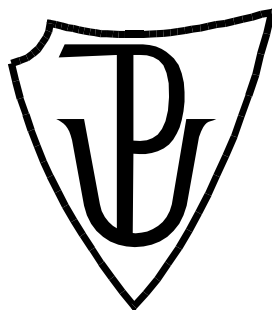


# UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Katedra biochemie



**Pigmentace trichomů sóji luštinaté *Glycine max* [L.]**

**Merr. jako důležitý agronomický znak**

## BAKALÁŘSKÁ PRÁCE

Autor:	<b>Adéla Franková</b>
Studijní program:	B1406 Biochemie
Studijní obor:	Biochemie
Forma studia:	Prezenční
Vedoucí práce:	<b>Mgr. Mária Škrabišová, Ph.D.</b>
Rok:	2021

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně s vyznačením všech použitých pramenů a spoluautorství. Souhlasím se zveřejněním bakalářské práce podle zákona č. 111/1998 Sb., o vysokých školách, ve znění pozdějších předpisů. Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, ve znění pozdějších předpisů.

V Olomouci dne .....

Ráda bych touto cestou poděkovala své vedoucí bakalářské práce, Mgr. Márii Škrabišové, Ph.D., za poskytnutí potřebných informací a literatury, za její odborné vedení, čas a trpělivost při psaní této práce. Také děkuji za grantovou podporu projektu IGA\_PrF\_2020\_013.

## Bibliografická identifikace

Jméno a příjmení autora	Adéla Franková
Název práce	Pigmentace trichomů sóji luštinaté <i>Glycine max</i> [L.] Merr. jako důležitý agronomický znak
Typ práce	Bakalářská
Pracoviště	Katedra biochemie
Vedoucí práce	Mgr. Mária Škrabišová, Ph.D.
Rok obhajoby práce	2021

### Abstrakt

Tato bakalářská práce se zabývá trichomy, jejich vlivem na albedo, biosyntetickou drahou flavonoidů, jejich regulací na úrovni transkripce a jejich účinky na lidské zdraví. V teoretické části jsem popsala publikované alely a geny zodpovědné za pigmentaci trichomů. Popsané mutace v těchto genech však nevysvětlují fenotyp u všech variet. Cílem praktické části je najít a ověřit kandidátní geny, které by dovysvětlily pigmentaci trichomů. K tomu jsem použila bioinformatické nástroje SNPviz v2.0, AccuTool, BLAST a Katalog alel sóji. Za použití zmíněných nástrojů jsem vyloučila původně navržený publikovaný gen *Glyma.16g016500*. Naopak jsem našla statisticky významnější mutace v genomických oblastech asociovaných s pigmentací, tyto geny jsem navrhla jako kandidátní.

Klíčová slova	Sója, <i>Glycine max</i> , trichomy, flavonoidy, albedo, přirozená genetická variabilita
Počet stran	64
Počet příloh	0
Jazyk	Český

## Bibliographical identification

Autor's name	Adéla Franková
Title	Pigmentation of trichomes in soybean <i>Glycine max</i> [L.] Merr. as an important agronomical trait
Type of thesis	Bachelor
Department	Department of biochemistry
Supervisor	Mgr. Mária Škrabišová, Ph.D.
The year of presentation	2021

### Abstract

This bachelor thesis deals with a study of trichomes of *Glycine max*, their effect to albedo, the biosynthesis of flavonoid, regulation of this biosynthesis at the transcriptional level and effect of flavonoids to human health. The theoretical part deals with published alleles and genes which are responsible for the pigmentation of trichomes. However, not all of the phenotypic variance is explained by published alleles and genes. The goal of the practical part is to find and verify candidates genes, which would explain the remaining phenotype. In this study there were used bioinformatical tools such as SNPviz v2.0, AccuTool, BLAST and Soybean Allele Catalogue. By using the mentioned tools, the proposed gene *Glyma.16g016500* was excluded. I found statistically more significant mutations in the genomic regions associated with pigmentation, these genes were proposed as candidates.

