

**Jihočeská univerzita v Českých Budějovicích**  
**Přírodovědecká fakulta**

**Příběh hladové housenky:**  
**anotace chuťových receptorů předivky**  
***Yponomeuta evonymella***

Bakalářská práce

**Kseniya Bobryshava**

Školitelka: RNDr. Martina Dalíková, Ph.D.

České Budějovice 2023

Bobryshava, K., 2023: Příběh hladové housenky: anotace chuťových receptorů předivky *Yponomeuta evonymella*. [The story of the hungry caterpillar: annotation of the gustatory receptors of the small ermine moth *Yponomeuta evonymella*. Bc. Thesis in Czech] – 48 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

### **Annotation**

Small ermine moths are important models for studying speciation and adaptation. *Yponomeuta* species underwent host shift which might be associated with changes in genes for sugar receptors. The present thesis focuses on sequencing and annotation of genes for sugar receptors in *Yponomeuta evonymella*. Although two sequencing approaches were used only a partial sequence of the region coding sugar receptors was retrieved. However, our results suggest interesting changes in these genes in *Y. evonymella*, such as putative recent duplication.

### **Finanční podpora**

Tento projekt byl financován grantem Grantové agentury České republiky reg. Č. 20-20650Y a grantem Studentské grantové agentury PřF JU (SGA 2022).

### **Prohlášení**

Prohlašuji, že jsem autorem této kvalifikační práce a že jsem ji vypracovala pouze s použitím pramenů a literatury uvedených v seznamu použitých zdrojů.

České Budějovice, 13. dubna, 2023

.....

Kseniya Bobryshava

## **Poděkování**

Nejprve bych chtěla poděkovat mojí školitelce Martině Dalíkové za cenné rady, odborné vedení mé práce a za to, že se vždy snažila najít na této práci něco, za co by mě pochválila. Velký dík patří také celé naší laboratoři, a především našemu vedoucímu za vytváření příjemné a přátelské atmosféry.

# Obsah

1. Úvod	1
1.1. Receptory	1
1.2. Chuťové receptory hmyzu	1
1.2.1. Evoluce chuťových receptorů	3
1.2.2. Typy chuťových receptorů a jejich funkce	4
1.3. Cukerné receptory	5
1.3.1. Stavba cukerných receptorů	6
1.3.2. Struktura genů pro cukerné receptory	6
1.3.3. Specifita cukerných receptorů	7
1.4. <i>Yponomeuta evonymella</i>	7
1.5. Historický přehled sekvenování	9
1.5.1. Sekvenování první generace	9
1.5.2. Sekvenování druhé generace	10
1.5.3. Sekvenování třetí generace	11
2. Cíle práce	12
3. Materiál a metody	13
3.1. Složení genomu <i>Y. evonymella</i>	13
3.2. Anotace genů Gr	14
3.3. Prohledávání (screening) BAC knihovny	14
3.3.1. BAC knihovna	14
3.3.2. Primer test	14
3.3.3. PCR I	15
3.3.4. PCR II	15
3.4. Izolace BAC DNA a ověření klonu	16
3.5. Chromosome walking	17
3.6. Úprava BAC DNA	17
3.7. Příprava knihovny a Oxford Nanopore sekvenování	17
3.8. Basecalling, Nanoplot, Nanofilt	18
3.9. Skládání sekvence BAC klonů a mapování genů	18

4.	Výsledky	20
4.1.	Skládání genomu <i>Y. evonymella</i>	20
4.2.	Mapování genů <i>P. xylostella</i> na genom <i>Y. evonymella</i>	21
4.3.	Prohledávání BAC knihovny	22
4.4.	Příprava BAC DNA pro ONT sekvenování	25
4.5.	ONT	26
4.6.	Mapování genů pro cukerné receptory	28
5.	Diskuze	31
6.	Závěr	35
7.	Literatura	36
8.	Přílohy	44

# 1. Úvod

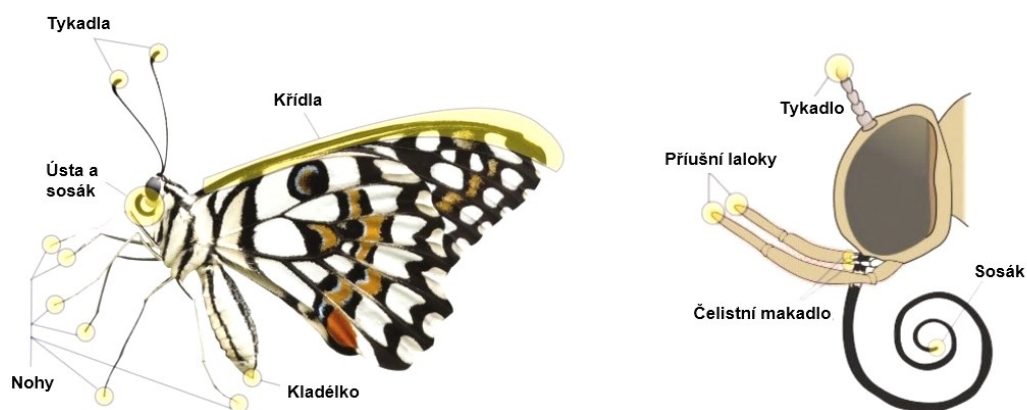
## 1.1. Receptory

Každý organismus vnímá svoje vnitřní a vnější prostředí pomocí receptorů, které jsou součástí tzv. smyslového ústrojí. To si můžeme představit jako složitý stroj, který přijímá informace z okolí a plní specifické funkce. Taková soustava může reagovat na mechanické, světelné, tepelné nebo chemické podněty, dle kterých dělíme receptory živočichů na čtyři skupiny mechano-, foto-, termo- a chemoreceptory (Vácha a kol, 2004).

Kontaktní (chuťová) a distanční (čichová) chemorecepce je důležitým smyslem pro všechny organismy (Vácha a kol, 2004). Chuť je komplexní chemosenzorický smysl určený k analýze potravy, která je nezbytná pro přežití. Kontaktní chemoreceptory savců umožňují rozlišovat pět chutí: sladká, kyselá, slaná, hořká a umami (vnímání kyseliny glutamové nebo jejích solí). Sladká chuť a umami jsou obecně považovány za „dobré“ a vyvolávají pozitivní reakci jedinců, opakem jsou například hořké a kyselé chutě, které nás varují před toxickými látkami nebo látkami se sníženým pH (Yarmolinsky a kol., 2009).

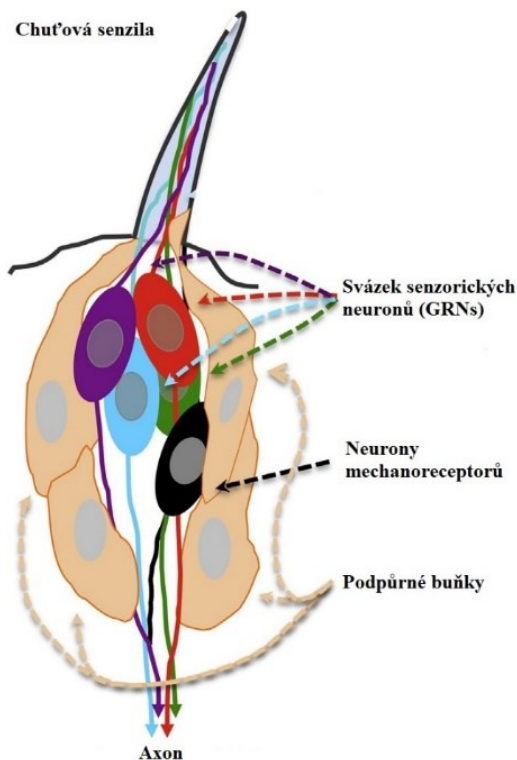
## 1.2. Chuťové receptory hmyzu

Na rozdíl od savců, kteří mají chuťové receptory specificky umístěny v chuťových pohárcích na jazyku a měkkém patře, hmyz má mnohem sofistikovanější chemosenzorický systém. Navzdory tomu, že obratlovci a hmyz si nezávisle na sobě vyvinuli odlišné anatomické a molekulární dráhy pro chuťové vjemy, existují jasné paralely v organizaci a kódování genů pro chuťové receptory (Yarmolinsky a kol., 2009). Hlavní podobností mezi savcím a hmyzím chuťovým systémem je to, že oba fungují na principu vyvolání specifických behaviorálních reakcí (Gordon & Scott, 2009). Na rozdíl od savců chuťové receptory hmyzu ovlivňují nejenom stravovací návyky jedince, ale také jeho rozmnožovací chování (Scott a kol., 2001). Oproti savcům je hmyz mnohem menší než jeho zdroje potravy a před požitím potravy se obvykle pohybuje po jejím povrchu. S tímto faktem nejspíš souvisí to, že hmyz má chuťové orgány nejenom v oblasti úst, ale také na nohou a křídlech (Obr. 1) (Yarmolinsky a kol., 2009).



**Obrázek 1: Lokalizace chuťových receptorů u motýlů.** Rozmístění chuťových receptorů na těle a hlavě motýla je vyznačeno žlutou barvou (převzato z Agnihotri a kol., 2016).

Základní jednotkou chuťového aparátu hmyzu je senzila (Obr. 2), uvnitř které se nachází svazek tři nebo čtyř sensorických neuronů (gustatory receptor neurons, GRNs). Dlouhou dobu byla specifická GRNs u hmyzu neznámá, tedy až na několik příkladů z výzkumu octomilky *Drosophila melanogaster* (např. Clyne a kol., 2000, Dahanukar a kol., 2007). Později však několik studií prokázalo, že ačkoliv se struktura senzila téměř neliší mezi druhy, GRNs uvnitř senzila mohou plnit odlišné funkce (Zhang a kol., 2013). Velké rozdíly ve zdrojích potravy mezi jednotlivými druhy hmyzu jsou pak způsobeny chuťovými receptory (gustatory receptor, Gr) na GRNs, množstvím GRNs v senzilech a jejich umístěním na těle (Agnihotri a kol., 2016). Dále se exprese Gr může lišit u dospělců a larev, což znamená, že stejný GRN může mít odlišný vliv na chování jedince v průběhu života (Ernst & Westerman, 2021). GRNs se rozdělují na několik druhů podle toho, jaké chuťové receptory exprimují na povrchu své membrány (Kwon a kol., 2014). Na základě molekulární fylogeneze a afinity ke specifickým ligandům byly Gr hmyzu klasifikovány do čtyř skupin: receptory citlivé na oxid uhličitý, D-fruktózu, cukry a hořké látky (Xu, 2020).



**Obrázek 2: Buněčná organizace chuťové senzily.** Senzila obsahuje čtyři neurony s chuťovými receptory (GRNs; červená, fialová, modrá a zelená barva), které jsou citlivé na odlišné chemické podněty, a jeden neuron mechanoreceptoru (černá barva). Dendrity se zanořují do dutiny naplněné endolymfou obsahující proteiny schopné vázat signální molekuly nerozpustné ve vodě (odorant binding proteins). Proteiny jsou exprimovány do dutiny z podpurných buněk (hnědá barva). Axony vedou přímo do mozku (převzato z Amrein, 2016).

### 1.2.1. Evoluce chuťových receptorů

Chuťové receptory jsou kódované multigenovou rodinou proteinů, jejichž evoluční dynamika je již dlouhou dobu středem pozornosti molekulárních biologů (Agnihotri et al., 2016). Mechanismem vzniku genových rodin jsou duplikace, ty jsou obecně jednou z nejdůležitějších hnacích sil v evoluci genů a najdeme je u všech organismů (Letunic, 2002). Jedná se o zdvojení určitého úseku DNA na stejném či jiném chromozomu, také se může duplikovat celý chromozom nebo dokonce i celý genom. Duplikace vzniká většinou zlomem a výměnou genetické informace mezi nesesterskými chromatidami. Mohou také vznikat v důsledku poruch rekombinace např. nerovnoměrným crossing-overem. Bez ohledu na jejich původ mohou mít duplikované geny tři různé osudy. Obě kopie mohou přetrvávat v genomu s dokonalou (nebo téměř dokonalou) sekvenční identitou, což může mít za následek vyšší úroveň exprese genového produktu. V druhém případě je jedna kopie potlačena negativní selekcí nebo nahromaděním bodových mutací, dokud se nestane pseudogenem. Třetí možností



je vznik náhodných mutací, které mohou způsobit to, že se alespoň jedna ze dvou kopií bude funkčně lišit, buď získáním nové funkční role, nebo specializací některého aspektu její předchozí funkce. Interakce a kombinace duplikace a přirozené selekce mohou vést ke vzniku genových rodin obsahujících od dvou do stovek více či méně podobných genů vykonávajících stejné nebo rozličné funkce (Sankoff, 2001).

Ukázalo se, že vývoj rodin hmyzích chemoreceptorových genů je značně dynamický a dochází k častým vznikům i zánikům genů, což pak hraje roli v adaptaci a speciaci jednotlivých druhů (Sánchez-Gracia a kol., 2009). Například Guo a kol., 2017, našli u *Bombyx mori* sedm nových kopií genu pro jeden z chuťových receptorů. Velikost genových rodin je tak u motýlů velmi variabilní. Například v genomu *B. mori* bylo nalezeno 65 genů Gr, (Wanner & Robertson, 2008), u *Danaus plexippus* 47 genů (Zhang a kol., 2011) a 25 genů u *Plutella xylostella* (You a kol., 2013). Z těchto a dalších výzkumů také vyplývá, že pro evoluci Gr jsou charakteristické četné tandemové duplikace, které vedou ke vzniku genových klastrů, a jak pozitivní, tak i negativní genová selekce (Engsontia a kol., 2014).

### **1.2.2. Typy chuťových receptorů a jejich funkce**

Jak již bylo zmíněno v kapitole 1.2. chuťové receptory hmyzu se dělí v závislosti na jejich ligandu na receptory citlivé na oxid uhličitý, D-fruktózu, hořké látky a cukry. Oxid uhličitý je plyn všudypřítomný v životním prostředí. CO<sub>2</sub> hraje velkou roli v životě hmyzu a jeho detekce má odlišný význam mezi druhy. Některé druhy mohou například používat detekci CO<sub>2</sub> pro orientaci v prostoru. Přesně tímto způsobem včely *Apis mellifera* jednoduše vyhledávají své úly podle vyšší koncentrace CO<sub>2</sub>, než je v atmosféře (Seeley, 1974) a komáři používají tyto receptory k lokalizaci hostitele (Chen a kol., 2021). Naproti tomu motýli používají určení koncentrace CO<sub>2</sub> obsaženého v nektaru k odhadu stáří a kvality květin (Thom a kol., 2004).

D-fruktóza obsažená v hemolymfě hmyzu je měřena specifickými chuťovými receptory. Tyto receptory se nachází v mozku a hrají velkou roli v regulaci stravování jedince. Receptory inhibují pocit hladu při nadměrné hladině D-fruktózy, a naopak stimulují krmení při jejím nedostatku (Miyamoto a kol., 2012). V roce 2011 Sato a kol. prokázali, že receptory citlivé na tento cukr jsou vysoce specifické napříč druhy a reagují pouze na D-fruktózu. Díky tomuto objevu byly D-fruktózové receptory klasifikovány zvlášť od cukerných receptorů.

Receptory citlivé na hořké látky umožňují hmyzu podobně jako u savců detekovat potenciálně škodlivé chemikálie. Hořké látky, se kterými hmyz přichází do kontaktu,

pocházejí z několika různých zdrojů. Například mnoho rostlin je produkuje ve formě sekundárních metabolitů, které slouží jako obranný mechanismus proti škůdcům. Další z těchto chemikálií často pocházejí z mikroorganismů, které kolonizují rozkládající se ovoce. „Hořké“ chemikálie zahrnují širokou škálu strukturně odlišných sloučenin, jako jsou alkaloidy, terpenoidy a fenoly (Amrein, 2016). Vnímání tohoto variabilního spektra chemikálií je pokryto rozsáhlou skupinou Gr, které jsou schopné vázat na sebe jednu nebo více signálních molekul. Chuťové receptory pro hořké látky jsou ukázkovým příkladem toho, jak flexibilní a rozmanité mohou chuťové receptory hmyzu být.

