfní, jsou rozprostřeny ve všech chromozomech, a to v mezigenových sekvencích a intronech, nepodílí se ani na fenotypu, ani na chování člověka. Podle délky repetece se dělí na minisatelity o velikosti dvanáct až několik set bází a mikrosatelity, které jsou 1–6 bází dlouhé. VNTR byly objeveny také u zvířat, rostlin a ostatních eukaryot, což výrazně rozšiřuje škálu použití této metody. Izolovaná DNA je štěpena restrikcí enzymem, který rozdělí sekvenci na jednom nebo druhém konci VNTR. Pro vytvoření fragmentů je však potřeba dostatečně celistvá DNA nejméně o 20 až 25 tisících bp. Tato skutečnost zapříčinila převážné nahrazení metody RFLP řetězovou polymerázovou reakcí, která dokáže amplifikovat i degradovanou DNA (Cabe a Capriccioso 2016).

Nejčastěji se metoda PCR-RFLP využívá v molekulární diagnostice dědičných chorob. Je-li známa podstata genu ovlivňujícího dědičné monogenní onemocnění, je možné navrhnout vhodný enzym pro detekci přítomnosti mutantní alely. K testům stačí mít díky polymerázové řetězové reakci jen malé množství vzorku DNA, takže je možno u některých onemocnění provádět i prenatální diagnostiku na buňkách plodu odebraných amniocentézou nebo biopsií choriových klků, případně i na buňce z oplozeného preimplantačního embrya. Tímto způsobem lze u lidí detekovat například srpkovitou anémii, hemofilii, cystickou fibrózu a další choroby (Snustad a Simmons 2009). Zdravotní stav zvířat je monitorován hlavně pro šlechtitelské účely. U skotu se nejčastěji testuje dědičná porucha imunity BLAD (Bovine Leukocyte Adhesion Deficiency), komplexní vertebrální malformace (CVM) holštýnského skotu a vloha pro chondrodysplasiu u plemena dexter. U arabských koní se diagnostikuje přenašečství choroby SCID (Severe Combined Immunodeficiency Disorder) (Českomoravská společnost chovatelů, a. s. 2016). Nejčastěji jsou však testováni chovní psi a kočky – zde existuje mnoho různých dědičných nemocí způsobených nesprávnými praktikami ve šlechtění (Slutsky et al. 2013).

Ve vztahu k užitečnosti se genetické testy u hospodářských zvířat používají ke zjištění předpokladů pro tvorbu svalových vláken, tučnost či kvalitu masa a mléka (Kuciel et al. 2004). Klíčovým genem pro výrobu vepřového masa se stal gen ryanodinového receptoru *RYR1*. Jeho mutace v recesivní alele (Obr. P3, P4) způsobuje prasečí stresový syndrom a bledé, měkké, vodnaté maso neboli PSE maso. Výskyt

problémového genotypu recesivního homozygota *nn* byl nejvýraznější hlavně v 60. letech 20. století, kdy závratnou rychlostí probíhalo šlechtění na masnou užitkovost. Postižená prasata vykazovala větší náchylnost ke stresu, trpěla maligní hypertermií, při uspání pomocí inhalačního narkotika halotanu projevovала netypický svalový spasmus, pH jejich masa hodinu po porážce kleslo i pod hodnotu 5,8, což mělo za následek snížení vaznosti masa, zvýšení jeho vodivosti a velmi světlé zbarvení. Metodou PCR-RFLP je možné určit genotyp prasete a odhalit tak mutaci C→T nukleotidu 1843 na cDNA, vedoucí k poruše fungování vápníkového kanálu v sarkoplazmatickém retikulu kosterního svalstva. Intenzivní selekcí se podařilo snížit výskyt recesivních homozygotů téměř na nulu (Marini et al. 2012).

Genotypování je důležité také pro potravinářství. Použitelnost potravin ke zpracování je determinována například geny proteosyntézy zásobních bílkovin u pšenice nebo geny mléčného kaseinu a sérového albuminu u dojníc (Kuciel et al. 2004). PCR-RFLP metoda lze využít i k určení přítomnosti masa různých druhů zvířat ve výrobku. Běžně se sice pro tento test užívá samotná PCR, nicméně té může chybět potřebná selektivita kvůli cross-reaktivitě, a tak se nabízí restriční analýza jako přijatelná alternativa. V těchto metodách je amplifikován mitochondriální gen cytochrom *b*, který je později štěpen například endonukleázou *HaeIII* v klasické PCR-RFLP anebo t-RFLP-PCR (polymorfismus délky terminálních restričních fragmentů). Vzniklé fragmenty jsou posléze porovnávány se standardem, přičemž při t-RFLP-PCR může být rozlišeno až čtyřicet různých druhů zvířat i při malých množstvích vzorků (Müller a Steinhart 2007).

### 3.4.5 Klonování DNA

Technologie rekombinantní DNA spočívá v izolaci části jakékoliv molekuly DNA, jejím následném připojení na samoreplikující se virový či plazmidový vektor a vložení do buněk, kde se bude množit (Wilson et al. 2012).

Lze klonovat celý genom organismu, nicméně je třeba jej rozštěpit na mnohem menší fragmenty, aby bylo možné je vložit do vektoru. Zde se uplatňují restriční endonukleázy štěpící fosfodiesterové vazby. Stejná restriktáza je použita na štěpení jak vektoru, tak genomové DNA (Obr. P5). Vzniknou tak komplementární konce, které by ideálně měly být lepivé, jelikož pak může jednodušeji proběhnout hybridizace.

Ke klonování jednoho specifického genu lze opět použít restriční enzymy nebo amplifikaci polymerázovou řetězovou reakcí *in vitro*. Stačí vybrat vhodné primery

(sekvence o 15–20 bázích), které se naváží na oba konce regionu, v němž se nachází gen, a odtud bude probíhat amplifikace.

Další možností získání požadovaných fragmentů je purifikace mediátorové RNA a její následné ošetření reverzní transkriptázou, která mRNA přeloží na komplementární DNA (cDNA), jež bývá často vhodnější pro účely klonování. Genomová DNA totiž na rozdíl od cDNA obsahuje introny, které přerušují kódující oblasti (Griffiths et al. 2012).

Důležitou součástí přípravy na tvorbu rekombinantní DNA je vhodný výběr vektoru. Vektory plazmidové či bakteriofágové mohou přijmout dárcovskou DNA velkou maximálně 25–30 kb. Pro přenášení větších fragmentů jsou používány fágomidy (zkonstruované hybridy DNA fága  $\lambda$  a DNA bakteriálního plazmidu) a umělé bakteriální chromozomy (BAC) (Griffiths et al. 2012). Také je nutné zajistit, aby se na vektoru nacházelo pouze jedna rozpoznávaná sekvence vybraného restričního enzymu. Toho se dosahuje vložením tzv. polylinkeru, neboli MCS (multiple cloning site), což je uměle vytvořená 50 bp dlouhá sekvence DNA obsahující restriční místa sedmi až osmi běžně užívaných enzymů. Důležité je vybrat taková místa, která se nenacházejí nikde v genetické informaci vektoru. Zařazení polylinkeru zajistí, že vektor nebude inzercí dárcovské DNA poškozen a že inzert se bude vždy nacházet na přibližně stejném místě (Clark a Pazdernik 2013).

Po získání fragmentů je třeba dárcovskou DNA připojit na vektorovou DNA. Mají-li obě DNA lepivé konce, stačí navodit příznivé fyziologické podmínky a dojde k hybridizaci, dále je třeba přidat DNA ligázu, která spojí obě DNA fosfodiesterovými vazbami a vznikne rekombinantní molekula. Mají-li fragmenty tupé konce, nezbyvá než lepivé konce na DNA připojit.

Pro amplifikaci rekombinantní molekuly DNA se využívá reprodukčních procesů prokaryot. Dárcovská DNA připojená na vektor je vložena do bakteriální buňky, kde je amplifikována stejně, jako se replikuje bakteriální genom (Obr. P6). Musí však být splněna základní podmínka: přítomnost počátku replikace DNA ve vektoru (Griffiths et al. 2012).

Druhou roli v klonování hrají restriční endonukleázy při detekci inzercí ve vektorech. Je totiž třeba určit, zda se zdařila replikace požadovaného rekombinantního vektoru. Samotný vektor (např. plazmid) lze identifikovat jednoduše tím, že hostitel díky vektoru získal rezistenci na antibiotika. Bohužel tento test neurčuje, zda je v replikovaných molekulách přítomna rekombinantní DNA. Většinou ani vložený gen nemá dostatečně rozlišitelné fenotypové projevy, takže je nutné použít jiné metody

detekce. První, méně sofistikovanou metodou, je rozmnožení mnoha samostatných bakteriálních kolonií, které obdržely vektor. Z nich je extrahována plazmidová DNA, která je posléze štěpena stejným restriční enzymem, který byl použit k původní tvorbě rekombinantní DNA. Nachází-li se dárcovská DNA v plazmidu, vzniknou dva fragmenty DNA. Nenachází-li se tam, vznikne z cyklické molekuly vektorové DNA molekula lineární. Restriční reakce jsou poté vyhodnoceny elektroforeticky. Z mnoha kolonií jsou touto prověřovací procedurou vybrány ty, které inzert obsahují. Méně pracnou metodu představuje použití plazmidů se dvěma geny pro resistenci na antibiotika, z nichž jeden obsahuje restriční místo. Pokud je do plazmidu vložena cizí DNA, je tento jeden gen inaktivován. Hostitelské buňky bez inzertu tedy budou rezistentní na oba druhy antibiotik, ale bakterie produkující rekombinantní plazmidy budou mít odolnost pouze na jedna antibiotika (Clark a Pazdernik 2013).

Zdárně amplifikovanou rekombinantní DNA je třeba vyjmout z bakteriální buňky ven. Byl-li vektorem fág, stačí získat jeho lyzát a izolovat DNA, kterou fág obsahoval. Byl-li vektorem plazmid, fágomid nebo BAC, pak se bakterie chemicky a mechanicky rozbijí a rekombinantní molekula je separována od mnohem většího bakteriálního chromozomu centrifugací, elektroforézou či dalšími odlučovacími technikami (Griffiths et al. 2012).

V případě, že je třeba identifikovat konkrétní gen z genomové knihovny, který se nachází spolu s jinými geny v jedné rekombinantní molekule, lze použít subklonování (Obr. P7). Po vyštěpení a rozdělení inzertu na jednotlivé geny následuje vložení každé sekvence zvlášť do nového vektoru. Po amplifikaci v hostitelském organismu je možné experimentálně určit, která rekombinantní molekula nese požadovaný gen (Russel 2010).

Díky objevu rekombinantní DNA došlo k průlomů v oblasti purifikace enzymů. Bez této technologie bylo možno získat enzymy z přírodních zdrojů, ovšem pouze v malých množstvích a o nevalné čistotě. Klonování umožnilo přesunout geny, jež kódují proteiny, do příhodných organismu, jako je *E. coli*, kde byly následně transkribovány a translatovány v samotné proteiny. Výsledné množství a čistota jsou mnohonásobně větší, stejně jako se zlepšila aktivita i stabilita purifikovaných enzymů. Co se naopak snížilo, je cena. Dnes snad již všechny komerčně dostupné endonukleázy pochází z naklonovaného genu exprimovaného z *E. coli* (Wilson et al. 2012).

### **3.4.6 Tvorba genomových knihoven**

Genomová knihovna představuje soubor rekombinantních klonů, v nichž je teoreticky obsažena alespoň jedna kopie každé sekvence DNA genomu, pro který je daná knihovna vytvořena. Ne vždy je možné klonovat všechny části genomu, ale pro zajištění co největšího pokrytí je většinou vytvořeno několik kopií každé rekombinantní molekuly (Griffiths et al. 2012).

Izolovaný genom daného organismu je štěpen restričními enzymy, jejichž specifická sekvence má čtyři až šest párů bází. Vznikají fragmenty o různých délkách, které se přes sebe překrývají. Toho lze dosáhnout použitím dvou různých restričních enzymů, nebo částečnou restriční reakcí jednoho enzymu, kdy je štěpena pouze náhodná část restričních míst díky omezení času inkubace nebo díky limitování množství použitého enzymu. Nyní je třeba vyselektovat pouze fragmenty správné velikosti, čehož lze dosáhnout gelovou elektroforézou. Po proběhnutí separace se část s požadovanými fragmenty jednoduše vyřízne z gelu a vyextrahuje. Dále je nutné vypočítat, kolik klonů bude třeba, aby v knihovně byla každá sekvence obsažena alespoň jednou. Následuje výběr vhodných vektorů, inzerce fragmentů do vektorů, integrace rekombinantních molekul do bakteriálního hostitele a množení rekombinantních molekul.

