

**Česká zemědělská univerzita v Praze**

**Fakulta agrobiologie, potravinových a přírodních zdrojů**

**Katedra genetiky a šlechtění**



**Hodnocení genetické variability odrůd máku setého  
(*Papaver somniferum* L.) pomocí molekulárních markerů**

Diplomová práce

Autor práce: Bc. Alžběta Zubíková

Obor studia: Biotechnologie a šlechtění rostlin

Vedoucí práce: Ing. Jakub Vašek, Ph.D.

© 2018 ČZU v Praze

### **Čestné prohlášení**

Prohlašuji, že svou diplomovou práci "Hodnocení genetické variability odrůd máku setého (*Papaver somniferum* L.) pomocí molekulárních markerů" jsem vypracovala samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autorka uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušila autorská práva třetích osob.

V Praze dne 13.4.2018

---

### **Poděkování**

Ráda bych touto cestou poděkovala vedoucímu mé diplomové práce Ing. Jakubu Vaškovi za mnoho cenných rad a věnovaného času.

Dále bych chtěla poděkovat Ing. Daniele Čílové a Mgr. Martině Melounové, Ph.D. za pomoc při práci v laboratoři a také za korekci textu.

# Hodnocení genetické variability odrůd máku setého (*Papaver somniferum* L.) pomocí molekulárních markerů

## Souhrn

Mák setý (*Papaver somniferum* L.) představuje jednu z nejdůležitějších olejnin světového zemědělství. Počátky jeho tradičního použití jsou známy již od 6. tisíciletí př. n.l. na území Středozeří a později jako významný zdroj opia v Evropě a Severní Africe. Dnes je mák znám jako hojně rozšířená plodina pěstovaná pro účely potravinářské, farmaceutické ale i technické. Během procesu domestikace bylo vyšlechtěno široké spektrum odrůd máku se značnou genetickou variabilitou.

Tato genetická variabilita byla zkoumána pomocí nově vytvořených mikrosatelitních markerů u vybraných genotypů. Dle hypotézy, která vychází z předpokladu, že pro registraci nové odrůdy máku je podmínkou její odlišnost od stávajících odrůd. Z tohoto důvodu by se jednotlivé odrůdy máku měly lišit jedna od druhé rovněž na úrovni DNA a tuto variabilitu lze detekovat vhodnými molekulárními markery. Klasické genetické markery představují i mikrosatelity sestávající se z 10 – 100 bází tvořených opakujícím se motivem o délce 1-6 bází.

Na základě bioinformatické analýzy dat bylo zpracováno 980 334 sekvencí o celkové délce 510 849 980 bází. Přičemž byly navrženy vhodné primerové páry pro 1 316 mikrosatelitních lokusů, z nichž bylo pro vlastní testování vybráno 80 potenciálních markerů.

Prostřednictvím agarózové a kapilární elektroforézy bylo provedeno testování na získané skupině odrůd máku spolu s optimalizací PCR reakce. Na konci analýzy bylo vybráno 8 nových kvalitních markerů (EST16, EST19, EST26, EST48c, EST51, EST53c, EST56b, EST61), které byly polymorfní, poskytovaly jednoznačně interpretovaný profil a zároveň poskytovaly reprodukovatelné výsledky.

Pomocí zmíněných nově vytvořených EST-SSR markerů se podařilo úspěšně identifikovat 17 odrůd máku s přesně definovanou velikostí a frekvencí alel na všech 8 lokusech.

**Klíčová slova:** *Papaver somniferum* L., mikrosatelity, EST-SSR, SNP, genetické profilování

# Genetic characterisation of poppy cultivars (*Papaver somniferum* L.) by molecular markers

## Summary

Poppy (*Papaver somniferum* L.) is one of the most important oilseed of the world agriculture. The beginning of its traditional use have been known since the 6th millennium BC. In the Mediterranean and later as a major source of opium in Europe and North Africa. Today, poppy is known as a very widespread crop grown for food, pharmaceutical and technical, purpose. During the domestication process, a wide variety of varieties with a high genetic variability were developed.

I investigated this genetic variability of poppy seed with newly created microsatellite markers in selected genotypes. According to a hypothesis based on the assumption that the registration of a new variety of poppy is a condition of its distinction from existing varieties. For this reason, the different poppy varieties should differ from one another at the DNA level and this variability can be detected by appropriate molecular markers..

Based on the bioinformatics data analysis, 980 334 sequences with a total length of 510 849 980 bases were processed. Appropriate primer pairs were designed for 1316 microsatellite loci, of which 80 potential markers were selected for self-testing.

An agarose and capillary electrophoresis assay was performed on the poppy varieties obtained, along with optimization of the PCR reaction. At the end of the analysis eight new high-quality markers (EST16, EST19, EST26, EST48c, EST51, EST53c, EST56b, EST61) were selected that were polymorphic, providing a clearly interpreted profile while providing reproducible results.

Using the newly created EST-SSR markers, 17 varieties of poppy have been successfully identified with precisely defined size and frequency alleles in all 8 loci.

**Keywords:** *Papaver somniferum* L., microsatellite, EST-SSR, SNP, profiling

# Obsah

<b>1</b>	<b>Úvod</b>	<b>1</b>
<b>2</b>	<b>Cíl práce</b>	<b>2</b>
2.1	Vědecká hypotéza	2
2.2	Cíle práce	2
<b>3</b>	<b>Literární rešerše</b>	<b>3</b>
<b>3.1</b>	<b>Taxonomické zařazení máku setého</b>	<b>3</b>
3.1.1	Biologie rodu <i>Papaver</i>	3
3.1.2	Zástupci a využití	4
<b>3.2</b>	<b>Mikrosatelity</b>	<b>7</b>
3.2.1	Klasifikace mikrosatelitů	7
3.2.2	Distribuce mikrosatelitů v genomu	8
3.2.3	Mutace u mikrosatelitů	9
<b>3.3</b>	<b>Genetické markery</b>	<b>11</b>
<b>3.4</b>	<b>Elektroforéza</b>	<b>16</b>
3.4.1	Gelová elektroforéza	17
3.4.2	Kapilární elektroforéza	18
<b>3.5</b>	<b>Problematika kvality dat při genotypizaci</b>	<b>20</b>
3.5.1	Nulové alely	22
3.5.2	Homoplazie	22
3.5.3	Přidání netemplátového nukleotidu	23
<b>4</b>	<b>Metodika</b>	<b>24</b>
<b>4.1</b>	<b>Rostlinný materiál</b>	<b>24</b>
<b>4.2</b>	<b>Izolace DNA</b>	<b>25</b>
<b>4.3</b>	<b>Bioinformatická analýza dat</b>	<b>26</b>
4.3.1	Publikované mikrosatelitní markery	26
4.3.2	Nově navržené mikrosatelitní markery	26
<b>4.4</b>	<b>Polymerázová řetězová reakce (PCR)</b>	<b>27</b>
4.4.1	Koncepce testování a optimalizace nových primerů	27
<b>4.5</b>	<b>Elektroforéza</b>	<b>28</b>
4.5.1	Agarózová elektroforéza	28
4.5.2	Kapilární gelová elektroforéza	28
4.5.2.1	Příprava vzorků	28
4.5.2.2	Nastavení podmínek separace	29
4.5.3	Sekvence	29
4.5.3.1	Purifikace před sekvenační reakcí	29
4.5.3.2	Sekvenační reakce	30
4.5.3.3	Purifikace po sekvenační reakci	30

4.5.3.4	Sekvenace.....	30
4.5.4	Statistická analýza a zpracování dat.....	31
<b>5</b>	<b>Výsledky.....</b>	<b>32</b>
5.1	Bioinformatická analýza dat .....	32
5.2	Optimalizace a testování navržených markerů pomocí agarózové elektroforézy.....	32
5.3	Testování markerů pomocí kapilární elektroforézy .....	33
5.3.1	Tvorba alelických žebříků.....	35
5.4	Vyhodnocení variability získaných markerů.....	37
5.4.1	Deskriptivní statistika .....	37
5.4.2	Vícerozměrná analýza pro identifikaci jednotlivých odrůd .....	38
<b>6</b>	<b>Diskuze.....</b>	<b>40</b>
6.1	Bioinformatická analýza.....	40
6.2	Optimalizace a testování navržených primerů.....	41
<b>7</b>	<b>Závěr .....</b>	<b>43</b>
<b>8</b>	<b>Literatura.....</b>	<b>44</b>
<b>9</b>	<b>Seznam použitých zkratk .....</b>	<b>51</b>
<b>10</b>	<b>Přílohy.....</b>	<b>52</b>

## Seznam obrázků

Obrázek 1 Mák setý ( <i>Papaver somniferum</i> L.)	5
Obrázek 2 Mák vlní ( <i>Papaver rhoeas</i> )	6
Obrázek 3 Schématické rozdělení tandemově se opakujících sekvencí v DNA	7
Obrázek 4 Mikrosatelity v genomu a jejich regulační funkce	9
Obrázek 5 Schéma replikačního sklouznutí	10
Obrázek 6 Pohled na změny využívání jednotlivých markerů na konci 20. století	11
Obrázek 7 Detekce amplifikovaného produktu o velikosti 264 bp	12
Obrázek 8 Snímek z autoradiogramu	12
Obrázek 9 Variabilita amplifikace a následná elektroforetická vizualizace u metody RAPD	13
Obrázek 10 Navrhnuté primery lemující oblast mikrosatelitu	14
Obrázek 11 Detekce SSR polymorfismu	14
Obrázek 12 Polymerní struktura agarózového gelu	16
Obrázek 13 Schéma průběhu gelové elektroforézy (upraveno dle Yilmaz et al., 2012)	17
Obrázek 14 Schéma kapilární elektroforézy	18
Obrázek 15 Schéma separace a detekce HPCE	18
Obrázek 16 (a) Snímek z agarózové elektroforézy s odlišnými bandy, (b) příklad velikostní homoplazie	23
Obrázek 17 Thermal Cycler C 1000	27
Obrázek 18 Cela SubCell GT pro agarózovou elektroforézu	28
Obrázek 19 Genetický analyzátor ABI PRISM 310	29
Obrázek 20 Četnost sekvence komplementárních motivů	32
Obrázek 21 Agarózový gel testování EST-SSR markerů na 8 odrůdách	33
Obrázek 22 Ukázka kvalitního profilu v případě markeru EST48c	34
Obrázek 23 Alelický žebřík 6 alel mikrosatelitního markeru EST48	36
Obrázek 24 Dendrogram příbuznosti 17 odrůd	39
Obrázek 25 Shluková analýza 17 odrůd	39

## Seznam tabulek

Tabulka 1 Rozdělení rodu <i>Papaver</i> do sekcí s vybranými zástupci	3
Tabulka 2 Koncentrace agarózového gelu pro separaci nukleových kyselin	17
Tabulka 3 Příčinné zdroje chyb	21
Tabulka 4 Seznam odrůd	24
Tabulka 5 Publikované primery	26
Tabulka 6 Program PCR reakce při testování navržených primerů	27
Tabulka 7 Teplotní profil sekvenační reakce	30
Tabulka 8 Nově navržené primery pro kvalitní fragmentační analýzu v kapilární elektroforéze	34
Tabulka 9 Testování 22 EST-SSR markerů na kapilární elektroforéze	35
Tabulka 10 Velikosti a frekvence alel EST-SSR markerů na 31 genotypch	37
Tabulka 11 Velikosti a frekvence alel EST-SSR markerů na 17 genotypch	38
Tabulka 12 Navržené primery EST01-40	52
Tabulka 13 Navržené primery EST01-40	53



# 1 Úvod

Mák setý (*Papaver somniferum* L.) je jednou z nejznámějších rostlin rodu *Papaver*. Jedná se o druhou nejdůležitější olejninu českého zemědělství a Česká republika představuje nejdůležitějšího světového pěstitele a exportéra konzumního máku ve světě. Mák setý ale nachází využití i ve farmaceutickém průmyslu hlavně díky obsahu velkého množství různých alkaloidních látek s širokým spektrem účinku a také díky latexu z nezralých makovic, které představuje významný zdroj opia.

Přesto, že se jedná o rostlinu tradičně využívanou od dob před našim letopočtem, z hlediska genetiky nebyl doposud osekvenován celý její genom a většina studií se zaměřuje na tvorbu sekundárních metabolitů. Známo je pouze 138 lokusů a 21 094 „syrových“ exprimovaných sekvenčních úseků neboli EST (NCBI, 2018). Takováto data lze získat analýzou DNA pomocí molekulárních markerů. Jedním z typů markerů využívaných pro tyto a další účely, jako jsou například studie genomu, identifikace a rozlišení odrůd rostlin, jsou mikrosatelity. Mikrosatelity disponují vysokou četností a variabilitou v genomu (Buschiazzo and Gemmell, 2006), ale také vykazují vysokou specifitu a reprodukovatelnost, díky nimž se výsledky dají poměrně dobře porovnat mezi laboratořemi (Powel *et al.*, 1996).

## **2 Cíl práce**

### **2.1 Vědecká hypotéza**

Studium genetické variability kultivarů máku setého (*Papaver somniferum* L.) vychází z předpokladu, že pro registraci nové odrůdy máku je podmínkou její odlišnost od stávajících odrůd. Z tohoto důvodu by se jednotlivé odrůdy máku měly lišit jedna od druhé rovněž na úrovni DNA a tuto variabilitu lze detekovat vhodnými molekulárními markery.

### **2.2 Cíle práce**

Z uvedené hypotézy vycházejí následující cíle práce, které jsou uvedené v následujících bodech:

- a. Provést bioinformatickou analýzu transkriptomických dat za účelem navržení nových EST-SSR markerů
- b. Provést testování EST-SSR markerů na získané skupině odrůd a případně optimalizovat PCR
- c. Vybrané markery analyzovat a vyhodnotit pomocí kapilární elektroforézy
- d. Otestovat rozlišovací schopnost finálních markerů pro identifikaci testovaných odrůd

## 3 Literární rešerše

### 3.1 Taxonomické zařazení máku setého

Jedním z nejznámějších máků je Mák setý (*Papaver somniferum* L.). Patří do čeledi Mákovitých (*Papaveraceae*) a do rodu Mák (*Papaver*), který zahrnuje okolo 110 dalších druhů. Máky můžeme nalézt ve Střední a Jihozápadní Asii, Střední a Jižní Evropě, Severní Africe a nejhojněji v Turecku, jež je místem původu mnoha druhů (Carolan *et al.*, 2006; Parmaksiz and Ozcan, 2011). Na základě morfologických a histologických znaků, jako je například barva květů, velikosti tyčinek nebo obsahu alkaloidů klasifikuje mnoho autorů máky do sekcí (Tabulka 1) (Tétényi, 1997; Carolan *et al.*, 2006).

Tabulka 1 Rozdělení rodu *Papaver* do sekcí s vybranými zástupci

Sekce	Zástupci
<i>Californicum</i>	<i>P. californium</i>
<i>Meconidium</i>	<i>P. ameniacum</i>
<i>Meconella</i>	<i>P. croceum</i> , <i>P. anomalum</i> , <i>P. nudicaule</i> , <i>P. radiocatum</i> , <i>P. alpinum</i> , <i>P. myobeanum</i>
<i>Argemonidium</i>	<i>P. apulum</i> , <i>P. argemone</i> , <i>P. pavonium</i> , <i>P. hybridum</i>
<i>Pilosa</i>	<i>P. polisum</i>
<i>Pseudopilosa</i>	<i>P. atlanticum</i> , <i>P. rupifragum</i>
<i>Horrida</i>	<i>P. aculeatum</i>
<i>Carinatee</i>	<i>P. macrostomum</i>
<i>Rhoeadium</i>	<i>P. dubium</i> , <i>P. rhoeas</i> , <i>P. commutantum</i>
<i>Macrantha /Oxytona Bernh</i>	<i>P. orientale</i> , <i>P. pseudorientale</i> , <i>P. bracteatum</i>
<i>Papaver</i>	<i>P. somniferum</i> , <i>P. glaucum</i> , <i>P. gracile</i> , <i>P. decaisnei</i> , <i>P. setigerum</i>

(Botany.cz, 2018)

#### 3.1.1 Biologie rodu *Papaver*

Mezi máky patří jednoleté, dvouleté i vytrvalé byliny převážně cizosprašného typu. Můžeme se ale setkat i se samosprašnými druhy, jako je například mák setý (*Papaver somniferum* L.) (Tétényi, 1997). Habitem se jedná převážně o byliny, výjimečně pak keře. Nejvýraznějším prvkem pro odlišení jednotlivých druhů máku je zbarvení květních obalů. Paleta barevných odstínů se pohybuje od bílé přes oranžovou, rudou až do tmavě fialové. Odlišit

od sebe jednotlivé druhy máku na základě barvy květu není tak těžké, odlišit od sebe ale jednotlivé druhy dle semen je vizuálně velice náročné, proto se k jejich identifikaci a dalším analýzám využívá například metody AFLP (Amplified Fragment Length Polymorphism) nebo mikrosatelitních markerů (Celik *et al.*, 2014). Genetickými analýzami lze máky rozlišovat i podle ploidie, tedy dle počtu homologních sad chromozomů. Je známo, že máky obsahují diploidní (2x) až oktaploidní (8x) druhy, se základním počtem chromozomů 11, 7 (Parmaksiz and Ozcan, 2011; Verma *et al.*, 2016) či 6 (Bara *et al.*, 2007).

### 3.1.2 Zástupci a využití

Máky jsou významné hlavně pro potravinářský a farmaceutický průmysl, díky vysokému obsahu alkaloidních látek a olejů, které se nachází v semenech rostlin, přičemž u některých druhů tvoří až 50 % semene (Sangwan *et al.*, 2000). Nejvýznamnějším zástupcem rodu *Papaver* je z hlediska obsahu alkaloidních látek mák setý (*Papaver somniferum* L.), který produkuje více než 80 různých alkaloidů. Morfin s analgetickými a narkotickými účinky, kodein s antitusickými účinky, papaverin s vazodilatačním a relaxačním účinkem na hladkou svalovinu, antitusický a protirakovinný noscapin a výchozí materiály pro polosyntetické opiáty – oxykodon, oxymorfon, eterfin, buprenorfin a nejdůležitějším tebainem (Celik *et al.*, 2014; Verma *et al.*, 2016), který slouží jako prekurzor pro kodein. Dalším druhem máků, který obsahuje vysoké množství tebainu (1-4 %) je mák listenatý (*Papaver bracteatum*) (Espinasse and Dosba, 1982). Velice známým je v našich podmínkách také mák vlčí (*Papaver rhoeas*), který tvoří součást zemědělských ekosystémů především jako plevel a hostitelská rostlina pro fytopatogenní organismy, ale zároveň představuje cenný zdroj pylu. Jedná se o druh s vysokou genetickou variabilitou, která se fenotypově projevuje rozmanitým zabarvením okvětních lístků (Kati *et al.*, 2013).