Keywords	Soybean, <i>Glycine max</i> , trichomes, flavonoids, albedo, natural genetic variation
Number of pages	64
Number of appendices	0
Language	Czech

## OBSAH

<b>1</b>	<b>ÚVOD .....</b>	<b>9</b>
<b>2</b>	<b>SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY .....</b>	<b>11</b>
2.1	Sója luštinatá .....	11
2.1.1	Agronomické znaky sóji .....	11
2.2	Trichomy .....	12
2.2.1	Pigmentace trichomů.....	13
2.2.2	Flavonoid-3'-hydroxylasa.....	14
2.2.3	MYB faktor .....	15
2.3	Flavonoidy .....	16
2.3.1	Biosyntéza flavonů/flavonoidů .....	16
2.3.2	Degradace chalkonů .....	19
2.3.3	Vliv konzumace pigmentů na lidské zdraví .....	20
2.3.4	Výhodné znaky spojené s šedými trichomy .....	22
2.4	Albedo plodin .....	23
2.4.1	Vliv barvy na albedo .....	24
2.5	Fenotypová změna jako důsledek přirozených genetických variant .....	24
2.6	GWAS .....	26
2.6.1	Identifikace alternativních mutací v <i>t</i> alele .....	28
2.6.2	Mutace v <i>Td</i> alele .....	30
<b>3</b>	<b>EXPERIMENTÁLNÍ ČÁST .....</b>	<b>31</b>
3.1	Návrh experimentů .....	31
3.2	Provedené experimenty .....	33
3.2.1	SNPViz v2.0.....	33
3.2.2	AccuTool.....	34
3.2.3	BLAST .....	35
3.2.4	Katalog alel sóji.....	35
3.2.5	Analýza zastoupení jednotlivých fenotypů pigmentace trichomů .....	36
3.3	Plánované experimenty .....	37
3.3.1	Potřebné chemikálie .....	37
3.3.2	Přístroje a vybavení.....	37
3.3.3	Roztoky .....	37
3.3.4	Reagencie pro molekulární biologii .....	37
3.3.5	Biologický materiál.....	38

3.3.6	Navržení primerů .....	40
3.3.7	Karty typu FTA®, Whatman™ .....	41
3.3.8	Příprava vzorku pro PCR .....	42
3.3.9	PCR .....	42
3.3.10	Elektroforéza v agarosovém gelu.....	44
3.3.11	Přečistění produktu PCR.....	45
3.3.12	Spektrofotometrické stanovení koncentrace DNA.....	45
3.3.13	Sekvenování Sangerovou metodou .....	46
3.3.14	Analýza výsledků pomocí programu BioEdit .....	46
<b>4</b>	<b>VÝSLEDKY A DISKUZE .....</b>	<b>47</b>
4.1	Analýza zastoupení jednotlivých fenotypů pigmentace trichomů variet sóji z databáze GRIN .....	47
4.2	Ověření případných alternativních <i>td</i> alel pomocí SNPviz v2.0 .....	50
4.3	Analýza přesnosti asociace fenotypu pro mutace v kandidátních genech pigmentace trichomů pomocí AccuTool .....	52
4.4	Identifikace genů kódujících tvorbu pigmentů v sóji pomocí Katalogu alel sóji. ....	54
4.4.1	Ověření kandidátního genu <i>Glyma.16g016500</i> pro tvorbu pigmentů v sóji. ....	55
<b>5</b>	<b>ZÁVĚR .....</b>	<b>57</b>
<b>6</b>	<b>LITERATURA .....</b>	<b>58</b>
<b>7</b>	<b>SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....</b>	<b>62</b>

## **CÍLE PRÁCE**

### **Teoretická část**

V teoretické části se věnuji rešerši na téma pigmentace trichomů: rozdílné akumulace jednotlivých pigmentů v odlišně zbarvených trichomech, biosyntetickým drahám těchto pigmentů a regulaci těchto drah na úrovni transkripce.

- trichomy a jejich vliv na albedo
- známé alely a geny a jejich funkční varianty týkající se různého zbarvení trichomů
- metabolická dráha flavonoidů

### **Praktická část**

Ověření přítomnosti identifikovaných alternativních alel u linií s různou intenzitou šedého zbarvení trichomů. Identifikace kandidátních genů zodpovědných za různé zbarvení trichomů u variet s dominantní alelou *T* a *Td*.