### 1.3. Cukerné receptory

Cukerné receptory umožňují hmyzu detekovat cukry v potravě (Slone a kol., 2007). Vědci věnují svoji pozornost těmto receptorům již od osmdesátých let minulého století. Cukry nejsou pro hmyz pouze zdrojem energie, ale také důležitým fagostimulantem. Jednotlivé cukry nebo jejich kombinace v rostlině dávají hmyzu signály o jejím složení a kvalitě a zároveň stimulují nebo inhibují jejich stravovací návyky (Agnihotri a kol., 2016). Je prokázáno, že exprese a molekulární složení genů cukerných receptorů se může měnit mezi druhy nebo mezi skupinami jedinců stejného druhu žijících v odlišném prostředí. Jedná se tedy o dynamický systém, který pomáhá přežít v nových podmínkách (Menken a kol., 1992; Ômura a kol., 2008; Wada-Katsumata a kol., 2013). Lze říci, že cukerné receptory mohou hrát velkou roli ve speciaci a adaptaci hmyzu. Zajímavým příkladem adaptace je šváb *Blattella germanica*, u kterého se vyvinula adaptivní behaviorální reakce na glukózu. V normálních podmínkách by se měla glukóza vázat na cukerné receptory tohoto švába a působit na něj jako fagostimulant. Právě proto byla glukóza dlouhou dobu přidávána do pastí na šváby. Wada-Katsumata a kol. (2013) prokázali, že u švábů žijících v domácnostech se vyvinula averze k látkám obsahujícím glukózu. Při porovnání jejich receptorů s chuťovými receptory divokých švábů bylo zjištěno, že u jedinců, kteří se mnohokrát setkali s nástrahou, došlo ke změnám na úrovni GRNs a místo cukerných receptorů se u nich při setkání s glukózou aktivovaly receptory pro hořké látky. Dalším příkladem jsou změny stravovacích návyků u motýlů, kteří se běžně stravují nektarem květin, jako je například motýl *Vanessa indica*. Ômura a kol. (2008) zkoumali cukerné receptory *V. indica* a *Argynnis hyperbius* patřících do čeledi Nymphalidae. Na rozdíl od *A. hyperbius* se *V. indica* nestravuje pouze nektarem, ale také mízou ze stromů a hniječím ovocem, které obsahují ethanol a kyselinu octovou. Ukázalo se, že ačkoliv obsah cukru ve fermentačních produktech je nižší, ethanol dokáže zvýšit citlivost cukerných receptorů *V. indica*. Když se pokusili ověřit reakci cukerných receptorů

*A. hyperbius*, zjistili, že ethanol a kyselina octová na ně mají opačný efekt a zcela inhibují stravování.

V této bakalářské práci se zabývám cukernými receptory předivky zhoubné neboli *Yponomeuta evonymella*, která slouží jako modelový organismus pro zkoumání speciace a adaptace motýlů. Podrobný popis role cukerných receptorů *Y. evonymella* je uveden v kapitole 1.4.

### 1.3.1. Stavba cukerných receptorů

Cukerné receptory jsou tvořené sedmi transmembránovými doménami s N terminálním koncem exponovaným na vnitřní straně cytoplazmatické membrány a C koncem na vnější straně (Zhang a kol., 2011). Existují dva odlišné názory, zda se jedná o receptory spřažené s G-proteinem podobné receptorům savců (Mombaerts, 1999) nebo ligandově řízené iontové kanály (Sato a kol., 2008). Modernější studie ale naznačují, že Gr mohou být obou typů. Ve své studii Zhang a kol. (2011) prokázali, že cukerné Gr motýlů jsou svojí strukturou podobné čichovým receptorům, o jejichž struktuře víme mnohem víc. V roce 2021 byl publikován článek, který potvrdil, že čichové receptory hmyzu tvoří tetramer spřažený s G-proteinem a zároveň mají iontový kanál umístěný uprostřed čtyř jednotlivých proteinů (del Marmol a kol., 2021).

Největší množství výsledků, tykajících se přímo chuťových receptorů hmyzu, bylo získáno výzkumem *D. melanogaster*. Jiao a kol. (2008) zkoumali cukerný receptor Gr64f citlivý na sacharózu, glukózu a maltózu a dospěli k závěru, že tento konkrétní receptor má strukturu heterodimeru. Podobné výsledky byly publikovány o rok později (Lee a kol., 2009), a potvrdily, že všechny cukerné receptory *D. melanogaster* jsou zpravidla heteromultimery. Existují i výjimky, například cukerný receptor Gr5a citlivý na trehalózu byl již v roce 2003 určen jako homodimer (Chyb a kol., 2003). Studie, které se zabývaly Gr motýlů, potvrzují, že struktura receptorů u Lepidoptera se nejspíš liší od *D. melanogaster* (Ozaki a kol., 2011; Sato a kol., 2011), a v současné době se vědci více přiklání k tomu, že se v případě motýlů jedná o homomultimery (Xu, 2020).

### 1.3.2. Struktura genů pro cukerné receptory

Engsontia a kol. (2014) na příkladu *B. mori* ukázali, že geny pro cukerné receptory tvoří klastr, kde maximální vzdálenost mezi jednotlivými geny je 100 kbp. U *B. mori* bylo nalezeno pět genů umístěných na 15. chromosomu (chr. 15), podobnou klastrovou strukturu má taktéž *P. xylostella* (Engsontia a kol., 2014). Lokalizace genů pro cukerné receptory

*P. xylostella* na chr. 15 byla ověřena pomocí nástroje BLAST proti složenému genomu (ilPluXylo3.1). Dle Engsontia a kol. (2014) jsou geny pro cukerné receptory *P. xylostella* v průměru tvořeny 387 aminokyselinami, počet exonů se pohybuje mezi 3-5 a průměrná délka exonu je 150 bp. Složení intronů je konzervované a moc se neliší mezi druhy.

### 1.3.3. Specifita cukerných receptorů

Všechny druhy řádu Lepidoptera můžeme rozdělit do tří skupin na základě specifčnosti jejich živných rostlin. V přírodě najdeme monofágní, oligofágní a polyfágní druhy. Takové rozdílné stravovací návyky jsou výsledkem ligandově specifických Gr (Tang a kol., 2014). Bylo pozorováno, že více než jeden typ Gr dokáže detekovat jeden typ cukru, a naopak více typů cukrů může aktivovat jeden druh Gr (Agnihotri a kol., 2016). U mnoha motýlů je iniciace stravování řízena aktivací několika receptorů zároveň. Například stravování *Manduca sexta* je indukováno kombinací inositolu s glukózou nebo sacharózou (Glendinning a kol., 2007). Příklady míry afinity některých cukrů k Gr motýlů jsou znázorněny v Tab. 1.

**Tabulka 1: Míra citlivosti mediální (M) a laterální (L) senzily na vybrané cukry.** Převzato z Agnihotri a kol., 2016, původní data pochází z Dahanukar a kol., 2007; Slone a kol., 2007; Jiao a kol., 2008; Wanner & Robertson, 2008; Jiang a kol., 2015.

Druh	<i>Pieris brassicae</i>		<i>Pieris rapae</i>		<i>Bombyx mori</i>	<i>Mamestra brassicae</i>
	M	L	M	L	L	L
Cukr						
D-xylóza	-	-	-	+	+	NA
D-ribóza	-	-	+	-	+	NA
D-fruktóza	+	-	++	-	+	+
D-galaktóza	-	-	-	-	+	+
D-glukóza	+	++	+	++	++	+
D-maltóza	-	-	-	+++	+++	+
Sacharóza	++	+++	++	+++	++++	+
Trehalóza	-	-	-	+	+	-

- necitlivost, + slabá citlivost, ++ mírná citlivost, +++ silná citlivost, NA - data nejsou dostupná.

## 1.4. *Yponomeuta evonymella*

Rod *Yponomeuta*, česky předivky, patří do poměrně malé čeledi Yponomeutidae (Lepidoptera, Ditrysia), a zahrnuje 76 druhů se širokou distribucí v jihovýchodní Asii, východní a západní Evropě, na západě Kanárských ostrovů, v Africe, Austrálii, na Novém Zélandu a v Severní Americe (Turner a kol., 2010). Pro tento rod je charakteristické monofágní a velmi zřídka oligofágní stravování na rostlinách z čeledi Celastraceae, Rosaceae nebo Salicaceae. Živné rostliny z čeledi Celastraceae jsou typické pro evolučně starší druhy rodu *Yponomeuta* pocházející z Asie. *Yponomeuta evonymella* se společně s několika dalšími

druhy vyskytuje na území západní Evropy (Povel, 1984) a zpravidla se stravuje na rostlinách z čeledi Rosaceae nebo Salicaceae (Turner a kol., 2010). Protože se druhy *Yponomeuta* morfologicky téměř neliší, dlouhou dobu byly rozlišovány pouze na základě určení živné rostliny. Bylo prokázáno, že chemické složení živné rostliny ovlivňuje stravování larev a ovipozici u dospělců. Zajímavé je, že i přesto, že se různé druhy *Yponomeuta* mohou vyskytovat ve stejné lokalitě, mezi druhy nedochází k páření. U těchto druhů tedy muselo dojít ke vzniku pre-zygotických reprodukčních bariér. Například aktivace sexuálního chování může být ovlivněna druhově specifickou výškou rostliny, na které obvykle dochází k ovipozici, nebo odlišnými vibračními signály a ovíjením křídel u různých druhů. Důležitou roli při páření také hrají druhově specifické samčí feromony (Lofstedt a kol., 1991). V přírodě je znám pouze jeden hybrid a to u druhů *Y. malinellus* a *Y. padellus* (Arduino, 1985). Ukázalo se, že páření je ovlivněno nejenom sexuálními feromony produkovanými samcem a samicí, ale i těkavými chemickými látkami uvolňovanými živnou rostlinou a detekovanými čichovými receptory předivek (van der Pers a kol., 1980; Deisig a kol., 2014). Existuje výzkum (Bakker a kol., 2008), který se více zabýval tímto jevem a jeho výsledky ukazují, že neexistuje prokazatelná spojitost mezi ovipozicí a živnou rostlinou. Zároveň je ale potřeba brát v potaz, že výzkum byl prováděn pouze na dvou druzích *Yponomeuta* a chování jedinců bylo pozorováno v klecích. Je vidět, že existuje určitý vztah mezi složením živné rostliny a chemoreceptory předivek, který se projevuje jinak během různých fází jejich života.

Samice kladou vajíčka do shluků až 100 kusech na stonky konkrétní živné rostliny, často v blízkosti pupenu nebo vedlejších větví. Ačkoliv se cukerné Gr neúčastní řízení ovipozice, je důležité, aby samice správně volily živnou rostlinu, která obsahuje vhodné fagostimulanty pro jejich larvy. Dle fyziologických experimentů je *Yponomeuta* citlivá převážně na izomery cukerných alkoholů obsažených v listech živných rostlin. Dulcitol v rostlinách čeledi Celastraceae působí jako fagostimulant, sorbitol, který má obdobný efekt, je obsažen v čeledi Rosaceae (Menken a kol., 1992). Malé množství dulcitolu bylo také nalezeno v Rosaceae a předpokládá se, že to mohlo usnadnit přesmyk na jinou živnou rostlinu (Roessingh a kol., 1999). Experimenty ukázaly, že *Y. evonymella* pozitivně reaguje jak na dulcitol, tak na sorbitol, ačkoliv reakce na sorbitol je mnohem výraznější. *Y. cagnagella*, která se stravuje na Celastaceae, reaguje pouze na dulcitol (Roessingh a kol., 1999). Tato fakta podporují teorii o vzniku nových druhů *Yponomeuta* v důsledku přesmyku na jinou živnou rostlinu. Roessingh a kol. (1999) uvádí, že receptory, které vážou oba cukerné alkoholy, se nachází na jednom druhu GRNs a jejich citlivost mezi druhy může být daná jednoduchými

modifikacemi genů Gr. Morfologické a funkční studium cukerných receptorů je velmi obtížné, jelikož jsou receptory exprimované pouze v několika neuronech nacházejících se v senzile připojené na galeu. Z důvodu náročného pitvání předivek podmíněného sezónností, malou velikostí larev a velmi nízkým výtěžkem materiálu neexistují moderní studie, které by se zabývaly cukernými receptory *Yponomeuta*. Anotace genu cukerných receptorů v genomu *Yponomeuta* může být jedním ze způsobů, jak se dozvědět víc o evoluci Gr a speciaci předivek, aniž bychom byli limitováni množstvím materiálu.

## 1.5. Historický přehled sekvenování

### 1.5.1. Sekvenování první generace

Zájem o sekvenování genomů roste již od 50. let minulého století, kdy byla odhalena struktura dvoušroubovice DNA a její funkce. První sekvenování bylo provedeno v roce 1968, kdy se Dr. Wu podařilo osekvenovat pouhých 13 nukleotidů bakteriofága lambda (Wu, 1972). I takové krátké sekvenování pojmenované „Primer extension method“ bylo obrovským pokrokem, na jehož základě byla následně vytvořena nová, dosud používaná metoda, Sangerovo sekvenování, která patří do tzv. sekvenování první generace (Sanger a kol., 1977). Při Sangerově sekvenování je využíván proces replikace DNA, kde k jednořetězcovému vlákně přisedá 15-25 bp dlouhý primer, který je komplementární k začátku sekvenovaného místa, dále probíhá syntéza DNA za přítomnosti čtyř 2'-deoxynukleosid 5'-trifosfátů (dNTP) a čtyř dideoxynukleotidů (ddNTP). Metoda prošla řadou modernizací a v současné době se používají dideoxynukleotidy značené odlišnými fluorofory. Jednotlivé ddNTP jsou v porovnání s dNTP v reakci zastoupeny jen v relativně malém množství, při jejich navázání dojde k zastavení syntézy DNA. Jelikož je v reakci velké množství stejných molekul, jejich syntéza je pozastavená v různých místech. Nakonec jsou molekuly DNA, které začínají v místě nasednutí primeru a končí navázáním ddNTP, rozdělené podle velikosti pomocí kapilárové elektroforézy a jednotlivé ddNTP jsou snímány laserem (Karger a Guttman, 2009).

Do stejné skupiny sekvenování první generace také patří Maxamovo–Gilbertovo sekvenování, které bylo vyvinuto Allanem Maxamem a Walterem Gilbertem v letech 1976–1977 (Maxam a Gilbert, 1977). Tato metoda je založena na principu chemického štěpení jednovláknové DNA ve čtyřech zkumavkách. V každé zkumavce je prováděno štěpení jen určitých typů bází G, A+G, C+T a C. Dále jsou reakce nanášeny na hustý polyakrylový gel a rozdělené podle velikosti. Odečtením pozice jednotlivých bází ve všech čtyřech reakcích pak lze stanovit sekvenci daného úseku.

Trvalo to téměř 13 let, než se vědcům podařilo oskenovat a složit lidský genom pomocí Sangerova sekvenování (The Human Genome Project, 1990-2003). Možnosti sekvenování první generace jsou velmi omezené. Pro porovnání délka lidského diploidního ženského genomu je 6,37 Gbp (Piovesan a kol., 2019) a maximální délka sekvence, kterou lze získat Sangerovým sekvenováním je 1,1 kbp (dle SEQme s.r.o). Další nevýhodou této metody je možnost zpracování pouze jednoho úseku DNA během jedné reakce.

## 1.5.2. Sekvenování druhé generace

Obrovský zájem o sekvenování způsobil velký pokrok v rozvoji sekvenační metodiky a přinesl nám řadu nových způsobů sekvenování, kterým dnes říkáme sekvenování dalších generací (Next Generation Sequencing – NGS). Cílem NGS bylo co nejvíce zrychlit a zlevnit sekvenování, a to prostřednictvím současného zpracování obrovského množství různých fragmentů DNA zachycených na jediném sekvenačním elementu, např. na destičce (Hu a kol., 2021). NGS můžeme dále rozdělit na dvě skupiny: historicky starší druhou generaci „sekvenování krátkých čtení“ a nejmodernější třetí generaci „sekvenování jediné molekuly“ (Adewale, 2020). Do první skupiny patří například Illumina, platforma poskytující sekvenování o délce čtení 100-450 bp, která dokáže sekvenovat až 900 miliard bází zároveň. Jelikož daná metoda nedokáže přesně určit nukleotid z analýzy pouze jedné molekuly, je nutno namnožit DNA pomocí polymerázové řetězové reakce (Polymerase Chain Reaction – PCR) (Hu a kol., 2021). Při sekvenování metodou Illumina jsou pomocí adaptérů jednotlivé nastříhané úseky DNA přichyceny na malou destičku. Vlastní sekvenační proces pak využívá mechanismu replikace DNA, kdy jsou do rostoucího řetězce zařazeny báze s navázanou fluorescenční barvou. Kamera sekvenátoru v každém jednotlivém kole syntézy řetězce DNA snímá signál z celé destičky a podle rozdílné fluorescence pozná, jaký nukleotid byl přidán u každého jednotlivého úseku.

Známou metodou, která je založena na Illumina sekvenování, ale umožňuje získat sekvenční informace z delších úseků, je příprava knihovny 10X Genomics. Během této metody je DNA naštěpená na dlouhé úseky. Molekuly jsou následně spolu s enzymy napojené na tzv. „Gel Beads“, které obsahují četné kopie unikátních primerů nesoucích barkód, tyto primery existují v milionech různých verzí. Spojení DNA, enzymů a „Gel Beads“ je uzavřeno v olejové kapénce a tvoří tzv. GEM (“Gel Beads Emulsion”). Uvnitř jedné kapénky dochází ke štěpení fragmentu DNA a přípravě sekvenační knihovny, fragmenty pocházející z jedné molekuly DNA tak kromě klasických adapterů nesou i barkód. Všechny „GEMs“ jsou pak smíchané do jedné knihovny připravené k sekvenování. Značení krátkých fragmentů DNA

barkódem usnadňuje následné skládání genomu. V roce 2019 firma Bio-Rad Laboratories, Inc. zažalovala 10X Genomics za použití patentované technologie, která využívá rozdělení vzorků dle mikrofluidity. Proto se tato metoda od roku 2020 již neprovádí z právních důvodů.

### **1.5.3. Sekvenování třetí generace**

Do třetí generace patří dva konkurenti: PacBio od společnosti Pacific Biosciences (PacBio) (Menlo Park, CA, USA) a platforma Oxford Nanopore Technologies (ONT) (Oxford, UK). Obě metody jsou založeny na principu sekvenování vysokomolekulárních úseků DNA. Střední délka čtení pro PacBio je 15 kbp a průměrná chybovost > 1 % (Wenger a kol., 2019) a pro ONT 30 kbp, > 5 % (Jain a kol., 2018).