Obrázek 1 Mák setý (*Papaver somniferum* L.)

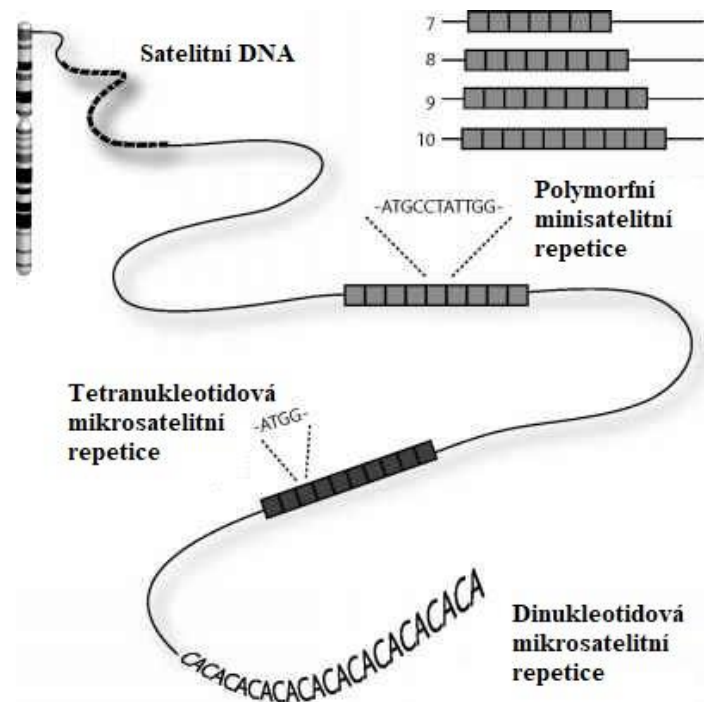




Obrázek 2 Mák vlčí (*Papaver rhoeas*)

## 3.2 Mikrosatelity

V DNA různých organismů, od kvasinek po obratlovce, se nachází tandemově se opakující sekvence nukleotidů (Obrázek 3) (Ellegren *et al.*, 2014; Kalia *et al.* 2011). Dle počtu nukleotidů se dělí na **satelity** (5 a 100< nukleotidů ), **minisatelity** (10-100 nukleotidů) a **mikrosatelity** (1-5 nukleotidů) (Bennett, 2000).



Obrázek 3 Schématické rozdělení tandemově se opakujících sekvencí v DNA

(*Genetic diversity*, 2018)

Mikrosatelity, rovněž známé jako SSR (simple sequence repeats) nebo STR (short tandem repeats) (Ellegren, 2004; Kalia *et al.*, 2011), jsou hojně rozšířené po celém genomu a díky vysoké míře mutací vykazují vysoký polymorfismus. Jejich analýza je poměrně jednoduchá (Galindo *et al.*, 2009), a proto představují významný zdroj rozmanitých markerů v řadě oblastí genetiky (Metzgar *et al.*, 2009; Guichoux *et al.*, 2011). Jejich využití je více rozvedeno v kapitole [3.3](#) Genetické markery.

### 3.2.1 Klasifikace mikrosatelitů

Mikrosatelity se dělí podle své délky motivu, typu repetice a umístění v genomu (Kalia *et al.*, 2011). Dle počtu nukleotidů jsou nejčastějšími typy mikrosatelitů mono- (př. AAA), di- (př. CACA), tri- (př. ATGATGATG) a tetranukleotidové repetice, ale můžeme se setkat i s penta – či hexanukleotidovými sekvencemi (Ellegren *et al.*, 2004).

Z hlediska uspořádání motivu se mikrosatelity dělí do následujících kategorií:

- a. **Dokonalé (perfect):** motiv dokonalého typu mikrosatelitu není přerušen žádnou bází (př. TATATATATATATATA)
- b. **Nedokonalé (imperfect):** u těchto mikrosatelitů se objevuje přerušení opakujícího se motivu jednou bází (př. TATATATACTATATATA)
- c. **Přerušené (interrupted):** v případě přerušeného motivu dochází k přerušení motivu jinou sekvencí bází (př. TATATAGGGTATATATA)
- d. **Složené (composite):** u tohoto typu dochází k sestavení dvou a více přiléhajících mikrosatelitů s odlišnou repeticí (př. TGTGTGTGCACACA) (Kalia *et al.*, 2011; Oliveira *et al.*, 2006)

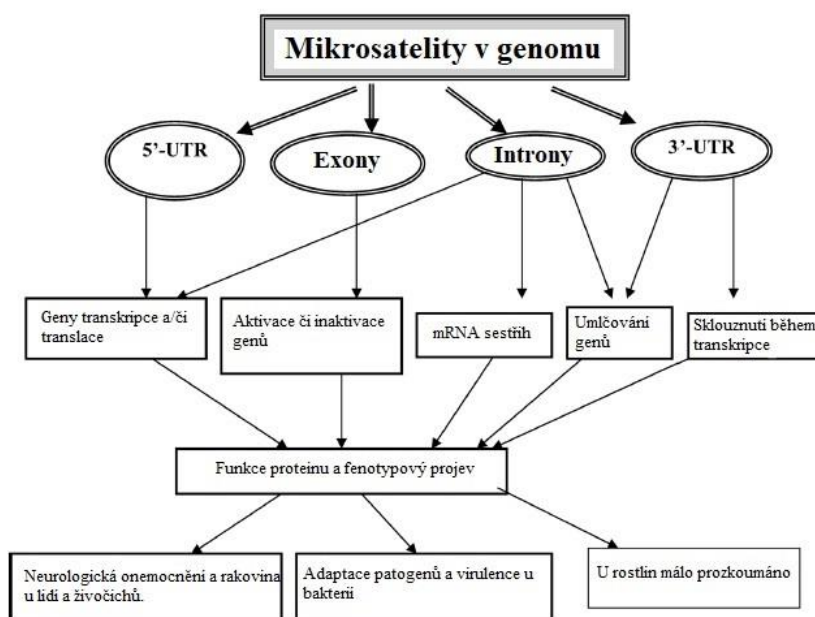
### 3.2.2 Distribuce mikrosatelitů v genomu

Přítomnost mikrosatelitů byla prokázána v různých organismech od kvasinek, rostlin až po obratlovce. Dle místa výskytu v buňkách se dělí na jaderné (nuSSRs), mitochondriální (mtSSRs) a chloroplastové (cpSSRs). Dle motivu je u rostlin nejčastějším mikrosatelitním motivem AT dinukleotid (Kalia *et al.*, 2011; Morgante *et al.*, 2002). Z trinuklotidových sekvencí je nejčastější sekvence CCG (z 35 %) a druhá nejčastější (z 33,8 %) sekvence AGC. Tato sekvence se ve vysoké míře objevuje i u hlodavců (z 57,1 %), u savců (42,9 %) a až z 60 % u hmyzu (Li *et al.*, 2004).

V jakých částech genomu a jakou regulační funkci mají je, zobrazeno v Obrázku 4. Nejčastěji se mikrosatelity vyskytují v nekódujících oblastech DNA (telomery, subtelomery, heterochromatin u centromer) (Metzgar *et al.*, 2000; Oliveira *et al.*, 2006). Nicméně velké množství mikrosatelitů se rovněž nachází v tzv. nepřekládaných oblastech (untranslated regions- 5'-UTR, 3'-UTR) a i v genech a intronech. V intronech se nachází hlavně mono – a dinukleotidové repetic, kdežto 5'-UTR oblasti obsahují více tripletových nukleotidových repetic a v 3'-UTR oblastech jsou spíše mononukleotidy (Kalia *et al.*, 2011).

Mikrosatelity v intronech, 5'-UTR a 3'-UTR oblastech mohou mít vliv na expresi genů. Prodlužováním a zkracováním mikrosatelitů v těchto oblastech, může zapříčinit posun čtecího rámce, což může způsobit ztrátu či změnu funkčnosti některých genů, fenotypové změny a vznik různých genetických onemocnění (Li *et al.*, 2004).





Obrázek 4 Mikrosatelity v genomu a jejich regulativní funkce

(upraveno dle Li *et al.*, 2004)

U člověka bylo zjištěno více než 40 různých neurologických, neurodegenerativních a neuromuskulárních poruch, vzniklých v důsledku expanze SSR. U rostlin byl význam mikrosatelitů potvrzen zejména v 5'-UTR oblastech. Například byl prokázán vztah tripletových SSR, konkrétněji (CCG)<sub>n</sub> v rRNA, na fertilitu u kukuřice nebo na obsah amylozy u rýže v korelaci s množstvím GA a CT repetice v 5'-UTR oblastech (Kalia *et al.*, 2011).

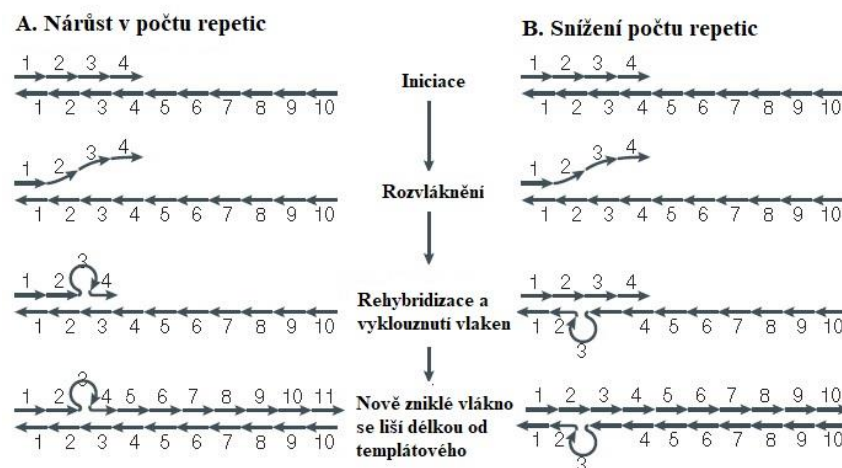
Mnoho studií prokázalo přítomnost mikrosatelitů v malé míře i v kódujících a regulačních oblastech, kde mohou způsobovat aktivaci/inaktivaci genů a mít tak vliv na vzniklý protein a fenotypový projev. V těchto oblastech – exonech – se nachází převážně trinukleotidové a hexanukleotidové mikrosatelity (Li *et al.*, 2004).

### 3.2.3 Mutace u mikrosatelitů

Je známo, že četnost mutací je u mikrosatelitů vyšší než v jiných částech genomu, a to v rozmezí  $10^{-2}$  -  $10^{-6}$  v lokusu na generaci u eukaryot (Oliviera *et al.*, 2006). Na variabilitu mikrosatelitů a jejich mutace má nesporně vliv mnoho faktorů – mutační mechanismy, struktura mikrosatelitů, selektivní vlivy, genomické a biologické pozadí (Buschiazzo and Gemmel, 2006).

Do mutačních mechanismů se řadí například chyba při rekombinaci, nerovnoměrný crossing over či sklouznutí DNA polymerázy během replikace DNA (Li *et al.*, 2004; Buschiazzo and Gemmel, 2006; Oliviera *et al.*, 2006; Kalia *et al.*, 2011).

Během replikace DNA může docházet k omezení pohybu DNA polymerázy. Replikace není plynulá a DNA polymeráza nesyntetizuje některé triplety ve více kopiích, které vyklouznou a vytvoří tzv. vlásenku, anebo může docházet k opačnému jevu. Kdy jsou vlásenky pro činnost DNA polymerázy těžko dostupné, a proto nedochází k jejich replikaci a vzniklý řetězec je o tyto úseky kratší (Ellegren *et al.*, 2004). Sklouznutí DNA polymerázy zapříčiní tedy prodloužení, či zkrácení sekvence (Obrázek 5). Naopak bodové mutace mají za následek přerušení delších opakujících se sekvencí. Oba tyto procesy ale mají vliv na množství a rozptýlení mikrosatelitů v genomu (Li *et al.*, 2004).



Obrázek 5 Schéma replikačního sklouznutí

(upraveno dle Ellegren *et al.*, 2004)

Proces replikace DNA doprovází řada opravných mechanismů, které se snaží minimalizovat vznik různých mutací. Jedním z nich je tzv. mismatch repair (MMR), který provádí korekci správného párování bází (Kalia *et al.*, 2004).

Z biologického hlediska může být variabilita mikrosatelitů ovlivněna pohlavím, věkem organismu, mutacemi v genech MMR systému a vnějším prostředím (Buschiazzo and Gemmel, 2006). Ke změnám v mikrosatelitech může tedy docházet i následkem vnějších podnětů jako je například UV radiace, gama záření, oxidativní stres apod. (Kalia *et al.*, 2011).

Na mutabilitu mikrosatelitů má vliv i umístění v genomu, zda se nachází v kódujících či nekódujících oblastech, na přítomnosti sousedních mikrosatelitů a nemalý vliv mají také sekvence, které mikrosatelity ohraničují (Buschiazzo and Gemmel, 2006).

Strukturní vlastnosti samotných mikrosatelitů, jako je délka motivu, délka alely, složení jednotlivých nukleotidů v motivu mikrosatelitů a jejich vnitřní architektura (zda je dokonalá/nedokonalá/přerušená či složená) má vliv na tendenci k mutacím, a tudíž zvýšení variability mikrosatelitů v genomu (Galindo *et al.*, 2009).

### 3.3 Genetické markery

Genetický marker, někdy také DNA marker, je snadno identifikovatelná sekvence DNA se známou polohou na chromozomu. Sleduje se na předem daném místě, kde označuje (markeruje) určitou část genu a lze jej detekovat pomocí analýzy DNA (Schlötterer, 2004; Lateef, 2015; Jiang, 2018).

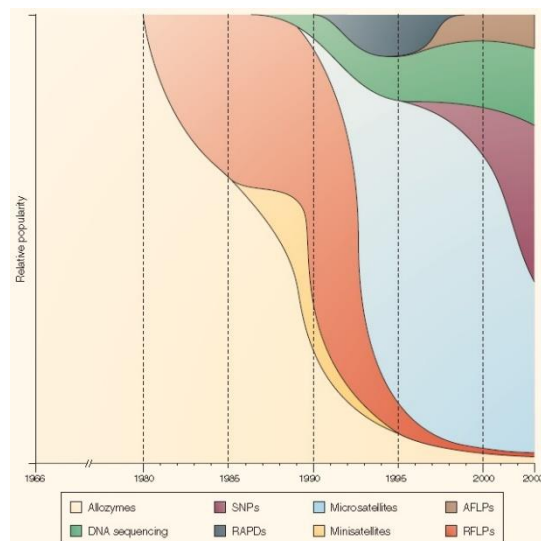
Dle použité metody Garrido-Cardenass *et al.* (2018) klasifikuje analýzu DNA do následujících kategorií:

1. Techniky bez PCR reakce, založené na hybridizaci – RFLP
2. Techniky založené na PCR reakci – RAPD, AFLP, SSR
3. Techniky založené na sekvenování - SNP

Podmínkou těchto metod je spolehlivost, jednoduchost a rychlost. Stejně tak by metody neměly být moc nákladné a k jejich užití by mělo stačit jen malé množství vzorku (Garrido-Cardenass *et al.*, 2018).

#### RFLP – Restriction Fragment Length Polymorphism

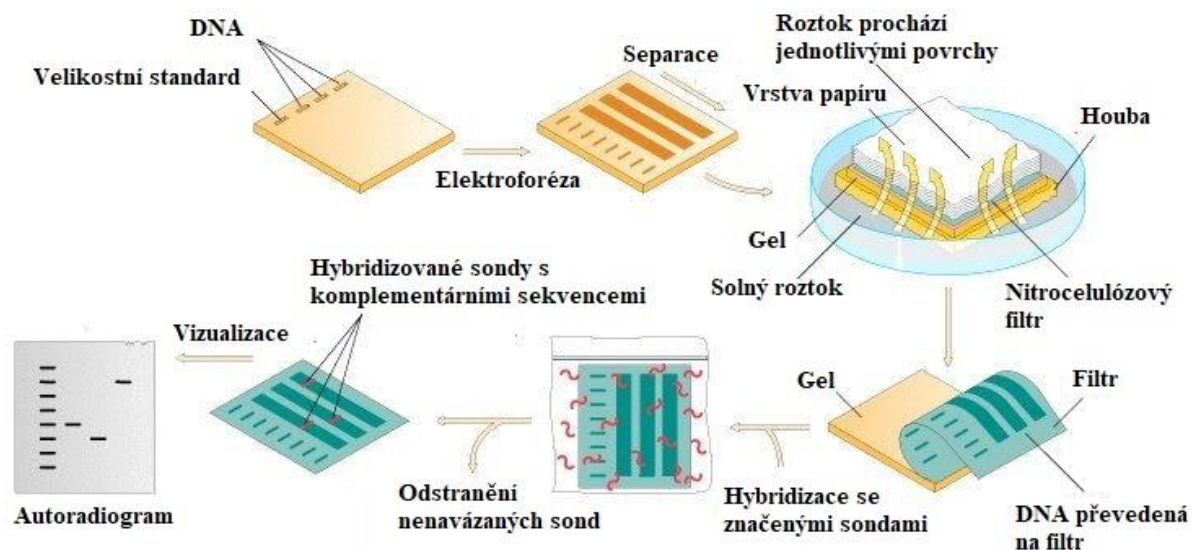
RFLP byla hojně využívána metoda v 80 a 90. letech minulého století (Obrázek 6) a její markery jsou známy jako molekulární markery první generace. Pomocí RFLP byly vytvořeny genetické mapy mnoha druhů kulturních plodin jako je například pšenice, rýže a kukuřice (Lateef, 2015).