## 1 ÚVOD

Sója luštinatá je jednoletá dvouděložná rostlina. Patří mezi nejrozšířenější olejninu na světě. *Glycine max* (L.) Merr., byla domestikována v Číně a Koreji z jejího předka *G. soja* Sieb. et Zucc. Představuje asi polovinu celosvětové produkce rostlinných olejů, ale slouží také k výrobě bionafty a mýdel. Hlavně je ale určena pro potravinářské účely, jako například pro výrobu sójového mléka, nattó, tempeh, tofu, sójové omáčky, sójové mouky, sójového masa nebo jako čistý bílkovinný izolát s celou řadou aplikací.

Je velmi žádaná v Japonsku, ale poptávka po sóji se celosvětově neustále zvyšuje. Ekonomicky výhodnější je sóju produkovat než ji dovážet. Ačkoli je sója přizpůsobená mírným i tropickým oblastem, měnící se klima vyžaduje lepší pochopení adaptace na specifické podmínky životního prostředí. Limitujícím faktorem pro pěstování v krajinách s velkým rozdílem délky denního a nočního světla je právě fakt, že sója je rostlina krátkého dne citlivá na fotoperiodu (Bu *et al.*, 2020). Kromě adaptace sóji na délku denního osvětlení (přesněji délky noci) je nejdůležitějším faktorem pro adaptaci sóji také náročnost na vláhu a teplotu (Bandillo *et al.*, 2017). Tyto faktory ovlivňuje především morfologie trichomů. Bylo zjištěno, že variety sóji s kudrnatým typem trichomů, žlutohnědými trichomy a linie s malou hustotou trichomů na plochu odrážejí záření více než variety se vzpřímenými či přilehlými trichomy, šedými trichomy a větší hustotou trichomů (Doughty *et al.*, 2011). Pokud by byly ve vyšších zemědělských šířkách vyměněny rostliny s trichomy za rostliny bez trichomů, došlo by teoreticky ke snížení teploty o ~0,40 °C (Doughty *et al.*, 2011).

Za šedou, žlutohnědou, světle žlutohnědou a téměř šedou barvu trichomů může obsah flavonů a hydroxylace na jejich B kruhu. Čím méně jich trichomy obsahují, tím je jejich barva méně výrazná. Řídí ji dominantní alela *T* (Toda *et al.*, 2002), která mění strukturu flavonů hydroxylací B kruhu a poskytuje deriváty luteolinu, což způsobuje žlutohnědé zbarvení trichomů. Alela *Td* podporuje také žlutohnědé zbarvení, produkuje více luteolinu a může kontrolovat množství flavonů (Bernard, 1975). Za světle žlutohnědou až šedou barvu je zodpovědná alela *td* s malým množstvím luteolinu. Recesivní alela *t* produkuje naopak deriváty apigeninu a šedé zbarvení (Yan *et al.*, 2020).

Ačkoli lokusy pigmentace trichomů jsou známy již z minulého století a *T* alela byla identifikována před necelými 20 lety, alely *Td* byly identifikovány metodou GWAS

(Genome-wide association study), což je metoda asociace fenotypu na genotyp. Využití této metody pro žlutohnědé, světle žlutohnědé, téměř šedé a šedé fenotypy poukázalo na asociaci na chromosomu 6 – kde se nachází gen *T* alely a chromosomu 3 – kde se nachází gen *Td* alely. Další (a statisticky významnější) asociace byly pozorovány na chromosomu 12 a 16, kde doposud žádné geny pro pigmentaci trichomů identifikovány nebyly.

## 2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

### 2.1 Sója luštinatá

Sója je důležitým zdrojem proteinů a olejů jak pro lidi, tak pro hospodářská zvířata. Obsah proteinů může být až 45 %. Bílkovinné složení bobů ovlivňuje i fyzikálně-chemické vlastnosti sóji, které jsou důležité především při jejím zpracování. Díky vysokému obsahu proteinů je v ní přítomen i například inhibitor trypsinu, antinutriční látky a dochází ke špatné stravitelnosti této luštěniny (Clarke a Wiseman, 2005). Sója je také, co se týče luštěnin, hlavním zdrojem fytoestrogenů – látek prospěšných jako prevence rakoviny (Herman *et al.*, 1995).

V České republice je registrováno 19 odrůd sóji. Z toho tři odrůdy jsou registrované nově (v roce 2020), a to konkrétně Marzena, Aurelina a Hana. Z dat o přihlášených množitelských ploch plyne, že v roce 2019 se nejčastěji pěstovala velmi raná až raná odrůda Silesia (Převzato z archivu ÚKZÚZ) pocházející z Kanady.