PacBio taktéž využívá replikaci DNA, při které jsou nukleotidy fluorescenčně značeny. Celý proces probíhá ve speciální komůrce veliké jen několik nanometrů, která umožňuje specifické šíření světla, a tedy zachycení nepatrného fluorescenčního signálu. ONT na druhou stranu je zcela nová metoda, která vůbec nevyužívá fluorescence ani jiného značení nukleotidů. Tato metoda používá nanopóry: póry ve dvojvrstvě lipidové membrány, které připomínají transmembránové iontové kanály biomolekul. Sekvenování se provádí v sekvenátoru, do kterého je vložena komůrka (flowcell), obsahující dvě komory vyplněné elektrolytem a rozdělené membránou. Mezi komory je přivedeno napětí, které vytváří elektrický gradient na membráně. Molekuly DNA jsou zachycené na jednotlivých pórech a jsou hnané na druhou stranu membrány. Při průchodu nukleotidů póry dochází ke změně napětí mezi membránami, a tyto změny jsou detekované a zaznamenávané velmi citlivým detektorem.

Udaondo a kol. (2021) jako jedni z několika provedli praktické porovnání těchto dvou technologií a dospěli k závěru, že sekvenování jediné molekuly může rozhodně nahradit sekvenování druhé generace, ale zároveň nejde jistě říci, zda je metoda PacBio lepší než ONT, nebo naopak.



## 2. Cíle práce

Rod *Yponomeuta* je oblíbeným modelovým organismem pro studium sympatrické speciace, adaptace a interakce hmyzu s hostitelem. Hlavní motivací této práce je hypotéza, že změny chuťových receptorů předivek vedly k přesmyku na jinou živnou rostlinu a ovlivnily tak vznik nových druhů. Cílem je získat kompletní sekvence genů pro cukerné receptory u předivky *Y. evonymella*, které poté mohou posloužit jako podklad pro funkční studie cukerných receptorů a objasnění jejich role ve vnímání sorbitolu a dulcitolu a změnách živných rostlin. K řešení bylo stanoveno několik dílčích cílů:

- Složení genomu *Y. evonymella* s použitím dostupných krátkých sekvencí
- Identifikace genů pro cukerné receptory ve složeném genomu *Y. evonymella* s pomocí známých sekvencí genů příbuzného druhu *P. xylostella* (Engsontia a kol., 2014)
- Vyhledávání genů pro cukerné receptory v genové knihovně *Y. evonymella*
- Sekvenování části genomu *Y. evonymella* obsahující potenciální geny pro cukerné receptory s použitím technologie ONT
- Anotace genů na osekvenované části genomu.

### 3. Materiál a metody

Všechny skripty použité v této práci jsou dostupné na stránkách Github. Název jednotlivých skriptů je dále uveden v textu ve formátu (github/“název“). <https://github.com/kbobryshava/Skripts>

#### 3.1. Složení genomu *Y. evonymella*

10X Genomics sekvenační čtení *Y. evonymella* byly staženy ve formátu cram ze stránek projektu Darwin tree of life.

[https://github.com/darwintreeoflife/darwintreeoflife.data/blob/master/species/Yponomeuta\\_evonymella/Yponomeuta\\_evonymella.md](https://github.com/darwintreeoflife/darwintreeoflife.data/blob/master/species/Yponomeuta_evonymella/Yponomeuta_evonymella.md)

Soubory byly převedeny do formátu FASTQ pomocí Samtools verze 1.11 (Danecek et al., 2021) (github/1.cram2fastq) a jejich kvalita byla zkontrolována v programu FastQC verze 0.11.5 (Andrews, 2010) (github/2.fastqc).

Genom byl složen v programu Supernova verze 2.1.1 (github/3.supernovarun, github/4.supernova\_format) se třemi odlišnými hodnotami parametru maxreads (Weisenfeld a kol., 2017). Hodnota byla určena pomocí návodu 10X Genomics (Danecek a kol., 2021), který je optimalizován pro savce, proto bylo rozhodnuto vyzkoušet několik možností pro nalezení optimálního nastavení pro Lepidoptera. Parametr maxreads=140M (hloubka čtení 27x) byl použit jako minimální možná hodnota pro složení genomu, maxreads=250M (hloubka čtení 56x) jako doporučená hodnota pro savce, maxreads=350M (hloubka čtení 68x) jako největší možná hodnota pro využití všech dostupných čtení z Darwin tree of life. Složené genomy byly poté vyhodnoceny v programech Busco verze 3.0.2 (Manni a kol., 2021) a Quast verze 4.6.3 (Mikheenko a kol., 2018) (github/5.busco+quast).

Ukázalo se, že všechny verze složeného genomu obsahují velké množství duplikací, proto byl použit program Purge-dups verze 1.0.1, jehož úkolem bylo odstranit duplikace mimo jiné na základě hloubky čtení ([https://github.com/dfguan/purge\\_dups](https://github.com/dfguan/purge_dups)) (github/7.purgedups). Jelikož se v tomto programu využívá mapování původních sekvencí na složení genom, 10X sekvenační čtení byly filtrovány na kvalitu a délku v programu Trimmomatic verze 0.39, při čemž byl zároveň odstraněn barkód a adaptéry Illumina (<https://github.com/usadellab/Trimmomatic>) (github/6.trimmomatic).

## 3.2. Anotace genů Gr

Pro nalezení genu Gr byly použity sekvence mRNA genů pro cukerné receptory *P. xylostella* (Engsontia et al., 2014), které byly namapovány na složený genom *Y. evonymella* v programu Minimap verze 2.17 (Li, 2018) (github/8.minimap2). Překrývající se sekvence byly označeny jako „exon“ a na každý exon byly navrženy primery v programu Geneious Prime 2021.0.3 (Tab. 2). Dále byly potenciální geny Gr vyhledávány v BAC (Bacterial artificial chromosome) knihovně *Y. evonymella*.

## 3.3. Prohledávání (screening) BAC knihovny

### 3.3.1. BAC knihovna

BAC knihovna *Y. evonymella* byla vytvořena komerční firmou AC Amplicon Express (Pullman, WA, USA). Samčí vysokomolekulární DNA byla částečně štěpena Hind III a vložena do bakteriálního vektoru pCC1BAC (Epicentre, Madison, WI, USA), který byl poté transportován do DH10B *Escherichia coli*. Knihovna obsahuje 20736 BAC klonů s průměrnou délkou insertu 125 kbp, které jsou rozděleny do 54 destiček. Prohledávání probíhalo metodou „Matrix technology: pooling and superpooling“ (verze manuálu září 2014). BAC klony byly rozděleny do 18 superpoolů a každý superpool byl následně rozdělen do 96 matrixpoolů. Uspořádání klonů v matrixpoolech odpovídalo jejich poloze v zásobních destičkách. Prohledávání bylo prováděno pomocí dvou kol polymerázové řetězové reakce (PCR – polymerase chain reaction) a jeho výsledkem bylo určení destičky, řady a sloupce, ve kterém se nachází požadovaný BAC klon.

### 3.3.2. Primer test

V prvním kroku byly určeny optimální teploty nasedání primerů a jejich specifčnost pomocí PCR. Jedna PCR reakce o objemu 10  $\mu$ l obsahovala 1x One Taq® Quick-Load® Reaction Buffer (NEB, Ipswich, MA, USA); 0,2 mM dNTPs; 2,5 U One Taq® Quick-Load® DNA Polymerase (NEB, Ipswich, MA, USA); 0,03  $\mu$ M forward a reverse primerů a 15-20 ng genomové DNA *Y. evonymella*. PCR reakce probíhala v Thermal cycler Biometra Tadvanced (Analytik Jena, Německo) za následujících podmínek: denaturace při 94 °C 5 minut, poté 35 cyklů při 94 °C 30 sekund, 54-62 °C 30 sekund, 68 °C 30 sekund, finální syntetická fáze při 68 °C 3 minuty. Výsledné PCR produkty byly separovány na 1,5% agarózovém gelu v 1x TAE a vizualizovány pomocí barvení ethidium bromidem v konečné koncentraci 0,5  $\mu$ g/ml. Pomocí těchto výsledků byla určena nejvhodnější teplota nasednutí pro jednotlivé primery a

nespecifické primery byly vyřazeny. PCR produkty byly dále přečištěny pomocí soupravy Wizard SV Gel & PCR Clean Up System (Promega, Madison, WI, USA) a odeslány na Sangerovo sekvenování (SEQme s. r. o.). Získané sekvence byly namapovány na exony Gr na složený genom *Y. evonymella*. Primery byly určeny jako sekvenčně specifické, pokud identita jejich PCR produktů vůči exonu, na který byly navrženy, byla větší než 85 % (Tab. 2)

### **3.3.3. PCR I**

První kolo screeningu probíhalo podobně jako primer test. Jako pozitivní kontrola byla použita genomová DNA *Y. evonymella* vyizolována pomocí cetyltrimetylamoniumbromidu (CTAB) dle Kreklová Monika (2020), do zbytku reakcí byl přidán 1 µl BAC DNA. Pro každou reakci bylo provedeno tři opakování, matrixpool byl zvolen pouze tehdy, kdy všechna tři opakování byla pozitivní.

### **3.3.4. PCR II**

Druhé kolo screeningu PCR II probíhalo na vybraných matrixpoolech podle výsledků PCR I. Matrixpool je rozdělen na několik částí. První tři vzorky odpovídají destičce, dalších osm vzorků řadě a dalších deset sloupci. Podle toho, jaké vzorky se jeví jako pozitivní na gelu lze určit souřadnice hledaného BAC klonu. Destičku určuje pořadí pouze jednoho pozitivního proužku (na obr. P1-P3) současně s číslem matrixpoolu. Řadu a sloupec lze zjistit z kombinace dvou pozitivních proužků pomocí klíče (Obr 3). V případě řady se musí v klíči shodovat písmeno, u sloupce číslo.

	1	2	3	4	5	6	7	8	9	10	11	12
A	P:1	R:A,B,C,D	C:1,2,3,4,5,6	C:5,11,17,23	P:1	R:A,B,C,D	C:1,2,3,4,5,6	C:5,11,17,23	P:1	R:A,B,C,D	C:1,2,3,4,5,6	C:5,11,17,23
B	P:2	R:E,F,G,H	C:7,8,9,10,11,12	C:6,12,18,24	P:2	R:E,F,G,H	C:7,8,9,10,11,12	C:6,12,18,24	P:2	R:E,F,G,H	C:7,8,9,10,11,12	C:6,12,18,24
C	P:3	R:I,J,K,L	C:13,14,15,16,17,18	Positive Control	P:3	R:I,J,K,L	C:13,14,15,16,17,18	Positive Control	P:3	R:I,J,K,L	C:13,14,15,16,17,18	Positive Control
D		R:M,N,O,P	C:19,20,21,22,23,24	Negative Control		R:M,N,O,P	C:19,20,21,22,23,24	Negative Control		R:M,N,O,P	C:19,20,21,22,23,24	Negative Control
E		R:A,E,I,M	C:1,7,13,19			R:A,E,I,M	C:1,7,13,19			R:A,E,I,M	C:1,7,13,19	
F		R:B,F,J,N	C:2,8,14,20			R:B,F,J,N	C:2,8,14,20			R:B,F,J,N	C:2,8,14,20	
G		R:C,G,K,O	C:3,9,15,21			R:C,G,K,O	C:3,9,15,21			R:C,G,K,O	C:3,9,15,21	
H		R:D,H,L,P	C:4,10,16,22			R:D,H,L,P	C:4,10,16,22			R:D,H,L,P	C:4,10,16,22	
	Plate Matrix	Row Matrix	Column Matrix		Plate Matrix	Row Matrix	Column Matrix		Plate Matrix	Row Matrix	Column Matrix	

**Obrázek 3: Klíč k určění polohy BAC klonů v knihovně** (převzat z Users Manual For Superpool systém constructed after January 1, 2008, Amplikon express). Červenou barvou jsou zvýrazněny výsledky uvedené v Obr. 6.

### 3.4. Izolace BAC DNA a ověření klonu

Vybrané BAC klony byly vysety na LB medium (1% pepton, 5% kvasničný extrakt, 1% NaCl) s 1,5% agarem a chloramfenikolem (25 mg/ml) a inkubovány přes noc při 37 °C. Pro každý BAC klon byly náhodně vybrány tři izolované kolonie. Jednotlivé bakteriální kolonie byly přeneseny na nové medium a použity pro PCR klon test. PCR reakce probíhala stejně jako v předchozích krocích s jednou výjimkou, jako templát nebyla použita vyzolovaná DNA, ale do připravené PCR reakce byla na jednu minutu ponořena plastová špička, na kterou byla nabrána část zkoumané kolonie. PCR produkty byly přečištěny pomocí soupravy Wizard SV Gel & PCR Clean Up System a 50 ng DNA bylo posláno na Sangerovo sekvenování.

Ověřené kolonie s pozitivním PCR testem byly přeneseny do 200 ml tekutého LB media s chloramfenikolem (25 mg/ml) a třepány asi 15 hodin při 37 °C a 200 RPM. Namnožené buňky byly ihned po třepání umístěny na led a centrifugovány 10 minut při 4 °C a 14000 RPM. BAC DNA byla izolována pomocí kitu NucleoBond Xtra Midi plasmid DNA purification kit (Macherey-Nagel, Francie) dle protokolu od výrobce. Koncentrace přečištěné

DNA byla změřena na Qubit 4 (Thermo Fisher Scientific, USA) pomocí kitu Qubit dsDNA BR Assay Kit (ThermoFisher Scientific, USA). PCR produkty byly přečištěny pomocí soupravy Wizard SV Gel & PCR Clean Up System a 50 ng DNA bylo posláno na Sangerovo sekvenování (SEQme s. r. o.) (Příloha I).

### **3.5. Chromosome walking**

Konce insertu BAC klonu 35:O21 byly osekvenovány pomocí Sangerova sekvenování (SEQme s. r. o.) s použitím univerzálních primerů M13. Sekvence byly zbaveny vektoru a byly na ně navrženy primery v Geneious Prime 2021.0.3 (Tab. 3). Dále následovalo druhé prohledávání BAC knihovny a izolace BAC DNA. PCR produkty byly přečištěny pomocí soupravy Wizard SV Gel & PCR Clean Up System a 50 ng DNA bylo posláno na Sangerovo sekvenování (SEQme s. r. o.) (Příloha I).

### **3.6. Úprava BAC DNA**

Pro odstranění lineárních molekul DNA byla použita Exonukleáza V (New England BioLabs, UK). Reakce o 50  $\mu$ l obsahovala 7,5-10  $\mu$ g DNA; 1x NEBuffer™ 4; 1 mM ATP; 20 U Exonukleázy V a celkový objem byl doplněn přidáním milliQ H<sub>2</sub>O. Reakce probíhala v Thermal cycler Biometra Tadvanced při 37 °C 1 hodinu a byla zastavena zahřátím na 70 °C po dobu 30 minut. Dále byly kruhové molekuly částečně naštěpeny pomocí restriční endonukleázy Afe I (New England BioLabs, UK). Pro štěpení 5  $\mu$ g DNA byla použita 1 U enzymu a 5  $\mu$ g BSA, reakce o objemu 50  $\mu$ l probíhala v Thermal cycler Biometra Tadvanced při 37 °C 10–90 minut a byla zastavena chladem při -20 °C nebo přidáním 2,5  $\mu$ l nanášejícího pufru (50% glycerol; 250 mM EDTA; 5,9 mM bromfenol). K odstranění veškerých úseků DNA kratších, než 25 kbp, byl použit Short read Eliminator Kit Family (Pacbio, USA) dle klasického protokolu od dodavatele.

### **3.7. Příprava knihovny a Oxford Nanopore sekvenování**

Knihovna byla připravena z BAC DNA o celkové koncentraci 1  $\mu$ g v objemu 50  $\mu$ l. Aby byly všechny molekuly DNA v knihovně zastoupeny ve stejné míře, vzorek obsahoval 250 ng BAC DNA každého klonu. Příprava sekvenační knihovny pro Oxford Nanopore probíhala dle protokolu „Ligation Sequencing gDNA protocol“ pomocí SQK-LSK109 Ligation Sequencing Kit a NEBNext Companion Module for Oxford Nanopore Technologies (New England BioLabs, UK). K sekvenování, které probíhalo na přístroji MinION Mk1B,

byla použita komůrka Flow Cell (R9.4.1) FLO-MIN106D. K řízení sekvenování byl použit software MinKNOW verze 22.08.9.

### **3.8. Basecalling, Nanoplot, Nanofilt**

Pro převedení dat z napětí do bází (basecalling) byl použit program Guppy verze 6.0.6 s využitím „high accuracy“ algoritmu (Sherathiya et al., 2021) (github/9.guppy). Dataset byl filtrován na délku čtení > 30 kb a kvalitu > Q15 v programu Nanofilt (<https://github.com/wdecoster/nanofilt>) a výsledky byly vizualizovány v programu Nanoplot (<https://github.com/wdecoster/NanoPlot>) (github/10.nanofilt+nanoplot).

### **3.9. Skládání sekvence BAC klonů a mapování genů**

Filtrovaná ONT sekvenační čtení byla složena v programu Flye verze 2.2.9 (Kolmogorov et al., 2019) (github/11.flye). Na složenou kruhové sekvence BAC klonů byly namapovány exony potenciálních Gr *Y. evonymella* v programu Minimap verze 2.17 (Li, 2018). Dále se výsledky zpracovávaly v Geneious Prime verze 2021.0.3, kde byl na sekvence taktéž namapován vektor.