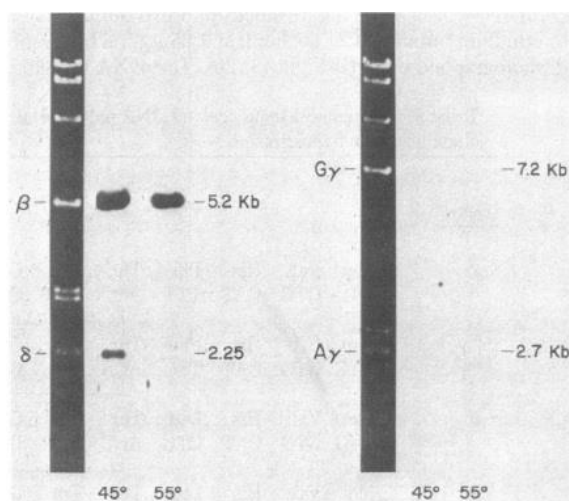
Obrázek 6 Pohled na změny využívání jednotlivých markerů na konci 20. století

(Schlötterer, 2004)

RFPL je založeno na štěpení molekuly DNA pomocí restričních endonukleáz, které mají schopnost rozeznat a rozštěpit specifické sekvence. Restrikcí se získají rozdílné fragmentů DNA od různých jedinců, které se následně separují pomocí gelové elektroforézy. Detekce se provádí pomocí hybridizace DNA sond, následně Southern blottem (Obrázek 7) (Garrido-Cardenas, 2018) a vizualizací v podobě autoradiogramu (Obrázek 8).



Obrázek 7 Detekce amplifikovaného produktu o velikosti 264 bp (Griffiths, 2000)



Obrázek 8 Snímek z autoradiogramu (Conner et al., 1983)

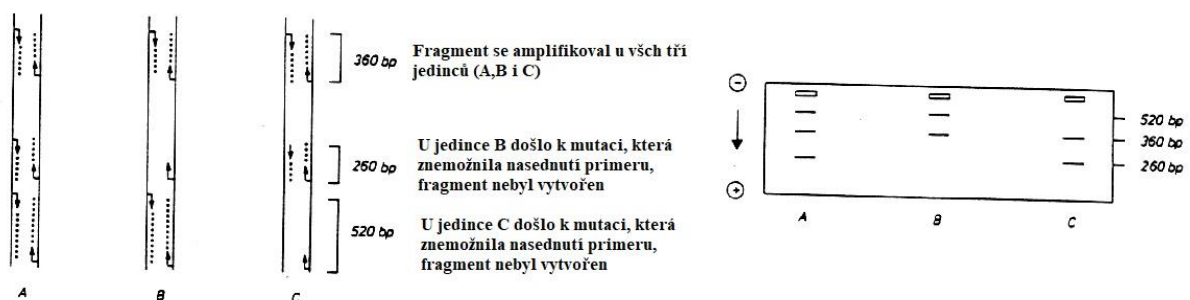
Výhodou RFLP je, že nevyžaduje znalosti o struktuře genomu, dále je předností kodominance, vysoká lokusová specifita a dobrá reprodukovatelnost. Na druhou stranu jde o časově náročný postup, vyžadující velké množství vzorku. Navíc každá bodová mutace musí být analyzována jednotlivě (Lateef, 2015).

## RAPD – Random Amplified Polymorphic DNA

Největší rozmach zažila tato metoda v 90. letech 20. století, kdy nacházela uplatnění hlavně v hodnocení variability populací a druhů, nebo fylogeneze. Později byla vystřídána modernějšími metodami (Lateef, 2015).

Jedná se o techniku s PCR, ve které nasedají krátké primery (zhruba o velikosti 10 nukleotidů) ve vhodné vzdálenosti a orientaci k homologním sekvencím DNA a amplifikuje se úsek mezi nimi (Obrázek 9). Následně se provede gelové elektroforéza a sleduje se polymorfismus fragmentů vzniklých na základě mutací.

Pozitivem RAPD je jednoduchost a absence specifických primerů. Negativem je špatná reprodukovatelnost, náročnost provedení a dominance markerů (Garrido-Cardenas, 2018).



Obrázek 9 Variabilita amplifikace a následná elektroforetická vizualizace u metody RAPD (Fér, 2009)

## AFLP – Amplified Fragment Length Polymorphism

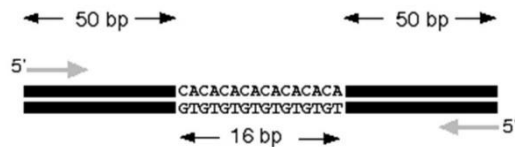
AFLP je kombinací RFLP a PCR s rozdílem, že u AFLP dochází ke štěpení pomocí dvou různých endonukleáz. Mezi rozštěpené konce DNA se ligují adaptory o známé sekvenci s lepivými konci. Navrhnu se vhodné primery a amplifikuje se sekvence mezi nimi. Následně proběhnou dvě PCR – preselektivní a selektivní, aby byly namnoženy pouze žádané fragmenty.

Následně jsou fragmenty separovány pomocí gelové elektroforézy a vizualizovány pomocí stříbra či fluorescenčního barviva (Lateef, 2015).

Metoda je poměrně časově náročná, ale její markery jsou vysoce polymorfní, spolehlivé a dobře reprodukovatelné (Garrido-Cardenas, 2018). Není třeba znalostí o struktuře genomu, a proto nachází AFLP největší uplatnění v tvorbě genetických map (Schlötterer, 2004).

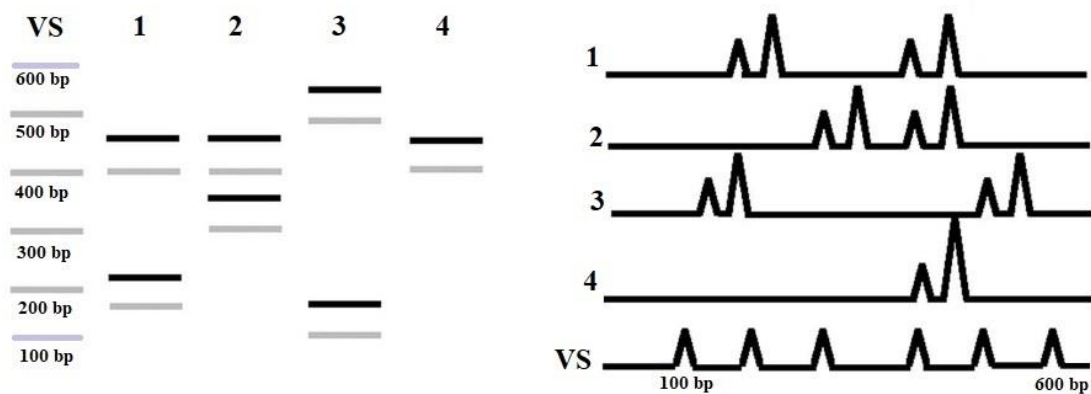
## SSR – Simple Sequence Repeats

Vzhledem k zaměření předložené diplomové práce je mikrosatelitům a jejich detailnějšímu popisu věnována kapitola 3.2 Mikrosatelity. Metoda využívající SSR markery začíná návrhnutím primerů na lemující oblasti mikrosatelitů (Obrázek 10). Mohou být fluorescenčně značené či neznačené. Vždy je navrhnut jeden pár primerů na jeden lokus. Následně probíhá PCR a vizualizace pomocí gelové elektroforézy nebo pomocí kapilární gelové elektroforézy (Guichoux *et al.*, 2011) (Obrázek 11).



Obrázek 10 Navrhnuté primery lemující oblast mikrosatelitu

(Křenek, 2018)



Výsledný gel z gelové elektroforézy

Vizualizace pomocí kapilární gelové elektroforézy

Obrázek 11 Detekce SSR polymorfismu

(Křenek, 2018)

SSR markery jsou rozděleny dle svého původu na genomické (g-SSRs) a kódující EST-SSRs (expressed sequence tags). V obou případech se jedná o markery, které disponují kodominantním charakterem, vysokým polymorfismem, lokusovou specifičností, multialelickým zastoupením, dobrou reprodukovatelností a poměrně snadným vyhodnocením (Taheri *et al.*, 2017). Tyto vlastnosti dělají z SSR všestranné markery nalézající uplatnění v mnoha oblastech jako je například genotypizace, tvorba genetických map, zkoumání genetické diverzity, určení paternity, forézní analýza, stanovení haplotypů, genetická diagnostika a studium struktury genomů, celogenomová asociační studie (GWAS Genome-Wide Association Study), asistované šlechtění, genové toky a míra crossing-overů (Varshney *et al.*, 2005; Guichoux *et*

*al.*, 2011; Kalia *et al.*, 2011; Antiqueira, 2013; Cvrčková *a kol.*, 2013; Gao *et al.*, 2015; Zargar *et al.*, 2015; Carucci *et. al.*, 2017; Li *et al.*, 2017; Taheri *et al.*, 2017).

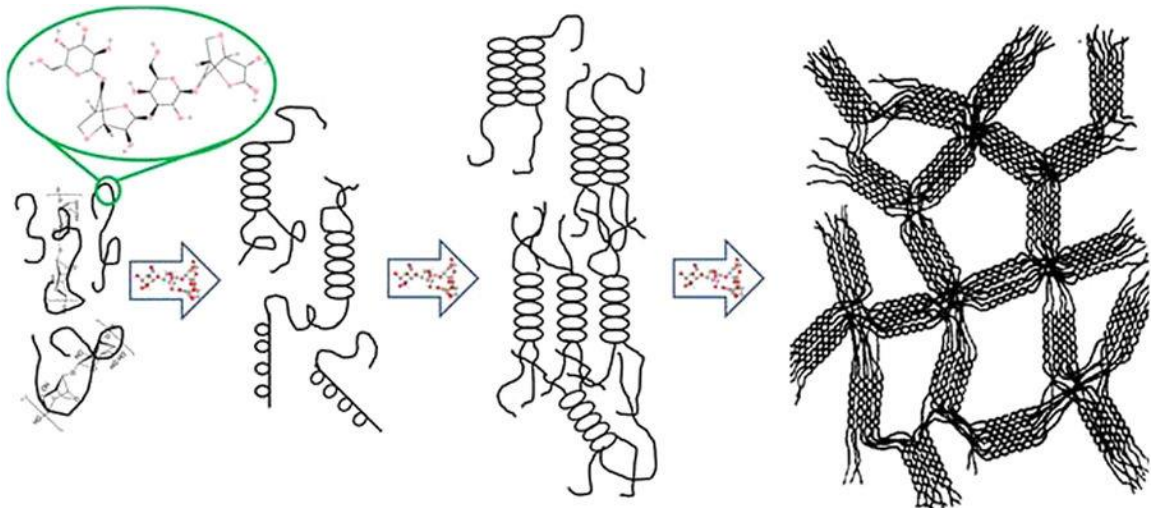
### **SNP – Single Nucleotide Polymorphisms**

SNP neboli jednonukleotidový polymorfismus je založen na záměně jednoho nukleotidu (A, T, C nebo G) v určitém místě DNA (Lateef, 2015). K záměně dochází bodovou mutací a dle typu se dělí na (i) transici (záměna C/T nebo G/A) či (ii) transverzí (C/G, A/T, C/A). Tyto změny v určité pozici jsou využívány jako velmi efektivní genetické markery prakticky u všech studovaných druhů živočichů u rostlin, hlavně díky hojnému zastoupení v celém genomu, vysoké heritabilitě a nízké míře mutací (Garrido-Cardenas, 2018). SNP se využívají například ke genotypizaci, určení paternity a příbuznosti (Kösoglu *et al.*, 2017), při expresi genů (Zeng *et al.*, 2017) nebo při hodnocení variability populací (Liu *et al.*, 2018).



### 3.4 Elektroforéza

Elektroforéza je separační metoda, která využívá schopnosti nabitých částic v elektrickém poli. Principem metody je pohyb nabitých částic (bílkovin, iontů) v elektrickém poli k elektrodě s opačnou polaritou vlivem stejnosměrného elektrického proudu. Rychlost pohybu je závislá na velikosti náboje, velikosti a konformaci molekul, které se pohybují v roztoku elektrolytu. V současné době se nejvíce používají gelové roztoky škrobový, agarózový (Obrázek 12) anebo polyakrylamidový, které jsou tvořeny složitou sítí polymerních molekul s póry, jimiž se molekuly pohybují. Je obecně známo, že menší částice se pohybují rychleji než větší a konformačně složitější částice rychleji než jednodušší, protože jsou kompaktnější a snadněji prochází póry v gelu (Lee *et al.*, 2012; Serwer, 1983; Yilmaz *et al.*, 2012). Na pohyb částic neboli mobilitu ( $m$ ) má vliv i mnoho dalších faktorů jako je teplota, pH a viskozita roztoku (Kašička, 1997).



Obrázek 12 Polymerní struktura agarózového gelu

(Zarrintaj *et al.*, 2017)

Rozdělení elektroforéz dle prostředí, ve které dochází k separaci částic:

- Kapilární zónová elektroforéza (CZE)
- Kapilární gelová elektroforéza (CGE)
- Elektroforéza v plynné fázi
- Iontová mobilní spektrometrie (IMS)
- Kapilární izotachoforéza (CITP)
- Kapilární izoelektrická fokusace (CIEF)
- Micelární elektrokinetická kapilární elektrochromatografie (MEKC)
- Kapilární elektrochromatografie (CEC)



### Typy elektroforézy využívané k separaci nukleových kyselin:

Pomocí elektroforézy lze rozdělit fragmenty nukleových kyselin. Nukleová kyselina nese záporný náboj díky záporně nabitým fosfátům, proto se v elektrickém poli pohybuje od záporného pólu (katody) ke kladnému (anodě) (Adler, 2013).

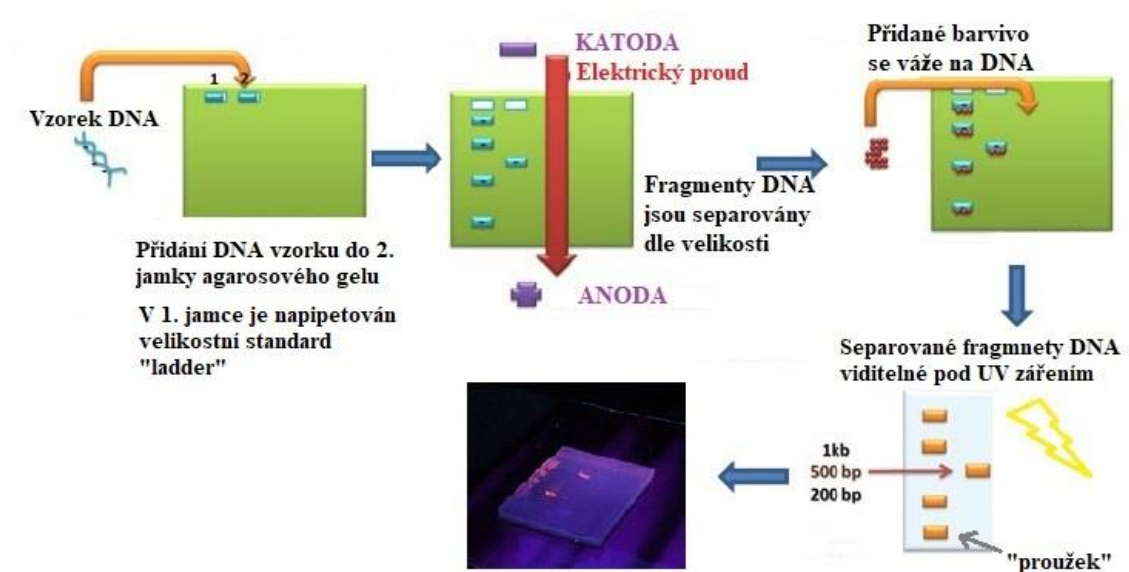
#### 3.4.1 Gelová elektroforéza

K separaci dochází v agarózovém gelu, jehož koncentrace se odvíjí od velikosti separovaných molekul (Tabulka 2) (Yilmaz *et al.*, 2012). Vzorky se postupně nanášou do jamek v gelu, které byly vytvořeny pomocí tzv. hřebínku. Následně probíhá samotná separace. Jak je znázorněno v Obrázku 13, při gelové elektroforéze se využívá velikostního standardu "ladder", který slouží pro odhad velikosti separovaných fragmentů. Jedná se o soubor velikostně definovaných fragmentů a nanáší se do jedné jamky v gelu paralelně k testovaným vzorkům (Lee *et al.*, 2012).

Koncentrace agarózového gelu (%)	Denaturovaná DNA/RNA (kp)
0,3	5-60
0,6	1-20
0,7	0,8-10
0,9	0,5-7
1,2	0,4-6
1,5	0,2-3
2,0	0,1-2

Tabulka 2 Koncentrace agarózového gelu pro separaci nukleových kyselin

(upraveno dle Yilmaz *et al.*, 2012)



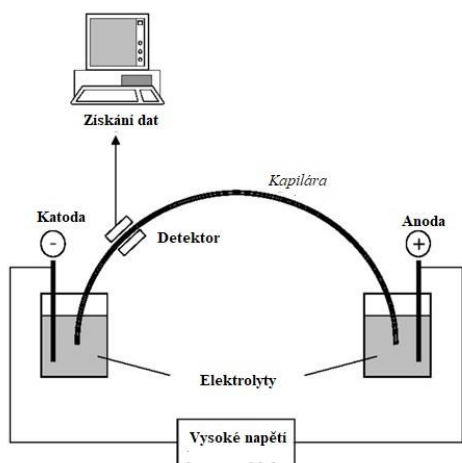
Obrázek 13 Schéma průběhu gelové elektroforézy (upraveno dle Yilmaz *et al.*, 2012)

K vizualizaci nukleové kyseliny v gelu se běžně využívá ethidium bromid (EtBr) anebo SYBR Green, po jejichž osvětlení UV zářením je možné fragmenty v podobě proužků na gelu pozorovat.

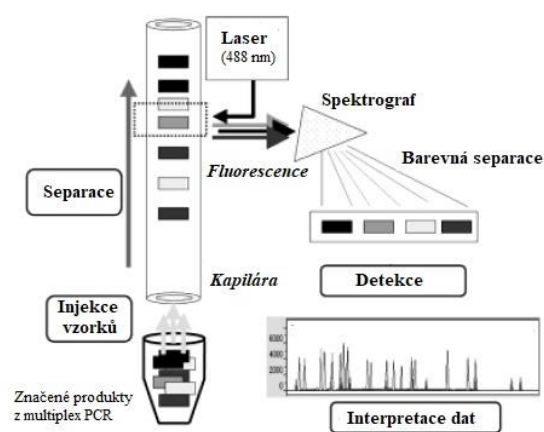
Gelová elektroforéza je univerzální metodou s širokým využitím, například při odhadu velikosti DNA fragmentů, při analýze produktů PCR, genetických analýzách (Yilmaz *et al.*, 2012), nebo při Sangrově metodě sekvenování (Sanger *et al.*, 1977). Dále je využívána také při oddělení DNA v Southern blotu a při oddělení RNA v Northern blotu (Yilmaz *et al.*, 2012).