#### 2.1.1 Agronomické znaky sóji

Lodyha sóji dorůstá do výšky 160 cm, je vztyčená a rozvětvená. Její kořeny pronikají do hloubky 0,5 až 2 m (Tvrzická 2016). Jednotlivé variety se od sebe liší v morfologii trichomů, velikosti zrna, doby kvetení, odolnosti, v obsahu bílkovin a oleje v bobech nebo v požadavcích na klima. Různá může být i doba zrání a to od 75 dní až po 200 a více. Záleží na velikosti (ty nejvyšší rostliny dozrávají jako poslední). Původní barva květů je fialová, která se vlivem domestikace změnila u variet s vysokým výnosem na bílou.

Zatímco variety pro většinovou produkci mají žlutá semena, a používají se pro získání oleje či mouky, variety vhodnější na vaření mají slámově žlutá, olivově žlutá nebo výjimečně existují dvoubarevná semena s jemnou oříškovou chutí. Variety pro průmyslové zpracování mají syrovou chuť a špatně se vaří. Černé sójové boby jsou výbornými antioxidanty, obsahují velký podíl polyfenolů a anthokyaninů. Deriváty fenylpropanoidové dráhy jsou zodpovědné za barvu obalu a pupku bobů sóji (Gillman *et al.*, 2011). Kultivary sóji se světlou barvou pupku se používají pro výrobu nattó a tofu. U domestikovaných variet se setkáváme navíc s černě nebo hnědě zbarvenými lusky, které mohou být v odstínech žluté, hnědé, černé a zelené, ale mohou

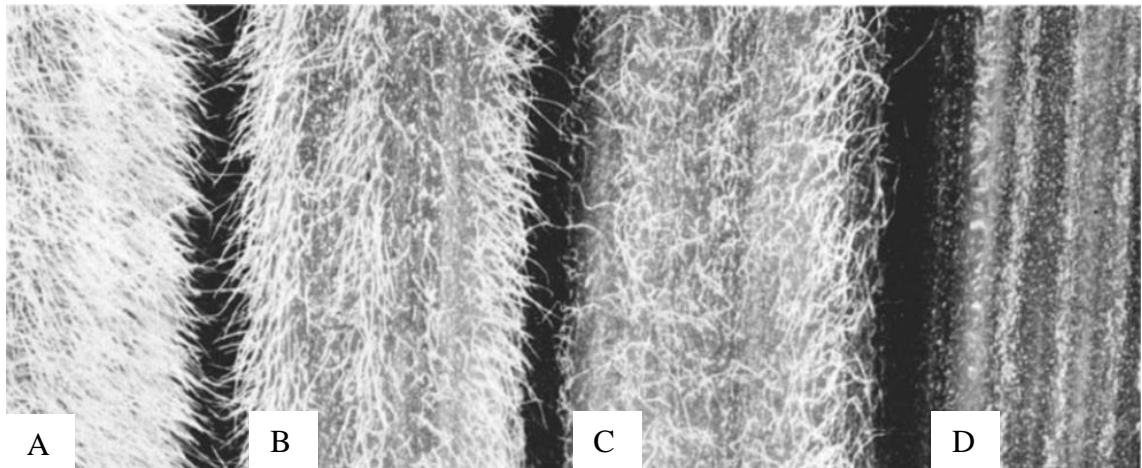
být i vícebarevné. Bílé boby neexistují, jsou to původně žluté boby, které vybledly stářím (Morse *et al.*, 1949).

Složené listy se liší tvarem, velikostí, barvou a dobou, po kterou je držen na stonku, než opadne. To se většinou děje v době, kdy je lusk zralý. List postupně zežloutne a nakonec opadne. Lusky jsou dlouhé 2 až 6 cm a obsahují 2 až 4 boby, lusky se přirozeně v době dozrávání otevírají a dochází k disperzi semen. Disperze semen je u elitních kultivarů eliminována, stejně tak jako prevalentní světlá barva lusků.