**Tabulka 2: Primery navržené pro screening BAC knihovny.** Zelenou barvou jsou zvýrazněné primery, které byly použité k vyhledávání. Specifičnost na gelu znamená, že PCR produkt obsahoval pouze jeden proužek a jeho velikost odpovídala délce úseku, na který tyto primery byly navržené. Teplota nasedání primerů odpovídá optimální teplotě na základě primer testu.

Gen Gr	Pořadí exonu	Název primerů	Forward primer (5'→3')	Reverse primer (5'→3')	Teplota nasednutí primerů (°C)	Velikost produktu (bp)	Specifičnost na gelu	Podobnost produktu vůči původní sekvenci
1	1	scr_e1_Gr1	ACTAAAGCCGTCGGAAAGCA	TGTTGTACGATGTGCCGTC	60	400	ne	
1	4	scr_e4_Gr1	CAGGAAGTTGAAGACCCACG	AGCACATCCTGTGCTACTG	54	160	ano	56,0%
1	5	scr_e5_Gr1	TAGAACAAGTCGGCCCTGTT	TTGGAGATGCGTCTATGCGT	60	271	ano	85,9%
2	2	scr_e2_Gr2	CTGAAAACAAACGAAGATCAGG	CCGTAGGAGAGGAATATGGC	58	155	ne	
2	5	scr_e5_Gr2	ACAGTATAATGACGCAAGTGA	ACAGTGTCTACAGCACGG	54	200	ano	57,0%
2	5	scr_e5_Gr2	TGTACGGGACTAATGGCTGC	AAGTTGGGTCTCAGTTGGCCC	60	180	ne	
3	1	scr_e1_Gr3	GGGTTCTTCTGCGGATCTCC	CCGCCGAAACGCTTTATTCC	62	180	ano	91,8%
3	3	scr_e3_Gr3	AGGCGAGAGGTCAGGTAGAA	CGAAAGTCGGACCGAGAGAG	62	160	ano	55,0%

**Tabulka 3: Primery navržené pro chromosome walking.** Zelenou barvou jsou zvýrazněné primery, které byly použity k vyhledávání.

BAC klon	M13 primer	Název primerů	Forward primer (5'→3')	Reverse primer (5'→3')	Teplota nasednutí primerů (°C)	Velikost produktu (bp)	Specifičnost na gelu
35:O21	forward	chrw_m13f_1	TGCGGAAACTTGTTTTGGCT	CCGAATTCAGTAGCGCCCTT	58	450	ano
35:O21	forward	chrw_m13f_2	TCACTCAGGTTTCCTCACGC	CCCAGGTTTCGATTCACAGTG	58	200	ano
35:O21	reverse	chrw_m13r_1	AGTGAACCTTCCGTGGAAGT	GTCCGAGGAAGAAAGAACGC	60	250	ne
35:O21	reverse	chrw_m13r_2	ACAAAACACACACAGGACATTACA	CCAACAGCTTAACGTGCCTT	58	300	ano



## 4. Výsledky

### 4.1. Skládání genomu *Y. evonymella*

V prvním kroku byl genom *Y. evonymella* složen *de novo* s použitím 10X Genomics sekvenačních čtení. K dispozici bylo 211,5 tisíc čtení dlouhých 151 bp. Celkově byly vytvořeny tři verze genomu s využitím odlišných hodnot parametru maxreads a tyto genomy byly porovnány mezi sebou. Celková délka genomů se pohybovala mezi 620 Mbp a 850 Mbp a hodnota N50 mezi 60 kbp a 200 kbp (Tab. 5) Komplettnost genomů byla posouzena prostřednictvím vyhledávání 1658 genů pro Insecta a později 5286 genů pro Lepidoptera z databáze BUSCO (Tab. 4). Na základě těchto výsledků byl pro další použití vybrán genom, který vykazoval největší komplettnost (90,3 % kompletních genů Insecta, N50 = 199,3 kbp). Vzhledem k velkému počtu duplikací v genomu (45,5 % genů Insecta) byla provedena deduplikace.

Pro ověření efektivity deduplikace byl původní složený genom porovnán s genomem po deduplikaci (Tab. 6, Tab. 7). Délka genomu klesla téměř o polovinu, ale podařilo se zachovat velké N50 (122,5 kbp) a 88,6 % kompletních ortologních genů pro Lepidoptera. Knihovna pro Lepidoptera byla k programu BUSCO přidána později, proto některé analýzy byly provedeny pouze s knihovnou Insecta.

**Tabulka 4: Výsledky analýzy prostřednictvím BUSCO databáze pro tři složené genomy.** Červenou barvou jsou znázorněny hodnoty duplikovaných genů, dle kterých bylo rozhodnuto provést deduplikaci výchozích dat.

<b>Insecta</b>	maxreads = 140M		maxreads = 250M		maxreads = 350M	
Počet kompletních BUSCO genů	1382	83,4%	1234	90,3%	1351	81,5%
Počet kompletních BUSCO genů v jedné kopii	1120	67,6%	612	44,8%	870	52,5%
Počet kompletních duplikovaných BUSCO genů	262	15,8%	622	45,5%	481	29,0%
Počet fragmentovaných BUSCO genů	173	10,4%	64	4,7%	205	12,4%
Počet chybějících BUSCO genů	103	6,2%	69	5,0%	102	6,2%
Celkový počet prohledaných BUSCO skupin	1658					

**Tabulka 5: Kvantitativní analýza složených genomů prostřednictvím programu QUAST.**

Quast	maxreads = 140M	maxread = 250M	maxread = 350M
Celková délka genomu (Mbp)	616,26	851,62	819,35
Počet kontigů	53146	39263	56574
Největší kontig (Mbp)	0,47	1,59	0,92
N50 (Kbp)	58,07	199,29	131,52

**Tabulka 6: Srovnání kompletnosti genomů (maxreads = 250M) před a po deduplikaci.** Červenou barvou jsou znázorněny kategorie duplikovaných genů.

<b>Insecta</b>	před deduplikací		po deduplikaci	
Počet kompletních BUSCO genů	1234	90,3%	1203	88,0%
Počet kompletních BUSCO genů v jedné kopii	612	44,8%	1197	87,6%
Počet kompletních duplikovaných BUSCO genů	622	45,5%	6	0,4%
Počet fragmentovaných BUSCO genů	64	4,7%	73	5,3%
Počet chybějících BUSCO genů	69	5,0%	91	6,7%
Celkový počet prohledaných BUCBO skupin	1367			
<b>Lepidoptera</b>	před deduplikací		po deduplikaci	
Počet kompletních BUSCO genů	4794	90,7%	4686	88,6%
Počet kompletních BUSCO genů v jedné kopii	2452	46,4%	4653	88,0%
Počet kompletních duplikovaných BUSCO genů	2342	44,3%	33	0,6%
Počet fragmentovaných BUSCO genů	157	3,0%	199	3,8%
Počet chybějících BUSCO genů	335	6,3%	401	7,6%
Celkový počet prohledaných BUCBO skupin	5286			

**Tabulka 7: Srovnání složeného genomu (maxreads = 250) před a po deduplikaci.**

Složený genom	před deduplikací	po deduplikaci
Celková délka genomu (Mbp)	851,62	549,59
Počet kontigů	39263	23347
Největší kontig (Mbp)	1,59	1,04
N50 (Kbp)	199,29	122,50

## 4.2. Mapování genů *P. xylostella* na genom *Y. evonymella*

Ve svojí studii Engsontia a kol. (2014) určili pět genů pro cukerné receptory *P. xylostella*: PxGr1, PxGr2, PxGr3, PxGr4 a PxGr44. Sekvence mRNA těchto genů byly namapovány na složený deduplikovaný genom *Y. evonymella*. Podařilo se najít dva kontigy, které nesou některé z těchto genů. Geny PxGr1 a PxGr2 byly namapovány na kontig 217727 dlouhý 39,34 kbp a gen PxGr3 na kontig 121737 dlouhý 45,26 kbp. Podařilo se namapovat celou mRNA *P. xylostella*, přičemž se sekvence rozdělily do čtyř až sedmi exonů (Tab. 8). *P. xylostella* a *Y. evonymella* jsou stále poměrně vzdálené druhy, proto identita mezi namapovanými úseky byla velmi nízká. Geny pro PxGr1 a PxGr2 mapovaly na jednu oblast a s velikou pravděpodobností se jednalo o jeden gen pro cukerný receptor. Jelikož PxGr jsou paralogní geny, nelze říci, jestli je namapovaný gen ortologem PxGr1 nebo PxGr2.

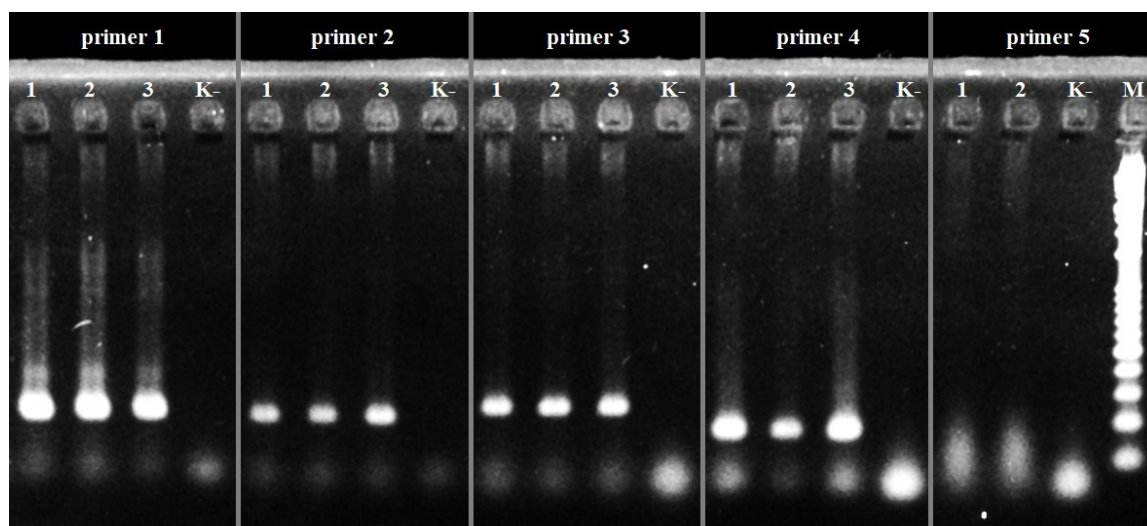
**Tabulka 8: Obecné informace o předpokládaných genech pro cukerné receptory nalezených ve složeném genomu *Y. evonymella*.** Exony *Y. evonymella* jsou delší z důvodů přítomnosti inzercí.

Kontig	Gen	Počet namapovaných exonů	mRNA <i>Y. evonymella</i> (bp)	mRNA <i>P. xylostella</i> (bp)	Směr
217727	PxGr1	5	1127	1098	reverse
217727	PxGr2	7	1193	1092	reverse
121737	PxGr3	4	1080	1050	reverse

### 4.3. Prohledávání BAC knihovny

Výsledky prvního mapování naznačují, že složený genom *Y. evonymella* není kompletní. Například dle výsledků BUSCO v genomu chybí 401 ortologní genů pro Lepidoptera a 199 genů je fragmentováno. Nepřítomnost genů pro cukerné receptory, které by odpovídaly PxGr4 a PxGr44 naznačuje, že by se tyto geny mohly vyskytovat v těch částech genomu, které se nedostaly do „assembly“. Zároveň bylo očekáváno, že všechny geny budou mapovat pouze na jeden kontig, protože se jedná o genový klastr. Výskyt genů na dvou kontizích může také napovídat, že složení genomu nebylo optimální. Proto bylo dále použito prohledávání BAC knihovny *Y. evonymella* a ONT sekvenování vybraných klonů k doplnění oblasti, ve které se potenciálně vyskytují geny pro cukerné receptory.

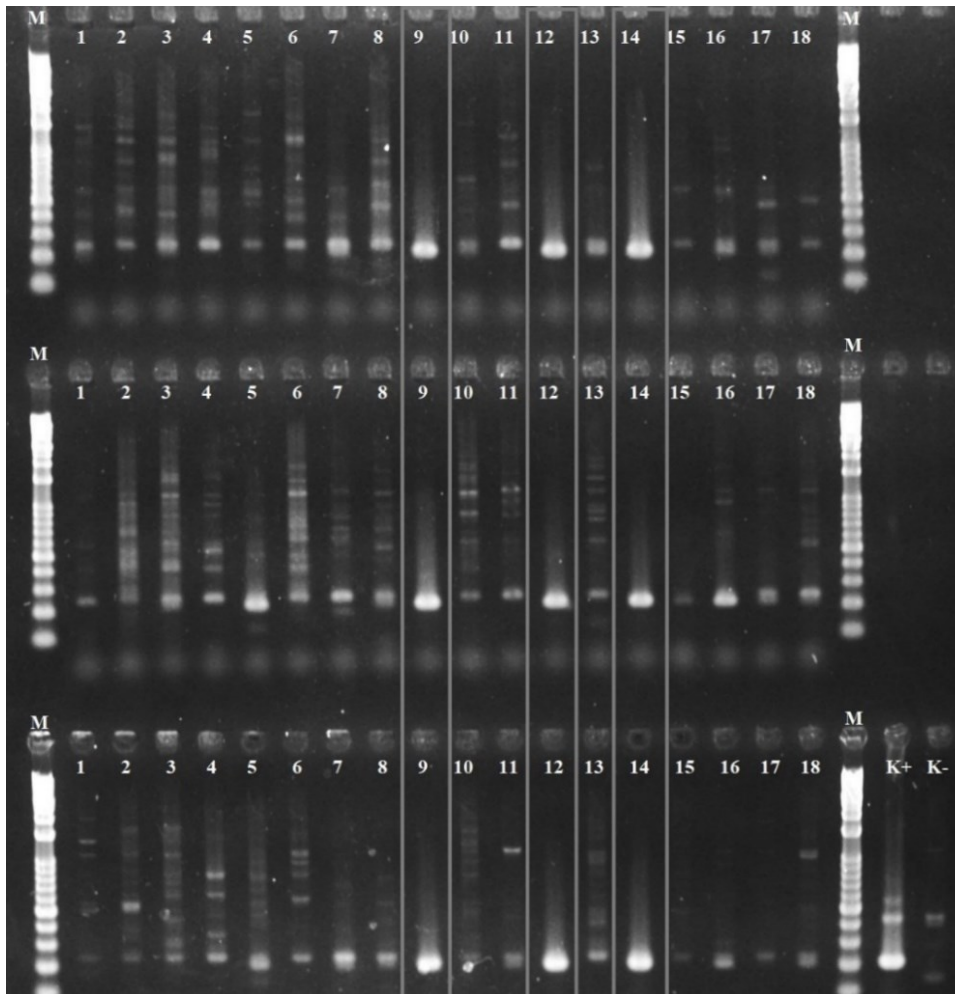
K prohledávání BAC knihovny bylo na genom *Y. evonymella* navrženo celkem osm primerů. Pro každý primer byl proveden primer test (Obr. 4) a zvolena teplota nasednutí primerů.



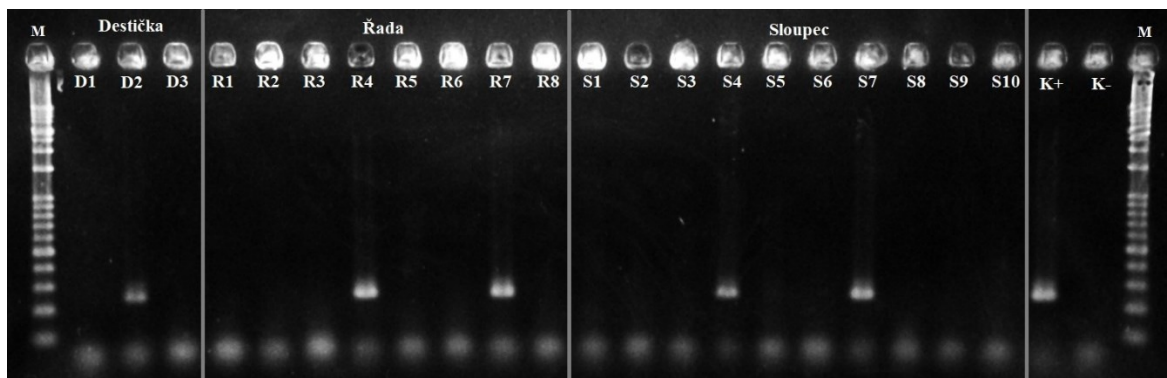
**Obrázek 4: Ukázka výsledků testování primerů.** PCR reakce byla provedena se třemi různými teplotami nasedání: 58 °C (1), 60 °C (2), 62 °C (3), K- je negativní kontrola, M je velikostní marker PCR BIO Ladder I (PCR Biosystems). Vzorek 3 pro primer 5 chybí na gelu z důvodu vyschnutí produktu v průběhu reakce.