### 3.4.2 Kapilární elektroforéza

Kapilární elektroforéza, nebo také vysokoúčinná kapilární elektroforéza (HPCE), je taktéž využívá k separaci rozdílné pohyblivosti fragmentů nukleových kyselin v elektrickém poli. Na rozdíl od gelové elektroforézy ale separace probíhá od opačných elektrod uvnitř křemičité kapiláry (Obrázek 14). Kapilára je naplněna opět agarózovým nebo polyakrylamidovým gelem (Heller, 2001). Průměr kapiláry se obvykle pohybuje okolo 50–100  $\mu\text{m}$ . Díky tomuto úzkému průměru kapilára umožňuje vyšší elektrické napětím a tím rychlejší dobu separace (Butler *et al.*, 2004). Na průběh separace má vliv ale i mnoho dalších faktorů jako je délka kapiláry, elektroosmotický tok, velikost napětí, teplota, pH a iontová síla elektrolytu (Dolník, 1999; Kašička, 1997).



Obrázek 14 Schéma kapilární elektroforézy



Obrázek 15 Schéma separace a detekce HPCE

(Butler *et al.*, 2004)

Celý proces gelové kapilární elektroforézy je plně automatizován od nasátí vzorků, separace, detekce laserem až po přenos dat do PC (Obrázek 15) (Butler *et al*, 2004). Detekce této elektroforézy je také kompatibilní s mnoha jinými metodami, jako je například hmotnostní spektrometrie, voltametrie, nebo s metodami založenými na absorbanci a fluorescenci (Heller, 2001).

Kapilární gelová elektroforéza je velmi užitečnou technikou při forezních a genetických analýzách, zejména při sekvenaci, ale i při fragmentačních analýzách, Westren blotu a mnoha dalších (Butler *et al*, 2004).

#### Srovnání metod

Velkou výhodou gelové elektroforézy je její jednoduchost a nižší finanční náročnost. Příprava agarozového gelu je snadná a relativně rychlá. Dalším pozitivem je, že se jedná o nedestructivní metodu. Je tedy možné vzorky z gelu opět vyextrahovat a použít pro další analýzu. Na druhou stranu je nevýhodou přírodní původ agarózy, jednotlivé šarže se od sebe mohou lišit, byť jsou dodány stejným výrobcem. To může zapříčinit mírnou odlišnost výsledků.

Oproti gelové elektroforéze je kapilární gelová elektroforéza mnohem rychlejší metoda s lepším rozlišením. Jedná se o vysoce účinnou techniku s malou spotřebou vzorků, malým množstvím odpadu a spotřeby chemikálií. Nicméně se jedná o metodu náročnou na obsluhu i finance.

### 3.5 Problematika kvality dat při genotypizaci

Během genotypizace se postupuje dle přesně definovaných protokolů, které se skládají z jednotlivých kroků. Počínající izolací DNA, stanovením její kvality či kvantity, PCR reakce, využití genetických markerů, až po vyhodnocování získaných dat. Na průběh genotypizace má vliv řada faktorů, které mohou zapříčinit vznik chyb (Tabulka 3). Snahou experimentálních protokolů je chybám předcházet a zajistit tak nejvyšší kvalitu dat a jejich reprodukovatelnost (Bonin *et al.*, 2004; Hoffman and Amos, 2005; Pompanon *et al.*, 2005). Důležité je pracovat s kvalitními vzorky, co nejlépe optimalizovat podmínky a zvolit vhodné markery. Všechny genetické markery jsou náchylné ke vzniku chyb. Předložená práce se zaměřuje na chybovost při využití SSR markerů (mikrosatelitů).

Chybovost SSR markerů v průběhu genotypizace může být způsobena nízkou kvalitou či kvantitou DNA, která se děje například při neinvazivním odběru vzorků. Chyba může nastat, ale i pokud máme dostatečnou kvalitu a kvantitu DNA. Například pokud proběhne mutace v místě nasedání primerů, amplifikace dané alely neproběhne a na výsledném elektroforeogramu se setkáme s tzv. nulovou alelou (podrobněji vysvětleno níže). Ani PCR reakce není zcela bezchybná. Často dochází k vynechání amplifikace jedné z alel u heterozygota a zesílení amplifikace alely druhé. Tento jev se nazývá „alelický dropout“ (z anglického allelic dropout) (Pompanon *et al.*, 2005).

Při vyhodnocení dělají problém i tzv. koktavé produkty (anglicky stutter bands), které jsou způsobené replikačním sklouznutím *Taq* polymerázy během PCR reakce. Nízká procesivita polymerázy a další aspekty mohou způsobit, že během elongačního cyklu PCR polymeráza nenasyntetizuje celou délku mikrosatelitního opakování. Výsledné fragmenty jsou o příslušný počet opakování kratší než skutečná alela a může tak docházet k záměně homozygota za heterozygota (Hoffman and Amos, 2005).

Pro kontrolu těchto a mnoha dalších chyb existují screeningové programy. Například SimWalk2, MENDEL (Sobel *et al.*, 2005), DROPOUT, Genecheck a další (Pompanon *et al.*, 2005).

Tabulka 3 Příčinné zdroje chyb

Zdroj dané chyby	Mechanismus vzniku chyby	Důsledek
<b>Interakce mezi molekulami DNA</b>		
DNA sekvence ohraničující marker	Selhání amplifikace následkem mutace v místě nasedání primerů	Nulová alela
DNA sekvence ohraničující marker	Inzerce či delece v amplifikovaném fragmentu	Velikostní homoplazie
DNA sekvence ohraničující marker	Preference amplifikace jedné alely u heterozygota	Alelický dropout
<b>Kvalita vzorku</b>		
Nízká kvalita či kvantita DNA	Amplifikace pouze alely u heterozygota	Alelický dropout
Nízká kvalita či kvantita DNA	Amplifikace kratší alely u heterozygota	Převaha výskytu krátkých alel
Kontaminace vzorku	Amplifikují se i kontaminace	Chybná alela
Nízká kvalita extrakce DNA	Limitovaná amplifikace vlivem inhibitorů	Alelický dropout
<b>Chemikálie a laboratorní vybavení</b>		
Nízká kvalita činidel	Snížení účinnosti amplifikace vlivem inhibitorů	Alelický dropout, chybné alely
Nízká kvalita činidel	Špatné označení a detekce fragmentů	Alelický dropout, chybné alely
Nízká přesnost a spolehlivost přístrojů	Špatné pipetování, odpařování během PCR, špatná detekce fluorescenčně značených primerů	Alelický dropout, chybná alela
Chyba Taq polymerázy	Sklouznutí polymerázy v prvních krocích PCR	Amplifikace nesprávné alely
Chyba Taq polymerázy	Přidání netemplátového nukleotidu	Amplifikace nesprávné alely
Nedostatečná specifita primerů	Amplifikace nespecifických fragmentů následkem špatného zabudování primerů	Amplifikace nesprávné alely
Nedostatečná specifita primerů	Nespecifické štěpení amplikonu	Amplifikace nesprávné alely
Elektroforetické artefakty	Zkreslení velikosti alel vlivem migrace v gelu (vysoká koncentrace PCR produktu, teplota)	Velikostní homoplazie, chybná alela
Elektroforetické artefakty	Nekonzistentní velikost jednotlivých alel mezi různými experimenty (například kapilární elektroforéza vs gelová elektroforéza)	Velikostní homoplazie, chybná alela
<b>Chyby způsobené lidským faktorem</b>		
Manipulace se vzorkem	Špatné označení vzorků	Chybná alela
Chyby vzniklé v průběhu experimentu	Kontaminace exogenní DNA nebo mezi vzorky	Chybná alela
Chyby vzniklé v průběhu experimentu	Použití nevhodného protokolu (nevhodná činidla, primery; koncentrace, naprogramování PCR)	Alelický dropout, chybná alela
Zpracování dat	Nesprávné či špatně zadané genotypy v databázích	Chybná alela
Zpracování dat	Nesprávné vyhodnocení a vizualizace fluorizační označených peaků	Chybná alela
Zpracování dat	Chyba v programech na analýzu dat	Chybná alela

(upraveno dle Pompanon et al., 2005)

### 3.5.1 Nulové alely

Následkem bodové mutace, inserce či delece dojde k pozměnění sekvence předcházející mikrosatelit (Gugerli *et al.*, 2008; Vieira *et al.*, 2016), která by měla být homologní pro PCR primer. Je tak znemožněna činnost polymerázy a daný úsek není amplifikován. Na finálním elektroforeogramu se nezobrazí žádný nebo jen velice slabý fragment, což může vést ke komplikacím při vyhodnocování dat (Callen *et al.*, 1993). Například při určení paternity, kdy mohou nulové alely způsobit chybné určení rodičů jedince.

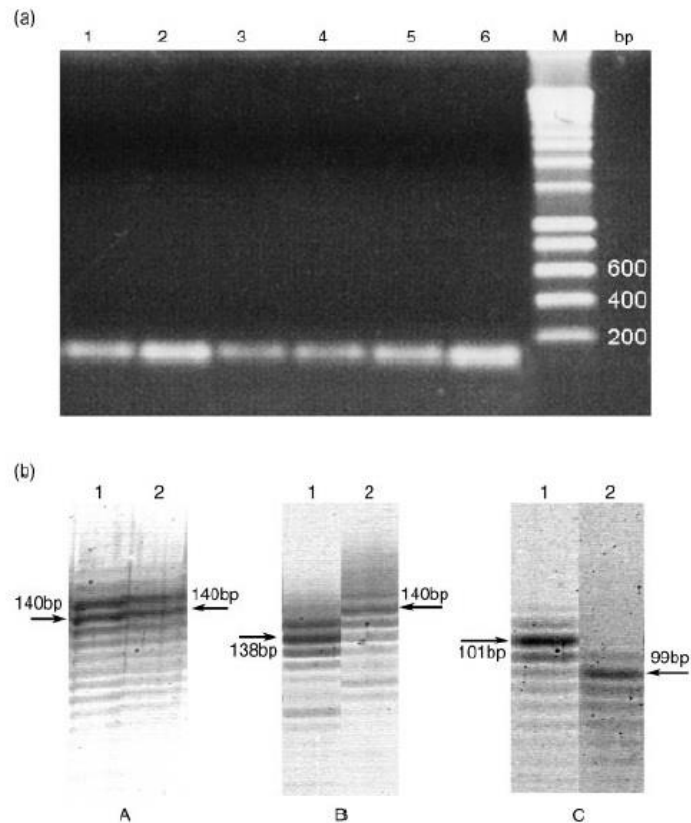
Při detekci může nastat problém jak u homozygota, tak i u heterozygota. V případě homozygota, který má nulovou alelu (nn, n je nulová alela), genotypizace selže. Pokud je jedinec heterozygotní s jednou nulovou alelou (in, kde i je normální nenulová alela), bude sledovaný genotyp těžko odlišitelný od skutečného homozygota (Wagner *et al.*, 2006). Z toho je patrné, že přítomnost nulových alel vede ke snížení hladiny pozorované heterozygotnosti (Lehmann *et al.*, 1996).

Přítomnost nulových alel může být zjištěna v populačních studiích výpočtem Hardy-Weinbergovy rovnováhy, kde se projeví jako nedostatek heterozygotů (za předpokladu, že pozorovaný nedostatek není jiného původu, tzv. Wahlundův efekt (Chapuis and Estoup, 2007). K nepřímému odhalení nulových alel se v dnešní době využívají různé počítačové programy, k přímému odhalení je nutná sekvenace PCR produktu sledovaného lokusu a úprava homologní sekvence nasedajícího primeru (Dakin and Avise, 2004).

### 3.5.2 Homoplazie

Termín převzatý z evolučních studií, který popisuje, že určitá vlastnost totožná u dvou a více druhů nemusí mít stejný původ, tj. nepochází od stejného předka. Příčinu je evoluční konvergence, paralelismus či reverze (Estoup *et al.*, 2002).

V případě genotypizace se setkáváme s podobným pojmem tzv. velikostní homoplazií, když jsou dvě alely považovány za identické, ale nevychází ze stejných genotypů (Obrázek 16) (Estoup *et al.*, 2002). Homoplazie se objevuje i při užití mikrosatelitních markerů v populačně genetických a fylogenetických studiích. Tyto studie jsou charakterizovány velkým počtem vzorků a mikrosatelitní alely jsou rozlišovány na základě fragmentů (tzv. elektromorf), pomocí gelové elektroforézy. Pokud ale v lemujících oblastech („the flanking regions“) proběhne substituce nukleotidů, přerušení či sloučení mikrosatelitových repetitiv, není možné výslednou velikost homoplazie rozlišit (Estoup *et al.*, 1995). K detekci sekvenční rozdílů lze následně využít například sekvenování (Liepelt *et al.*, 2001).



Obrázek 16 (a) Snímek z agarózové elektroforézy s odlišnými bandy, (b) příklad velikostní homoplazie  
(Liepelt *et al.*, 2001)

### 3.5.3 Přidání netemplátového nukleotidu

Na konci prodlužovací fáze během PCR reakce má *Taq* polymeráza tendenci přidat netemplátový nukleotid, většinou se jedná o adenin, na 3' konec nově syntetizovaného vlákna. Tento jev je relativně častý. Na výsledném elektroforeogramu z gelové elektroforézy vytváří fragment a v případě kapilární elektroforézy nadbytečný pík, což komplikuje vyhodnocení velikosti příslušného fragmentu a ztěžuje detekci alel (Pompanon *et al.*, 2005).

## 4 Metodika

### 4.1 Rostlinný materiál

Byly získány vzorky v podobě semen z genobanky Opava, Ústředního kontrolního a zkušebního ústavu zemědělského a z pokusné stanice Červený Újezd. Pro potřeby analýz byly vypěstovány rostliny jednotlivých odrůd ve skleníku. V růstové fázi DC25, kdy měly rostliny dva až tři pravé listy byl získán materiál pro analýzu DNA. Seznam testovaných odrůd uvádí Tabulka 4.

Tabulka 4 Seznam odrůd

Číslo	Název odrůdy	Zdroj materiálu
1	FLORIAN	(Genobanka Opava)
2	GERLACH	(Genobanka Opava)
3	ROSEMARIE	(Genobanka Opava)
4	OPAL	(Genobanka Opava)
5	SOKOL	(Genobanka Opava)
6	POSTOMI	(Genobanka Opava)
7	ORFEUS	(Genobanka Opava)
8	MALSAR	(Genobanka Opava)
9	MAJOR	(Genobanka Opava)
10	ALBÍN	(Genobanka Opava)
11	KÉK DUNA	(Genobanka Opava)
12	RACEK	(Genobanka Opava)
13	ZENO	(Genobanka Opava)
14	KOSMOSZ	(Genobanka Opava)
15	AMETISZT	(Genobanka Opava)
16	MARIANNE	(Genobanka Opava)
17	BERGAM	(Genobanka Opava)
18	REDY	(Genobanka Opava)
19	DANISH FLAG	(maloobchod)
20	BUDDHA	(Genobanka Opava)
22	KORNEUBURGER	(Červený Újezd)
23	AKVAREL	(Červený Újezd)
24	ORBIS_II	(ÚKZÚZ)
25	OPEX_II	(ÚKZÚZ)
26	ONYX_II	(ÚKZÚZ)
27	LAZUR	(Genobanka Opava)
28	BERGAM_II	(ÚKZÚZ)
29	RACEK_II	(ÚKZÚZ)
30	OPÁL_II	(ÚKZÚZ)
31	MARATON	(Genobanka Opava)
32	OREL	(Genobanka Opava)



## 4.2 Izolace DNA

Izolace byla provedena pomocí DNeasy Plant Mini Kit (Qiagen) dle manuálu výrobce. Z každé rostliny bylo naváženo 150 mg a následně byla provedena homogenizace vzorku rozmělněním v třecích miskách pomocí tloučku a tekutého dusíku. Homogenizované vzorky byly převedeny do zkumavek spolu s přidáním 400  $\mu\text{l}$  pufru AP1 a 4  $\mu\text{l}$  roztoku RNase A. Následně byla směs ve zkumavkách homogenizována na vibrační třepačce a uloženy na 10 minut do termobloku teplotou 65 °C, kde byly každé tři minuty lehce promíchány. Po vyjmutí z termobloku bylo do zkumavek napipetováno 130  $\mu\text{l}$  lyzačního pufru P3 a byly ponechány na 5°C na ledu po dobu 5 min. Následky byly zkumavky centrifugovány při 20 000 x g (14, 000 rpm) po dobu 5 minut. Vzniklý supernatant byl převeden do nových zkumavek s filtrem a znovu centrifugován po dobu 2 minut. Po centrifugaci se odstranila vrchní část s filtrem a k supernatantu bylo přidáno 650  $\mu\text{l}$  pufru AW1. Vzniklá směs byla lehce promíchána pomocí pipety a polovina byla převedena do nových kolonek s filtrem a opět byla provedena centrifugace při 8000 rpm po dobu 1 minuty. Po centrifugaci byl odstraněn odstředěný supernatant. Na filtr byl pomocí pipety přenesen zbytek směsi, která byla opět centrifugována. V této fázi byla DNA zachycena na filtru. V dalším kroku bylo ke směsi přidáno 500  $\mu\text{l}$  pufru AW2 s ethanolem, který zbavil vzorky nepolárních látek, barviv a dalších nežádoucích látek a následně se provedla centrifugace při 8000 rpm po dobu 1 minuty. Tento krok se provedl dvakrát za sebou a následně byla provedena ještě jedna centrifugace bez přidání AW2 pufru. Vždy byl odstraněn odstředěný supernatant a dále se pracovalo s filtrem, na kterém byla navázána DNA vzorků. Posledním krokem bylo oddělení DNA z filtru do připravených zkumavek. Bylo přidáno 100  $\mu\text{l}$  AE pufru na filtr ve zkumavkách a po 5 minutách při laboratorní teplotě proběhla opět centrifugace při 8000 rpm po dobu 1 minuty. Po izolaci byla provedena kvantifikace izolované DNA na UV spektrometru S-111107 NanoPhotometer (Implen). Zároveň s kvantifikací byla ověřena i čistota získaných vzorků na základě poměru absorbancí  $A_{260}/A_{280}$ . Získaná DNA byla naředěna na 10  $\text{ng}\cdot\mu\text{l}^{-1}$ . Nakonec byla ověřena vysokomolekularita DNA pomocí separace v gelové elektroforéze.

## 4.3 Bioinformatická analýza dat

### 4.3.1 Publikované mikrosatelitní markery

K porovnání nově navržených EST-SSR markerů bylo rovněž testováno 15 genomických SSR markerů, které publikoval Celik *et al.* (2014) a Kati *et al.* (2013).