## 2.2 Trichomy

Většina variet sóji má stonek, listy a lusky pokryté jemnými chloupky – trichomy (Obr. 1). Jejich sklon je buď částečně přilehlý, vztyčený nebo přilehlý k čepeli listu. Pro tyto fenotypy byly identifikovány dvě alely na chromosomech 12 a 13 a to *Pa1/pa1* a *Pa2/pa2* (Bernard, 1975a). Lee *et al.* (1999) vyvinuli dvě populace (segregující každá jen jednu alelu pro přilehlý fenotyp trichomů) pro zjištění, který fenotyp (vzpřímený nebo přilehlý) je dominantní a který recesivní. Když pěstovali linie  $F_{2,4}$  pro potvrzení fenotypu rostlin  $F_2$ , ukázalo se, že jen pár rostlin mělo zcela přilehlé nebo úplně vztyčené trichomy. Zároveň se ale potvrdilo, že mezi dvěma fenotypy je velká možnost variací. Na pozadí *pa1pa1* by genotypy byly *Pa2Pa2* pro domněle homozygotní dominantní, *pa2pa2* pro domněle homozygotní recesivní a *Pa2pa2* pro domněle heterozygotní.

Pokud je rostlina lysá, je zároveň menší a kratší a nemá tolik bobů jako variety s trichomy. Na druhou stranu je ale odolnější vůči hmyzím škůdcům, proto je pěstována v oblasti Orientu (Morse *et al.*, 1949). Trichomy jsou prospěšné v oblastech s nepravidelnými srážkami, dochází díky nim ke snížení transpirace rostlin.



Obr. 1 Stonky sóji luštinaté s různými druhy trichomů. A: husté, B: normální hustota, C: kudrnaté, D: bez trichomů. (Převzato z Bernard a Singh, 1969)

### 2.2.1 Pigmentace trichomů

Za pigmentaci trichomů stejně jako u všech ostatních pletiv je zodpovědný klíčový enzym flavonsynthasa (*FNS*). Flavony zodpovědné za pigmentaci pletiv jsou syntetizovány prostřednictvím dvou kopií genů *FNS II*, a to *FNS II-1* (*Glyma.12g067000*) a *FNS II-2* (*Glyma.12g067100*) na chromosomu 12. *FNS II* převádí flavanony na flavony (Yan *et al.*, 2020). Různá aktivita enzymů zapojených do biosyntetické dráhy flavonů má za následek různé koncentrace celé řady derivátů flavonů. V trichomech sóji bylo izolováno, kvantifikováno (vysokoúčinnou kapalinovou chromatografií) a identifikováno 6 druhů flavonoidů (zodpovědných za pigmentaci). Jsou to apigenin, apigenin 7-O-glukosid (oba získány jako světle žluté prášky), luteolin (získán jako žlutý prášek), luteolin 7-O-glukosid, kaempferol 3-O-rutinosid a quercetin 3-O-rutinosid (Iwashina *et al.*, 2006). V ostatních částech/orgánech rostliny byla identifikována celá řada dalších derivátů těchto látek (hlavně v semenech a květech).

Za různou pigmentaci jednotlivých pletiv jsou zodpovědné následující alely. Pro barvu obalu semene sóji jsou to alely *T*, *W1*, *I*, *R* a *O*. Barva květů je řízena alelami *W1*, *W3*, *W4*, *Wm* a *Wp* a barva trichomů alelami *T* (na chromosomu 6) a *Td* (chromosom 3) (Toda *et al.*, 2002). Bernard (1975a) uvedl, že barva trichomů je řízena dvěma lokusy, *T/t* a *Td/d*. Žlutohnědé odpovídá *T- Td-*, světle žlutohnědé *T- tdt* a šedé *T tdt* a *tt tdt*. Recesivní alela *t* poskytuje šedou barvu trichomů bez ohledu na přítomnost alely lokusu *Td* (*tt TdTd* nebo *tt tdt*). Na rozdíl od *T* genu ovlivňuje *Td* gen pouze barvu trichomů, nikoli barvu květů, semen a hypokotylu. Alela *Td* kódující transkripční faktor MYB byla mapována 1,1 cM (centimorganů) daleko od markeru směrem ke konci

chromosomu 3 (*Glyma.06g258700*) (Yan *et al.*, 2020). Ani jeden z *FNS* však neodpovídají *Td* alele. *Td* alela kóduje transkripční faktor řídicí expresi obou *FNS* genů.