Problematická manipulace s knihovnami větších genomů kvůli velkému počtu klonů může být vyřešena tvorbou chromozomových knihoven. Vznikají buď rozdělením velké genomové knihovny na několik menších knihoven obsahujících genetický kód jednotlivých chromozomů, nebo separací jednotlivých chromozomů například průtokovou cytometrií a následnou fragmentací a klonováním chromozomové sekvence (Russel 2010).

### **3.4.7 Tvorba genomových knihoven pro sekvenování genomů**

Pro tvorbu knihovny určené k sekvenování genomu je třeba použít neupravenou genomovou DNA, jelikož obsahuje geny v původní formě s introny a regulačními sekvencemi. Dále je možné vytvořit knihovnu cDNA, která je díky výlučnému obsahu oblastí genomu kódující proteiny velmi vhodná pro zkoumání specifických genů (Griffiths et al. 2012).

Metody sekvenování první generace, vynalezené ve 20. století, používají restriční endonukleázy k fragmentaci sekvenované DNA. Sangerova metoda, jiným názvem dideoxy sekvenace, využívá restričními enzymy štěpenou DNA integrovanou do

bakteriální genomové knihovny. Po následné extrakci jsou plazmidy spolu s primery, DNA polymerázou a fluorescenčně značenými deoxynukleotidy (dNTP, jiná barva pro každý dNTP – A je zelený, G je černý, C modrý a T červený) vloženy do zkumavky, v níž se nachází menší množství dideoxynukleotidů (ddNTP), které ukončují tvorbu vláknů DNA. Výsledkem reakce jsou fragmenty o různých délkách, které jsou zakončené těmito ddNTP. Po elektroforéze a autoradiografii vznikne soubor fluorescentních proužků, z nichž se odvozuje sekvence DNA (Russel 2010).

Metody sekvenování nové generace oproti tradičnímu sekvenování nepoužívají mikrobiální hostitele pro amplifikaci DNA a jsou zároveň levnější i rychlejší. Nová generace využívá fragmentaci genomové DNA, a to buď nespecifickým štěpením restričními enzymy (Clark a Pazdernik 2013), působením enzymu transposázy, která zároveň fragmentuje a vkládá adaptérové sekvence do DNA (Head et al. 2014), nebo fyzikálními metodami jako je rozbití ultrasonickými vlnami a nebulizace (opakované protlačení DNA skrz zúžený prostor stlačeným dusíkem či vzduchem). Kvůli rozdílným sekvencím musí být na konec fragmentů přidána známá sekvence (adaptér), na kterou mohou nasednout primery pro následnou amplifikaci při PCR. Milióny vzniklých fragmentů jsou poté sekvenovány paralelně v jednom kroku (Clark a Pazdernik 2013).

Nejrozšířenější metodou druhé generace je sekvenování 454 neboli Roche, které využívá tzv. pyrosekvenování. Další metody, pro které je možné použít fragmentaci DNA restričními endonukleázami, jsou Illumina/Solexa a SOLiD/Applied Biosystems (Clark a Pazdernik 2013).

Existují také novější metody sekvenování, které dokáží zpracovat celou nefragmentovanou molekulu DNA. To však znamená, že nevyužívají restriční enzymy. Mezi tyto metody patří SMRT sekvenování, Heliscope Single Molecule Sequencer a Nanopore sekvenování (Clark a Pazdernik 2013).



## 4 MATERIÁL A METODIKA

Praktická část mé bakalářské práce spočívá v rozpoznání polymorfizmů u 3 různých genotypů zadaných genů, jejich zpracování webovými programy pro vyhledávání restrikčních míst a restrikční analýzu DNA, nalezení a navržení vhodné restrikční endonukleázy pro testování polymorfizmu metodou PCR-RFLP a následném porovnání funkčnosti a přehlednosti daných programů.

### 4.1 Geny

Obdržela jsem k analýze geny *MC1R* a *MATP*, které byly sekvenovány z koňského genomu. Jednalo se celkem o čtyři sekvence ke každému genu – jednu pro recesivního a jednu pro dominantního homozygota, dvě pro heterozygota. V programu Sequence Scanner v1.0 jsem identifikovala polymorfizmy charakterizující jednotlivé alely.

#### 4.1.1 *MC1R*

Melanokortinový receptor 1, situovaný na 3. chromozomu koně, je genem exprimovaným ve velkém množství v melanocytech. Hraje významnou roli při syntéze melaninu. Je-li gen aktivován, což znamená, že alespoň jedna alela genu je dominantní, syntetizuje se v kůži a srsti zvířete eumelanin, který představuje černý či hnědý pigment. U recesivních homozygotů je zase syntetizován faeomelanin zodpovědný za červenou či žlutou barvu srsti. Lokus tohoto genu se nazývá Extension (rozšiřující) a alely se označují písmeny E, e (Thiruvenkadan, Kandasamy a Panneerselvam 2008).

#### 4.1.2 *MATP*

Na 21. chromozomu koně se nachází gen *MATP* (Membrane-associated transporter protein), nebo také *SLC45A2*, který zodpovídá za zředění zbarvení srsti. Alely C a C<sup>cr</sup> mohou tvořit tři genotypy – CC pro plnou pigmentaci, CC<sup>cr</sup> pro ředění faeomelaninu do žluta a C<sup>cr</sup>C<sup>cr</sup> pro ředění obou typů melaninu do světlé krémové barvy srsti, nicméně ovlivněna je také barva kůže a očí (Mariat, Taourit a Guérin 2003).

### 4.2 Počítačový software a webové aplikace

U každého genotypu obou genů jsem si vybrala část sekvence s polymorfizmem a osmi až deseti nejbližšími přilehlými nukleotidy z každé strany. Tyto výseky jsem pak analyzovala ve třech bezplatných internetových programech – Webcutter 2.0, WatCut a NEBcutter V2.0.

#### 4.2.1 Webcutter 2.0

Program Webcutter 2.0, který je k bezplatnému použití na webové stránce <http://rna.lundberg.gu.se/cutter2/>, představuje nástroj pro restriční mapování nukleotidových sekvencí. Webcutter nabízí automatické hledání a vložení sekvencí z GenBank, která náleží Národnímu centru pro biotechnologické informace (NCBI), dále nabízí zvýraznění oblíbených enzymů tučně či barevně, hledání míst potenciálně vystavených tiché mutaci, analyzování restričních map degenerovaných sekvencí (tzn. obsahujících mnohoznačné nukleotidy jako N, Y, R atd.), výběr mezi lineární a cyklickou sekvencí, odkazy na informace o enzimech v databázi restričních enzymů REBASE (Heiman 1997).

Do vstupní části programu jsem zadala název sekvence, vlastní sekvenci jsem vepsala do připraveného políčka, vybrala jsem analýzu lineární sekvence a zobrazení mapy restričních míst i tabulky – u prvního genu seřazenou dle názvu enzymů, u druhého genu dle pořadí bp štěpení. K použití jsem zvolila všechny enzymy v databázi, které štěpí sekvenci pouze jednou.

#### 4.2.2 WatCut

Projekt WatCut, dostupný z internetové adresy <http://watcut.uwaterloo.ca/template.php>, je program pro analýzu sekvencí DNA restričními enzymy. WatCut nabízí tři nástroje využitelné pro analýzu tichých mutací, restriční analýzu a analýzu SNP-RFLP. Program používá informace z databáze restričních enzymů REBASE, což mu umožňuje využít všechny komerčně dostupné enzymy Typu II. WatCut vznikl díky Michaelu Palmerovi a představuje pouze jeho koníček, takže program není aktualizován příliš často a nemusí vždy fungovat.

Funkce hledání tichých mutací slouží ke vnesení nového restričního místa do oligonukleotidu, aniž by se změnila sekvence kódovaného proteinu. Díky tomu program nalezne všechna místa vytvořená mutacemi.

Nástroj pro restriční analýzu dovoluje zobrazit všechna restriční místa v zadané sekvenci DNA, přičemž výsledky lze zobrazit v grafické formě, jako jednoduchou tabulku nebo pouze v textové formě.

Analýza SNP-RFLP rychle nalezne restriční místa vhodná pro SNP genotypování pomocí PCR-RFLP (Palmer 2014).

Vstupní část aplikace pro restriční analýzu vyžaduje pouze zadání sekvence, případně názvu pro sekvenci. Ve výstupní části se nachází další možnosti analýzy.

Výrazně praktičtější mi však pro SNP polymorfismus, jaký se vyskytuje u zadaných genů, připadá aplikace pro analýzu SNP-RFLP, kam lze zadat i několik sekvencí, v nichž stačí vyznačit polymorfismus v hranatých závorkách. Z první aplikace není výsledek na první pohled zřejmý a je třeba zjišťovat další informace na webových stránkách z odkazů.

#### **4.2.3 NEBcutter V2.0**

Program NEBcutter V2.0, dostupný z webové stránky <http://nc2.neb.com/NEBcutter2/>, vytvoří zevrubný výpis restričních enzymů, které štěpí zadanou sekvenci DNA. Výstupem programu je mapa restričních enzymů s ukázkou místa štěpení a rozpoznávané sekvence enzymů. Výstup je propojen odkazy s databází restričních enzymů REBASE, která je denně aktualizována. Program také zaznamenává všechna místa potenciálně zasažená methylací DNA. Stejně jako Webcutter 2.0 nabízí NEBcutter hledání v databázi GenBank a výběr lineární a cyklické DNA (Vincze, Posfai a Roberts 2003).

Do programu NEBcutter V2.0 jsem zadala ručně sekvenci DNA, zvolila jsem, že sekvence je lineární, a vybrala jsem k použití enzymy se všemi komerčně dostupnými specifitami. Dále jsem zvolila název sekvence a určila jsem, že sekvence je pouze fragmentem. Další četné možnosti nastavení jsem nevyužila.

## 5 VÝSLEDKY A DISKUZE

### 5.1 Sequence Scanner v1.0

Z programu Sequence Scanner v1.0 jsem zjistila, že alely genu *MCIR* se liší díky bodové mutaci, kterou je substituce cytosinu za tymin (Obr. P8). U genu *MATP* program odhalil, že alely se liší díky výměně guaninu za adenin (Obr. P9).

### 5.2 Webcutter 2.0

Na následujících obrázcích (Obr. 1, 2) je znázorněn výstup z programu Webcutter 2.0. U genu *MCIR* program našel šest vhodných endonukleáz, které představují typický příklad izoschizomer. Pro gen *MATP* program našel dokonce šest vhodných endonukleáz, jež štěpí alelu obsahující adenin. Jedná se o dvě trojice izoschizomer – *MseI*, *Tru1I*, *Tru9I* a *Sse9I*, *TspEI*, *Tsp509I*.

| <a href="#">Graphic map</a>   <a href="#">Table by enzyme name</a> |            | Enzyme | No.  | Positions | Recognition |                           |
|--|------------|--------|------|-----------|-------------|---------------------------|
|  |            | name   | cuts | of sites  | sequence    |                           |
| BsuRI  |            | BspMI  | 1    | 19        | acctgc      | <a href="#">More info</a> |
| HaeIII   |            | BsuRI  | 1    | 4         | gg/cc       | <a href="#">More info</a> |
| CfrI   | BspMI      | CfrI   | 1    | 2         | y/ggccr     | <a href="#">More info</a> |
| ctggccgtgttcgacctgctg  | base pairs | CviJI  | 1    | 4         | rg/cy       | <a href="#">More info</a> |
| gaccggcacacagctggacgac   | 1 to 21    | EaeI   | 1    | 2         | y/ggccr     | <a href="#">More info</a> |
| EaeI   |            | HaeIII | 1    | 4         | gg/cc       | <a href="#">More info</a> |
| CviJI  |            | PalI   | 1    | 4         | gg/cc       | <a href="#">More info</a> |
| PalI   |            |        |      |           |             |                           |

| <a href="#">Graphic map</a>   <a href="#">Table by enzyme name</a> |               | Enzyme  | No.  | Positions | Recognition |                           |
|--|---------------|---------|------|-----------|-------------|---------------------------|
|  |               | name    | cuts | of sites  | sequence    |                           |
| BsuRI  |               | BspMI   | 1    | 19        | acctgc      | <a href="#">More info</a> |
| HaeIII   |               | BsuRI   | 1    | 4         | gg/cc       | <a href="#">More info</a> |
| CfrI   | TthHB8I BspMI | CfrI    | 1    | 2         | y/ggccr     | <a href="#">More info</a> |
| ctggccgtgttcgacctgctg  | base pairs    | CviJI   | 1    | 4         | rg/cy       | <a href="#">More info</a> |
| gaccggcacacagctggacgac   | 1 to 21       | EaeI    | 1    | 2         | y/ggccr     | <a href="#">More info</a> |
| EaeI   | TaqI          | HaeIII  | 1    | 4         | gg/cc       | <a href="#">More info</a> |
| CviJI  |               | PalI    | 1    | 4         | gg/cc       | <a href="#">More info</a> |
| PalI   |               | TaqI    | 1    | 11        | t/cga       | <a href="#">More info</a> |
|  |               | TthHB8I | 1    | 11        | t/cga       | <a href="#">More info</a> |