Tabulka 5 Publikované primery

Publikovaný název	primery F (5'3')	primery R (5'3')	repetice	zdroj
psgSSR-05	TGTTTTAGCTGCTTAGCTGT	AAATGGAAATGGAACAGATG	CATCTG	Celik et al. (2014)
psgSSR-06	AATCCCTCTCTTCTTACCT	GGGGATATTTTATGTCTTGG	AACA	Celik et al. (2014)
psgSSR-21	AGGACCAGGATCTACAGATTT	GTATTGTTTGTGGGTTGAAAG	TA	Celik et al. (2014)
psgSSR-23	CTGATTGAGTACCTTTCATGC	GAGCCTAATGAGTATTGAAGC	TGTC	Celik et al. (2014)
psgSSR-27	TACGTCTCGAAATACCAAATG	CTCAAGGACTTCATGGAGATT	TA	Celik et al. (2014)
psgSSR-30	AAGAAGTGGACGTGATGATAA	GTCAAACATCTGAGTGATGGT	AACA	Celik et al. (2014)
psgSSR-34	GCACCAACTTCTATGAAGAAT	CATCATTTTCATGGACATACC	TGG	Celik et al. (2014)
psgSSR-37	TCTGAACTGAGATCAACAGAGA	GATTAGTACGAAGGGGATTGT	GAA	Celik et al. (2014)
psgSSR-47	GACGAGATTTCTTGAGAACT	ATGTCTTCAGCTTAGGGTTC	AGA	Celik et al. (2014)
psgSSR-68	TGGGTTTTCTTCTATACCACA	CCCATTACTCTCAAGCTAAT	TTCT	Celik et al. (2014)
psgSSR-76	TGGATACCTCTCAAGTAAGCA	CATGCAGATCATAAAAATCC	AATG	Celik et al. (2014)
psgSSR-98	CCTTATTTCGGATATGGTT	CAGAATCTCGCAGCAATAAT	TGTTG	Celik et al. (2014)
PMS037	ACTGATACTACTTCTCTCCACC	TCGAAGAGCCTGTATTGAATC	TCT	Kati et al. (2013)
PMS051	GGAATCTCGTGGCATTCAATTAC	GAATCTTCTCCAAACACATCGAAC	GAA	Kati et al. (2013)
PMS073	TCTTCTGCATAAGGAGCATGAG	TGATGATATCTTGGGAAGATTGG	TTCT	Kati et al. (2013)

### 4.3.2 Nově navržené mikrosatelitní markery

Vzhledem k tomu, že dosud nebyl publikován kompletní genom máku (*Papaver somniferum* L.), tak byla využita hrubá sekvenční data z transkriptomických analýz máku. Tato data byla stažena z SRA archivu Národního centra pro biotechnologické informace (NCBI). Získaná data byla převedena do tzv. fasta formátu a vlastní návrhy EST-SSR markerů bylo provedeno pomocí softwaru GMATA (Wang and Wang, 2016), jehož parametry byly nastaveny na vyhledání trinukleotidových repetitiv s 12 a více opakováními. Navržené primery byly zkontrolovány pomocí online programu Primer3 (Rozen and Skaletsky, 2000). Tento software byl rovněž použit k návrhu nových markerů pro vybrané lokusy v případě, že původní primery neposkytovaly amplikony požadované kvality.

## 4.4 Polymerázová řetězová reakce (PCR)

### 4.4.1 Koncepce testování a optimalizace nových primerů

Testování potenciálních markerů bylo provedeno v několika krocích. Prvním krokem bylo ověření, jestli dochází k amplifikaci při teplotě 60 popř. 57 °C na 4 genotypech. V druhém kroku bylo provedeno opětovné testování na 8 odrůdách při vhodné annealační teplotě. Pro vybrané markery byly zakoupeny fluorescenčně značené primery k dalšímu testování na kapilární elektroforéze. Jako fluorofor sloužila modrá barva 6FAM.

Amplifikace potenciálních markerů probíhala po jednotlivých lokusech a vlastní složení reakční směsi o objemu 10 µl bylo následující: 1x Multiplex PCR Plus Kit (Qiagen), 20 ng DNA, 0,2 µM F a R primer, 3 µl ddH<sub>2</sub>O. Teplotní profil PCR je uveden v tabulce 8, přičemž po finální elongaci následovalo chlazení vzorků na 12°C. Amplifikace probíhala v termocykleru C-1000 (Bio-Rad) nebo T Gradient Thermocykler (Biometra) (Obrázek 17).

*Tabulka 6 Program PCR reakce při testování navržených primerů*

	Doba	Opakování	Teplota (°C)
<b>Před-denaturace</b>	900 s	1x	95
<b>Denaturace</b>	30 s	35x	94
<b>Annelace</b>	90 s	35x	60 nebo 57
<b>Elongace</b>	60 s	35x	72
<b>Finální elongace</b>	600 s	1x	72

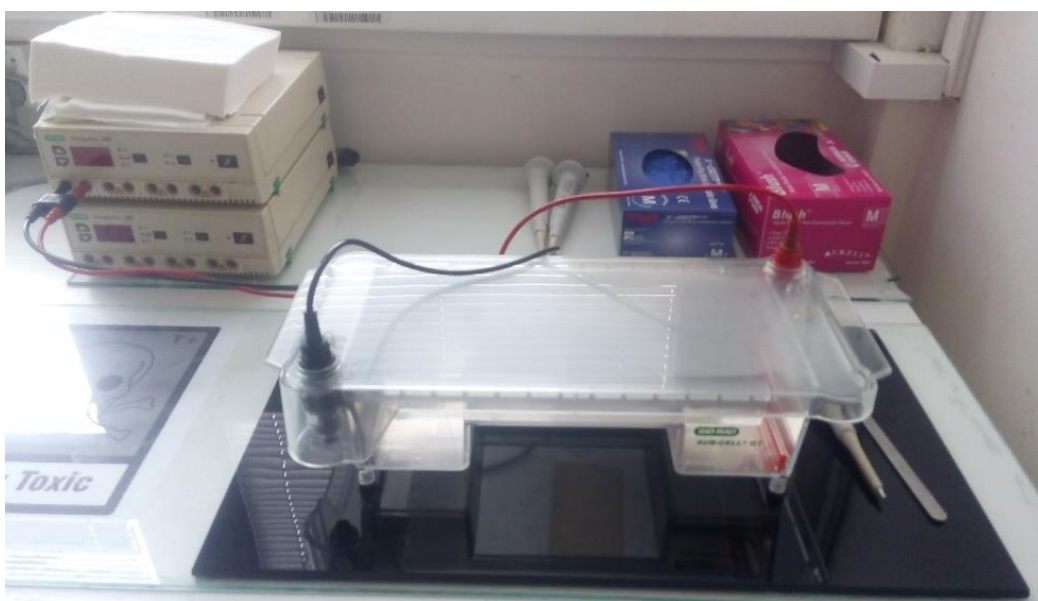


*Obrázek 17 Thermal Cycler C 1000*

## 4.5 Elektroforéza

### 4.5.1 Agarózová elektroforéza

Separace genomické DNA probíhala v 1% agarózovém gelu a testování EST-SSR markerů bylo provedeno v 1,5% agarózovém gelu. Separace byla provedena v cele SubCell GT (Bio-Rad) naplněné 1xTBE pufrem při napětí  $4 \text{ V} \cdot \text{cm}^{-1}$  (Obrázek 18). K orientačnímu stanovení velikosti produktů byl použit velikostní standard GeneRuler 100 bp nebo GeneRuler 100 bp plus (Thermo Fischer Scientific).



*Obrázek 18 Cella SubCell GT pro agarózovou elektroforézu*

### 4.5.2 Kapilární gelová elektroforéza

#### 4.5.2.1 Příprava vzorků

Před vlastní separací v kapilární elektroforéze byl  $1 \mu\text{l}$  PCR produktu naředěn vodou v poměru 1:50-120 dle příslušného vzorku a markeru. Dále byl  $1 \mu\text{l}$  ředěného PCR produktu smíchán s  $12 \mu\text{l}$  Hi-Di formamidu (Life Technologies) a  $0,2 \mu\text{l}$  hmotnostního standardu GeneScan LIZ600 (Life Technologies). Získaná směs byla ve finále denaturována při teplotě  $95^\circ\text{C}$  po dobu 5 minut.

#### 4.5.2.2 Nastavení podmínek separace

Fragmentační analýza byla provedena v genetickém analyzátoru ABI PRISM 310 (Life Technologies) (Obrázek 19). K vlastní separaci byla použita 47 cm kapilára vyplněná POP4 polymerem (Life Technologies). Separace probíhala při teplotě 60°C, která snižuje pravděpodobnosti vzniku sekundárních struktur DNA, v 1x „running“ pufru. Doba nástřiku se pohybovala v rozmezí 3-5 s při napětí 3-15 kV.



Obrázek 19 Genetický analyzátor ABI PRISM 310

#### 4.5.3 Sekvenace

Byla provedena přímá sekvenace PCR produktů pomocí klasické dideoxy terminační metody dle Sanger *et al.* (1977).

##### 4.5.3.1 Purifikace před sekvenační reakcí

Nejprve bylo nutné jednotlivé vzorky připravit, vyčistit a kvantifikovat. U všech zkoumaných markerů bylo nutné vyříznout skalpelem jednotlivé fragmenty z 1,5% agarózového gelu a izolovat je pomocí komerčního kitu GeneJET Gel Extraction Kit (Thermo Fisher Scientific). Kvantifikace probíhala pomocí UV spektrofotometru NanoPhotometer P300 od firmy Implen.

### 4.5.3.2 Sekvenační reakce

Reakční směs o celkovém objemu 20 µl obsahovala následující složky:

- 10-40 ng DNA
- 1x sekvenační pufr (Life Technologies)
- 3,2 µM F či R primeru
- 2 µl sekvenační směsi BigDye Terminátor v3.1 (Life Technologies)

Pro sekvenační reakci byl použit následující teplotní program dle  $T_A$  použitého primeru:

*Tabulka 7 Teplotní profil sekvenační reakce*

	Doba	Opakování	Teplota (°C)
<b>Před-denaturace</b>	60 s	1x	96
<b>Denaturace</b>	10 s	25x	96
<b>Annealing</b>	5 s	25x	50
<b>Elongace</b>	240 s	25x	60

### 4.5.3.3 Purifikace po sekvenační reakci

Čištění vzorků po sekvenační reakci se skládalo z následujících kroků:

- Ke vzorkům se přidalo 2 µl glykogenu (Fermentas), 50 µl 96% ethanolu a 2 µl 3M octanu sodného
- Následně se směs promíchala a ponechala 15 minut při laboratorní teplotě
- Poté proběhla centrifugace 30 min. při 13 000 rpm
- Opatrné odebrání roztoku s ponecháním bílé sraženiny ve zkumavce a přidání 200 µl 70% ethanolu
- Následně centrifugace 15 min. při 13 000 rpm
- Předchozí dva kroky proběhly ve dvou opakování
- Sušení vzorků při 50 °C 20 min.

V posledním kroku bylo do každé zkumavky přidáno 12 µl Hi-Di formamidu (Life Technologies) a proběhla denaturace vzorků při 95°C 5 min. v termocykleru C1000 (Bio-Rad).

### 4.5.3.4 Sekvenace

Vlastní sekvenace proběhla pomocí genetického analyzátoru ABI PRISM 310 (Life technologies). Separace probíhala v 61 cm dlouhé křemičité kapiláře s POP6 polymerem v 1x Running pufru (Life Technologies). V Data Collection Software 3.1.0 byl nastaven KB\_310\_POP6\_{BDv3}50Std Dye Set a jako základní modul byl zvolen Seq POP6 (1mL)E.

Doba nástřiku byla nastavena na 15 až 30 s při napětí 2,5 kV. Separace probíhala při 12,2 kV a teplotě 50°C po dobu 40-110 minut dle předpokládané velikosti separovaných fragmentu. K zajištění kvality a konzistence výsledků byly všechny fragmenty sekvenovány oboustranně u více genotypů.

#### **4.5.4 Statistická analýza a zpracování dat**

Zhodnocení bioinformatické analýzy dat a deskriptivní statistika výsledků testovaných markerů byla zpracována v podobě tabulek a grafů v programu MS Excel 2016. K určení počtu alel a jejich frekvence a identifikaci jednotlivých genotypů byl použit program Cervus 3.0.3 (Kalinowski *et al.*, 2007). Předběžná analýza rozlišovací schopnosti finálních markerů byla provedena pomocí shlukové analýzy (CLU) na základě UPGMA metody a analýzy hlavních komponent (PCA) v programu STATISTICA v13.3.721.0. Shluková analýza byla provedena na základě  $D_A$  distanční matice (Nei *et al.*, 1983) vytvořené v programu Populations ver 1.2.3.0 (vyvinutý O. Langellou, volně dostupný na <http://bioinformatics.org/~tryphon/populations>).

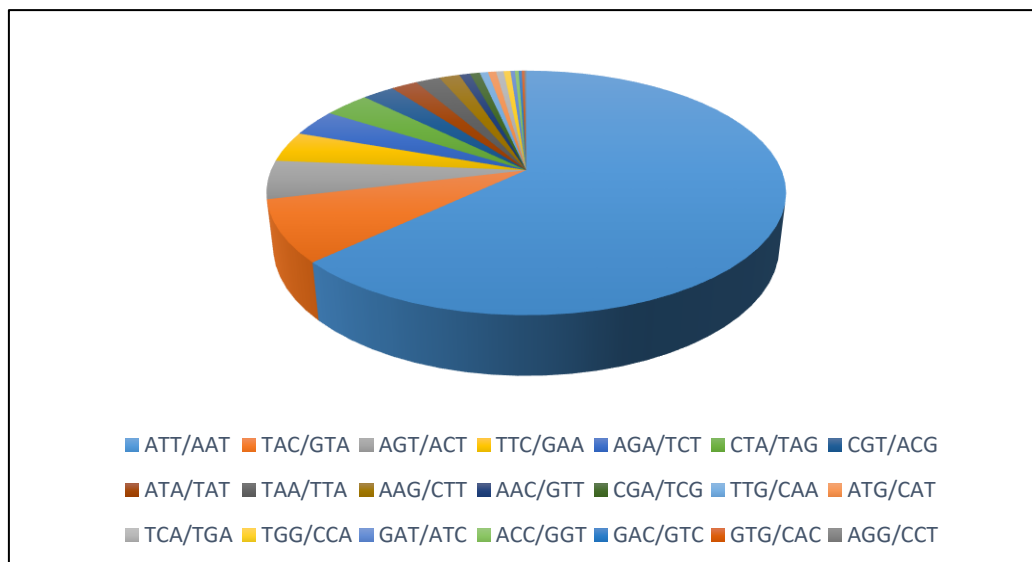
## 5 Výsledky

### 5.1 Bioinformatická analýza dat

Během analýzy bylo zpracováno 980 334 sekvencí o celkové délce 510 849 960 bází, přičemž bylo nalezeno 2 149 SSR lokusů s deseti a více trinukleotidovými opakováními. Pro 1 316 SSR lokusů byly navrženy vhodné páry primerů, z nichž bylo pro vlastní testování vybráno 45 potenciálních markerů s 12 až 14 repeticemi a 35 potenciálních markerů s 15 a více repeticemi.

Na základě analýzy transkriptomických dat bylo detekováno 21 komplementárních motivů z 64 možných (Obrázek 1). Z toho nejvíce převládaly motivy ATT/TAA (63 %), TAC/GTA (8,14 %) a AGT/ACT (5,07 %). Ostatní typy repetic měly frekvenci nižší než 5 %.

Obrázek 20 Četnost sekvenčně komplementárních motivů



### 5.2 Optimalizace a testování navržených markerů pomocí agarózové elektroforézy

Z 80 navržených EST-SSR markerů (Tabulka 12 a 13) úspěšně amplifikovalo 66 markerů (82,5 %) při  $T_a = 57^\circ\text{C}$  nebo  $60^\circ\text{C}$ . Pro další testování v kapilární elektroforéze nebo sekvenaci byly vybrány pouze kvalitně amplifikující markery u 8 testovaných odrůd (Obrázek 21). Tímto způsobem bylo vybráno 34 markerů (42,5 %).



Obrázek 21 Agarózový gel testování EST-SSR markerů na 8 odrůdách

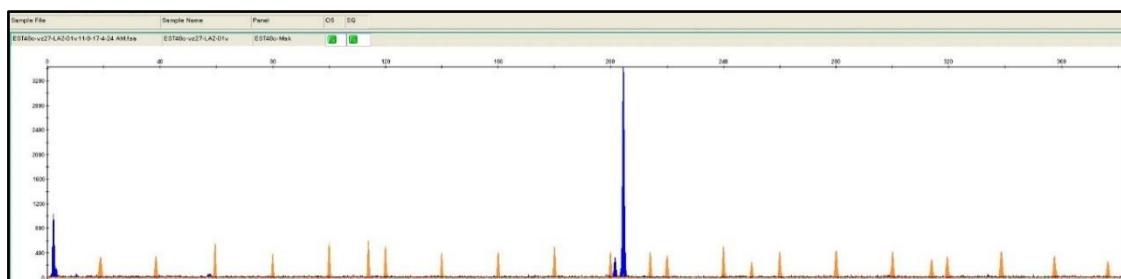


Legenda: h.s. - hmotnostní standard GeneRuler 100 bp DNA Ladder, Akv – Akvarel, Bud – Buddha, Dan – Danish Flag, Kor – Korneuburger, Mar – Maraton, Ope – Opex, Orb – Orbis, Ore – Orel, Ony – Onyx, Tat – Tatranský. V druhém kroku PCR sloužil Maraton jako pozitivní kontrola.

### 5.3 Testování markerů pomocí kapilární elektroforézy

Fragmentační analýza na kapilární elektroforéze proběhla v několika krocích. V 1. kroku byly potenciální markery testovány na 4 genotypech. Na základě těchto výsledků bylo vybráno 22 markerů (27,5 %) pro testování na celé kolekci získaných odrůd a pro některé byly navrženy nové primery zlepšující kvalitu získaného signálu (Tabulka 8). V případě markeru EST26 a EST51 byly navrženy primery se speciální koncovou sekvencí 5' GTTTCTTT '3 („pig tail“) usnadňující přidávání netemplátového A polymerázou. Nakonec bylo vybráno 8 (10%) kvalitních markerů, které byly polymorfní a poskytovaly jednoznačně interpretovatelný profil (Obrázek 22) a zároveň poskytovaly reprodukovatelné výsledky. Jedná se markery EST16, EST19, EST26, EST48c, EST51, EST53c, EST56b, EST61 (Tabulka 9).