Selekce BAC klonů probíhá v několika krocích: PCR I slouží k selekci matrixpoolů (Obr. 5), PCR II k určení přesné polohy BAC klonu v knihovně (Obr. 6). Výsledkem první PCR byla selekce celkem šesti matrixpoolů (MP 3, 9, 12, 14, 16,17), ve kterých by se potenciálně mohly vyskytovat úseky genomu nesoucí geny pro cukerné receptory. Výsledkem druhé PCR bylo určení tří BAC klonů, jeden pro PxGr1 a dva pro PxGr3 (Tab. 9). Lokalizace těchto klonů v knihovně byla určena podle klíče od dodavatele (kapitola 3.3.4., Obr. 3).

Jelikož je známo, že geny pro cukerné receptory tvoří genový klastr, kde vzdálenost mezi geny nepřesahuje 100 kbp, byla použita metoda „Chromosome walking“. Tato metoda umožňuje najít v BAC knihovně klony nesoucí části DNA, které navazují na původní úsek z obou stran. Dle dodavatele by velikost DNA vložené do bakteriálního plazmidu měla v průměru odpovídat 125 kbp. DNA před klonováním je nastříhána na náhodných místech, proto překryvy mezi molekulami mohou být desítky až tisíce bází. Pomocí chromosome walking lze teoreticky prodloužit hledaný úsek DNA až o 250 kbp, ale v praxi bývá prodloužení kratší. Pro chromosome walking byl použit pouze jeden BAC klon 35:O21. Vyizolovaná BAC DNA byla sekvenována s univerzálními primery M13. Sekvence pro primer M13 forward byla dlouhá 838 bp, pro M13 reverse 1020 bp. Dále následovalo druhé kolo prohledávání knihovny s použitím primerů navržených na oba konce tohoto klonu (Tab. 3). Během prohledávání se podařilo vyselektovat tři klony, jeden úplně nový klon 52:A24 a dva, které se shodovaly s klony pro PxGr3 z prvního kola prohledávání (Tab. 9).



**Obrázek 5: Ukázka výsledků PCR I.** Pro PCR II byly zvoleny pouze matrixpools, které byly pozitivní ve všech třech opakováních. Na obrázku je výsledek první PCR s použitím primerů pro P<sub>x</sub>Gr1 exon 5, šedými rámečky jsou znázorněny proužky, které byly určeny jako pozitivní. PCR II byla v tomto případě prováděna na matrixpoolech 9, 12 a 14. Čísla uvedená na gelu odpovídají pořadí matrixpoolů v superpoolu, K<sup>+</sup> je pozitivní kontrola (genomová DNA), K<sup>-</sup> je negativní kontrola, M je velikostní marker PCR<sub>BIO</sub> Ladder I (PCR Biosystems).



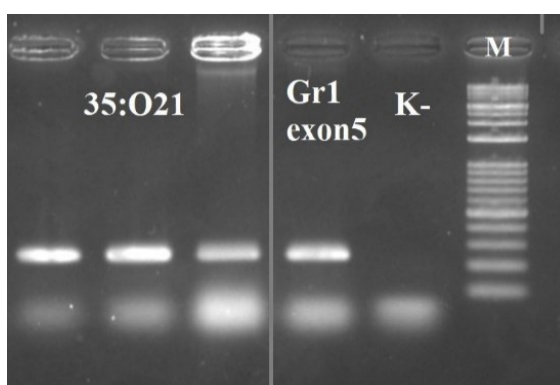
**Obrázek 6: Ukázka výsledků PCR II.** Na obrázku je výsledek druhé PCR prováděné na matrixpoolu 12 s použitím primerů pro P<sub>x</sub>Gr1 exon 5. Podstatou tohoto kroku je určení polohy BAC klonu v knihovně pomocí klíče (Obr. 3). Pozitivní proužky na gelu odpovídají číslu destičky (2), stejnému písmenu z 2. sloupce klíče (O) a stejnému číslu z 3. sloupce klíče (21). D1-D3 jsou vzorky pro určení destičky, R1-R8 pro určení řady a S1-S10 pro určení sloupce. K<sup>+</sup> je pozitivní kontrola (genomová DNA), K<sup>-</sup> je negativní kontrola, M je velikostní marker PCR<sub>BIO</sub> Ladder I (PCR Biosystems).

**Tabulka 9: Výsledky prohledávání BAC knihovny.** Umístění vyhledávaných BAC klonů v knihovně.

Metoda	Primery	Destička	Řáda	Sloupec	Název
screening	scr_e5_Gr1	35	O	21	35:O21
screening	scr_e1_Gr3	26	F	13	26:F13
screening	scr_e1_Gr3	41	M	4	41:M4
chr. walking	chrw_m13f_1	26	F	13	26:F13
chr. walking	chrw_m13f_2	41	M	4	41:M4
chr. walking	chrw_m13r_2	52	A	24	52:A24

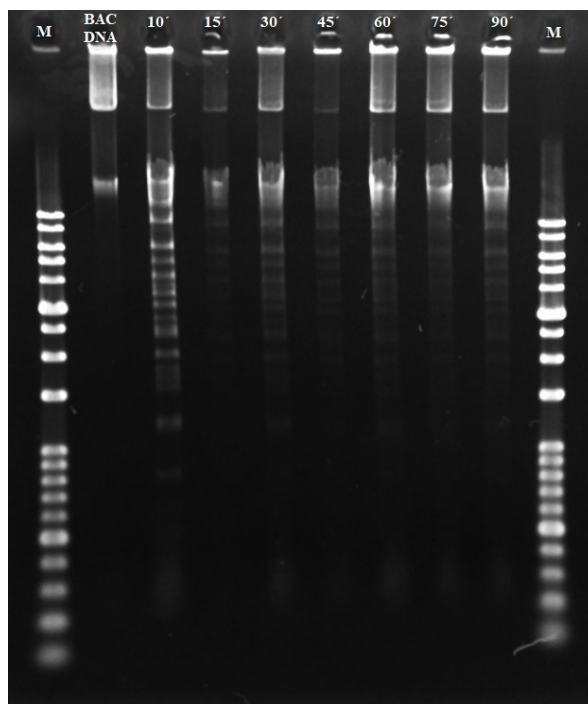
#### 4.4. Příprava BAC DNA pro ONT sekvenování

Celkově byla izolována DNA čtyř BAC klonů. Před izolací byly vybrány tři kolonie pro každý BAC klon a ověřeny na přítomnost genů pro cukerné receptory pomocí PCR (Obr. 7). Koncentrace DNA se po izolaci pohybovala mezi 244-352 ng/μl.



**Obrázek 7: Ukázka výsledků PCR pro kontrolu BAC klonů.** Gr1 exon5 je pozitivní kontrola, K- je negativní kontrola a M je velikostní marker PCRBIOLadder I (PCR Biosystems).

Před úpravou DNA bylo potřeba provést optimalizaci metody pro použití enzymu Afe I. Vzorek BAC DNA byl štěpen po dobu 10–90 minut a poté byla reakce zastavena buď přemístěním vzorku do -20 °C nebo přidáním nanášecího pufru. Všechny reakce byly následně naneseny na gel a neštěpená DNA byla rozdělena podle velikosti v elektrickém poli (Obr. 8). Jelikož pro moje účely bylo třeba pouze částečné štěpení, byly vzorky štěpeny po dobu 10 minut. Zároveň bylo zjištěno, že vzorky, do kterých byl přidán nanášecí pufr, nebyly kompatibilní s protokolem ONT. Finální reakce byly tedy po 10 minutách zastaveny okamžitým zchlazením vzorku na -20 °C. Jelikož během uprav docházelo k velkým ztrátám DNA (Tab. 10), byla určena optimální počáteční koncentrace vzorků minimálně 200 ng/μl (12 000 ng). Výsledná koncentrace připravené knihovny byla 10,2 ng/μl (153 ng).



**Obrázek 8: Optimalizace doby štěpení BAC DNA pomocí Afe I.** V první jamce je nanesen neštěpený vzorek DNA, v dalších jamkách jsou vzorky, které byly štěpeny 10–90 minut. M je velikostní marker PCRBIOLadder I (PCR Biosystems).

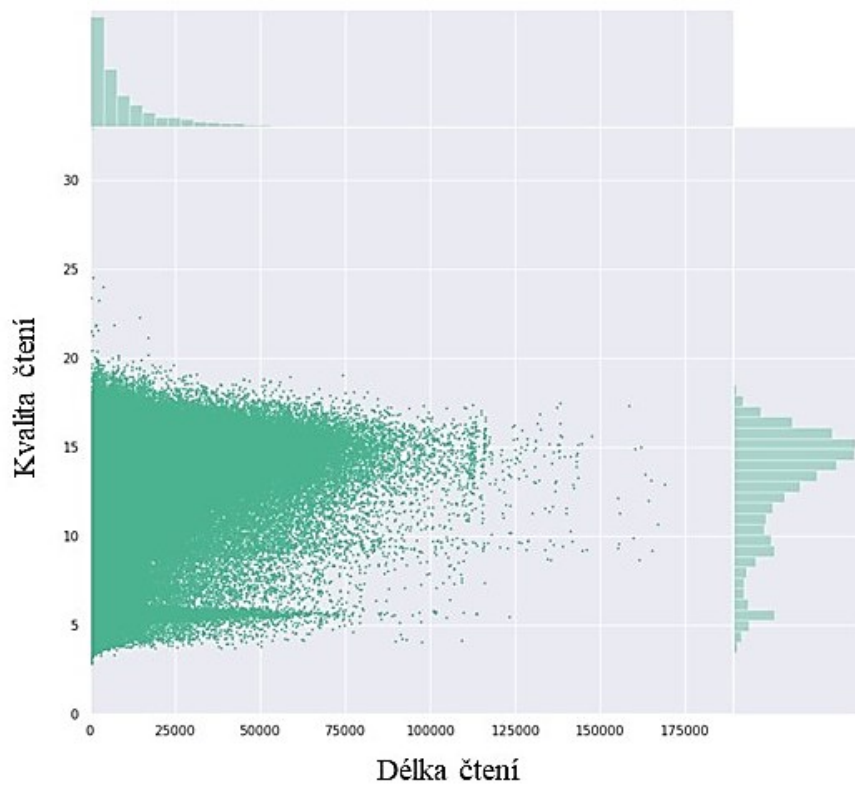
**Tabulka 10: Výsledné množství BAC DNA v jednotlivých krocích úprav.**

BAC klon	Vyizolovaná DNA (ng)	Štěpení Exonukleázou V	Selekce dlouhých fragmentů SRE (ng)	Objem DNA v knihovně (μl)
35:O21	14640	5150	3880	3,2
26:F13	21120	8900	900	13,9
41:M4	14640	6450	945	13,2
52:A24	19980	8150	635	19,7

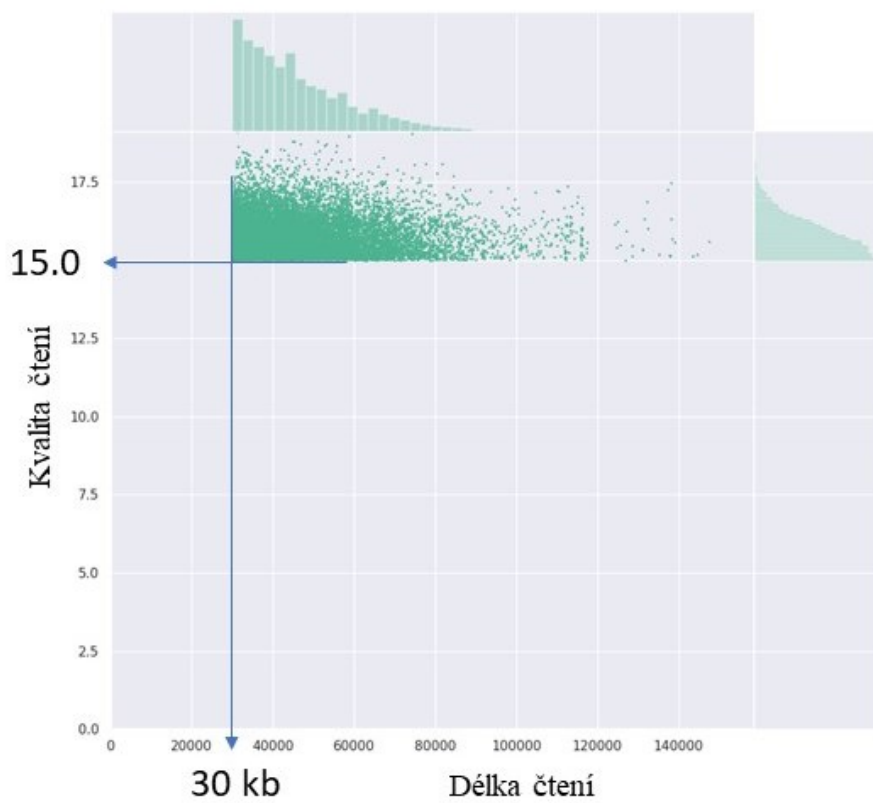
## 4.5. ONT

Komůrka před sekvenováním obsahovala 1558 funkčních pórů a sekvenování probíhalo po dobu 18 hodin a 35 minut při počátečním napětí 200 mV. Během ONT sekvenování se podařilo získat 4,88 Gb dat, která dohromady tvořila 518 tisíc čtení s odhadovanou hodnotou  $N_{50} = 26,04$  kbp. Dataset byl filtrován na délku čtení  $> 30$  kbp a kvalitu  $> Q_{15}$  v programu Nanofilt a filtrovaná data odpovídala hloubce čtení cca. 100x (Obr. 9).

A)



B)



Obrázek 9: Rozložení kvality a délky sekvencí v datasetu A) před filtrací, B) po filtraci. ONT grafy jsou vygenerované programem Nanoplot.



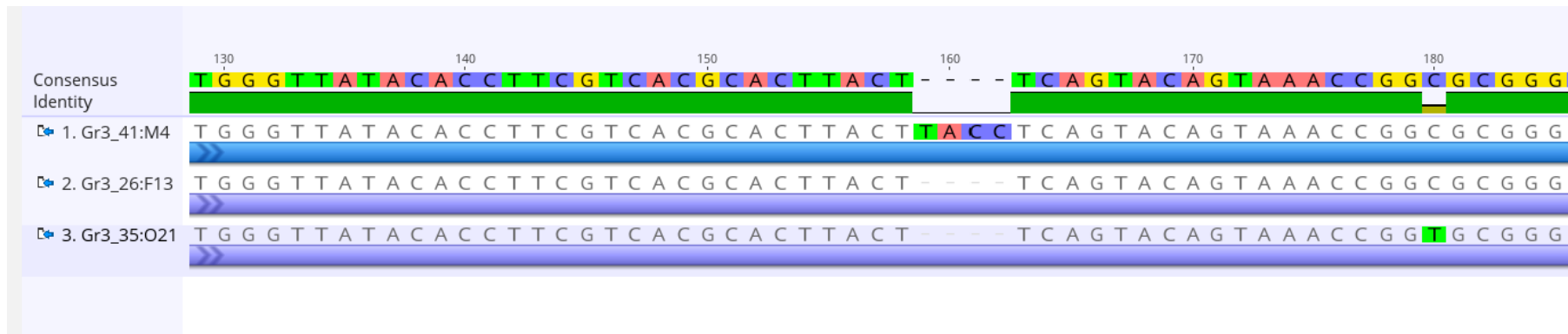
## 4.6. Mapování genů pro cukerné receptory

Sekvence získané ONT sekvenováním se podařilo složit do pěti kruhových molekul DNA. Ke každému kontigu byl přiřazen klon pomocí mapování PCR produktů použitých na prohledávání knihovny (Příloha I) a určena jeho velikost (Tab. 11). Kontig 8 byl pomocí BLASTu určen jako genomová DNA *E. coli*.