Obr. 1 Výstup z programu Webcutter 2.0 pro restriční analýzu obou alel genu *MCIR*  
(upraveno, Vilímová 2016)

[Graphic map](#) | [Table by site position](#)

|                                 |            |
|---------------------------------|------------|
| TspEI                           |            |
| Tru9I                           | BsoFI      |
| Tru1I                           | ItaI       |
| ggttctctttt <b>a</b> attttgctgc | base pairs |
| ccaagagaa <b>a</b> ctaaaacgacg  | 1 to 21    |
| MseI                            | Fsp4HI     |
| Sse9I                           |            |
| Tsp509I                         |            |

| Cut site | Enzyme name | Recognition sequence |                           |
|----------|-------------|----------------------|---------------------------|
| 9        | Tru1I       | t/taa                | <a href="#">More info</a> |
| 9        | MseI        | t/taa                | <a href="#">More info</a> |
| 9        | Tru9I       | t/taa                | <a href="#">More info</a> |
| 10       | Sse9I       | /aatt                | <a href="#">More info</a> |
| 10       | TspEI       | /aatt                | <a href="#">More info</a> |
| 10       | Tsp509I     | /aatt                | <a href="#">More info</a> |
| 18       | ItaI        | gc/ngc               | <a href="#">More info</a> |
| 18       | Fsp4HI      | gc/ngc               | <a href="#">More info</a> |
| 18       | BsoFI       | gc/ngc               | <a href="#">More info</a> |

[Graphic map](#) | [Table by site position](#)

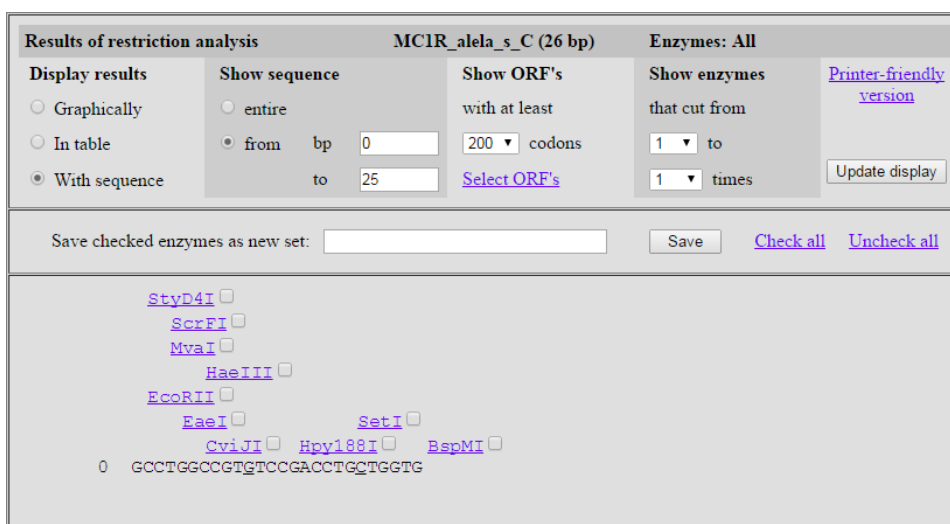
|                                 |            |  |  |
|---------------------------------|------------|--|--|
|                                 | BsoFI      |  |  |
|                                 | ItaI       |  |  |
| ggttctctttt <b>g</b> attttgctgc | base pairs |  |  |
| ccaagagaa <b>a</b> ctaaaacgacg  | 1 to 21    |  |  |
|                                 | Fsp4HI     |  |  |

| Cut site | Enzyme name | Recognition sequence |                           |
|----------|-------------|----------------------|---------------------------|
| 18       | ItaI        | gc/ngc               | <a href="#">More info</a> |
| 18       | Fsp4HI      | gc/ngc               | <a href="#">More info</a> |
| 18       | BsoFI       | gc/ngc               | <a href="#">More info</a> |

Obr. 2 Výstup z programu Webcutter 2.0 pro restriční analýzu obou alel genu *MATP* (upraveno, Vilímová 2016)

### 5.3 WatCut

Výsledky z aplikace pro restriční analýzu programu WatCut (Obr. 3) nebyly pro její nepřehlednost a špatnou intuitivnost použity. Z aplikace pro SNP-RFLP (Obr. 4) jsou výsledky na první pohled patrné, nicméně se na výstupu zobrazují i enzymy, které by štěpily sekvenci jen při mutaci. Z endonukleáz navržených aplikací pro SNP-RFLP vyhovují *Sse9I* a *Tru9I* pro gen *MATP*, *Hpy188I* a *TaqI* pro gen *MC1R*.



Obr. 3 Výstup z programu WatCut pro restriční analýzu jedné alely genu *MC1R*

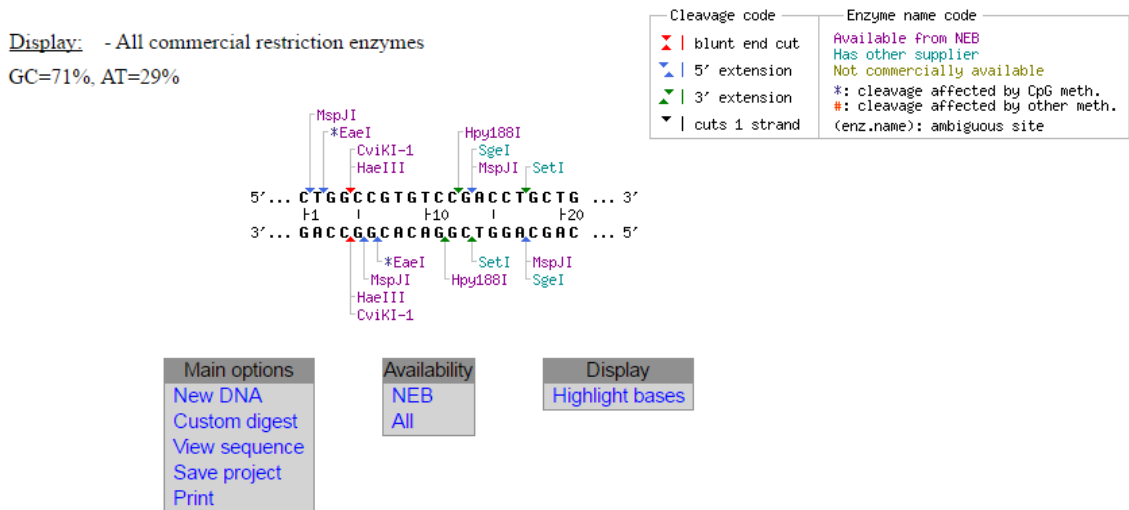
| Results of SNP-RFLP analysis  |   |  |   | Enzymes: All                               |                               |
|---|---|--|---|--|-------------------------------|
| Mutations: up to <input type="text" value="1"/>   |   | Distance to SNP <input type="text" value="3"/> | Sort by <input type="text" value="enzyme"/> | <input type="button" value="Update page"/> | <a href="#">Print version</a> |
| Save checked enzymes as new set: <input type="text"/>   |   |  | <input type="button" value="Save"/>         | <a href="#">Check all</a>                  | <a href="#">Uncheck all</a>   |
| <b>my_working_snp (#1)</b>  |   | GGTTCTCTTT [G/A] ATTTTGCTGC                    |   |  |                               |
| Enzyme  | recognizes                                      | cleaves at                                     | SNP site / mutations:                       | Base changes                               |                               |
| <input type="checkbox"/> <a href="#">HinfI</a>  | G <sup>^</sup> ANTC                             | 1  | GGTTCTCTTT G ATTC <b>T</b> GTCTGC           | 1  |                               |
| <input type="checkbox"/> <a href="#">PfeI</a>   | G <sup>^</sup> AWTC                             | 1  | GGTTCTCTTT G ATTC <b>T</b> GTCTGC           | 1  |                               |
| <input type="checkbox"/> <a href="#">Sse9I</a>  | <sup>^</sup> AATT                               | 0  | GGTTCTCTTT A ATTT <b>T</b> GCTGC            | 0  |                               |
| <input type="checkbox"/> <a href="#">Tru9I</a>  | T <sup>^</sup> TAA                              | -1   | GGTTCTCTTT A ATTT <b>T</b> GCTGC            | 0  |                               |
| These enzymes recognize the sequence within 25 bp of the polymorphic site and were excluded:<br>AbaSI, BisI, BlnI, BstV1I, FspEI, MspJI, TseI                             |   |  |   |  |                               |
| <b>my_working_snp (#2)</b>  |   | CTGGCCGTGT [C/T] CGACCTGCTG                    |   |  |                               |
| Enzyme  | recognizes                                      | cleaves at                                     | SNP site / mutations:                       | Base changes                               |                               |
| <input type="checkbox"/> <a href="#">AfiI</a>   | CCN <sub>5</sub> <sup>^</sup> N <sub>2</sub> GG | 7  | CTGGCCGTGT C CGACCTG <b>C</b> GG            | 1  |                               |
| <input type="checkbox"/> <a href="#">Bpu14I</a>   | TT <sup>^</sup> CGAA                            | 1  | CTGGCCGTGT T CGA <b>A</b> CTGCTG            | 1  |                               |
| <input type="checkbox"/> <a href="#">Hpy188I</a>  | TCN <sup>^</sup> GA                             | 2  | CTGGCCGTGT C CGACCTG <b>C</b> TG            | 0  |                               |
| <input type="checkbox"/> <a href="#">MmeI</a>   | TCCRACN <sub>20</sub> <sup>^</sup>              | 25   | CTGGCCGTGT C CGACCTG <b>C</b> TG            | 0  |                               |
| <input type="checkbox"/> <a href="#">TaqI</a>   | T <sup>^</sup> CGA                              | 1  | CTGGCCGTGT T CGACCTG <b>C</b> TG            | 0  |                               |
| <input type="checkbox"/> <a href="#">Tth111I</a>  | GACN <sup>^</sup> N <sub>2</sub> GTC            | -4   | CTGA <b>C</b> CGTGT C CGACCTG <b>C</b> TG   | 1  |                               |
| These enzymes recognize the sequence within 25 bp of the polymorphic site and were excluded:<br>AbaSI, BceAI, BspMI, CviJI, EaeI, FspEI, HaeIII, LpnPI, MspJI, SetI, SgeI |   |  |   |  |                               |

Obr. 4 Výstup z programu WatCut pro analýzu SNP-RFLP (upraveno, Vilímová 2016)

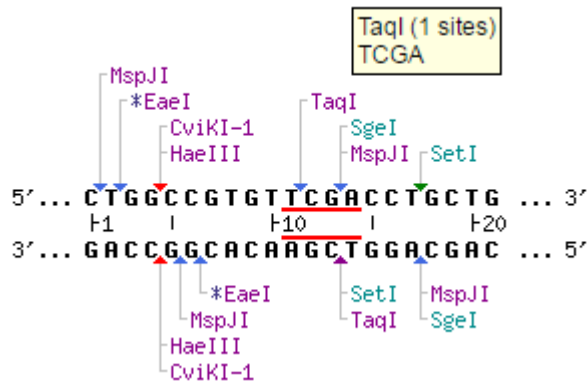
## 5.4 NEBcutter V2.0

Ve výstupu programu NEBcutter V2.0 se zobrazí vždy mapa enzymů. K přehlednosti přispívá barevné rozlišení, tabulka vysvětlivek a možnost barevně zvýraznit jednotlivé báze (Obr. 5). Najetí myší na název endonukleázy ukáže počet štěpných míst a rozpoznávanou sekvenci enzymu, přičemž je ještě podtržena tato část celkové sekvence (Obr. 6).