Obrázek 22 Ukázka kvalitního profilu v případě markeru EST48c



Tabulka 8 Nově navržené primery pro kvalitní fragmentační analýzu v kapilární elektroforéze

Marker	Sekvence primeru	Typ primeru
<b>EST26*</b>	gtttcttCCGTTGTAAAATACCGACTGC	R
<b>EST37b</b>	CCTTTCCATTTCACCATC	F
<b>EST37b</b>	TATCCGATGCGAAGGTCTTG	R
<b>EST48c</b>	TCGTTCACTGAGTTCTGATATGA	F
<b>EST48c</b>	ACCGATGAAGTCATGTTTGC	R
<b>EST48b</b>	ACCGATGAAGTCATGTTTGC	R
<b>EST51*</b>	gtttcttCGACATGATGCCAATTCAAG	F
<b>EST53b</b>	GACAAAGAAACCAAGCCAAGA	R
<b>EST53c</b>	TCAAGACAAAGAAACCAAGCCA	R
<b>EST56b</b>	GGGTTTCTGAGCTTCTCCG	R
<b>EST58b</b>	TGAGATAAAATGATACATGCAATACC	R
<b>EST58c</b>	ACCAGTAATGTTTGTTGCTCGA	R

Tabulka 9 Testování 22 EST-SSR markerů na kapilární elektroforéze

Marker	Testování na kapilární elektroforéze		Komentář
	1.test*	2. test**	
<b>EST10</b>	+	+	vyřazen
<b>EST16</b>	+	+	<b>OK</b>
<b>EST19</b>	+	+	<b>OK</b>
<b>EST26</b>	+	+	<b>OK</b>
<b>EST36</b>	+	-	vyřazen
<b>EST37b#</b>	+	+	vyřazen
<b>EST39</b>	+	-	vyřazen
<b>EST41</b>	+	+	vyřazen
<b>EST42</b>	+	-	vyřazen
<b>EST45</b>	+	-	vyřazen
<b>EST48c#</b>	+	+	<b>OK</b>
<b>EST51</b>	+	+	<b>OK</b>
<b>EST53c#</b>	+	+	<b>OK</b>
<b>EST56b#</b>	+	+	<b>OK</b>
<b>EST58b#</b>	+	+	vyřazen
<b>EST61</b>	+	+	<b>OK</b>
<b>EST62</b>	+	+	vyřazen
<b>EST66</b>	+	-	vyřazen
<b>EST69</b>	+	+	vyřazen
<b>EST70</b>	+	+	vyřazen
<b>EST71</b>	+	+	vyřazen
<b>EST74</b>	+	+	vyřazen

\* 4 vzorky

\*\* 31 genotypů (single)

# Použity nové primery navržené po sekvenaci těchto lokusů

### 5.3.1 Tvorba alelických žebříků

Pro zvýšení kvality analýzy a standardizaci výsledků pro možnost mezilaboratorních srovnání byly vytvořeny pro každý z 8 finálních markerů tzv. alelické žebříky. Každý žebřík byl vytvořen smícháním vhodných genotypů dohromady reprezentujících všechny detekované alely daného lokusu (Obrázek 23a). Pro účely standardizace bylo rovněž zapotřebí jednoznačně definovat každou alelu pomocí sekvenování (Obrázek 23b,c). Z časových důvodů bylo zatím osekvenováno pouze 6 alel lokusu EST48.



## 5.4 Vyhodnocení variability získaných markerů

### 5.4.1 Deskriptivní statistika

Ke statistickému zpracování výsledků byly vytvořeny dva datové soubory. První soubor (S31) zahrnuje všech 31 odrůd testovaných na 8 EST-SSR markerech. Nicméně řadu odrůd se nepodařilo genotypovat pro všechny lokusy, a proto byl vytvořen druhý datový soubor (S17) se 17 odrůdami, pro které byl získán celkový profil, tj. byly genotypovány ve všech sledovaných lokusech. Z těchto důvodů jsou prezentovány výsledky pro oba datové soubory.

Celkem bylo detekováno 50 (S31) či 45 (S17) alel v 8 lokusech, přičemž průměrný počet alel na lokus činil 6,25 (S31) a 5,63 (S17) a počet alel na lokus se pohyboval v rozmezí 5-8 (S31) a 4-7 (S17). Detailní informace ohledně velikosti a frekvenci alel pro každý dataset uvádí tabulka 10 a 11.

Testování 15 publikovaných genomických markerů pro srovnávací účely ukázalo, že úspěšně amplifikovalo 14 z 15 markerů (93,3 %). Dalším testováním v kapilární elektroforéze bylo zjištěno, že 5 (33,3 %) markerů vytváří řadu nspecifických amplikonů a tudíž je nepoužitelných pro markerovací účely a zbylých 9 markerů (60 %) je monomorfních. Přímé srovnání variability a rozlišovací schopnosti nových a publikovaných markerů tudíž nebylo možné.

Tabulka 10 Velikosti a frekvence alel EST-SSR markerů na 31 genotypech

EST16		EST19		EST26		EST48	
Alela	Frekvence	Alela	Frekvence	Alela	Frekvence	Alela	Frekvence
268	0,0250	191	0,2857	118	0,2258	186	0,6290
271	0,1750	203	0,0536	121	0,0645	189	0,1290
277	0,250	206	0,2500	127	0,0968	192	0,0323
280	0,4750	209	0,1429	142	0,4839	195	0,0645
286	0,0750	212	0,2143	145	0,0968	198	0,0806
		227	0,0179	148	0,0323	204	0,0645
		230	0,0179				
		233	0,0179				
EST51		EST53		EST56		EST61	
Alela	Frekvence	Alela	Frekvence	Alela	Frekvence	Alela	Frekvence
221	0,0741	177	0,0645	145	0,0323	221	0,0323
226	0,5370	180	0,5806	148	0,1290	224	0,1290
229	0,3333	183	0,1452	151	0,1935	227	0,1935
232	0,0185	186	0,0968	154	0,2742	230	0,2742
235	0,0307	189	0,0323	157	0,0161	233	0,0161
		192	0,0323	160	0,3548	236	0,3548
		198	0,0323				
		201	0,0161				

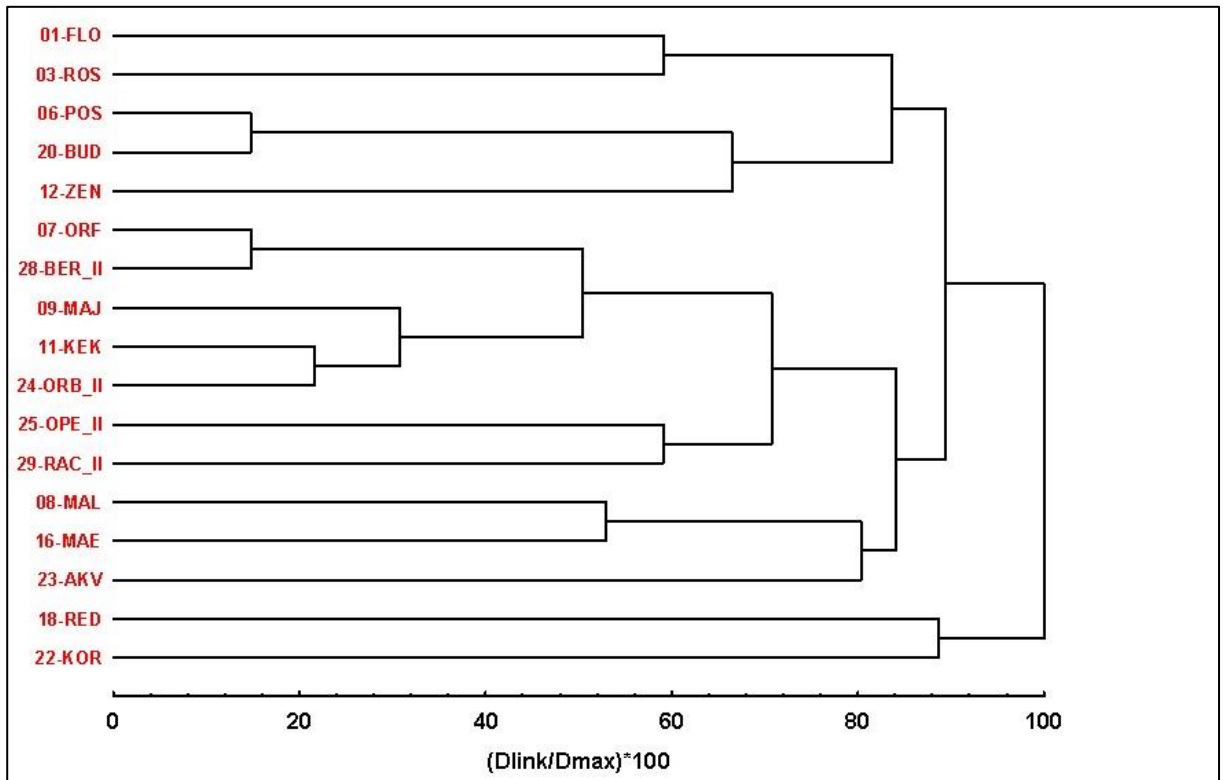
Tabulka 11 Velikosti a frekvence alel EST-SSR markerů na 17 genotypech

EST16		EST19		EST26		EST48	
Alela	Frekvence	Alela	Frekvence	Alela	Frekvence	Alela	Frekvence
268	0,0145	191	0,0683	118	0,2353	186	0,0013
271	0,0440	203	0,0133	121	0,1176	189	0,0002
277	0,0745	206	0,0543	127	0,1176	192	0,0001
280	0,1223	209	0,0268	142	0,4118	198	0,0001
286	0,0291	212	0,0683	145	0,0588	204	0,0001
		230	0,0133	148	0,0588		
		233	0,0133				
EST51		EST53		EST56		EST61	
Alela	Frekvence	Alela	Frekvence	Alela	Frekvence	Alela	Frekvence
221	0,0588	180	0,5294	145	0,0588	221	0,0588
226	0,5000	183	0,2059	148	0,1176	224	0,0157
229	0,4118	186	0,1176	151	0,1765	227	0,0237
232	0,0294	192	0,0588	154	0,2647	230	0,0398
		198	0,0588	157	0,0294	233	0,0078
		201	0,0294	160	0,3529	236	0,0480

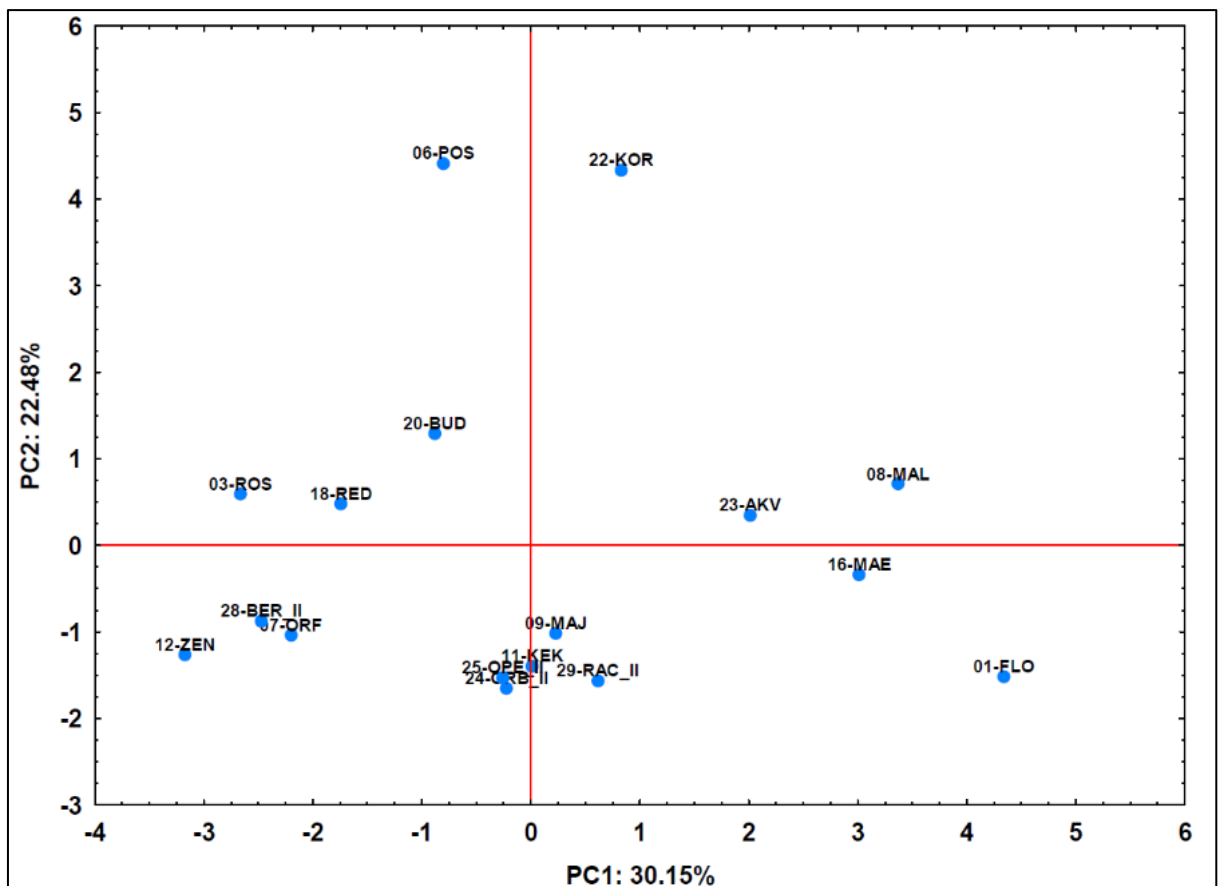
#### 5.4.2 Vícerozměrná analýza pro identifikaci jednotlivých odrůd

Předběžné testování rozlišovací schopnosti nových EST-SSR markerů bylo provedeno pomocí shlukové analýzy (CLU) a analýzy hlavních komponent (PCA). K analýze byl použit pouze dataset S17, jelikož PCA neumí pracovat s chybějícími hodnotami, a navíc by docházelo ke zkreslení rozlišovací schopnosti získaných markerů. Jak ukazují výsledky v podobě dendrogramu (Obrázek 24) a 2D grafu komponentního skóre (Obrázek 25), tak se podařilo jednoznačně odlišit od sebe všech 17 odrůd.

Obrázek 24 Dendrogram příbuznosti 17 odrůd



Obrázek 25 Shluková analýza 17 odrůd



## 6 Diskuze

### 6.1 Bioinformatická analýza

Jak ukázaly výsledky této práce, spojení bioinformatické analýzy sekvenčních dat spolu s laboratorním ověřením a testováním *de novo* navržených EST-SSR markerů, umožňuje nové markery vyvinout poměrně rychle a efektivně. V rámci experimentální části se začalo s ověřením již publikovaných genomických SSR markerů dle studií Celik *et al.* (2014) a Kati *et al.* (2013). Bohužel během testování na kapilární elektroforéze bylo zjištěno, že 5 z nich vytváří řadu nespecifických produktů a 9 z nich je monomorfních a tudíž nepoužitelných pro odlišení odrůd. Proto se dále pracovalo s 80 EST-SSR sekvencemi, pro které byly navrženy vhodné primery jako potenciální *de novo* markery, nakonec 8 markerů vykazovalo vysoký polymorfismus a reprodukovatelnost, což jsou nezbytné vlastnosti kvalitních markerů pro odlišení odrůd máku setého (*Papaver somniferum* L.). Úspěšnost EST-SSR při detekci a identifikaci máku setého a jemu příbuzných druhů uvádí také studie Şelale *et al.* (2013), která testovala 93 EST-SSR markerů na 37 genotypch a shledala 67 EST-SSR markerů vhodných k dalším analýzám.

Mnoho další autorů využívá právě EST-SSR pro popis a identifikaci řady rostlinných druhů (Chung *et al.*, 2000; Yu *et al.*, 2004; Poncet *et al.*, 2006; Lee *et al.*, 2011), hlavně díky rychlosti, efektivnosti a jednoduchosti této technologie. Velkou výhodou je i nižší nákladovost a dobrá dostupnost charakterizovaných EST sekvencí, kterých je u máku setého (*Papaver somniferum* L.) známo již 21 094 (NCBI, 2018).

Jak uvádí Varshney *et al.* (2005), EST-SSR markery vyvinuté pro jeden daný druh, mohou být dobře reprodukovatelné i pro jiné příbuzné druhy, přičemž tato přenosnost je v rámci genomických SSR markerů velmi omezená. Na druhou stranu genomické SSR markery vykazují vyšší míru polymorfismu, než EST-SSR a mají tedy lepší schopnost odlišit i genotypy s vysokou podobností. Nicméně pro genomické SSR je nezbytné disponovat dostatečným množstvím sekvenčních dat a jedná se o finančně náročnější techniku.



## 6.2 Optimalizace a testování navržených primerů

U navržených EST-SSR markerů proběhla testovací a optimalizační fáze poměrně rychle a jednoduše za použití simplex PCR, jejíž úspěšnost se dá přisuzovat kvalitně izolované DNA a použití speciálního typu reakční směsi tzv. kitu (nebo také master mixu). Výhodou takovýchto PCR mixů je zajištění správné koncentrace jednotlivých komponent reakční směsi, a tudíž snížení či eliminace chyb způsobených například pipetováním. Neopomenutelným kladem těchto reakčních směsí je vysoká robustnost systémů, způsobená hlavně používáním speciálních aditiv. V případě použitého kitu od Qiagenu se jedná o blíže nespecifikovaný „MP faktor“ a „Q-solution“, které usnadňují amplifikaci problematických úseků bohatých na GC páry a napomáhají stabilizovat jednotlivé primery. Další zlepšení představuje používání Hot-Start polymerázy (Dabney and Meyer, 2011). Hot Start polymeráza je modifikovaným typem *Taq* polymerázy, která má reverzibilně blokovanou enzymatickou aktivitu. V prvním cyklu PCR dochází k odblokování této aktivity a tím je zajištěno, že nedojde k nespecifické amplifikaci DNA během prvních několika cyklů, což je klíčové z hlediska kvality. Dříve, kdy nebyly dostupné podobné kity, bylo nutné přidávat polymerázu do reakční směsi samostatně, zpravidla v posledním kroku přípravy směsi pro PCR. Jelikož i za pokojové teploty docházelo k nespecifické amplifikaci díky aktivitě polymerázy, bylo nutné pracovat co nejrychleji a výsledek amplifikace byl nejistý. Z těchto hledisek se použitý kit osvědčil pro minimalizaci chyb a urychlení procesů analýzy navzdory vyšším finančním nákladům na jednu PCR reakci.