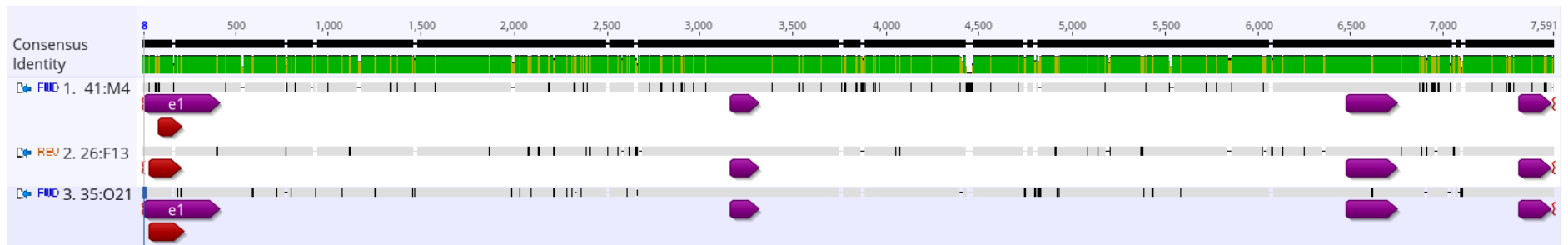
**Tabulka 11: Informace o kontizích.**

Kontig	Klon	Velikost (kbp)	Hloubka čtení
1	35:O21	198,41	77x
2	52:A24	116,79	329x
3	41:M4	113	451x
4	26:F13	111,08	367x
8	genom <i>E. coli</i>	4572,1	14x

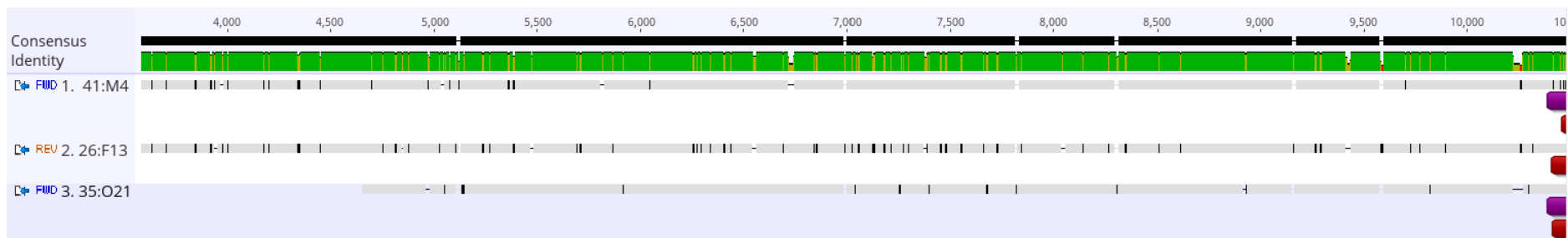
Dále byly namapovány exony genů Gr1, Gr2 a Gr3 vyextrahované ze složeného genomu *Y. evonymella* na všechny čtyři kontigy (Příloha II). Zároveň byly také namapovány geny PxGr4 a PxGr44 *P. xylostella*, které nebyly nalezeny ve složeném genomu. Gen Gr2 se znovu namapoval na oblast shodnou s Gr1 (dále je tento gen uváděn jako Gr1/2). Geny PxGr4 a PxGr44 se opět nepovedlo namapovat. Navzdory výsledkům chromosome walking 26:F13 a 41M4 se nepodařilo spojit s 35:O21 ani mezi sebou. To samé platilo i pro klon 52:A24. Tři klony 35:O21, 26:F13 a 41:M4 obsahovaly gen pro Gr3. Celková identita sekvencí těchto BAC klonů v oblasti od prvního do posledního exonu Gr3 se pohybovala v rozmezí 97,2 % až 98,4 %. V oblasti exonu 1 se 26:F13 a 35:O21 se lišily od odpovídajícího úseku na 41:M4 delecí čtyř nukleotidů (Obr. 10). V oblastech dalších exonů Gr3 se BAC klony mezi sebou liší pouze přítomností několika jednonukleotidových polymorfismů (Single Nucleotide Polymorphisms – SNP), přičemž exon2 je u všech klonů zcela identický (Obr. 11). Oblast dlouhá 5000 bp ve směru 5' od prvního exonu má mezi klony vysokou podobnost 97,8 % (Obr. 12), kdežto podobnost oblastí ve směru 3' od posledního exonu prudce klesá (Obr. 13). Již po 1500 bp pozorujeme pokles identity na 76,9% a po 5000 bp je identita sekvencí pouhých 54,2 %.



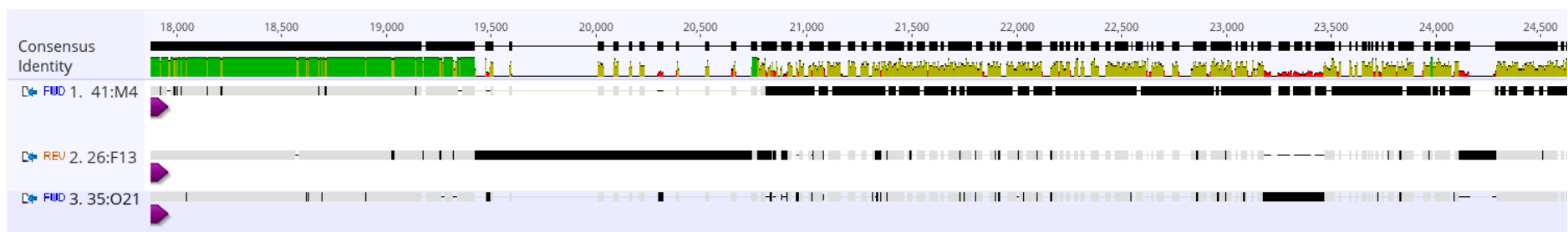
**Obrázek 10: Porovnání sekvencí DNA *Y. evonymella* odpovídajících produktu pro exon 1 Gr3.**



**Obrázek 11: Porovnání genů Gr3 mezi třemi BAC klony. Fialovou barvou jsou označeny namapované exony, červenou PCR produkty, v šedé oblasti se sekvence shodují, v černé se liší.**



**Obrázek 12: Porovnání okolních oblastí ve směru 5' od prvního exonu. V šedé oblasti se sekvence shodují, v černé se liší.**



**Obrázek 13: Porovnání okolních oblastí ve směru 3' od posledního exonu. V šedé oblasti se sekvence shodují, v černé se liší.**

## 5. Diskuze

### 5.1. Složení genomu *Y. evonymella*

Pro složení genomu *Y. evonymella* byl využit program Supernova verze 2.1.1 (Weisenfeld a kol., 2017). Tento software byl optimalizován na lidském genomu a všechna doporučení pro jeho použití platí pro savce. Savčí genomy jsou zpravidla větší než motýlí, a mohou se lišit v distribuci repetit. V práci Weisenfeld a kol. (2017) upozorňují na to, že při skládání jiných genomů než savčích, je třeba optimalizovat množství DNA, které je přidáváno do reakce pro vytvoření „10X Genomics Linked-Reads“. Dále doporučují vyzkoušet různé nastavení hloubky čtení. Hloubka, která se používá při skládání savčích genomů je 56x, to odpovídá parametru maxreads = 250 M použitému v této práci. Při porovnání složených genomů s různou hloubkou čtení v této práci, se doporučené nastavení ukázalo jako optimální i pro předivky. Avšak nejde to říci jednoznačně, je vidět, že složený genom není zcela kompletní, což by kromě parametru maxreads mohlo být ovlivněno i přípravou sekvenační knihovny (Weisenfeld a kol., 2017). Jelikož se ale jednalo o převzatá sekvenační data, detaily tohoto postupu nám nejsou známy. Je důležité zmínit, že samotný projekt Darwin tree of life používá primárně PacBio a ONT data, ke kterým mají doplňkové sekvence zpracované 10X Genomics (Bishop et al., 2021). Většinou se jedná o kombinaci dvou až čtyř metod zároveň. *Y. evonymella* je jeden z mála organismů, u kterého jsou zatím dostupná pouze 10X Genomics data.

### 5.2. Složení BAC klonů

Během skládání sekvence BAC klonů byl taktéž složen celý genom *E. coli*, který by se neměl vyskytovat v připravené sekvenační knihovně. Existují dva kroky, ve kterých mohlo dojít k chybě. Při správném průběhu experimentu by se měla veškerá bakteriální genomová DNA odstranit již během izolace plazmidové (BAC) DNA. Soupravy pro danou izolaci fungují na principu alkalické lyze (Sambrook & Russell, 2001). Nejdříve se provádí lyze buněk v alkalickém prostředí, při které je celý jejich obsah uvolněn do roztoku, během tohoto kroku dochází k denaturaci nukleových kyselin a proteinů. Nukleoidová DNA a bakteriální proteiny precipitují při neutralizaci roztoku, kdežto plazmidová DNA díky kratší délce renaturuje a zůstává tak v roztoku. Precipitát obsahující proteiny a chromosomální DNA se odstraní centrifugací a plazmidová DNA se ze supernatantu vysráží isopropanolem a promyje ethanolem. Správně by mělo dojít k odstranění vysrážené DNA dřív, než dojde k její renaturaci a zpětnému rozpuštění. Pro tento krok platí pravidlo, čím kratší je molekula DNA, tím rychleji

dochází k renaturaci. To znamená, že pokud by genomová DNA *E. coli* byla během izolace nalámána na krátké úseky, mohlo by dojít k tomu, že by zůstala v roztoku spolu s plazmidovou DNA. Taková DNA by se vyskytovala ve vzorku v lineární podobě, a proto by měla být odstraněna Exonukleázou V. Exonukleáza V byla zakoupena speciálně pro tento konkrétní experiment a vzhledem k tomu, že enzym byl nový a uchovával se v podmínkách doporučených výrobcem, je velmi malá pravděpodobnost, že by měl sníženou reaktivitu. Uvažujeme-li, že enzymatické štěpení proběhlo v pořádku, pak to znamená, že se vzorku nacházela kruhová DNA *E. coli*. Zároveň ale lze říci, že kontaminace knihovny neměla velký vliv na její kvalitu. V Tab. 11 je vidět, že nejmenší hloubka čtení u BAC klonů byla 77x a největší 451x, u genomové DNA *E. coli* pouze 14x.

Sekvenovaná knihovna byla tvořena tak, aby všechny klony byly zastoupeny ve stejném množství, proto je zvláštní, že hloubka čtení 35:O21 je pětikrát nižší, než u ostatních klonů. Tento jev lze vysvětlit tím, že klon 35:O21 je nejdelší, a to znamená, že se ve vzorku se stejným množstvím DNA (250 ng) nacházelo méně kopií tohoto klonu. Ale také lze předpokládat, že tato hloubka čtení byla ovlivněna tím, že vzorek byl kontaminován DNA *E. coli*. Při izolaci DNA měl klon 35:O21 nejnižší výtěžek a DNA tohoto klonu se nepodařilo získat v dostatečném množství ani během několika opakování izolace. Pro sekvenování byla použita DNA z několika izolací, což mohlo vést ke zvýšenému riziku kontaminace DNA *E. coli*.

### 5.3. Anotace genů pro cukerné receptory

BAC klony byly rozlišeny pomocí mapování PCR produktů (Příloha I). Dále na tyto klony byly namapovány exony genů Gr vyextrahované ze složeného genomu *Y. evonymella*. Zajímavým objevem této práce byl výskyt tří kopií genu Gr3, přičemž se tento gen lišil v počtu nukleotidů v exonu 1 mezi klonem 41:M4 a dvojicí klonů 26:F13 a 35:O21. Vzhledem k tomu, že exony jsou určeny na základě mRNA *P. xylostella*, která není kompletní (neobsahuje ani jednu z nepřekládaných oblastí ani počáteční metionin), nelze přesně říci, jaký konkrétní vliv má tato inserce na sekvenci kódovaného proteinu. Zároveň ale víme, že již první anotovaný exon pochází z kódující oblasti, lze tedy usuzovat, že inserce se projeví buď posunem čtecího rámce, a tedy výraznou změnou většiny aminokyselin, anebo vznikem stop kodonu a vznikem nefunkčního proteinu. Je tedy málo pravděpodobné, že pozorované rozdílné varianty genu Gr3 jsou důsledkem velké vnitrodruhové variability, která by se mohla v BAC knihovně objevit díky tomu, že na její vytvoření byly použity desítky jedinců z jedné populace. Přestože se celková identita anotovaných exonů a intronů pohybuje v rozmezí 97,2 % až 98,4 %, což zhruba

odpovídá nukleotidové diversitě zjištěné např. u druhu *Heliconius melpomene* (Martin et al., 2016) nebo počtu variabilních nukleotidů u *P. xylostella* (Perry et al., 2020), homologie mezi sekvencemi se ztrácí necelých 1,5 kbp od posledního anotovaného exonu, což dále naznačuje, že hypotéza vnitrodruhové variability je málo pravděpodobná. Dalším vysvětlením rozdílů v anotovaných Gr3 může být jejich recentní duplikace. Tuto hypotézu nelze s dostupnými daty dostatečně testovat. Jednou z možností, jak potvrdit možný výskyt duplikace, je provedení BAC-FISH (Fluorescenční *in situ* hybridizace se sondou vyrobenou z bakteriálního umělého chromosomu). Pomocí této metody lze namapovat vyselektované BAC klony na chromozomy *Y. evonymella*. Geny Gr se pravděpodobně budou nacházet na chr. 15 podobně jako geny pro cukerné receptory *P. xylostella*. Pokud se ukáže, že všechny BAC klony hybridizují na jeden chromozom, lze ověřit, jestli se jedná o chr. 15 pomocí specifických genových markerů (Běhálková, 2019). Pokud se tyto klony budou nacházet příliš blízko u sebe, mohou hybridizovat na stejnou oblast i kdyby se jednalo o duplikace. Je to dáno strukturou chromosomů, ve kterých je DNA hodně kondenzována a dvě oblasti, na které mapují klony, se mohou překrývat. Řešením tohoto problému by bylo použití tzv. fiber-FISH, při které se molekula DNA natáhne na sklíčko a má pak lineární podobu (Ersfeld, 2004).

Přestože geny pro Gr mají společného předka a jeví se jako ortologní pro některé skupiny, jejich sekvence se mohou lišit o 10–95 % mezi druhy (Engsontia a kol., 2014). Již během mapování genů *P. xylostella* na složený genom *Y. evonymella* bylo zjištěno, že tyto druhy mají příliš velkou mezidruhovou variabilitu. Například geny pro PxGr1 a PxGr2 byly namapovány na stejnou oblast a nebylo možné určit, o jaký paralog se jedná. Pro přesné určení genů pro cukerné receptory by byl nejvíce vhodný transkriptom *Y. evonymella*. Existují dvě možnosti sekvenování mRNA pomocí Illumina nebo ONT. Během Illumina sekvenování je mRNA překládána do cDNA (complementary DNA), nevýhodou tohoto kroku je riziko ztráty málo zastoupených transkriptů (Zucha et al., 2020), mezi které patří geny Gr. ONT na druhou stranu lze použít pro sekvenování originálních molekul RNA. Jedním z důvodů, proč víme tak málo o cukerných receptorech *Yponomeuta*, je to, že jsou jejich geny exprimovány pouze na několika neuronech. Proto je velmi obtížné vyizolovat jejich transkript nebo vytvořit preparáty pro hybridizaci genů přímo na tkáni (Dr. Peter Roessingh, Institute for Biodiversity and Ecosystem Dynamics, Amsterdam, osobní sdělení). Pro sekvenování RNA metodou Illumina je potřeba 25-1000 ng RNA (Illumina Stranded mRNA Prep protocol, Illumina, Inc., USA), pro sekvenování ONT 50 ng polyadenylované RNA nebo 500 ng veškeré RNA (Direct RNA Sequencing protocol, Oxford Nanopore, UK). Nejvhodnějším materiálem pro studium

cukerných receptorů by mohla být ústa larev *Y. evonymella*, ve kterých se nacházejí chuťové senzily aktivně exprimující geny Gr na svých neuronech (Latorre Estivalis et al., 2022). Příprava vzorku z čelistních makadel, který by měl dostačující množství materiálu k izolaci doporučeného množství RNA, by byla příliš časově a finančně náročná. Tento problém by se dal vyřešit kompromisem, tedy izolací RNA z hlav larev, což by usnadnilo preparování a zároveň by to stále zvýšilo pravděpodobnost výskytu transkriptů pro cukerné receptory oproti izolaci RNA například z celého organismu.

Existuje také několik způsobů, jak lze pokračovat v sekvenování oblasti kódující cukerné receptory. První je velmi efektivní, ale finančně náročný, jedná se o vyhledávání genů v celém genomu *Y. evonymella* složeném z dlouhých čtení s vysokou přesností (např. PacBio HiFi sequencing). Druhou možností je obohacení o okolní sekvence již známých genů pomocí tzv. „Cas9 target enrichment“, které by mohlo být provedeno z jednoho jedince (McDonald a kol., 2021). Přítomnost více než dvou variant okrajových sekvencí jednoho Gr genu by ukazovala na nedávnou duplikaci. Je to také dostupný způsob, jak dostat chybějící úseky DNA pro složení lineární sekvence obsahující geny Gr aniž by bylo potřeba sekvenovat celý genom.

## 6. Závěr

Cílem této práce bylo anotovat geny pro cukerné receptory u předivky *Y. evonymella*. Nejdříve byl složen genom *Y. evonymella* z dat dostupných na stránkách projektu Darwin tree of life. Dále byly na tento genom namapovány známé geny pro cukerné receptory referenčního druhu *P. xylostella*. Z pěti známých genů se podařilo namapovat pouze tři, přičemž dva z nich se nacházely ve stejné oblasti. Kvůli nekompletnosti genomu byla oblast s pravděpodobným výskytem genů Gr doplněna pomocí sekvenování klonů z BAC genomové knihovny metodou ONT. Kombinování několika metod sekvenování je běžné a používá se například při skládání genomů v Darwin tree of life.

Na základě mapování byly vybrány dva geny, které se jevily jako geny pro cukerné receptory a na jejich exony byly navrženy primery. Pomocí těchto primerů bylo provedeno prohledávání BAC knihovny *Y. evonymella*. Během prohledávání se podařilo vyselektovat tři BAC klony. Dále bylo rozhodnuto provést chromosome walking za účelem prodloužení finální oblasti s možným výskytem genů Gr. Celkově byla sekvenována DNA čtyř BAC klonů. Sekvence byly složeny do pěti kruhových molekul, které odpovídaly plazmidové DNA BAC klonů a genomu *E. coli*.

Jednotlivé kontigy byly přiřazeny k BAC klonům pomocí mapování PCR produktů získaných během primer testu. Dále byly na klony namapovány exony genů pro cukerné receptory vyextrahované ze složeného genomu *Y. evonymella*. Podařilo se namapovat exony odpovídající genu Gr1/2 a tři kopie genu pro Gr3. Gen Gr3 na klonech 35:O21 a 26:F13 oproti 41:M4 měl delecii čtyř nukleotidů na prvním exonu a přestože samotné geny měly mezi sebou podobnost v průměru 97,8 %, tyto tři BAC klony nebylo možné spojit dohromady. Všechny tyto poznatky poukazují na možnou recentní duplikace genů pro cukerné receptory. Ačkoliv experimentálně získaná data nejsou dostatečná pro určení kompletní sekvence genů, tato práce přinesla motivaci pro pokračování ve výzkumu cukerných receptorů *Y. evonymella*.