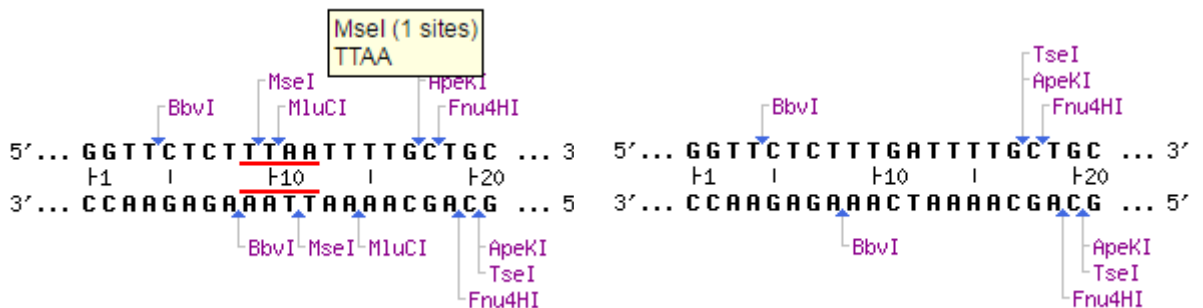
Co se týče výsledků, našel program NEBcutter V2.0 pro štěpení genu *MC1R* endonukleázu *Hpy188I* (na alele s cytosinem v místě polymorfizmu) a *TaqI* (na alele s tyminem v místě polymorfizmu). Pro gen *MATP* (Obr. 7) byla nalezena štěpná místa pro enzymy pouze na alele s adeninem v místě polymorfizmu a k restrikci byly navrženy endonukleázy *MluCI* a *MseI*.



Obr. 5 Kompletní výstup z programu NEBcutter V2.0 pro restriční analýzu jedné alely genu *MC1R*



Obr. 6 Výstup z programu NEBcutter V2.0 pro restriční analýzu druhé alely genu *MC1R*



Obr. 7 Výstup z programu NEBcutter V2.0 pro restriční analýzu obou alel genu *MATP*

## 5.5 Shrnutí a diskuze

Ačkoliv každý z programů poskytl veskrze rozdílné návrhy restričních enzymů, nastaly ve výsledcích i shody. Pro testování polymorfizmu genu *MC1R* bych z nabízených možností volila na prvním místě endonukleázu *TaqI* (T|CGA) a jako druhou variantu enzym *Hpy188I* (TCN|GA). *TaqI* sice vyžaduje vyšší teplotu inkubace, na druhou stranu má však specifitější rozpoznávanou sekvenci (New England Biolabs 2016). Co se týče výběru restriktázy pro zjištění polymorfizmu genu *MATP*, rozhodla bych se pro *MseI* (T|TAA), jelikož je dobře dostupná, stejně jako *MluCI*, což by byla moje druhá volba. Mírnou nevýhodou *MluCI* však je, že ji nelze příhodně inaktivovat termicky a bylo by nutné použít kolonu, případně fenolovou nebo chloroformovou extrakci (New England Biolabs 2016).

Porovnání softwaru je uvedeno v tabulce 1:

Tab. 1 Porovnání funkcí webových programů

|  | Webcutter 2.0 | WatCut | NEBcutter V2.0 |
|--|---------------|--------|----------------|
| <b>Vložení sekvence</b>                                |               |        |                |
| vypsání ručně  | +             | +      | +              |
| soubor z počítače                                      | +             | +      | +              |
| z GenBank  | +             | –      | +              |
| <b>Pojmenování sekvence</b>                            | +             | +      | +              |
| <b>Kruhová sekvence</b>                                | +             | –      | +              |
| <b>Tichá mutace</b>                                    | +             | –      | –              |
| <b>Zobrazení výsledků</b>                              |               |        |                |
| mapa   | +             | +      | +              |
| tabulka  | +             | –      | –              |
| <b>Výběr enzymů</b>                                    |               |        |                |
| všechny  | +             | +      | +              |
| všechny komerčně dostupné                              | –             | –      | +              |
| dle počtu štěpení                                      | +             | +      | –              |
| dle délky restr. místa                                 | +             | –      | +              |
| ruční výběr  | +             | –      | +              |
| <b>Uložení projektů</b>                                | –             | +      | +              |
| <b>Uvedení rozpoznávané sekvence nalezených enzymů</b> | +             | –      | +              |
| <b>Indikace místa štěpení</b>                          | +             | –      | ++             |
| <b>Propojení s REBASE</b>                              | – + (*)       | +      | +              |
| <b>Barevné zvýraznění</b>                              | –             | –      | +              |
| <b>Přehlednost</b>                                     | +             | –      | ++             |

\* odkazy nefungují



Dle mého názoru je nejlepším programem NEBcutter V2.0, a to díky výborné přehlednosti, nejširším možnostem při zadávání sekvence, propojení s webovými stránkami New England Biolabs ([www.neb.com](http://www.neb.com)) a REBASE, ale hlavně kvůli neustálé aktualizaci, která zbylým dvěma programům očividně chybí.

Na závěr jsem se v diskuzi chtěla věnovat rozporu v psaní názvů restričních endonukleáz kurzívou. Pod společným ustanovením pro úpravy nomenklatury (Roberts et al. 2003) jsou podepsaná zvučná jména jako Dryden, Kobayashi, Murray, Pingoud, Raleigh, Rao, Wilson a další. Úpravou se řídí i Loenen (př. Loenen et al. 2014a) a známí autoři učebních textů, jako jsou Pritchard a Korf (Pritchard a Korf 2013) nebo Alberts (Alberts et al. 2015). Mezi autory užívající kurzívu v názvech restričních enzymů patří světově uznávaní Snustad a Simmons (Snustad a Simmons 2009), dále také Griffiths (Griffiths et al. 2012) nebo Clark (Clark a Pazdernik 2013). V českém prostředí tradice užívání kurzívy nebyla porušena a stojí si za ní například genetici z Mendelovy Univerzity v Brně nebo z Ústavu molekulární genetiky AV ČR v Praze (Novotný et al. 2012). Nicméně zde také nejsou všichni autoři jednotní, viz Kočárek (Kočárek 2008).

Dle mého názoru by se měl úzus psaní názvů restričních enzymů sjednotit, i když je to jistě obtížný úkol. Jaká varianta bude zvolena, není příliš podstatné. Vydavatelstvím časopisů vyhovuje klasický font, čtenářům se zase lépe v textu hledají názvy psané kurzívou. Myslím si, že hlavní je, aby kvůli těmto rozdílům nedocházelo k nedorozuměním a nepochopení dané problematiky.

## 6 ZÁVĚR

Tato bakalářská práce se zabývá problematikou restričních endonukleáz. V literární rešerši jsou tyto štěpící enzymy charakterizovány a rozděleny na jednotlivé typy. Dále text pojednává o sekvenční specifitě a o vlivech podmínek reakce. Nakonec jsou popsána nejběžnější využití enzymů v molekulární genetice a diagnostice.

Restriční endonukleázy jsou enzymy štěpící DNA uvnitř její molekuly, které pracují společně s methyltransferázami v bakteriálních buňkách na ochraně genetické informace. Rozdělují se do čtyř typů a doposud jich bylo charakterizováno více než 4 tisíce a teoreticky nalezeno přes 38 tisíc. Endonukleázy jsou velmi přesné v rozpoznávání specifické sekvence, ale vlivem reakčních podmínek se mohou stát promiskuitními. Pro správný průběh restriční reakce je třeba sledovat teplotu, pH, kofaktory, složení soli, přítomnost stabilizátorů, správný objem a dobu trvání reakce.

Endonukleázy jsou důležitým pilířem molekulární genetiky. Představují levný, rychlý a přesný analytický nástroj. Využití nacházejí například v mapování genomu, chromozomů či jednotlivých genů, uplatňují se také při tvorbě rekombinantní DNA a v genetických testech, jako je určování paternity a diagnostika dědičných chorob lidí i zvířat. Mohou být také součástí forenzních metod k určení původu biologických materiálů. V potravinářství najdou využití při kontrole kvality potravin nebo obsahu masa různých druhů zvířat v produktech. V chovu zvířat mohou sloužit k určování genotypu jako jeden z nástrojů pro selekci a šlechtění.

Praktická část práce je založena na práci se softwarem pro analýzu restričních míst, a to konkrétně s programy Webcutter 2.0, WatCut a NEBcutter V2.0. Pro zadané geny *MATP* a *MC1R* byla úspěšně nalezena štěpná místa a navrženy vhodné restriční enzymy pro testování polymorfizmů (SNP) – pro *MC1R* endonukleáza *TaqI* a pro *MATP* endonukleáza *MseI*. V porovnání programů se jako nejlepší umístil NEBcutter V2.0.

## 7 POUŽITÁ LITERATURA

- ALBERTS, Bruce, Alexander JOHNSON, Julian LEWIS, David Owen MORGAN, Martin C RAFF, Keith ROBERTS a Peter WALTER. *Molecular biology of the cell*. Sixth edition. New York, NY: Garland Science, 2015. ISBN 978-0-8153-4432-2.
- BITINAITE, J., D. A. WAH, A. K. AGGARWAL a I. SCHILDKRAUT. FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci U S A* [online]. 1998, 95(18), s. 10570-5 [cit. 2016-04-11]. Dostupné z: <http://www.ncbi.nlm.nih.gov/pubmed/9724744>
- CABE, P. R. a R. P. CAPRICCIOSO. DNA fingerprinting. *Salem Press Encyclopedia of Health* [online]. 2016, 4 s. [cit. 2016-03-27]. Dostupné z: <http://eds.b.ebscohost.com/eds/detail/detail?sid=13d93c40-6a47-4f61-a5d7-9ee33169ae8c%40sessionmgr114&vid=7&hid=111&bdata=Jmxhbmc9Y3Mmc2l0ZT1lZHMtOGI2ZQ%3d%3d#AN=94416449&db=ers>
- CLARK, David P a Nanette Jean PAZDERNIK. *Molecular biology*. 2nd ed. Waltham, MA: Academic Press, 2013. ISBN 7143142.
- Laboratoř imunogenetiky: cmsch.cz. ČESKOMORAVSKÁ SPOLEČNOST CHOVATELŮ, A. S. *Českomoravská společnost chovatelů, a. s.* [online]. Hradištko, 2004, 2016 [cit. 2016-04-15]. Dostupné z: <http://www.cmsch.cz/laborator-imunogenetiky/>
- DAI, Shutao a Y. LONG. Genotyping Analysis Using an RFLP Assay. *Methods Mol Biol*. [online]. New York: Springer Science+Business Media, 2015, 1245, s. 91-99 [cit. 2016-03-27]. DOI: 10.1007/978-1-4939-1966-6\_7. ISBN 10.1007/978-1-4939-1966-6\_7. Dostupné z: [http://link.springer.com/10.1007/978-1-4939-1966-6\\_7](http://link.springer.com/10.1007/978-1-4939-1966-6_7)
- GRIFFITHS, A. J. F., J. H. MILLER, D. T. SUZUKI, R. C. LEWONTIN a W. M. GELBART. *Introduction to Genetic Analysis*. 7. vyd. New York: Freeman, W. H. & Company, 2000. ISBN 0716735202.
- GRIFFITHS, A. J. F., S. R. WESSLER, R. C. LEWONTIN, W. M. GELBART, J. H. MILLER a D. T. SUZUKI. *Introduction to Genetic Analysis*. 8. vyd. New York: Freeman, W. H. & Company, 2005. ISBN 0716749394.
- GRIFFITHS, A. J., S. R. WESSLER, S. B. CARROLL a J. DOEBLEY. *Introduction to genetic analysis*. 10. vyd. New York: Freeman, W. H. & Company, 2012. ISBN 1429272775.
- HEAD, S. R., H. KIYOMI KOMORI, S. A. LAMERE, T. WHISENANT, F. VAN NIEUWERBURGH, D. R. SALOMON a P. ORDOUKHANIAN. Library construction for next-generation sequencing: Overviews and challenges. *Biotechniques* [online]. 2014, 56(2), s. 61–77 [cit. 2016-04-11]. DOI:

- 10.2144/000114133. Dostupné z: <http://www.biotechniques.com/BiotechniquesJournal/2014/February/Library-construction-for-next-generation-sequencing-Overviews-and-challenges/biotechniques-349889.html>
- HEIMAN, Max. *Webcutter 2.0* [online]. 1997 [cit. 2016-03-03]. Dostupné z: <http://bio.biomedicine.gu.se/cutter2/>
- KAMPS-HUGHES, N., A. QUIMBY, Z. ZHU a E. A. JOHNSON. Massively parallel characterization of restriction endonucleases. *Nucleic Acids Research*. 2013, 41(11): e119 [cit. 2016-03-01]. DOI: 10.1093/nar/gkt257. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt257>
- KELLY, T. J. Jr a SMITH, H. O. A restriction enzyme from *Hemophilus influenzae*: II. Base sequence of the recognition site. *J. Mol. Biol.* 1970, 51(2), s. 393–409. [cit. 2016-04-15]. DOI: 10.1016/0022-2836(70)90150-6. ISSN 00222836. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/0022283670901506>
- KIM, Y., J. GRABLE, R. LOVE, P. GREENE a J. ROSENBERG. Refinement of EcoRI endonuclease crystal structure: a revised protein chain tracing. *Science* [online]. 1990, 249(4974), s. 1307-1309 [cit. 2016-03-10]. DOI: 10.1126/science.2399465. ISSN 1095-9203. Dostupné z: <http://www.sciencemag.org/cgi/doi/10.1126/science.2399465>
- KOČÁREK, Eduard. *Genetika: obecná genetika a cytogenetika, molekulární biologie, biotechnologie, genomika*. 2. vyd. Ilustrace Jan Maget. Praha: Scientia, 2008. Biologie pro gymnázia (Scientia). ISBN 978-80-86960-36-4.
- KUCIEL, J., J. BEDNÁŘ, T. URBAN. *Genetika zemědělských produktů: Vybrané kapitoly k přednáškám*. 1. vyd. Brno: Mendelova zemědělská a lesnická univerzita v Brně, 2004, 137 s. ISBN 80-7157-767-7.
- LOENEN, W. A. M., D. T. F. DRYDEN, E. A. RALEIGH, G. G. WILSON a N. E. MURRAY. Highlights of the DNA cutters: a short history of the restriction enzymes. *Nucleic Acids Research*. 2014a , 42(1): s. 3-19. DOI: 10.1093/nar/gkt990. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt990>
- LOENEN, W. A. M., D. T. F. DRYDEN, E. A. RALEIGH a G. G. WILSON. Type I restriction enzymes and their relatives. *Nucleic Acids Research*. 2014b , 42(1): s. 20-44. DOI: 10.1093/nar/gkt847. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt847>
- LOENEN, W. A. M. a E. A. RALEIGH. The other face of restriction: modification-dependent enzymes. *Nucleic Acids Research*. 2014c, 42(1): s. 56-69. DOI: 10.1093/nar/gkt747. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt747>

- MARIAT, Denis, Sead TAOURIT a Gérard GUÉRIN. A mutation in the MATP gene causes the cream coat colour in the horse. *GENETICS SELECTION EVOLUTION* [online]. BIOMED CENTRAL LTD, 2003, 35(1), s.119-133 [cit. 2016-02-29]. DOI: 10.1051/gse:2002039. ISBN 10.1051/gse:2002039. ISSN 0999193X. Dostupné z: <http://www.gse-journal.org/10.1051/gse:2002039>
- MARINI, S. J., L. S. VANZETTI, V. S. BORELLI, A. O. VILLAREAL, G. D. DENEGRI, G. A. COTTURA, D. PANICHELLI, P. SILVA, D. CAMPAGNA, N. SPINER, J. C. BRUNORI a R. FRANCO. RYR1 gene variability and effect on meat pH in Argentinean hybrids swines. *In Vet.* [online]. 2012, 14(1), s. 19-23 [cit. 2016-03-30]. Dostupné z: [http://www.fvet.uba.ar/publicaciones/archivos/vol\\_14\\_n1\\_2012/Articulo-2.pdf](http://www.fvet.uba.ar/publicaciones/archivos/vol_14_n1_2012/Articulo-2.pdf)
- MCCLARIN, J. A., C. A. FEDERICK, B. C. WANG, P. GREENE, H. W. BOYER, J. GRABLE a J. M. ROSENBERG. Structure of the DNA-Eco RI endonuclease recognition complex at 3 Å resolution. *Science* [online]. 1986, 234(4783), s. 1526-41 [cit. 2016-03-10]. Dostupné z: <http://science.sciencemag.org/content/234/4783/1526.long>
- MÜLLER, A. a H. STEINHART. Recent developments in instrumental analysis for food quality. *Food Chemistry* [online]. 2007, 102(2), s. 436-444 [cit. 2016-03-30]. DOI: 10.1016/j.foodchem.2006.10.015. ISSN 03088146. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0308814606007990>
- MURRAY, N. E. Type I Restriction Systems: Sophisticated Molecular Machines (a Legacy of Bertani and Weigle). *Microbiology and Molecular Biology Reviews*. 2000, 64(2): s. 412-434. Dostupné z: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC98998/>
- Cleaved Amplified Polymorphic Sequences (CAPS). NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. *National Center for Biotechnology Information* [online]. Bethesda MD, USA, 2014b-09-24 [cit. 2016-04-11]. Dostupné z: <http://www.ncbi.nlm.nih.gov/probe/docs/techcaps/>
- NEW ENGLAND BIOLABS. *New England Biolabs* [online]. 2016 [cit. 2016-03-22]. Dostupné z: <https://www.neb.com/>
- NOVOTNÝ, I., K. PODOLSKÁ, M. BLAŽÍKOVÁ, L. SHIVAYA VALÁŠEK, P. SVOBODA a D. STANĚK. Nuclear LSM8 affects number of cytoplasmic processing bodies via controlling cellular distribution of Like-Sm proteins. *Molecular Biology of the Cell* [online]. 2012, (23), s. 3776-3785 [cit. 2016-04-02]. Dostupné z: <http://www.iem.cas.cz/miranda2/export/sitesavcr/data.avcr.cz/lifesci/uem/research/publications/microscopy-unit/files/novotny-2012.pdf>

- Help. PALMER, Michael. UNIVERSITY OF WATERLOO. *WatCut* [online]. University of Waterloo, Ontario, Canada, 2014-04-17 [cit. 2016-04-17]. Dostupné z: <http://watcut.uwaterloo.ca/template.php?%20act=help>
- PINGOUD, A., M. FUXREITER, A. PINGOUD, W. WENDE. Type II restriction endonucleases: structure and mechanism. *Cellular and Molecular Life Sciences* [online]. Birkhauser-Verlag, 2004, 62(6) [cit. 2014-04-27], s. 685-707. ISSN: 1420-682X. Dostupné z: <http://link.springer.com/article/10.1007%2Fs00018-004-4513-1>
- PINGOUD, A., G. G. WILSON a W. WENDE. Type II restriction endonucleases--a historical perspective and more. *Nucleic Acids Research*. 2014, 42(12): s. 7489-7527. DOI: 10.1093/nar/gku447. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gku447>
- PRITCHARD, D a Bruce R KORF. *Medical genetics at a glance*. 3rd ed. Chichester, England: Wiley-Blackwell, 2013.
- Applications and Reaction Conditions For Restriction Enzymes. Promega Restriction Enzyme Resource. *Promega Corporation* [online]. 2016 [cit. 2016-03-14]. Dostupné z: <http://worldwide.promega.com/resources/product-guides-and-selectors/restriction-enzyme-resource/applications-and-reaction-conditions-for-restriction-enzymes/#substrate>
- RAO, D. N., D. T. F. DRYDEN a S. BHEEMANAİK. Type III restriction-modification enzymes: a historical perspective. *Nucleic Acids Research*. 2013, 42(1): s. 45-55. DOI: 10.1093/nar/gkt616. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt616>
- REVEL, H. R. Restriction of nonglycosylated T-even bacteriophage: properties of permissive mutants of *Escherichia coli* B and K12. *Virology*. 1967, 31(4), s. 688-701.
- ROBERTS, R. J., M. BELFORT, T. BESTOR, A. S. BHAGWAT, T. A. BICKLE, J. BITINAITE, R. M. BLUMENTHAL, S. KH. DEGTYAREV, D. T. F. DRYDEN, K. DYBVIK, K. FIRMAN, E. S. GROMOVA, R. I. GUMPORT, S. E. HALFORD, S. HATTMAN, J. HEITMAN, D. P. HOMBY, A. JANULAITIS, A. JELTSCH, J. JOSEPHSEN, A. KISS, T. R. KLAENHAMMER, I. KOBAYASHI, H. KONG, D. H. KRÜGER, S. LACKS, M. G. MARINUS, M. MIYAHARA, R. D. MORGAN, N. E. MURRAY, V. NAGARAJA, A. PIEKAROWICZ, A. PINGOUD, E. RALEIGH, D. N. RAO, N. REICH, V. E. REPIN, E. U. SELKER, P. SHAW, D. C. STEIN, B. L. STODDARD, W. SZYBALSKI, T. A. TRAUTNER, J. L. VAN ETTEN, J. M. B. VITOR, G. G. WILSON a SHUANG-YONG XU. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Research*. 2003, 31(7), s. 1805-1812. DOI: 10.1093/nar/gkg274. ISSN 13624962. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkg274>

- ROBERTS, R. J. *REBASE - The Restriction Enzyme Database* [online]. 1994, 2016-01-29a [cit. 2016-01-29]. REBASE Statistics. Dostupné z: <http://rebase.neb.com/rebase/statlist.html>
- ROBERTS, R. J. *REBASE - The Restriction Enzyme Database* [online]. 1994, 2016-01-29b [cit. 2016-01-29]. REBASE Statistics – Enzymes by Year. Dostupné z: <http://rebase.neb.com/rebase/rebenzyyearbar.html>
- RUSSELL, P. J. *iGenetics: a molecular approach*. 3rd ed. San Francisco: Benjamin Cummings, 2010. ISBN 0321569768.
- SEEMAN, N. C., J. M. ROSENBERG a A. RICH. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc Natl Acad Sci U S A*. [online]. 1976, 73(3), s. 804-808 [cit. 2016-03-10]. Dostupné z: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC336007/>
- SLUTSKY, Jeffrey, Karthik RAJ, Scott YUHNKE, Jerold BELL, Neale FRETWELL, Ake HEDHAMMAR, Claire WADE a Urs GIGER. A web resource on DNA tests for canine and feline hereditary diseases. *The Veterinary Journal* [online]. 2013, 197(2), s. 182-187 [cit. 2016-04-15]. DOI: 10.1016/j.tvjl.2013.02.021. ISSN 10900233. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S109002331300083X>
- SMITH, H.O. and NATHANS, D. A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J. Mol. Biol.* 1973, 81(3), s.419-423.
- SMITH, R. M., PERNSTICH, C. a HALFORD, S. E. TstI, a Type II restriction-modification protein with DNA recognition, cleavage and methylation functions in a single polypeptide. *Nucleic Acids Res.* [online]. 2014, 42(9), s. 5809–5822 [cit. 2016-02-08]. DOI: 10.1093/nar/gku187. ISBN 10.1093/nar/gku187. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gku187>
- SNUSTAD, D a Michael J SIMMONS. *Genetika*. Vyd. 1. Překlad Jiřina Relichová. Brno: Masarykova univerzita, 2009. ISBN 978-80-210-4852-2.
- STOLL, Vincent S. a John S. BLANCHARD. [4] Buffers: Principles and practice. *Methods in Enzymology* [online]. Elsevier Inc., 1990, 182, s. 24-38 [cit. 2016-03-14]. DOI: 10.1016/0076-6879(90)82006-N. ISBN 10.1016/0076-6879(90)82006-N. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/007668799082006N>
- Buffer Activity Chart. TAKARA BIO COMPANY. *Clontech Laboratories, Inc.: Life science tools and reagents* [online]. 2016 [cit. 2016-03-18]. Dostupné z: [http://www.clontech.com/US/Products/Molecular\\_Biology\\_Tools/Restriction\\_Enzymes/Buffer\\_Activity\\_Chart](http://www.clontech.com/US/Products/Molecular_Biology_Tools/Restriction_Enzymes/Buffer_Activity_Chart)

- Restriction enzyme buffer set* [online]. THERMO FISHER SCIENTIFIC INC. Carlsbad, CA, 2010 [cit. 2016-03-18]. Dostupné z: [https://tools.thermofisher.com/content/sfs/manuals/bufferset\\_man.pdf](https://tools.thermofisher.com/content/sfs/manuals/bufferset_man.pdf)
- THIRUVENKADAN, A.K., N. KANDASAMY a S. PANNEERSELVAM. Coat colour inheritance in horses. *Livestock Science* [online]. 2008, 117(2-3), s. 109-129 [cit. 2016-02-29]. DOI: 10.1016/j.livsci.2008.05.008. ISSN 18711413. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S1871141308001376>
- VASU, K., E. NAGAMALLESWARI, M. ZAHRAN, P. IMHOF, S.-Y. XU, Z. ZHU, S.-H. CHAN a V. NAGARAJA. Increasing cleavage specificity and activity of restriction endonuclease KpnI. *Nucleic Acids Research*. 2013, 41(21): s. 9812-9824. DOI: 10.1093/nar/gkt734. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt734>
- VINCZE, T., J. POSFAI, R. J. ROBERTS. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research*. 2003, 31(13): s. 3688-3691. DOI: 10.1093/nar/gkg526. ISSN 1362-4962. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkg526>
- WEI, H., C. THERRIEN, A. BLANCHARD, S. GUAN a Z. ZHU. The Fidelity Index provides a systematic quantitation of „star“ activity of DNA restriction endonucleases. *Nucleic Acids Research* [online]. 2008, 36(9), e50 [cit. 2016-02-25]. DOI: 10.1093/nar/gkn182. ISSN 0305-1048. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkn182>
- WILSON, Geoffrey G., Hua WANG, Daniel F. HEITER a Keith D. LUNNEN. Restriction Enzymes in Microbiology, Biotechnology and Biochemistry. *Encuentro*. 2012, 93: s. 19-48 [cit. 2016-02-25]. ISSN 0424-9674. Dostupné z: <http://www.uca.edu.ni/2/images/Revista-Encuentro/Revistas/e93/art-2.pdf>