Během PCR v experimentální práci byla dále sledována a optimalizována specifita primerů při různých anelačních teplotách. Tato teplota se standardně pohybuje mezi 45–60°C. Pokud je ale teplota v PCR programu nastavena příliš vysoko, může se stát, že primery vůbec nenasednou k templátové DNA a amplifikace fragmentů neproběhne. Nebo naopak pokud by byla teplota nízká, než je daná pro určitý primer, mohlo by docházet k nespecifické amplifikaci nežádoucích fragmentů. Jako optimální  $T_a$  v této práci byla vyhodnocena teplota 57 nebo 60°C, s čímž se shoduje i studie Şelale *et al.* (2013), kde byla použita teplota 55-60°C.

V rámci optimalizace amplifikace bylo pro další testování vybraných mikrosatelitních markerů navrženo několik nových primerů. Konkrétně u EST26 a EST51 byly navrženy nové primery se speciální koncovou sekvencí („pig tail“), která usnadňuje přidávání netemplátového A polymerázou. Pokud během PCR reakce dojde k připojení netemplátového A pouze k části fragmentů, dochází ke komplikacím při hodnocení výsledků (Smith *et al.*, 1995), a proto se v mnoha laboratořích nastavuje v programu PCR delší finální elongace, aby se u všech vzniklých produktů došlo k přidání netemplátového A a docílilo se jednotných profilů

v genetickém analyzátorů (Butler, 2012). Clarke *et al* (2001) uvádí, že vznik netemplátového nukleotidu může také souviset se sklouznutím DNA polymerázy. Nejčastěji se jedná o adenin, ale Browstein *et al.* (1996) se během své práce setkal také s přidáváním netemplátového jiného nukleotidu, konkrétně cytosinu. Problém s netemplátovým nukleotidem opět eliminovat pomocí pigtail primerů. Dále Clarke *et al.* 2001 uvádí, že lze těmto chybám předcházet optimalizací PCR a použitých pufrů.

Některé laboratoře řeší tento problém používáním polymerázy, které neumí přidávat netemplátové A na konec sekvence. Mezi takovéto patří např. Hot Start Fusion II polymeráza nebo Herculase II Fusion polymeráza (Dabney and Meyer, 2012).

Použití pigtailových primerů vedlo zlepšení kvality výsledného profilu u EST-SSR markerů 26 a 51, které byly spolu s 6 dalšími (EST16, EST19, EST48c, EST53c, EST56b, EST61) vybrány jako nově vytvořené kvalitní EST-SSR markery.

V rámci lepší kvality a reprodukovatelnosti získaných výsledků byly úspěšně vytvořeny pro každý z 8 finální markerů alelické žebříky, které slouží k zjednodušení interpretace při porovnávání dat v jiných laboratořích, či při identifikaci genotypů. Vzhledem k omezeným časovým možnostem bylo osekvenováno pouze 6 alel lokusu EST48c, a proto bych doporučila v další krocích osekvenování ostatních alel nově navržených markerů. Dále by bylo vhodné vytvoření a optimalizace multiplexové PCR pro zjednodušení identifikace na větším množství genotypů.

## 7 Závěr

Závěry předložené diplomové práce vychází ze stanovených cílů práce, které se podařilo ve všech bodech splnit. Stručné shrnutí poznatků je uvedeno v následujících bodech:

- Byla provedena bioinformatická analýza dat pomocí genetických databází
- Následně byla testována funkce nově navržených EST-SSR markerů na získané skupině odrůd. A dále byla provedena optimalizace podmínek pro jejich úspěšnou amplifikaci.
- Podařilo se vybrat a optimalizovat 8 ESR-SSR markerů s přesně známou velikostí a frekvencí alel na 8 lokusech 17 odrůd máku setého.
- Posledním bodem ze stanovených cílů práce bylo otestování rozlišovacích schopností nově vytvořených EST-SSR markerů, které byly zhodnoceny jako účinným nástroj pro identifikaci, hodnocení a další studie odrůd máků setého, čímž byla potvrzena hypotéza předložené práce

Doporučení pro praxi a další výzkum

Zde popsání mikrosatelitní markery představují nově vyvinuté nástroje pro identifikaci a detekci odrůd máku setého, popř. aplikovatelné i na druhy příbuzné, proto je možné doporučit jejich využití v oblastech potravinářského, farmaceutického a technického průmyslu.

## 8 Literatura

- Adler, R.** 2013. Electroforesis. Salem Press Encyclopedia of Science. Research Startes. 5 str.
- Antiqueira, L.M.O.R.** 2013. Application of Microsatellite Molecular Markers in Studies of Genetic Diversity and Conservation of Plant Species of Cerrado. J. Plant Sci. 1. 1–5.
- Bara, I.I., Bara, C.I., Capraru, G., Truta, E.** 2007. The possible ways of speciation in Papaveracea family. Sectiunea Genetica si Biologie Moleculara. 8. 223-233
- Bennett, P.** 2000. Demystified ... Molecular Pathology. 53 (4). 177-183.
- Bonin, A., Bellemain, E., Eidesen, B., Pompanon, F., Brochmann, C., Taberlet, P.** 2004. How to track and assess genotyping errors in population genetics studies. Molecular Ecology. 13. 3261–3273.
- Botany.cz. Herbář** [online]. [cit. 2018-04-09]. Dostupné z <<https://botany.cz/cs/rubrika/herbar/p/page/4/>>
- Browstein, M.J., Carpten J.D., Smith, J.R.** 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Bio Techniques. 20 (6). 1004-6, 1008-10.
- Butler, J. M.** 2012. Advanced topics in forensic DNA typing: methodology. Elsevier/Academi Press. London. 680 s. ISBN: 012375136
- Butler, J.M., Buel, E., Crivellente, F., McCord, B.R.** 2014. Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. Electroforesis. 25. 1397–1412.
- Buschiazzo, E., Gemmell, N. J.** 2006. The rise, fall and renaissance od microsatellites in eukaryotic genomes. BioEssays. 28. 1040-1050.
- Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C., Sutherland. G. R.** 1993. Incidence and Origin of "Null" Alleles in the (CA)<sub>n</sub> Microsatellite Markers. Am. J. Hum. Genet. 52. 922-927.
- Carolan, J.C., Hook, I.L.I., Chase, M.W., Kdereit, J.W., Hodkinson, T.R.** 2006. Phylogenetics of Papaver and Related Genera Based on DNA Sequences from ITS Nuclear Ribosomal DNA and Plastid trnL Intron and trnL–F Intergenic Spacers. Annals of Botany. 98. 141-155.
- Carucci, F., Garramone, R., Aversano, R., Carputo, D.** 2017. SSR markers distinguish traditional Italian bean (*Phaseolus vulgaris* L.) landraces from Lamon. Czech J. Genet. Plant Breed., 53 (4). 168–171.

- Clarke, L.A., Rebelo, C.S., Goncalves, J., Boavida, M.G, Jordan, P.** 2001 PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences. *Mol Pathol.* 54(5). 351–353.
- Celik, I., Gultekin, V., Allmer, J., Doganlar, S., Frary, A.** 2014. Development of genomic simple sequence repeat markers in opium poppy by next-generation sequencing. *Molecular breeding.* 34 (2). 323-334.
- Conner, B.J., Reyes, A.A., Morin, CH., Itakura, K., Teplitz, R.L., Wallance, R.B.** 1983. Detection of sickle cell beta globin allele by hybridization with synthetic oligonucleotides. *Proc. Natl.Acad.Sci.* 80. 278-282.
- Cvrčková, H., Máchová, P., Dostál, J., Malá, J.** 2013. Hodnocení genetické diverzity vybraných populací smrku ztepilého pomocí mikrosatelitových markerů. *Zprávy lesnického výzkumu.* 58 (3). 273–279.
- Dabney, J., Meyer, M.** 2012. Length and GC-biases during sequencing library amplification: A comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechnologies.* 52 87-94.
- Dakin, E. E., Avise, J. C.** 2004: Microsatellite null alleles in parentage analysis. *Heredity* 93: 504-509.
- Dolník, V.** 1999. DNA sequencing by capillary electrophoresis (review). *J. Biochem. Biophys. Methods.* 41. 103–119.
- Ellegren, H.** 2004. Microdatellites: Simple sequence with complex evolution. *Nature Review Genetics.* 5. 435-444
- Estoup A., Jarne, P., Cornuet, J. M.** 2002. Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology.* 11 (9). 1591–1604.
- Estoup, A., Tailliez, C., Cornuet, J.M., Solignac, M.** 1995. Size homoplasmy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae ). *Molecular Biology and Evolution.* 12 (6). 1074–1084.
- Espinasse, A., Dosba, F.** 1982. Cytological analysis of hybrids between *Papaver somniferum* L. and *Papaver bracteatum* Lindl.; phylogenetic relationships between the two species. *Agronomie.* 2 (3). 281–286.
- Fér, T.** RAPD I. - princip metody [online]. DNA Laboratoř katedry botaniky PřF

UK v Praze. 23. ledna 2009. [cit. 2018-04-09]. Dostupné z <[https://botany.natur.cuni.cz/dna/index.php?option=com\\_content&view=article&id=48:rapd-princip&catid=34:rapd&Itemid=56](https://botany.natur.cuni.cz/dna/index.php?option=com_content&view=article&id=48:rapd-princip&catid=34:rapd&Itemid=56)>

**Galido, C.L., McIver, L.J., McCormick, Skinner, M.A., Xie, Y., Gelhausen, R.A., Ng, K., Kumar, N.M., Garner, H.R.** 2009. Global Microsatellite Content Distinguishes Humans, Primates, Animals, and Plants. *Mol.Bio.Evol.* 26 (12). 2809-2819.

**Gao, H., Jiang, K., Geng, Y., Chen, X.-Y.** 2012. Development of microsatellite primers of the largest seagrass, *Enhalus acoroides* (Hydrocharitaceae). *Am. J. Bot.* 99. 99–101.

**Garrido-Cardenas, J.A., Mesa-Valle, C., Manzano – Agugliaro, F.** 2018. Trends in plant research using molecular markers. *Planta.* 247 (3). 543–557.

**Genetic diversity.** Satellite DNA [online]. Fossil Hunters. 18. března 2018. [cit. 2018-04-09]. Dostupné z <<https://www.fossilhunters.xyz/genetic-diversity/satellite-dna.html>>

**Gugerli, F., Brodbeck, S., Holderegger, R.** 2008. Insertions–Deletions in a Microsatellite Flanking Region May Be Resolved by Variation in Stuttering Patterns. *Plant Mol Biol Rep.* 26. 255–262.

**Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoitevin, C., Malausa, T., Revardel, E., Salin, F., Petit, R. J.** 2011. Current trends in microsatellite genotyping. *Molecular ecology resources.* 11 (4). 591-611.

**Griffiths,** Analyse of DNA by the Southern Blot technique. Newfoundland and Labrador's University. 2014. [cit. 2018-04-09]. Dostupné z <<https://www.mun.ca/biology/scarr/Gr12-18.html>>

**Heller, Ch.** 2001. Principles of DNA separation with capillary electrophoresis. *Electrophoresis.* 22. 629–643.

**Hoffman, J.I., Amos, W.** 2005. Microsatellite genotyping errors: detection approaches common sources and consequences for paternal exclusion. *Molecular Ecology.* 14. 599–612.

**Chapuis, M-P., Estoup, A.** 2007. Microsatellite Null Alleles and Estimation of Population Differentiation. *Molecular Biology and Evolution.* 24 (3). 621-631.

**Chmelová, D., Ondrejovič, M., Havrlentová, M., Kraic, J.** 2018. Evaluation of polar polyphenols with antioxidant activities in *Papaver somniferum* L. *Journal of Food and Nutrition Research.* 51 (1). 98–107.

**Chung, A. M., Staub, J. E., J. F. Chen.** 2006. Molecular phylogeny of *Cucumis* species as revealed by consensus chloroplast SSR marker length and sequence variation. *Genome* 49, 219—229.

- Jiang, G.** 2018. Encyclopedia of Applied Plant Sciences (Second Edition). Reference Models in Life Sciences. 2. 207-214.
- Kalia, R.K., Rai, M.K., Kalia, S., Singh, R., Dhawan, A.** 2011. Microsatellite markers: An overview of the recent progress in plants. *Euphytica*, 177. 309–334.
- Kalinowski, S. T., Taper, M. L., Marshall, T. C.** 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16 (5), 1099-1106.
- Kati, V. Le Corre, V., Michel, S., Jaffrelo, L., Poncet, Ch.** 2013. Isolation and Characterisation of 11 Polymorphic Microsatellite Markers in *Papaver rhoeas* L. (Corn Poppy), a Major Annual Plant Species from Cultivated Areas. *Int. J. Mol. Sci.* 14. 470–479.
- Kašička, V.** 1997. Teoretické základy a separační principy kapilárních elektromigračních metod. *Chemické listy*. 91. 320-329.
- Kösoglu, K., Yumuk, S., Aydin, Y., Evcı, G., Uncuoglu, A.A.** 2017. Use of SNP markers by KASP assay for MAS studies in sunflower against *Plasmopara halstedii*. *Tübitak*. 41. 480-489.
- Křenek.** Laboratoř růstových regulátorů [online]. Univerzita Palackého v Olomouci & Ústav experimentální botaniky AV ČR, v.v.i. [cit. 2018-04-09]. Dostupné z <<https://www.ncbi.nlm.nih.gov/>>.
- Lateef, D. D.** 2015. DNA Marker Technologies in Plants and Applications of Crop Improvements. *Journal of Biosciences and Medicines*. 3 (5). 7-18.
- Lehmann, T., Hawley, W. A., Collins, F. H.** 1996. A Evolution of Evolutionary Constraints on Microsatellite Loci Using Null Alleles. *Genetics Society of America*. 144. 1155-1163.
- Lee, P.Y., Costumbrado, J., Hsu, C.** 2012. Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp*. 20 (62). 3923.
- Lee, E. J., G. N. Jin, K. L. Lee, S. M. Han, Y. H. Lee, Yang, M. S.** 2011: Exploiting expressed sequence tag databases for the development and characterization of gene-derived simple sequence repeat markers in the opium poppy (*Papaver somniferum* L.) for forensic applications. *J. Forensic Sci.* 56, 1131—1135.
- Li, H., Liang, X.-Y., Zou, S.-J., Liu, S., De Clercq, P., Ślipinski, A., Pang, H.** 2017. New EST-SSR markers reveal strong genetic differentiation in native and introduced populations of the mealybug destroyer *Cryptolaemus montrouzieri*. *Biological Control*, 109. 21-26.
- Li, Y. C., Korol, A. B., Fahima, T., Nevo, E.** 2004: Microsatellites within genes: structure, function and evolution. *Molecular Biology and Evolution* 21: 991-1007.

- Liepelt, S., Kuhlenkamp, V., Andzidei, M., Vendrmin, G.G., Ziegenhagen, B.** 2001. Piafalls in determining sice homoplasly of microsatellite loci. *Molecular Ecology Notes*. 1. 332–335.
- Liu, S., Feuersteun, U., Leusink, W., Schulze, S., Aasp, T., Studer, B., Becker, H.C., Dehmer, K.J.** 2018. DArT, SNP, and SSR analyses of genetic diversity in *Lolium perenne* L. using bulk sampling. *BMC Genetics*. 19 (10). 1-13.
- Metzgar, D., Bytof, J., Wills, C.** 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res*. 10, 72–80.
- Morgante, M., Hanafey, M., Powell, W.** 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Geneticts*. 30. 194-200.
- Mishra, B.K., Mishra, R., Jena, S.N., Shukla, S.** 2016. Gene actions for yield and its attributes and their implications in the inheritance pattern over three generations in opium poppy (*Papaver somniferum* L.). *Journal od Genetics*. 95 (3). 705–717.
- National Center for Biotechnology Information (NCBI)** Biomolecular and biomedical literature databases [online]. [cit. 2018-04-09]. Dostupné z <<https://www.ncbi.nlm.nih.gov/>>.
- Nei, M., Tajima, F., Tateno, Y.** 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, 19 (2), 153-170.
- Oliviera, E. J., Pádua, J. G., Zucchi, M. I., Vencovsky, R., Viéra, M. L. C.** 2006. Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology*. 29 (2). 294-307.
- Parmakinsiz, I., Özcan, S.** 2011. Morpfological, chemical, and molecular analyses of Turkish *Papaver* accessions (Sect. *Oxytona*). *Tübitak*. 35. 1-16.
- Pompanon, F., Bonin, A., Bellemain, E., Taberlet, P.** 2005. Genotyping errors: cause, consequences and solutions. *Nature Reviews Genetics*. 6. 847–859.
- Poncet, V., Rondeau, M, Tranchant, C, Cayrel, A., Hamon, S. Kochko, A., Hamon, P.** 2006. SSR mining in coffee tree EST databases: potential use of EST–SSRs as markers for the *Coffea* genus. *Mol. Gen. Genomics* 276, 436—449.
- Powel, W., Machráy, G.C. Provan, J.** 1996. Polymorphism reveal by simple sequence repeats. *Trends in plant science*. 1 (7). 215-222.
- Rozen, S., Skaletsky, H. J.** 2000. Primer3 on the WWW for general users and for biologist programmers. In: **Krawetz, S., Misener, S.** (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, Humana Press, Totowa, NJ, 365-386. Dostupne na <<http://frodo.wi.mit.edu/primer3/>>.