## 7. Literatura

Adewale, B. A. (2020). Will long-read sequencing technologies replace short-read sequencing technologies in the next 10 years? *African Journal of Laboratory Medicine*, 9(1).

<https://doi.org/10.4102/ajlm.v9i1.1340>

Agnihotri, A. R., Roy, A. A., & Joshi, R. S. (2016). Gustatory receptors in Lepidoptera: chemosensation and beyond. *Insect molecular biology*, Vol. 25, Issue 5, pp. 519–529. Blackwell Publishing Ltd. <https://doi.org/10.1111/imb.12246>

Amrein, H. (2016). Mechanism of Taste Perception in *Drosophila*. In *Chemosensory Transduction: The Detection of Odors, Tastes, and Other Chemostimuli* 245–269. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-801694-7.00014-7>

Andrews, S. (2010). FASTQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics.

Arduino, P. B. L. (1985). Reproductive isolation and genetic divergence between the small ermine moths *Yponomeuta padellus* and *Y. malinellus* (Lepidoptera: Yponomeutidae). *Atti della Accademia Nazionale dei Lincei Memorie Classe di Scienze Fisiche Matematiche e Naturali Sezione IIIA* 18(2): 33-61.

Bakker, A. C., Roessingh, P., & Menken, S. B. J. (2008). Sympatric speciation in *Yponomeuta*: no evidence for host plant fidelity. *Entomologia Experimentalis et Applicata*, 128(1), 240–247. <https://doi.org/10.1111/j.1570-7458.2008.00684.x>

Běhálková, V., 2019: Testing the synteny of genes in *Yponomeuta evonymella*. Mgr. Thesis, in English – 34 pp., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Bishop, G., Ebdon, S., Lohse, K., & Vila, R. (2021). The genome sequence of the small tortoiseshell butterfly, *Aglais urticae* (Linnaeus, 1758). *Wellcome Open Research*, 6, 233. <https://doi.org/10.12688/wellcomeopenres.17197.1>

Chen, S. T., Kowalewski, J., & Ray, A. (2021). Prolonged activation of carbon dioxide-sensitive neurons in mosquitoes. *Interface Focus*, 11(2), 20200043.

<https://doi.org/10.1098/rsfs.2020.0043>

Chyb, S., Dahanukar, A., Wickens, A., & Carlson, J. R. (2003). *Drosophila* Gr5a encodes a taste receptor tuned to trehalose. *Proceedings of the National Academy of Sciences*, 100(suppl\_2), 14526–14530. <https://doi.org/10.1073/pnas.2135339100>

Clyne, P. J., Warr, C. G., & Carlson, J. R. (2000). Candidate Taste Receptors in *Drosophila*. *Science*, 287(5459), 1830–1834. <https://doi.org/10.1126/science.287.5459.1830>

Dahanukar, A., Lei, Y. T., Kwon, J. Y., & Carlson, J. R. (2007). Two Gr Genes Underlie Sugar Reception in *Drosophila*. *Neuron*, 56(3), 503–516. <https://doi.org/10.1016/j.neuron.2007.10.024>

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2). <https://doi.org/10.1093/gigascience/giab008>

Deisig, N., Dupuy, F., Anton, S., & Renou, M. (2014). Responses to Pheromones in a Complex Odor World: Sensory Processing and Behavior. *Insects*, 5(2), 399–422. <https://doi.org/10.3390/insects5020399>

del Marmol, J., Yedlin, M. A., & Ruta, V. (2021). The structural basis of odorant recognition in insect olfactory receptors. *Nature*, 597(7874), 126–131. <https://doi.org/10.1038/s41586-021-03794-8>

Engsontia, P., Sangket, U., Chotigeat, W., & Satasook, C. (2014). Molecular evolution of the odorant and gustatory receptor genes in lepidopteran insects: Implications for their adaptation and speciation. *Journal of Molecular Evolution*, 79(1–2), 21–39. <https://doi.org/10.1007/s00239-014-9633-0>

Ernst, D. A., & Westerman, E. L. (2021). Stage- and sex-specific transcriptome analyses reveal distinctive sensory gene expression patterns in a butterfly. *BMC Genomics*, 22(1). <https://doi.org/10.1186/s12864-021-07819-4>

Ersfeld, K. (n.d.). Fiber-FISH: Fluorescence & In Situ Hybridization on Stretched DNA. *Parasite Genomics Protocols* 395–402. Humana Press. <https://doi.org/10.1385/1-59259-793-9:395>

Glendinning, J. I., Jerud, A., & Reinherz, A. T. (2007). The hungry caterpillar: An analysis of how carbohydrates stimulate feeding in *Manduca sexta*. *Journal of Experimental Biology*, 210(17), 3054–3067. <https://doi.org/10.1242/jeb.004671>

Gordon, M. D., & Scott, K. (2009). Motor Control in a *Drosophila* Taste Circuit. *Neuron*, 61(3), 373–384. <https://doi.org/10.1016/j.neuron.2008.12.033>

Guo, H., Cheng, T., Chen, Z., Jiang, L., Guo, Y., Liu, J., Li, S., Tanaii, K., Asaoka, K., Kadono-Okuda, K., Arunkumar, K. P., Wu, J., Kishino, H., Zhang, H., Seth, R. K., Gopinathan, K. P., Montagné, N., Jacquin-Joly, E., Goldsmith, M. R., ... Mita, K. (2017). Expression map of a complete set of gustatory receptor genes in chemosensory organs of *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, 82, 74–82. <https://doi.org/10.1016/j.ibmb.2017.02.001>

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, 82(11), 801–811. <https://doi.org/10.1016/j.humimm.2021.02.012>

Jain, M., Koren, S., Miga, K. H., Quick, J., Rand, A. C., Sasani, T. A., Tyson, J. R., Beggs, A. D., Dilthey, A. T., Fiddes, I. T., Malla, S., Marriott, H., Nieto, T., O’Grady, J., Olsen, H. E., Pedersen, B. S., Rhie, A., Richardson, H., Quinlan, A. R., ... Loose, M. (2018). Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology*, 36(4), 338–345. <https://doi.org/10.1038/nbt.4060>

Jiao, Y., Moon, S. J., Wang, X., Ren, Q., & Montell, C. (2008). Gr64f Is Required in Combination with Other Gustatory Receptors for Sugar Detection in *Drosophila*. *Current Biology*, 18(22), 1797–1801. <https://doi.org/10.1016/j.cub.2008.10.009>

Karger, B. L., & Guttman, A. (2009). DNA sequencing by CE. *ELECTROPHORESIS*, 30(S1), S196–S202. <https://doi.org/10.1002/elps.200900218>

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, 37(5), 540–546. <https://doi.org/10.1038/s41587-019-0072-8>

Kreklová M (2020) Analýza pohlavních chromozomů modrásků (*Lycaenidae*) [Analysis of sex chromosomes of blue butterflies (*Lycaenidae*). Mgr. Thesis, in Czech] – 58 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

Kwon, J. Y., Dahanukar, A., Weiss, L., & Carlson, J. (2014). A map of taste neuron projections in the *Drosophila* CNS. *J Biosci.* 2014 Sep; 39(4): 565–574. <https://doi.org/10.1007/s12038-014-9448-6>

Latorre Estivalis, J. M., Traverso, L., Pontes, G., & Lorenzo, M. G. (2022). The antennal transcriptome of *Triatoma infestans* reveals substantial expression changes triggered by a blood meal. *BMC Genomics*, 23(1), 861. <https://doi.org/10.1186/s12864-022-09059-6>

Lee, Y., Moon, S. J., & Montell, C. (2009). Multiple gustatory receptors required for the caffeine response in *Drosophila*. *Proceedings of the National Academy of Sciences*, 106(11), 4495–4500. <https://doi.org/10.1073/pnas.0811744106>

Letunic, I. (2002). Common exon duplication in animals and its role in alternative splicing. *Human Molecular Genetics*, 11(13), 1561–1567. <https://doi.org/10.1093/hmg/11.13.1561>

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34(18), 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>

Lofstedt, C., Herrebout, W. M., & Menken, S. B. J. (1991). Sex pheromones and their potential role in the evolution of reproductive isolation in small ermine moths (*Yponomeutidae*). *Chemoecology*, 2(1), 20–28. <https://doi.org/10.1007/BF01240662>

Manni, M., Berkeley, M. R., Seppey, M., & Zdobnov, E. M. (2021). BUSCO: Assessing Genomic Data Quality and Beyond. *Current Protocols*, 1(12). <https://doi.org/10.1002/cpz1.323>

Martin, S. H., Möst, M., Palmer, W. J., Salazar, C., McMillan, W. O., Jiggins, F. M., & Jiggins, C. D. (2016). Natural Selection and Genetic Diversity in the Butterfly *Heliconius melpomene*. *Genetics*, 203(1), 525–541. <https://doi.org/10.1534/genetics.115.183285>

McDonald, T. L., Zhou, W., Castro, C. P., Mumm, C., Switzenberg, J. A., Mills, R. E., & Boyle, A. P. (2021). Cas9 targeted enrichment of mobile elements using nanopore sequencing. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-23918-y>

Menken, S. B. J., Herrebout, W. M., & Wiebes, J. T. (1992). (YPONOMEUTA): Their Host Relations and Evolution. *Annu, Rev, Entomol*, Vol. 37.

Miyamoto, T., Slone, J., Song, X., & Amrein, H. (2012). A fructose receptor functions as a nutrient sensor in the *Drosophila* brain. *Cell*, 151(5), 1113–1125. <https://doi.org/10.1016/j.cell.2012.10.024>

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, 34(13), i142–i150. <https://doi.org/10.1093/bioinformatics/bty266>

Mombaerts, P. (1999). Seven-Transmembrane Proteins as Odorant and Chemosensory Receptors. *Science*, 286(5440), 707–711. <https://doi.org/10.1126/science.286.5440.707>

Maxam, A. M., & Gilbert, W. (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Sciences*, 74(2), 560–564. <https://doi.org/10.1073/pnas.74.2.560>

Ômura, H., Honda, K., Asaoka, K., & Inoue, T. A. (2008). Tolerance to fermentation products in sugar reception: Gustatory adaptation of adult butterfly proboscis for feeding on rotting foods. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 194(6), 545–555. <https://doi.org/10.1007/s00359-008-0330-6>

Ozaki, K., Ryuda, M., Yamada, A., Utoguchi, A., Ishimoto, H., Calas, D., Marion-Poll, F., Tanimura, T., & Yoshikawa, H. (2011). A gustatory receptor involved in host plant recognition for oviposition of a swallowtail butterfly. *Nature Communications*, 2(1), 542. <https://doi.org/10.1038/ncomms1548>

Perry, K. D., Keller, M. A., & Baxter, S. W. (2020). Genome-wide analysis of diamondback moth, *Plutella xylostella* L., from Brassica crops and wild host plants reveals no genetic structure in Australia. *Scientific Reports*, 10(1), 12047. <https://doi.org/10.1038/s41598-020-68140-w>

Piovesan, A., Pelleri, M. C., Antonaros, F., Strippoli, P., Caracausi, M., & Vitale, L. (2019). On the length, weight and GC content of the human genome. *BMC Research Notes*, 12(1), 106. <https://doi.org/10.1186/s13104-019-4137-z>

Povel, G. D. E. (1984). The identification of the European small ermine moths, with special reference to the *Yponomeuta padellus*-complex (Lepidoptera, Yponomeutidae). *Proc. K. Ned. Akad. Wet. Ser. C* 87: 149-80

Roessingh, P., Hora K. H., van Loon J. J. A., & Menken S. B. J. (1999). Evolution of gustatory sensitivity in *Yponomeuta* caterpillars. *Journal of Comparative Physiology A* volume 184, 119–126

Sambrook, J., & Russell, D. W. (2001). *Molecular cloning : a laboratory manual*. Cold Spring Harbor (3 vol.).

Sánchez-Gracia, A., Vieira, F. G., & Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity*, Vol. 103, Issue 3, 208–216. <https://doi.org/10.1038/hdy.2009.55>

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12), 5463–5467. <https://doi.org/10.1073/pnas.74.12.5463>

Sankoff, D. (2001). Gene and genome duplication. *Current Opinion in Genetics & Development*, 11(6), 681–684. [https://doi.org/10.1016/S0959-437X\(00\)00253-7](https://doi.org/10.1016/S0959-437X(00)00253-7)

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B., & Touhara, K. (2008). Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature*, 452(7190), 1002–1006. <https://doi.org/10.1038/nature06850>

Sato, K., Tanaka, K., & Touhara, K. (2011). Sugar-regulated cation channel formed by an insect gustatory receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 108(28), 11680–11685. <https://doi.org/10.1073/pnas.1019622108>

Seeley, T. D. (1974). Atmospheric carbon dioxide regulation in honey-bee (*Apis mellifera*) colonies. *Journal of Insect Physiology*, 20(11), 2301–2305. [https://doi.org/10.1016/0022-1910\(74\)90052-3](https://doi.org/10.1016/0022-1910(74)90052-3)

Scott, K., Brady Jr, R., Cravchik, A., Morozov, Pavel, Rzhetsky, A., Zuker, C., & Axel, R. (2001). *A Chemosensory Gene Family Encoding Candidate Gustatory and Olfactory Receptors in Drosophila*. Columbia University Irving Medical Center.

Sherathiya, V. N., Schaid, M. D., Seiler, J. L., Lopez, G. C., & Lerner, T. N. (2021). GuPPy, a Python toolbox for the analysis of fiber photometry data. *Scientific Reports*, 11(1), 24212. <https://doi.org/10.1038/s41598-021-03626-9>

Slone, J., Daniels, J., & Amrein, H. (2007). Sugar Receptors in *Drosophila*. *Current Biology*, 17(20), 1809–1816. <https://doi.org/10.1016/j.cub.2007.09.027>

Tang, Q. B., Zhan, H., Cao, H., Berg, B. G., Yan, F. M., & Zhao, X. C. (2014). Central projections of gustatory receptor neurons in the medial and the lateral sensilla styloconica of *Helicoverpa armigera* larvae. *PLoS ONE*, 9(4). <https://doi.org/10.1371/journal.pone.0095401>

Thom, C., Guerenstein, P. G., Mechaber, W. L., & Hildebrand, J. G. (2004). Floral CO<sub>2</sub> Reveals Flower Profitability to Moths. *Journal of Chemical Ecology*, 30(6), 1285–1288. <https://doi.org/10.1023/B:JOEC.0000030298.77377.7d>

Turner, H., Lieshout, N., van Ginkel, W. E., & Menken, S. B. J. (2010). Molecular phylogeny of the small ermine moth genus *Yponomeuta* (Lepidoptera, Yponomeutidae) in the Palearctic. *PLoS ONE*, 5(3). <https://doi.org/10.1371/journal.pone.0009933>

Udaondo, Z., Sittikankaew, K., Uengwetwanit, T., Wongsurawat, T., Sonthirod, C., Jenjaroenpun, P., Pootakham, W., Karoonuthaisiri, N., & Nookaew, I. (2021). Comparative Analysis of PacBio and Oxford Nanopore Sequencing Technologies for Transcriptomic Landscape Identification of *Penaeus monodon*. *Life*, 11(8), 862.

<https://doi.org/10.3390/life11080862>

Vácha M., Fellnerová I., Bičík V., Petrásek R., Šimek V. (2010): Srovnávací fyziologie živočichů. Skripta. 2. upravené vydání. Masarykova univerzita Brno (ISBN 978-80-210-3379-5).

van der Pers, J. N. C., Thomas, G., & den Otter, C. J. (1980). Interactions between plant odours and pheromone reception in small ermine moths (Lepidoptera: Yponomeutidae). *Chemical Senses*, 5(4), 367–371. <https://doi.org/10.1093/chemse/5.4.367>

Wada-Katsumata, A., Silverman, J., & Schal, C. (2013). Changes in Taste Neurons Support the Emergence of an Adaptive Behavior in Cockroaches. *Science*, 340(6135), 972–975. <https://doi.org/10.1126/science.1234854> "

Wanner, K. W., & Robertson, H. M. (2008). The gustatory receptor family in the silkworm moth *Bombyx mori* is characterized by a large expansion of a single lineage of putative bitter receptors. *Insect Molecular Biology*, 17(6), 621–629.

<https://doi.org/10.1111/j.1365-2583.2008.00836.x>

Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2017). Direct determination of diploid genome sequences. *Genome Research*, 27(5), 757–767.

<https://doi.org/10.1101/gr.214874.116>

Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P.-C., Hall, R. J., Concepcion, G. T., Ebler, J., Functammasan, A., Kolesnikov, A., Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C.-S., Phillippy, A. M., Schatz, M. C., Myers, G., DePristo, M. A., ... Hunkapiller, M. W. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, 37(10), 1155–1162. <https://doi.org/10.1038/s41587-019-0217-9>

WU, R. (1972). Nucleotide Sequence Analysis of DNA. *Nature New Biology*, 236(68), 198–200. <https://doi.org/10.1038/newbio236198a0>

Xu, W. (2020). How do moth and butterfly taste?—Molecular basis of gustatory receptors in Lepidoptera. In *Insect Science*, Vol. 27, Issue 6, pp. 1148–1157.