## 8 SEZNAM ZKRATEK

|            |  |
|------------|--|
| A          | adenin   |
| AdoMet     | S-adenosylmethionin (také SAM)   |
| AMK        | aminokyselina  |
| ATP        | adenosintrifosfát  |
| BAC        | umělý bakteriální chromozom  |
| BLAD       | Bovine Leukocyte Adhesion Deficiency   |
| bp         | pár bází   |
| BSA        | bovine serum albumin   |
| C          | cytosin  |
| cDNA       | komplementární DNA   |
| CVM        | Complex vertebrae Malformatioon (komplexní vertebrální malformace)                     |
| DNA        | deoxyribonukleová kyselina   |
| ddNTP      | dideoxynukleotid   |
| dNTP       | deoxynukleotid   |
| DTE        | dithioerytriol   |
| DTT        | dithiotreitol  |
| EDTA       | ethylendiamintetraoctová kyselina  |
| EOP        | efektivita výsevu  |
| FI         | Fidelity Index   |
| G          | guanin   |
| <i>hsd</i> | <i>host specificity determinant</i> (určující člen specifity hostitele)                |
| kb         | kilobáze   |
| kDa        | kilodalton   |
| Mb         | megabáze   |
| mM         | jednotka molární koncentrace, odpovídá jednotce mmol/dm <sup>3</sup>                   |
| MT         | methyltransferáza  |
| N          | nespecifický nukleotid (zahrnuje A, C, G, T)   |
| PCR        | polymerázová řetězová reakce (polymerase chain reaction)                               |
| PSE        | Pale, Soft, Exudative (bledé, měkké, vodnaté)  |
| R          | purinové báze (A, T)   |
| R-M        | restrikčně – modifikační   |
| RES        | restrikční endonukleáza  |
| RFLP       | restriction fragment length polymorphisms (polymorfizmus délky restrikčních fragmentů) |
| RNA        | ribonukleová kyselina  |
| S          | báze se silným párováním (G, C)  |
| SCID       | Severe Combined Immunodeficiency Disorder  |
| SMRT       | Single-Molecule Real-Time  |
| SNP        | single nucleotide polymorphism (jednonukleotidový polymorfizmus)                       |
| T          | tymin  |
| TRD        | target recognition domain (doména rozpoznání c íle)                                    |
| UV         | ultrafialové (ultraviolet)   |
| VNTR       | variable number tandem repeats (tandemové repetice o variabilním počtu)                |
| W          | báze se slabým párováním (A, T)  |
| Y          | pyrimidinové báze (G, C)   |

## 9 SEZNAM OBRÁZKŮ A PŘÍLOH

### Obrázky

|        |  |    |
|--------|--|----|
| Obr. 1 | Výstup z programu Webcutter 2.0 pro restrikční analýzu obou alel genu <i>MCIR</i> .....              | 43 |
| Obr. 2 | Výstup z programu Webcutter 2.0 pro restrikční analýzu obou alel genu <i>MATP</i> .....              | 44 |
| Obr. 3 | Výstup z programu WatCut pro restrikční analýzu jedné alely genu <i>MCIR</i> . 44                    |    |
| Obr. 4 | Výstup z programu WatCut pro analýzu SNP-RFLP.....   | 45 |
| Obr. 5 | Kompletní výstup z programu NEBcutter V2.0 pro restrikční analýzu jedné alely genu <i>MCIR</i> ..... | 46 |
| Obr. 6 | Výstup z programu NEBcutter V2.0 pro restrikční analýzu druhé alely genu <i>MCIR</i> .....           | 46 |
| Obr. 7 | Výstup z programu NEBcutter V2.0 pro restrikční analýzu obou alel genu <i>MATP</i> .....             | 46 |

### Přílohy

|         |  |    |
|---------|--|----|
| Tab. P1 | Charakteristika restrikčních enzymů, jejich genů a podjednotek.....                            | 58 |
| Tab. P2 | Charakteristika R–M enzymů Typu I.....   | 58 |
| Tab. P3 | Změna AMK v proteinové struktuře S domény enzymu <i>MmeI</i> mění rozpoznávanou sekvenci ..... | 59 |
| Tab. P4 | R–M systémy Typu III .....   | 59 |
| Tab. P5 | Přehled modifikací vyvolávajících štěpení u enzymů Typu IV .....                               | 59 |
| Obr. P1 | Princip RFLP .....   | 60 |
| Obr. P2 | Tvorba restrikční mapy .....   | 60 |
| Obr. P3 | Modelové vyhodnocení genotypů prasat PCR-RFLP pro <i>RYRI</i> .....                            | 61 |
| Obr. P4 | Snímek elektroforetické separace po PCR-RFLP genu <i>RYRI</i> .....                            | 61 |
| Obr. P5 | Tvorba rekombinantní molekuly DNA .....  | 61 |
| Obr. P6 | Amplifikace rekombinantní DNA.....   | 62 |
| Obr. P7 | Subklonování.....  | 62 |
| Obr. P8 | Určení polymorfizmu genu <i>MCIR</i> v programu Sequence Scanner v1.0 .....                    | 63 |
| Obr. P9 | Určení polymorfizmu genu <i>MATP</i> v programu Sequence Scanner v1.0 .....                    | 64 |

## 10 PŘÍLOHY

### A. Doplnující tabulky

Tab P1. Charakteristika restrikčních enzymů, jejich genů a podjednotek  
(Loenen et al. 2014a)

| Type     | Type I   | Type II  | Type III  | Type IV   |
|----------|--|--|---|---|
| Features | Oligomeric REase and MTase complex<br>Require ATP hydrolysis for restriction<br>Cleave variably, often far from recognition site<br>'DEAD-box' translocating REase<br>bipartite DNA recognition domain | Separate REase and MTase or combined<br>REase~MTase fusion<br>Cleave within or at fixed positions close to recognition site<br>Many different subtypes | Combined REase+MTase complex<br>ATP required for restriction<br>Cleave at fixed position outside recognition site<br>'DEAD-box' REase | Methylation-dependent REase<br>Cleave at variable distance from recognition site<br>Cleave m6A, m5C, hm5C and/or other modified DNA<br>Many different types |
| Example  | e.g. EcoKI   | e.g. EcoRI   | e.g. EcoPII   | No 'typical' example  |
| Genes    | <i>hsdR, hsdM, hsdS</i>  | e.g. <i>ecorIR, ecorIM</i>   | e.g. <i>ecoPIIM, ecoPIIR</i>  | e.g. <i>mcrA, mcrBC, mrr</i>  |
| Subunits | ~135, ~62 and ~52 kDa  | ~31 and ~38 kDa for EcoRI  | ~106 and ~75 kDa for EcoPII   | Unrelated proteins  |
| Proteins | REase: 2R + 2M + S<br>MTase: 2M + S (±2R)  | Orthodox REase: 2R<br>Orthodox MTase: M  | REase: 1 or 2 R + 2M<br>MTase: 2M (±2R)   | Varies  |
| REBASE   | 104 enzymes, 47 genes cloned, 34 genes sequenced, 5140 putatives   | 3938 enzymes, 633 genes cloned, 597 sequenced, 9632 putatives  | 21 enzymes, 19 genes cloned & sequenced, 1889 putatives   | 18 enzymes & genes cloned, 15 sequenced, 4822 putatives   |

Tab. P2 Charakteristika R–M enzymů Typu I (Loenen et al. 2014b)

| Enzyme               | Family | Recognition sequence | Me-interval |
|----------------------|--------|----------------------|-------------|
| EcoBI                | IA     | TGA N8 TGCT          | 8           |
| EcoKI                | IA     | AAC N6 GTGC          | 8           |
| EcoAI                | IB     | GAG N7 GTCA          | 9           |
| StyLTHI              | IA     | GAG N6 RTAYG         | 8           |
| StySPI               | IA     | AAC N6 GTRC          | 8           |
| EcoDI                | IA     | TTA N7 GTCY          | 8           |
| EcoDXXI              | IC     | TCA N7 RTTC          | 8           |
| EcoRI24I             | IC     | GAA N6 RTCG          | 7           |
| EcoEI                | IB     | GAG N7 ATGC          | 9           |
| CfrAI                | IB     | GCA N8 GTGG          | 9           |
| EcoprrI              | IC     | CCA N7 RTGC          | 8           |
| StySKI               | IB     | CGAT N7 GTTA         | 9           |
| StySBLI              | ID     | CGA N6 TACC          | 6           |
| NgoAV <sup>a</sup>   | IC     | GCA N8 TGC           | 8           |
| Eco377I              |        | GGA N8 ATGC          | 9           |
| Eco585I <sup>b</sup> |        | GCC N6 TGCG          | ?           |
| Eco646I              |        | CCA N7 CTTC          | ?           |
| Eco777I              |        | GGA N6 TATC          | ?           |
| KpnBI                | IE     | CAAA N6 RTCA         | 7           |
| KpnAI                | ID     | GAA N6 TGCC          | 6           |
| StySEAI              |        | ACA N6 TYCA          | 6           |
| StySGI               |        | TAAC N7 RTCG         | 9           |
| Eco394I              |        | GAC N5 RTAAY         | 7           |
| Eco826I              |        | GCA N6 CTGA          | 7           |
| Eco851I              |        | GTCA N6 TGAY         | 6           |
| Eco912I              |        | CAC N5 TGGC          | 6           |
| CsaII                |        | CCAC N6 CTC          | 8           |
| VbrI                 |        | AGHA N7 TGAC         | 7           |
| VbrII                |        | CTAG N6 RTAA         | 8           |
| CjeFII               |        | CAAY N6ACT           | 9           |
| CjeFIV               |        | TAAY N5 TGC          | 6           |
| BceSVI               |        | TAAG N7 TGG          | 8           |
| EcoGIV               |        | CCAC N8 TGAY         | 9           |
| MpuII                |        | GA N7 TAY            | 7           |
| SauMW2I              |        | CCAY N5 TTAA         | ?           |
| SauMW2II             |        | CCAY N6 TGT          | ?           |
| SauN315I             |        | ATCN5CCT             | 9           |
| SauN315II            |        | CCAY N6 GTA          | 8           |

a) Podjednotka S enzymu NgoA V je zkrácená, což způsobuje symetrickou specifitu

b) V první polosekvenci rozpoznávaného místa Eco585I se nevyskytuje adenin, což naznačuje, že místo něj je methylován cytosin.

**Sloupec 1:** název enzymu (izoschizomery nejsou uvedeny); tučné písmo označuje dobře charakterizované systémy

**Sloupec 2:** rodina, do které enzym patří (je-li známa)

**Sloupec 3:** uveden pouze jeden řetězec rozpoznávané sekvence (orientace 5' - 3'); číslo v rozpoznávané sekvenci indikuje počet nespécifických pbází mezi dvěma polosekvencemi;

báze zvýrazněné tučně jsou methylované.