- Sanger, F., Nicklen, S., Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci USA. 74 (12). 5463–5467.
- Sangwan, R. S., Yadav, U., Sangwan, N. S.** 2000. Isolation of Genomic DNA from Defatted Oil Seed Residue of Opium Poppy (*Papaver somniferum*). Plant Molecular Biology Reporter. 18. 25-270.
- Şelale, H., Çelik, İ., Gültekin, V., Allmer, J., Doğanlar, S., Frary, A.** 2013. Development of EST-SSR markers for diversity and breeding studies in opium poppy. Plant Breeding. 132 (3). 344-351.
- Serwer, P.** 1983. Agarose gel: Properties and use for electrophoresis. Verlag Chemie GmbH. 4. 375–382.
- Schlötterer, C.** 2004. The evolution of molecular markers – just a matter of fashion. Nature Reviews Genetics. 5. 63–69.
- Smith, R.J., Carpten, J.D., Brown, M.J, Qhosh, S., Magnusson, V.L., Gilbert, D.A., Trent, J.M., Collins, F.S.** 1995. Approach to Genotyping Errors Caused by Nontemplated Nucleotide Addition by *Taq* DNA Polymerase. Genome Research. 5. 312-317.
- Sobel, E., Papp, J.C., Lange, K.** 2005. Detection and Integration of Genotyping Errors in Statistical Genetics. American Society of Human Genetics. 70. 496-508.
- Storfer, A., Easman, J.M., Spear, S. F.** 2017. Modern Molecular Methods for Amphibian Conservation. BioScience. 59 (7). 559–571.
- Taheri, S., Abdullah, T.L., Yusop, M.R., Hanafi, M.M., Sahebi, M., Azizi, P., Shamshiri, R.** 2018. Mining and Development of Novel SSR Markers Using Next Generation Sequencing (NGS) Data in Plants. Molecules. 23. 399.
- Tétényi, P.** 1997. Opium Poppy (*Papaver somniferum*): Botany and Horticulture. Research Institute for Medicinal Plants. 19. 373-408.
- TIBCO Software Inc.** (2017). Statistica (data analysis software system), version 13. <http://statistica.io>.
- Varshney, R. K., Graner, A., Sorrells, M. E.** 2005. Genic microsatellite markers in plants: features and applications. Trends in Biotechnology. 23 (1). 48-55.
- Verma, N., Jena, S. N., Shukla, S., Yadav, Kusum, Y.** 2016. Genetic diversity, population structure and marker trait associations for alkaloids content and licit opium yield in India-wide collection of poppy (*Papaver somniferum* L.). Plant Gene 7. 26-41.
- Vieira, M.L.C., Santini, L., Diniz, A.L., Munhoz, C.D.F.** 2016. Microsatellite markers: What they mean and why they are so useful. Genet. Mol. Biol. 2016. 39. 312–328.

- Wang, X., Wang, Le.** 2016. Gmata: An Intergrated Software Packeage for Genome-Scale SSR Mining, Marker Development and Viewing. *Frontiers in Plant Science*. 7. 1350.
- Wagner, A. P., Creel, S., Kalinowski, S. T.** 2006. Estimating relatedness and relationships using microsatellite loci with null alleles. *Heredity*. 97. 336-345.
- Yilmaz, M., Ozic, C, Gok, I.** 2012. Principles of Nucleic Acid Separation by Agarose Gel Electrophoresis. [www.intechopen.com](http://www.intechopen.com).
- Yu, J. K., Dake, T. M., Singh, S., Benschler, D., Li, W., Gill, B., Sorrells M. E.** 2004. Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. *Genome* 47, 805—818.
- Zarrintaj, P., Bakhshandeh, B. Rezaeian, Heshmatian, B., Ganjali, M.R.** 2017. Pentamer Platform as a Potential Candidate for Neural Tissue Engineering. *Scientific Reports*. 7(1):17187
- Zargar, S.M., Raatz, B., Sonah, H., Bhat, J.A., Dar, Z.A., Agrawal, G.K., Rakwal, R.** 2015. Recent advances in molecular marker techniques: Insight into QTL mapping, GWAS and genomic selection in plants. *J. Crop Sci. Biotechnol.* 18. 293–308.
- Zeng, P., Zhou, X., Huang, S.** 2017. Prediction of gene expression with cis-SNPs using mixed models and regularization methods. *BMC Genomics*. 18. 368-379.

## 9 Seznam použitých zkratk

AFLP = Amplified Fragment Length Polymorphism (délkový polymorfismus amplifikovaných fragmentů)

bp = base pair (párů bází)

DNA = deoxyribonucleic acid (Deoxyribonukleová kyselina)

EST = Expressed Sequence Tags (exprimované sekvenční úseky)

EST-SSR = Expressed Sequence Tags-Simple Sequence Repeats (mikrosatelitní exprimované úseky)

gSSR = genomic SSR (genomický mikrosatelit)

kb = kilobase (kilobáze)

MMR = Mismatch Repair (system) (opravný systém)

NCBI = National Center for Biotechnology Information

nt = nucleotide (nukleotid)

PCR = Polymerase Chain Reaction (Polymerázová řetězová reakce)

RAPD = Randomly Amplified Polymorphic DNA (náhodná amplifikace polymorfnní DNA)

RFLP = Restriction Fragment Length Polymorphism (polymorfismus délky restrikčních fragmentů)

SNP = Single Nucleotide Polymorphism (jednonukleotidový polymorfismus)

SSR = Simple Sequence Repeat (sekvenčně jednoduché repetice)

STR = Short Tandem Repeat (krátká tandemová repetice)

T<sub>a</sub> = annealing temperature (teplota přisedání primerů)

UTR = Untranslated region (3' či 5' nepřekládaná oblast)

## 10 Přílohy

Tabulka 12 Navržené primery EST01-40

Marker	Motiv	Repetice	Velikost	F primer (5' - 3')	R primer (5' - 3')
EST01	TAC	17	178 bp	AACCGTAAACCGTAAACCGTAA	AACGAACGAACGAACGAAAC
EST02	TCG	19	314 bp	CCTTACCCCTTACCCTTTACCC	GAACGACGAACGAACGAAC
EST03	TAC	20	315 bp	CCATTGTTCTGCCTTCTTCA	CGGTAAACGGTAAACGGTAAA
EST04	ACT	20	319 bp	AACCCGTAAACCCGTAAACC	GGTAGGTAGTGTGTGTGAGAGAGAG
EST05	CGT	15	286 bp	CCGTACCTACCTACCTACCTACC	CGGTAAACGTAAACGAACGAA
EST06	TCT	15	296 bp	GAAAGACCCCTCCTCCTCT	CCTTCTCGAGCTTTGTTTG
EST07	ACT	25	221 bp	ACTACTAACACACACACACACT	CGGTAAACGGTAAACGTAGTAGGTAG
EST08	AGA	15	199 bp	GATGGAGAAGAGAGAGGTGAGG	TTCGTTCTTCTTCTTCGTTTC
EST09	TAG	19	274 bp	CTACACTTCCCATCAACATCTTC	AGAGAGAGAGAGAGAGAGACGAC
EST10	TTC	15	188 bp	CCTAAGCAAATCAAACAACCTGC	GGATGAGCAAGAGAGCAAGG
EST11	TTC	17	300 bp	AACCCTGCTAACATCCCTCA	CTACTCCCTCGTGTGCTTGG
EST12	ATA	16	262 bp	CAAAAGTGAATTTGGACAGG	CGGCTTAAACCCTTTTCGTT
EST13	TTC	19	258 bp	TCAACGACTGATGGCTGTTC	ACGTTACGGGAAAGGGAAAC
EST14	TCT	27	285 bp	AAGACCCTCCTCCTCCTCT	CCTTAACCTTACCCTTCTTACG
EST15	ATT	23	311 bp	TCTCCTTCTTGTCTTCTTCTTC	CTCGTGTGCCTTGGCAGT
EST16	TCT	15	283 bp	CATCCAATCACCAACACAG	TGGCATCCAACCAAGTTACA
EST17	TAG	15	303 bp	TGTGGTCTGAGGATGATGTT	CTCTCTCTCTCTGTCGTCGT
EST18	CTA	20	237 bp	CACACACACAATAACACAACACA	GAGAAGAAGAAGAAGAACGAACG
EST19	CAT	15	227 bp	TTTCTTCTCCCGTTCTTCA	CGGTGATGATGATGGTTGTT
EST20	TCT	17	248 bp	AATCCCAGTCGTGGTGGTAG	TGCCTTGGCAGTCTCAGTAA
EST21	TTC	23	310 bp	CGCTATATCCTCCGCAATCA	CTATCCCTCGTGTGCCTTG
EST22	TCT	17	281 bp	CTTCATAAGCATGCGACCA	CCTTGGCAGTCTCAGCAGTT
EST23	TCT	18	313 bp	ATTTCCCAAGCTGTCTTCC	CGTAACCAACGACTTACTTAGGTTT
EST24	CTT	22	315 bp	GCCTCCAAATCTCTTTTCC	TGTCGTTCTCTCTCCATCTT
EST25	GTA	18	241 bp	CTCCTCCGCAATTGTTGTT	CCGTTTACCGTTTACCGTTT
EST26	TCT	15	132 bp	GTGAGGAGGACGAGCTTTTG	CCGTTGTAATAACCGACTGC
EST27	TTC	15	136 bp	CGAAGCTGAAGCCATTATCTG	CTCGTGTGCCTTGGCAGT
EST28	TCT	17	300 bp	ATTGTGGGGTCTTTGTGAA	ACCTACCGAAACCGACTTAA
EST29	TCT	21	178 bp	TTGATTTTCGAGTGGAAATG	TGGCAGTCTCAAGGAAACAA
EST30	TTC	20	111 bp	TCTTCTCGCTGTTCTCGATG	TATACCCTCGTGTGCCTTGG
EST31	CTT	22	280 bp	GCCTCCAAATCTCTTTTCC	TTTTCTTCGTACGCCAAACC
EST32	TTC	15	214 bp	TGAATGAGATTAGCGTCGTTG	CTTGGCAGTCTCAGTTCAG
EST33	ACG	15	281 bp	TCGTCAATCCAGATGCAAAA	TGTGTGTGTTACGTTTCGTCGT
EST34	TTC	15	190 bp	AAGCAGGTCTATCATCCACACTC	CGTACGAAACGACCGTACCT
EST35	AGT	26	309 bp	ACACCACCACCAACAGACAA	CAACGGACTAACGGAACACC
EST36	TCT	13	116 bp	GCAGCAGTACCACCAAATCC	TTGTGAACGGAGGAGTCTT
EST37	ATT	12	236 bp	GGTGGTTTTCTTTCCATT	AGCAGCAGCAGCAACAACATA
EST38	ATC	12	213 bp	GCCACCATCTTTCATCAC	GGGTAAGGGGTTAGGGGTAA
EST39	TTC	12	247 bp	TTGACACTCCACGCAAGAAA	TACCCACTTGTGCTGCTTGA
EST40	CGT	12	209 bp	TAACGAAGGACGGAGGGTAG	GGAACGGACGACGACTAAAC

Tabulka 13 Navržené primery EST01-40

<b>EST41</b>	<b>GAT</b>	<b>13</b>	<b>261 bp</b>	<b>ATTGATGGCCAGTGAGTTC</b>	<b>TTTCAACTTTCACCCCTCTCA</b>
<b>EST42</b>	<b>AAC</b>	<b>12</b>	<b>190 bp</b>	<b>TCATCTTCAATGGCGAGAGA</b>	<b>AGATTTCGTCGGTTCAACACC</b>
<b>EST43</b>	<b>TCT</b>	<b>12</b>	<b>171 bp</b>	<b>GGCTGCTGCTTCTTTCATCT</b>	<b>CTATCCCCTGTGTGCCTTG</b>
<b>EST44</b>	<b>CTT</b>	<b>14</b>	<b>280 bp</b>	<b>ATTTCTCCTCCCGCTAAGA</b>	<b>CGAGTACGTTACACGGGAGTT</b>
<b>EST45</b>	<b>CTT</b>	<b>14</b>	<b>106 bp</b>	<b>ACAAACTCCGCGAGTCAAT</b>	<b>GAGACGGCTTTGTGATGGAT</b>
<b>EST46</b>	<b>TGA</b>	<b>13</b>	<b>293 bp</b>	<b>TTGTGGATGTGATGGGCTAA</b>	<b>GTCCTCTCCAGTCTCCACCA</b>
<b>EST47</b>	<b>CTT</b>	<b>13</b>	<b>238 bp</b>	<b>CCAACCCCTGTTGTTTCATT</b>	<b>TTCTCATTCCCAGAAATCACC</b>
<b>EST48</b>	<b>TCT</b>	<b>12</b>	<b>274 bp</b>	<b>GCTCTGACTCGTGAGAAGCTA</b>	<b>GATGACGCATTGCCAAGATT</b>
<b>EST49</b>	<b>TTC</b>	<b>12</b>	<b>296 bp</b>	<b>TGGAGTGGGACAAACAGAGA</b>	<b>TGTTTGAAATCCACCACCAA</b>
<b>EST50</b>	<b>TTC</b>	<b>12</b>	<b>279 bp</b>	<b>CCTAAGCAAATCAAACAACGTC</b>	<b>TTGTCTCTCCCGTTTCATCC</b>
<b>EST51</b>	<b>TTA</b>	<b>14</b>	<b>225 bp</b>	<b>CCCAATTGTGGTTCTTTTG</b>	<b>CGACATGATGCCAATTCAAG</b>
<b>EST52</b>	<b>CTT</b>	<b>12</b>	<b>217 bp</b>	<b>AAGCTCTTCGACTCACTGC</b>	<b>AGGAAAGGTGCCGTCTAATG</b>
<b>EST53</b>	<b>TTC</b>	<b>12</b>	<b>149 bp</b>	<b>TCAATACCCACAAAAGGAGGA</b>	<b>CATCAGCAAATGCCATGAGT</b>
<b>EST54</b>	<b>CAT</b>	<b>13</b>	<b>289 bp</b>	<b>CATCCATCCATCCATCCATC</b>	<b>AAACGTTGCGAAAACCGTTGA</b>
<b>EST55</b>	<b>TCA</b>	<b>13</b>	<b>277 bp</b>	<b>CATCCATCCATCCATCATCA</b>	<b>GTTTCGAAAACCGTTGAACCT</b>
<b>EST56</b>	<b>TTC</b>	<b>13</b>	<b>240 bp</b>	<b>GGCTGCTGCTTCTTTTCATC</b>	<b>GGCCCACTATAGGCAAACGTC</b>
<b>EST57</b>	<b>CTT</b>	<b>12</b>	<b>286 bp</b>	<b>GAGACTGAGGCGATTACGC</b>	<b>AGTTCGAGGAGGCCGTTATT</b>
<b>EST58</b>	<b>CTT</b>	<b>12</b>	<b>283 bp</b>	<b>TCAACTCCACTCCATCA</b>	<b>CATCCCCAGCTGTACTGGAT</b>
<b>EST59</b>	<b>TTG</b>	<b>14</b>	<b>260 bp</b>	<b>TGGGGTTTGATTGCTAAAAGG</b>	<b>AAACCAACAAGATCACAAGC</b>
<b>EST60</b>	<b>AGT</b>	<b>12</b>	<b>259 bp</b>	<b>TCAGTTCGCTTTTGTTCCTTC</b>	<b>CGTTCGTTCTTTACTTCTTTTACG</b>
<b>EST61</b>	<b>TTC</b>	<b>12</b>	<b>235 bp</b>	<b>GGCTGCTGCTTCTTTTCATC</b>	<b>ATAGGGCAAACGCTGCTA</b>
<b>EST62</b>	<b>AAT</b>	<b>12</b>	<b>270 bp</b>	<b>GGCAGTTAAACGAACCTGGA</b>	<b>ATCAAACACCAACACGCTCA</b>
<b>EST63</b>	<b>CGT</b>	<b>13</b>	<b>258 bp</b>	<b>CAAATACATTCCTCCGCTCTG</b>	<b>ACGACGACGACGAGAGAGA</b>
<b>EST64</b>	<b>TCT</b>	<b>13</b>	<b>265 bp</b>	<b>CACAGCTTTTCAAGCACCAA</b>	<b>AACCAACCAACGCAAAGAAC</b>
<b>EST65</b>	<b>CTT</b>	<b>12</b>	<b>287 bp</b>	<b>GAGACTGAGGCGATTACGC</b>	<b>GGAGTTCAGGAGGCCGTTAT</b>
<b>EST66</b>	<b>CTT</b>	<b>14</b>	<b>135 bp</b>	<b>GCACCAATCAATCACCACCT</b>	<b>TTCATGATTCCGGCTGTTCTG</b>
<b>EST67</b>	<b>TTC</b>	<b>14</b>	<b>242 bp</b>	<b>GGATATTTTGGGGTTACAGTTG</b>	<b>GATTGCCTCTGTCGTCCATT</b>
<b>EST68</b>	<b>TTC</b>	<b>12</b>	<b>295 bp</b>	<b>GGAGTGGGGACAAACAGAGA</b>	<b>CCTGTTAGAAATCCACCACCA</b>
<b>EST69</b>	<b>CCA</b>	<b>12</b>	<b>178 bp</b>	<b>TGGAGAATCCAGAAGCACAA</b>	<b>ATGAAGGGCAAACGAACAC</b>
<b>EST70</b>	<b>CTT</b>	<b>14</b>	<b>170 bp</b>	<b>TTGAGCTTTCGGGTACCATC</b>	<b>AATTGGAGTGTGGGAACAGC</b>
<b>EST71</b>	<b>TTC</b>	<b>12</b>	<b>198 bp</b>	<b>AACCCCTCTAAGGATGTGGA</b>	<b>TACCCACTGTGTGCTTGA</b>
<b>EST72</b>	<b>CTA</b>	<b>13</b>	<b>144 bp</b>	<b>CGTTACGTTTCGTTCTTCTTCTTC</b>	<b>GTAGTGTGGTGGTGGTGGTG</b>
<b>EST73</b>	<b>ATT</b>	<b>13</b>	<b>231 bp</b>	<b>CCCAGGTGAAACTTCTCTG</b>	<b>CTATCCCCTGTGTGCCTTG</b>
<b>EST74</b>	<b>TCT</b>	<b>12</b>	<b>236 bp</b>	<b>GAATACGATGGCGTTTCACC</b>	<b>TGGCAGTCTAGATCTCCTCCTC</b>
<b>EST75</b>	<b>GTA</b>	<b>12</b>	<b>137 bp</b>	<b>GTCGTTGATGTTACGCGATG</b>	<b>CCTCCTCTCTCTCTCTCTCTCG</b>
<b>EST76</b>	<b>TTC</b>	<b>12</b>	<b>258 bp</b>	<b>CCTTCTCCCAACTCACAAA</b>	<b>TGGCAGTCTCAGAAACCCTAA</b>
<b>EST77</b>	<b>TTC</b>	<b>14</b>	<b>145 bp</b>	<b>GGCTGCTGCTTCTTTCATCT</b>	<b>CCTATCCCCTGTGTGCCTTG</b>
<b>EST78</b>	<b>TCT</b>	<b>12</b>	<b>244 bp</b>	<b>GAATACGATGGCGTTTCACC</b>	<b>CGTGTGCCTTGGCAGTCTA</b>
<b>EST79</b>	<b>TCA</b>	<b>12</b>	<b>186 bp</b>	<b>TCAGACTTGCCACTGAAAGC</b>	<b>CCTCTCGGATCCAGGTAGTG</b>
<b>EST80</b>	<b>TCT</b>	<b>14</b>	<b>234 bp</b>	<b>GCACATGTTCTTAGTGTGATGGA</b>	<b>GGTTACGTATACTGTCCCCTTG</b>