<https://doi.org/10.1111/1744-7917.12718>

Yarmolinsky, D. A., Zuker, C. S., & Ryba, N. J. P. (2009). Common Sense about Taste: From Mammals to Insects. In *Cell*, Vol. 139, Issue 2, pp. 234–244. Elsevier B.V.

<https://doi.org/10.1016/j.cell.2009.10.001>

You, M.S., Yue, Z., He, W.Y., Yang, X.H., Yang, G., Xie, M. a kol. (2013) A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics*, 45, 220–225.

Zhang, H. J., Faucher, C. P., Anderson, A., Berna, A. Z., Trowell, S., Chen, Q. M., Xia, Q. Y., & Chyb, S. (2013). Comparisons of Contact Chemoreception and Food Acceptance by Larvae of Polyphagous *Helicoverpa armigera* and Oligophagous *Bombyx mori*. *Journal of Chemical Ecology*, 39(8), 1070–1080. <https://doi.org/10.1007/s10886-013-0303-2>

Zhang, H.-J., Anderson, A. R., Trowell, S. C., Luo, A.-R., Xiang, Z.-H., & Xia, Q.-Y. (2011). Topological and Functional Characterization of an Insect Gustatory Receptor. *PLoS ONE*, 6(8), e24111. <https://doi.org/10.1371/journal.pone.0024111>

Zucha, D., Androvic, P., Kubista, M., & Valihrach, L. (2020). Performance Comparison of Reverse Transcriptases for Single-Cell Studies. *Clinical Chemistry*, 66(1), 217–228. <https://doi.org/10.1373/clinchem.2019.307835>



## 8. Přílohy

**Příloha I:** Výsledky Sangerovo sekvenování během klon testu.

>scr13\_26:F13\_produk

```
GCCGGAATATCGCACGAACTCCTAACACAACCTCTCTCCATTGCTCGTKGGGSGAC
CTTAAGCCTTCTTATCAGGCCCATAGTCACCGACCAGGWCTCACATCCAAGTCAT
CACTGGCAACACACATGGGTTATACACCTTCGTCACGCACTTACTTCAGTACAGT
AAACCGGCGCGGGCACGGAATAAA
```

>scr13\_41:M4\_produk

```
TCCTGACACRACTCTCTCCAKTGCTCGTTGGGCGACSTTGAGCCTKCTKATCAGAC
CCATAGTCAGCGACCAGGTCTCACATCCMAGTCATCACTGGCAACACACATGGGT
TATACACCTTCGWCACGCACTTACTTACCTCAGTACAGTAAACCGGCGCGGGCAC
GGAATAAAGCGTTTCGGCGGA
```

>scr51\_35:O21\_produk

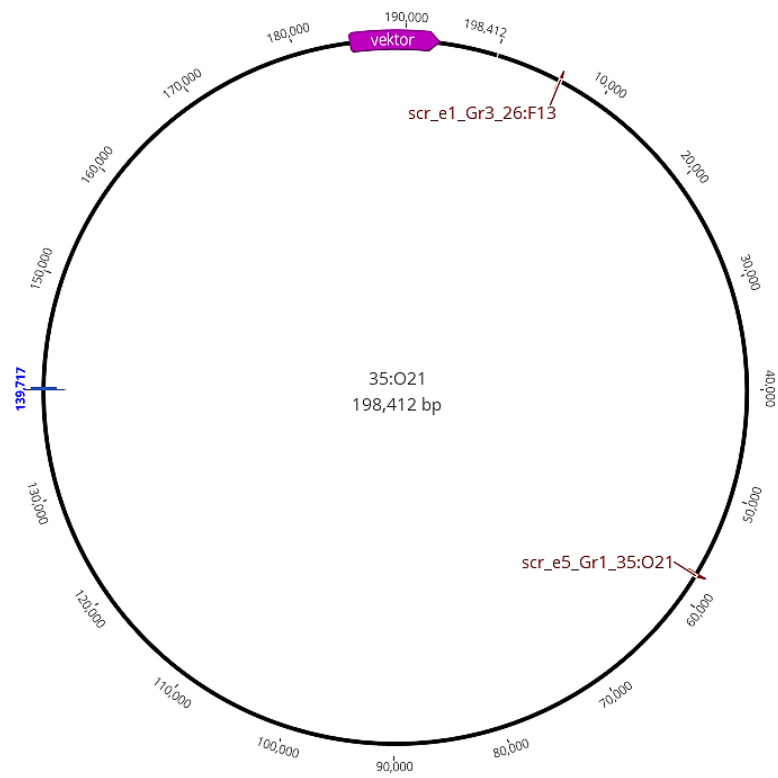
```
ATGCTATTGACTGGACTGACTGAMATYGRGAAKAYGATGCTTTACTARGARACA
ATATCWCCATTTCCWWTTTCACGKAACTACTACTTACTTGCAATTTTCAKYGAT
GKATCRWATTAMACTTTATACATGCATATTGTTGTTAKAAACARCTGACCTATCA
GKGAAACMAACGCRTACACSCWTCTCCA
```

>chrw\_52:A24\_produk

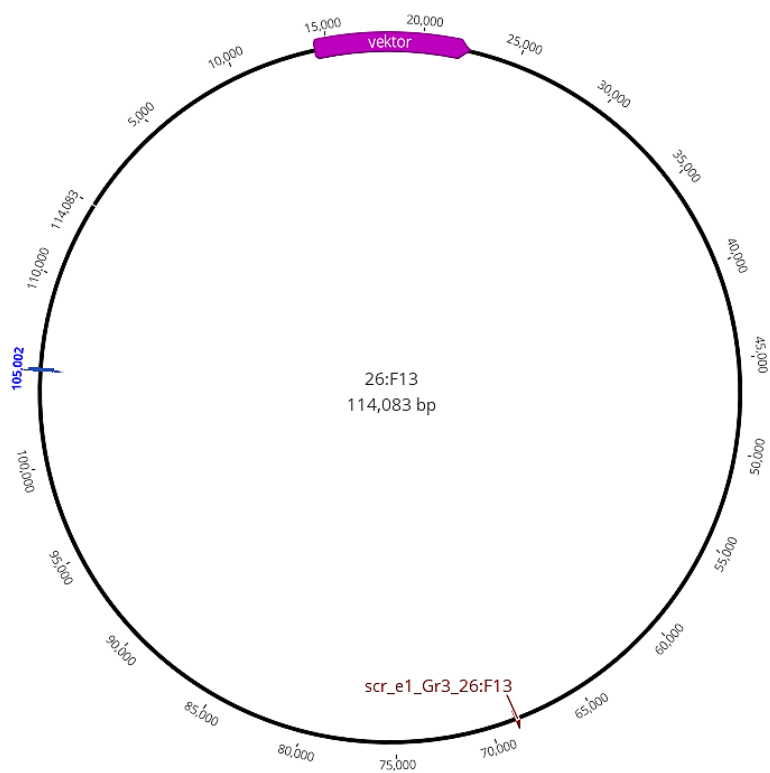
```
ARCCAGCTTCGATCGTTGATGTACTASCGTTAGGCATAKACGGMGACCCTTCCAT
GCWCTGAACGCGCCCTGTGTTGCTTAACCTTCGTGATACCTTAAGATCACTCAGG
TTACGGGGCTACGAGCGATTCCGAAAACCCGGAAAATTTTATTTATTGTAAACGA
ATTTTAAGGAATTTTTTACAAAAATGAGTTCAATGCAAGKATGATTAGCCGAATA
TTATCWTTGCTGGAGAMWTTTTTCCGAGYGGATCCTCGTTTGKMAWGGCCTGTG
AGGYTKTTGGTA
```

**Příloha II: Anotace cukerných receptorů a úseků použitých pro screening na DNA BAC klonů.**

A)

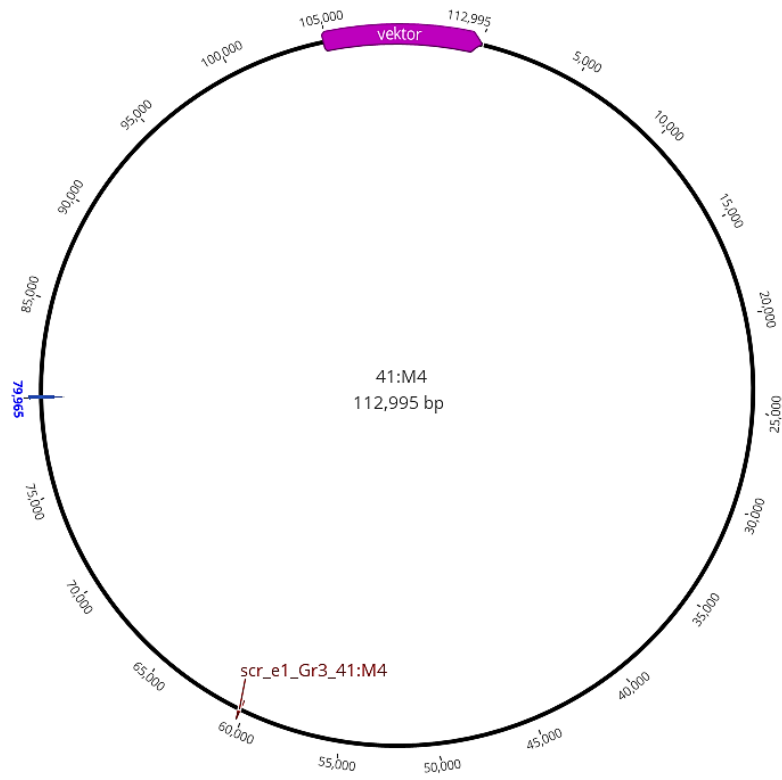


B)

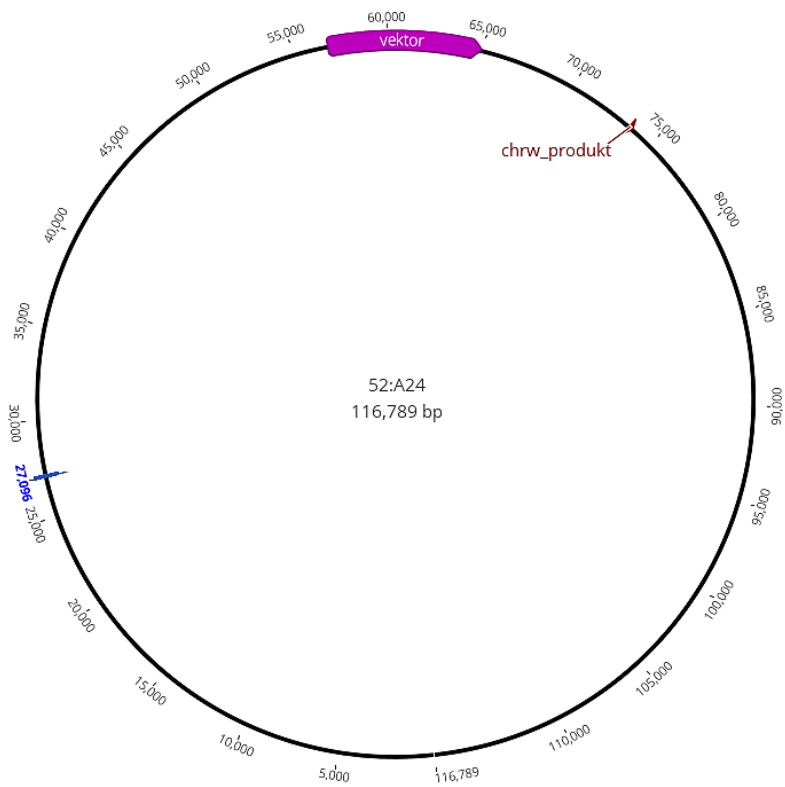


**Obrázek 1: PCR produkty namapované na BAC klony A) 35:O21 obsahuje produkt pro Gr1 a Gr3, B) 26:F13 obsahuje produkt pro Gr3.**

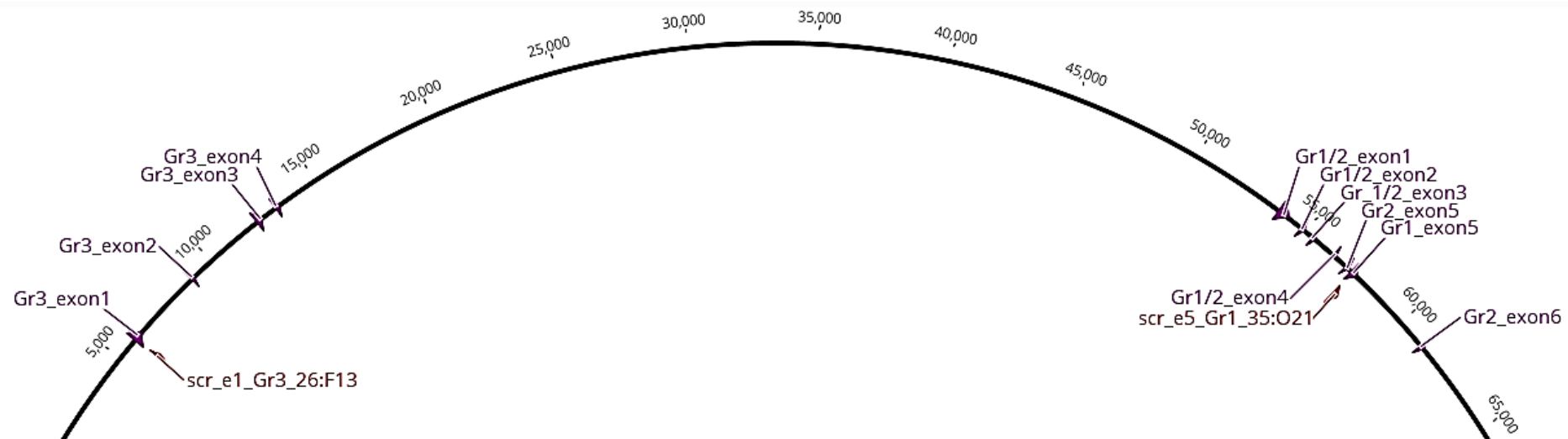
C)



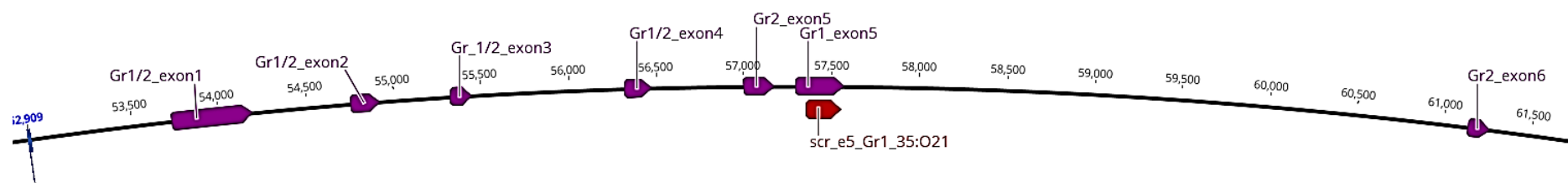
D)



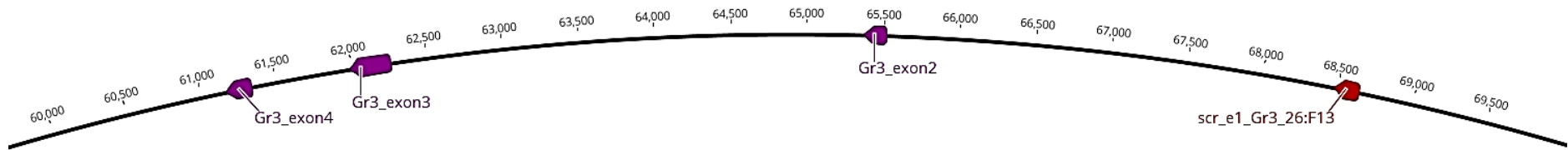
**Obrázek 2: PCR produkty namapované na BAC klony C) 41:M4 obsahuje produkt pro Gr3, D) 52A24 obsahuje pouze produkt pro chromosome walking (chrw\_52:A24\_produkut)**



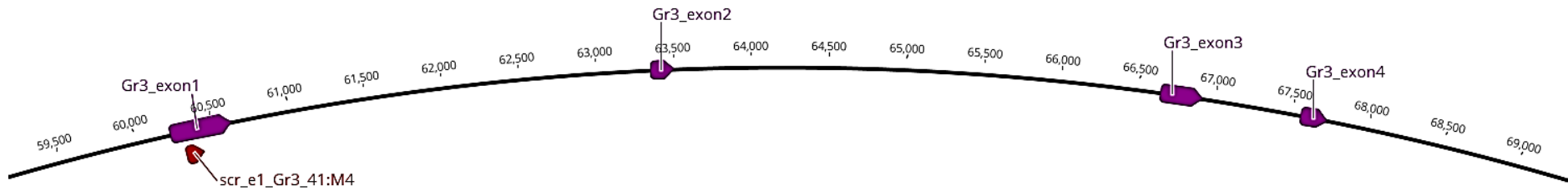
Obrázek 3: Mapování exonů genů pro cukerné receptory na klon 35:O21.



Obrázek 4: Mapování exonů genů pro cukerné receptory na klon 35:O21, obrázek je přiblížen na gen pro Gr1/2.



Obrázek 5: Mapování exonů genu pro cukrný receptor Gr3 na klon 26:F13.



Obrázek 6: Mapování exonů genu pro cukrný receptor Gr3 na klon 41:M4.