**Sloupec 4:** počet párů bází mezi dvěma methylovanými bázemi (enzymy ze stejné rodiny mají obdobný počet)

Tab. P3 Změna AMK v proteinové struktuře S domény enzymu *MmeI* mění rozpoznávanou sekvenci (Pingoud, Wilson a Wende 2014)

| Pozice zarovnání AMK | AMK         | Rozpoznávaná sekvence |
|----------------------|-------------|-----------------------|
| E806....R808         | Glu ... Arg | TCCRAC 20/18          |
| K806 ... D808        | Lys ... Asp | TCCRAG 20/18          |

Tab. P4 R–M systémy Typu III (Rao, Dryden a Bheemanaik 2013)

| Enzyme  | Source  | Recognition sequence   | Subunit composition           | Cofactor requirement for restriction | Cofactor requirement for methylation                        |
|---------|---|------------------------|-------------------------------|--------------------------------------|---|
| EcoP15I | Plasmid P15B in <i>E. coli</i> 15T <sup>-</sup> | 5'-CAGC <b>A</b> G-3'  | R <sub>2</sub> M <sub>2</sub> | Mg <sup>2+</sup> , ATP, (SAM?)       | SAM, Mg <sup>2+</sup>                                       |
| EcoP1I  | Prophage P1                                     | 5'-AG <b>A</b> CC-3'   | R <sub>2</sub> M <sub>2</sub> | Mg <sup>2+</sup> , ATP, (SAM?)       | SAM, Mg <sup>2+</sup> , Ca <sup>2+</sup> , Mn <sup>2+</sup> |
| StyLTI  | <i>Salmonella typhimurium</i>                   | 5'-CAG <b>A</b> G-3'   | R <sub>2</sub> M <sub>2</sub> | Mg <sup>2+</sup> , ATP, SAM          | SAM, Mg <sup>2+</sup>                                       |
| HinfIII | <i>Haemophilus influenzae</i>                   | 5'-CG <b>A</b> AT-3'   | R <sub>2</sub> M <sub>2</sub> | Mg <sup>2+</sup> , ATP               | SAM, Mg <sup>2+</sup>                                       |
| PstII   | <i>Providencia stuartii</i>                     | 5'-CA <b>A</b> TCAG-3' | R <sub>2</sub> M <sub>2</sub> | Mg <sup>2+</sup> , ATP, GTP, CTP     | SAM, Mg <sup>2+</sup>                                       |
| LlaFI   | <i>Lactococcus lactis</i>                       | ND                     | ND                            | Mg <sup>2+</sup> , ATP               | ND  |
| BceSI   | <i>Bacillus cereus</i>                          | ND                     | ND                            | Mg <sup>2+</sup> , ATP               | SAM   |
| PhaBI   | <i>Pasteurella haemolytica</i>                  | ND                     | ND                            | ND                                   | ND  |
| NgoAXP  | <i>Neisseria gonorrhoeae</i>                    | 5'-CC <b>A</b> CC-3'   | ND                            | ND                                   | SAM, Mg <sup>2+</sup>                                       |
| HP0593  | <i>Helicobacter pylori</i>                      | 5'-GC <b>A</b> G-3'    | ND                            | ND                                   | SAM, Me <sup>2+</sup>                                       |

ND, not determined.

**A** denotes site of methylation.

Tab. P5 Přehled modifikací vyvolávajících štěpení u enzymů Typu IV (Loenen a Raleigh 2014c)

| Protein         | m5C  | hm5C | ghm5C | m4C  | m6A  | PT |
|-----------------|------|------|-------|------|------|----|
| EcoKMcra        | (+)  | (+)  | (-)   | NT   | (-)  | NT |
| ScoA3Mcra       | +    | NT   | NT    | NT   | (+)  | +  |
| EcoKMcraBC      | +    | +    | -     | (+)  | -    | NT |
| BanUMcraB       | (+)  |      |       |      |      |    |
| BanUMcraB3      | (+)  |      |       |      |      |    |
| EcoKMrr         | (+)  | (-)  | (-)   | (-?) | (+)  | NT |
| BanUMrr         | (+)  |      |       |      | (+?) |    |
| ScoA3Mrr        | (-)  |      |       |      | (+)  |    |
| ZmoMrr          | (+?) |      |       |      | (+?) |    |
| SauUSI          | +    | +    | -     | -    | -?   | NT |
| SauNewI         | (+)  |      |       |      |      |    |
| SepRPMcraR      | (+)  |      |       |      |      |    |
| ScoA3I          | (+)  |      |       |      |      |    |
| PvuRts1I family | +/-  | +    | +     | NT   | -    | NT |
| GmrSD           | -    | -    | +     | NT   | -    | NT |
| ScoA3II+III     | (+)  |      |       |      | (-)  |    |

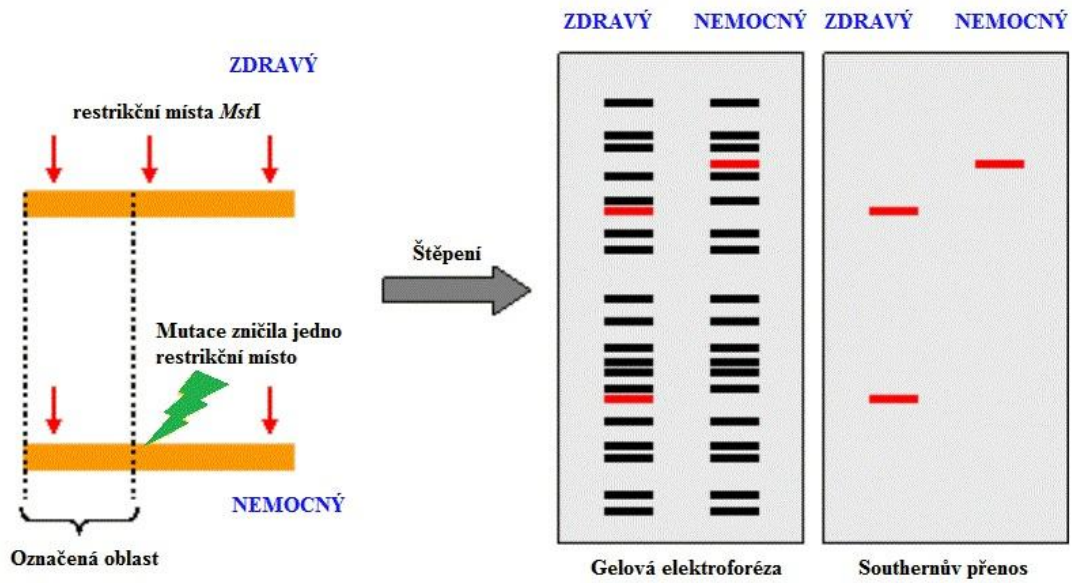
+/-: na substrátech - je alespoň 100 násobně menší aktivita než na substrátech + (-), (+): založeno na *in vivo* restrikci fágové infekce nebo plasmidové transformace s odpovídající konfigurací hostitelského hybridu; *in vitro* nebyly výsledky restrikce zaznamenány

(+?): ve studii nebylo odlišeno, zda je rozpoznávána modifikace m5C či m6A

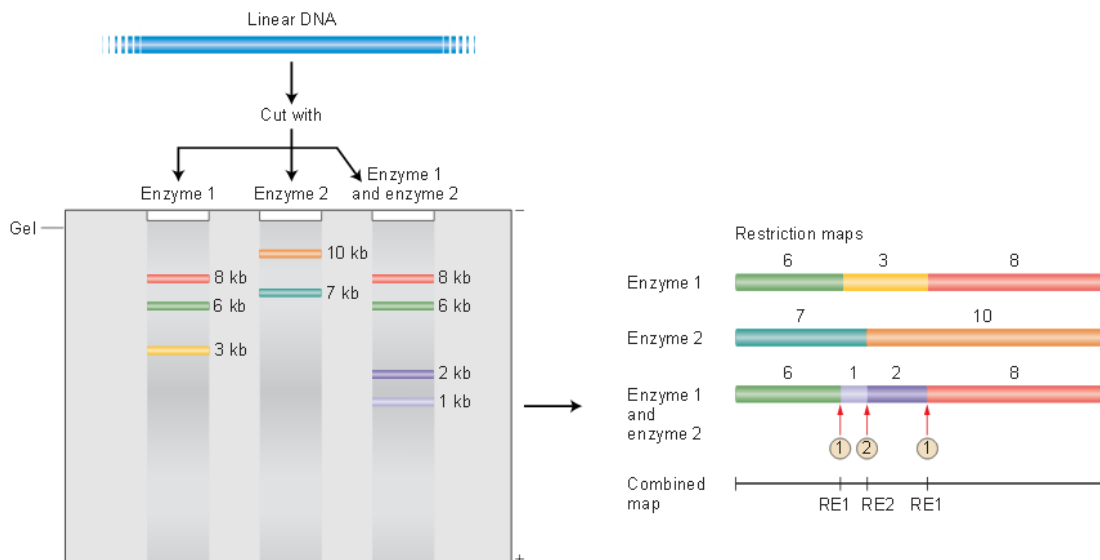
(-?): sekvence s m6A nebyly štěpeny, nicméně bylo testováno pouze malé množství modifikovaných sekvencí

NT: netestováno

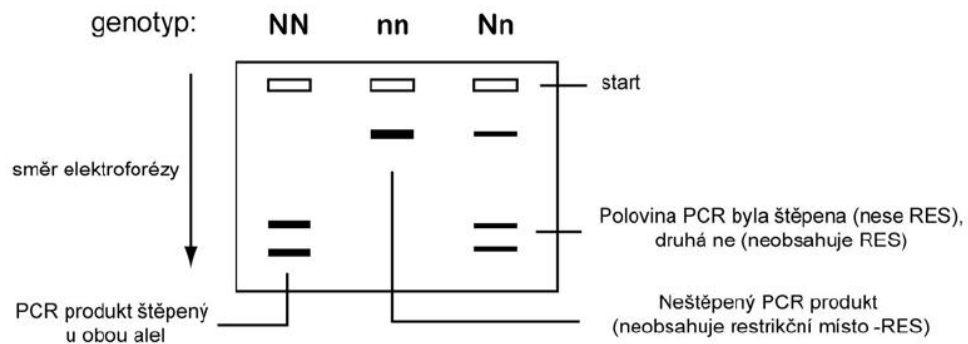
## B. Doplnující obrázky



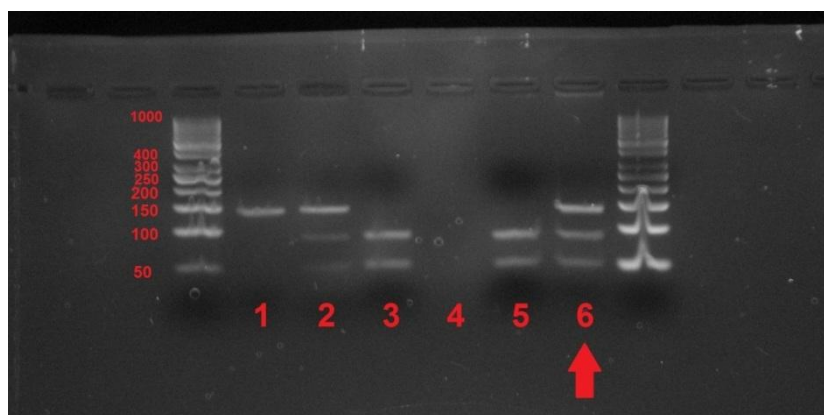
Obr. P1 Princip RFLP, převzato z National Center for Biotechnology Information 2016  
(<http://www.ncbi.nlm.nih.gov/genome/probe/doc/TechRFLP.shtml>)



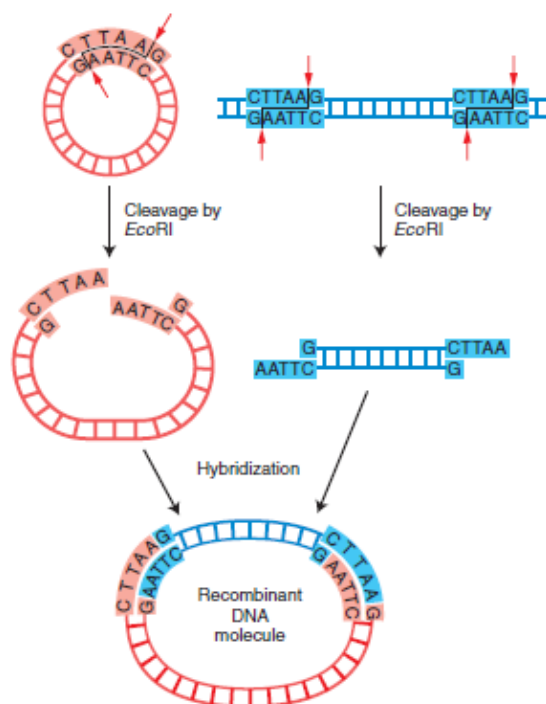
Obr. P2 Tvorba restrikční mapy (Griffiths et al. 2005)



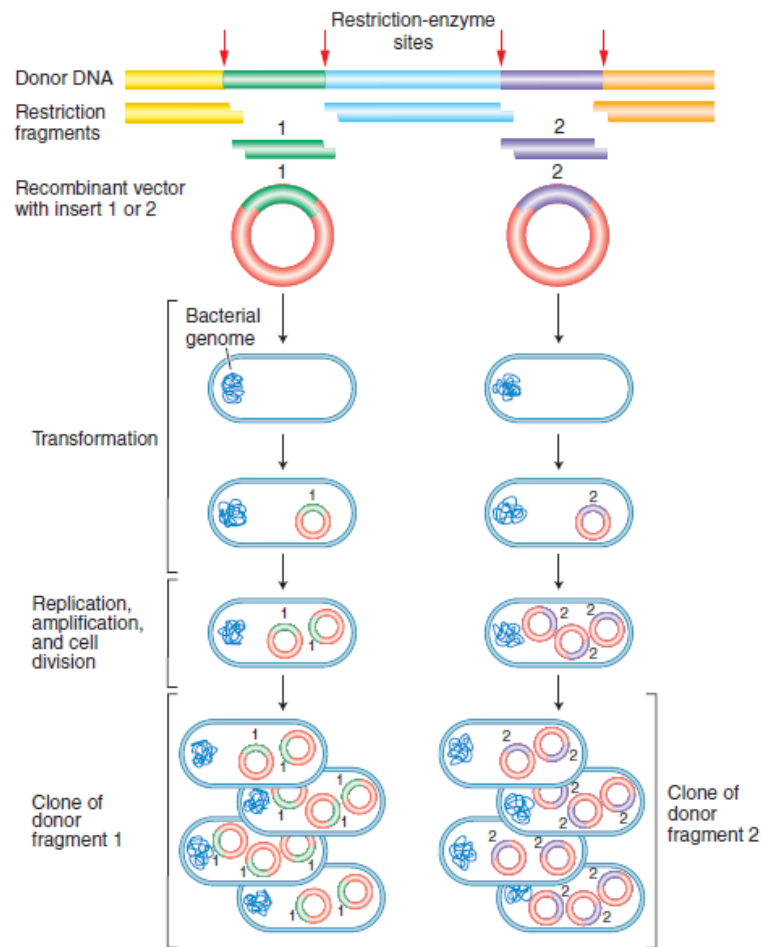
Obr. P3 Modelové vyhodnocení genotypů prasat PCR-RFLP pro *RYR1*  
(převzato z protokolu ze cvičení předmětu Genetika Z, letní semestr 2013/2014)



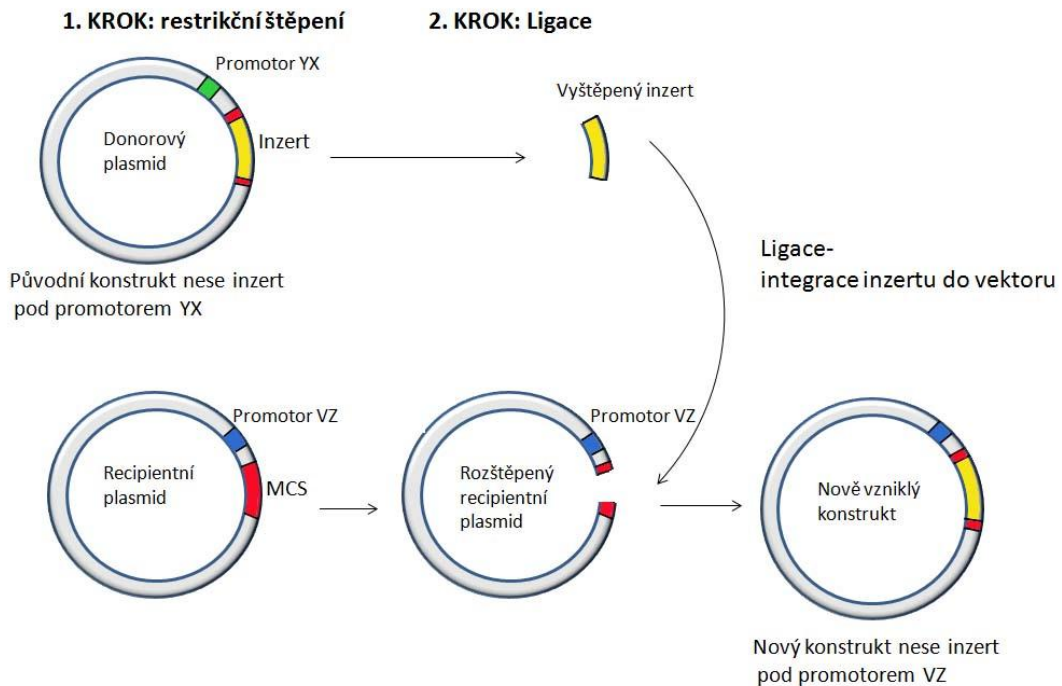
Obr. P4 Snímek elektroforetické separace po PCR-RFLP genu *RYR1* (získáno ve  
cvičení předmětu Genetika Z, letní semestr 2013/2014)



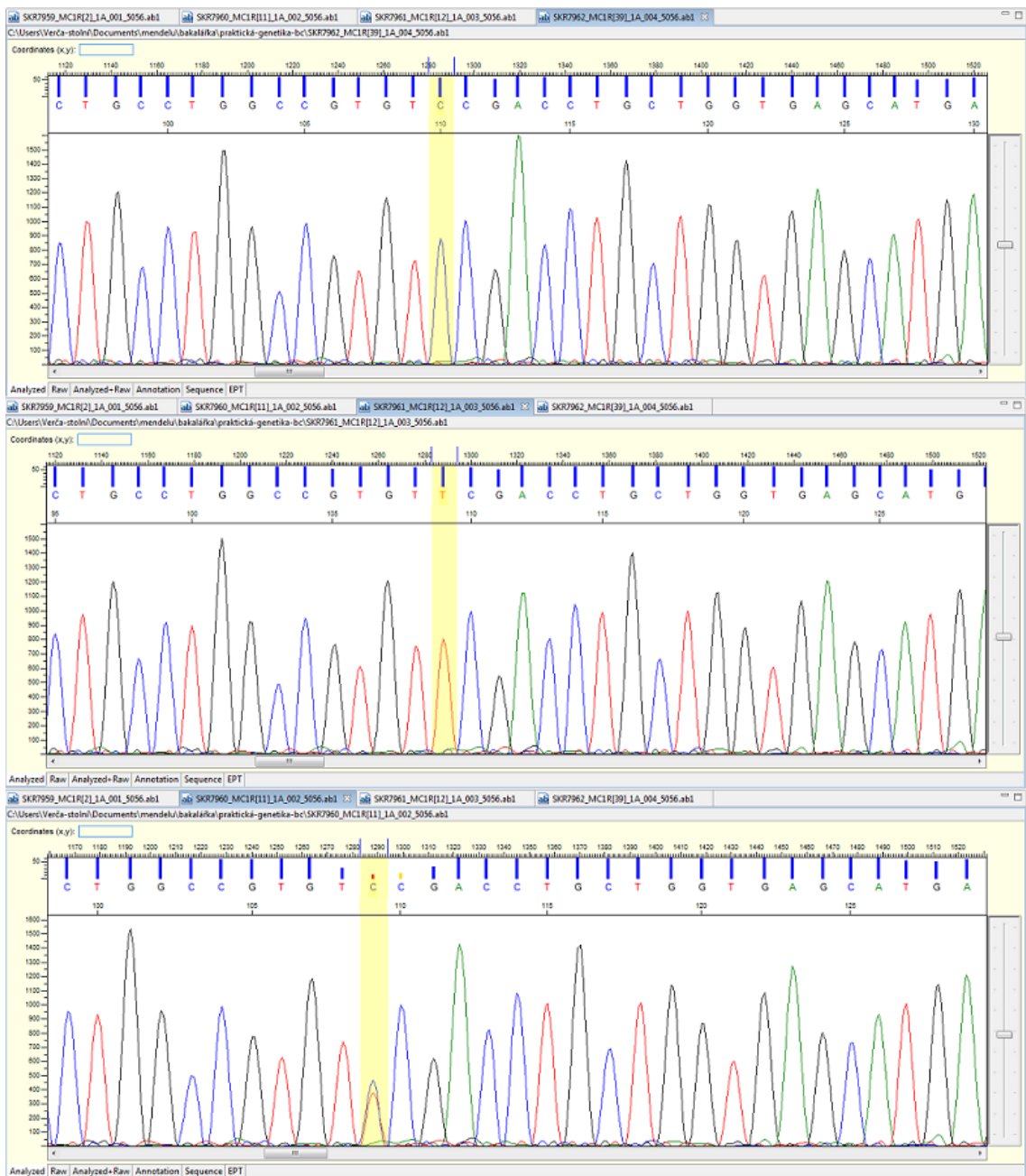
Obr. P5 Tvorba rekombinantní molekuly DNA (Griffiths et al. 2012)



Obr. P6 Amplifikace rekombinantní DNA (Griffiths et al. 2012)

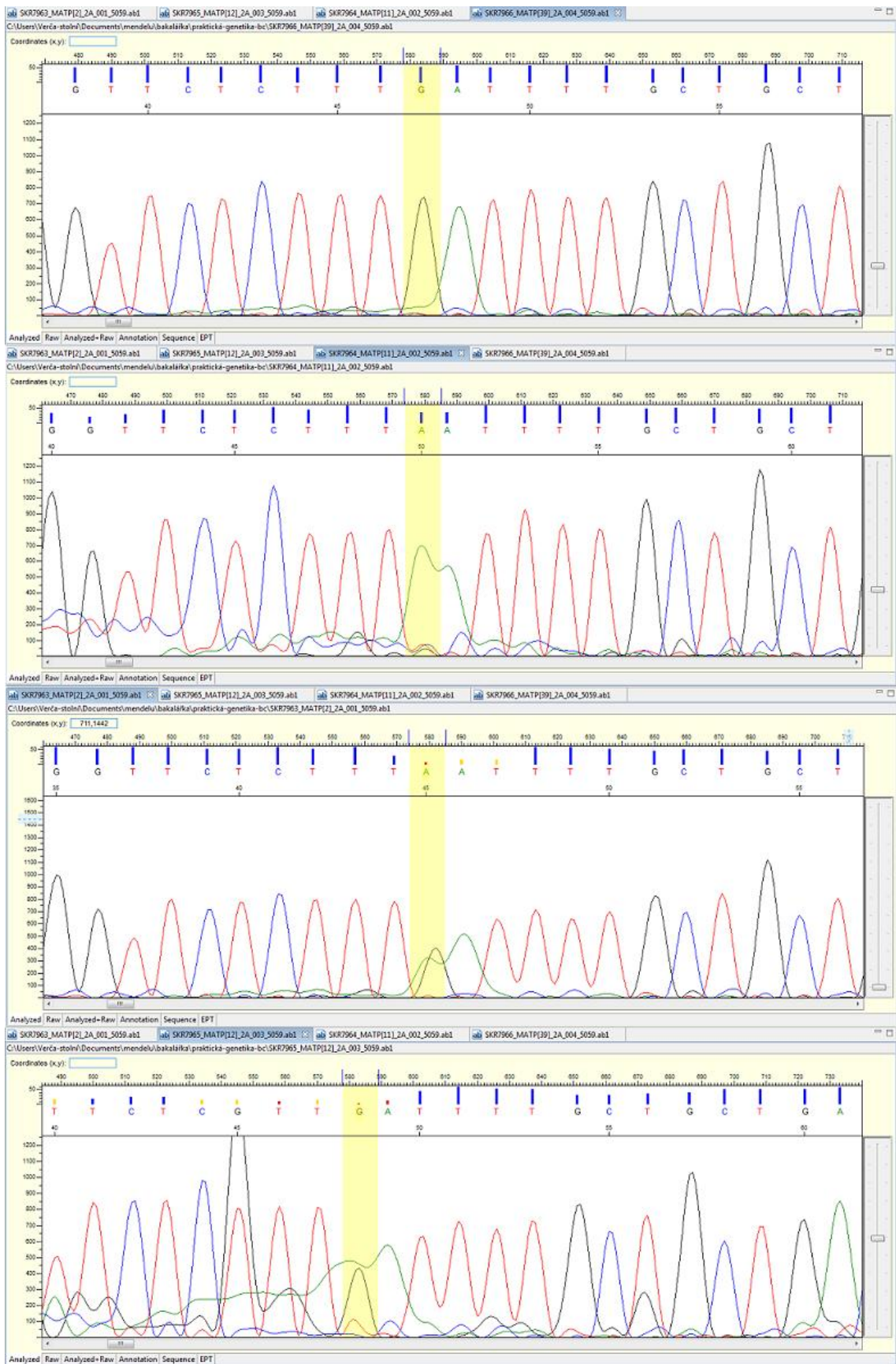


Obr. P7 Subklonování (převzato z <http://labguide.cz/tradicni-klonovani-subklonovani/>)



Obr. P8 Určení polymorfizmu genu *MC1R* v programu Sequence Scanner v1.0





Obr. P9 Určení polymorfizmu genu *MATP* v programu Sequence Scanner v1.0