*T* alela (gen *Glyma.06g202300*), která kóduje flavonoid-3'-hydroxylasu (F3'H), mění barvu trichomů přítomností nebo absencí flavonů. *Td* alela fenotyp ředí. Dominantní *Td* alela poskytuje vyšší množství derivátů luteolinu, zatímco recesivní alela *td* generuje menší množství derivátů luteolinu. *Td* alela dává žlutohnědou a recesivní alela *td* světle žlutohnědou (nebo šedou) barvu trichomů. To naznačuje, že alela *Td* může regulovat obsah flavonů. Dominantní *T* alela trichomů generuje deriváty luteolinu a žlutohnědou barvu, recesivní alela *t* produkuje deriváty apigeninu a šedou barvu. Alela *Td* neovlivňuje množství antokyanů, flavonol glykosidů nebo dihydroflavonol v okvětních lístcích (Iwashina *et al.*, 2007).

### 2.2.2 Flavonoid-3'-hydroxylasa

Aktivita enzymu byla poprvé demonstrována v mikrosomální kultuře rostlinných buněk rodu *Happlopappus* (Fritsch a Grisebach, 1975) a detekována v extraktu rostlin. Poprvé byl enzym izolován a charakterizován u petunie (Brugliera *et al.*, 1999). Jeho homologem je gen *sf3'h1* (Toda *et al.*, 2002).

Enzym je důležitým regulátorem antokyaninové dráhy v semenech sóji. V černých sójových bobech odklání biosyntézu oranžových a modrých antokyanů tak, že vzniká červený kyanidin-3-O-glukosid (Choung *et al.*, 2001). Jeho hlavní funkce je při hydroxylaci B kruhu flavonoidů na pozici 3', která vede ke vzniku pigmentů na bázi kyanidinu. Je závislý na monooxygenase mikrosomální cytochrom P450, který vyžaduje kofaktor NADPH (Forkmann, 1991).

Za pigmentaci trichomů je zodpovědná F3'H kódována *Glyma.06g202300*, která byla identifikována jako *T* alela. Bylo nalezeno 46 SNP spojených s barvou trichomů. Překrývají oblast přibližně 2 Mbp (milion párů bází) na chromosomu 6, překrývají *T* lokus. Delecí jedné báze v genu *Glyma.06g202300*, dojde ke změně barvy ze žlutohnědé na šedou (Obr. 2) (Wen *et al.*, 2015).



Obr. 2 Pigmentace trichomů. A: černé lusky se žlutohnědými trichomy, B: šedé trichomy. (Převzato z databáze GRIN/U.S. National Germplasm System)

### 2.2.3 MYB faktor

Existuje velké množství transkripčních faktorů R2R3 MYB podílejících se na biosyntéze flavonoidů. Proteiny MYB se také účastní biosyntézy anthokyaninů, proantokyanidiny a flavonolů (Liu *et al.*, 2015). Mutace proteinů způsobí vlivem předčasného vložení stop kodonu neschopnost exprese genu *FNS II*. To způsobí snížení obsahu flavonu a barva trichomů není tak výrazná (Yan *et al.*, 2020).

Kultivary a linie mající téměř šedé, šedé nebo světle žlutohnědé trichomy s alelou *td* mají v genu předčasné stop kodony. Yan *et al.* (2020) zjistili porovnáním sekvencí genomů souvislost mezi předčasnými stop kodony v *Glyma.03g258700* (transkripční faktor R2R3 MYB) a recesivní alelou *td*. Ve srovnání se žlutohnědou linií trichomů produkovala téměř izogenní linie s alelou *td* extrémně malé množství transkriptů *Glyma.03g258700*, *FNS II-1* a *FNS II-2* u trichomů. Promotor *FNS II-1* a *FNS II-2* sdílí regulační prvky pro vazbu proteinů MYB. Tyto výsledky naznačují, že divoký typ (WT) proteinu *Glyma.03g258700* se může vázat na promotor genů *FNS II* a regulovat jejich expresi, což vede ke zvýšenému obsahu flavonu a hlubší barvy trichomů. Naproti tomu mutovaný protein *Glyma.03g258700* může selhat při zvýšení exprese genů *FNS II*, což vede ke snížení obsahu flavonu a zředění barvy trichomů.

## 2.3 Flavonoidy

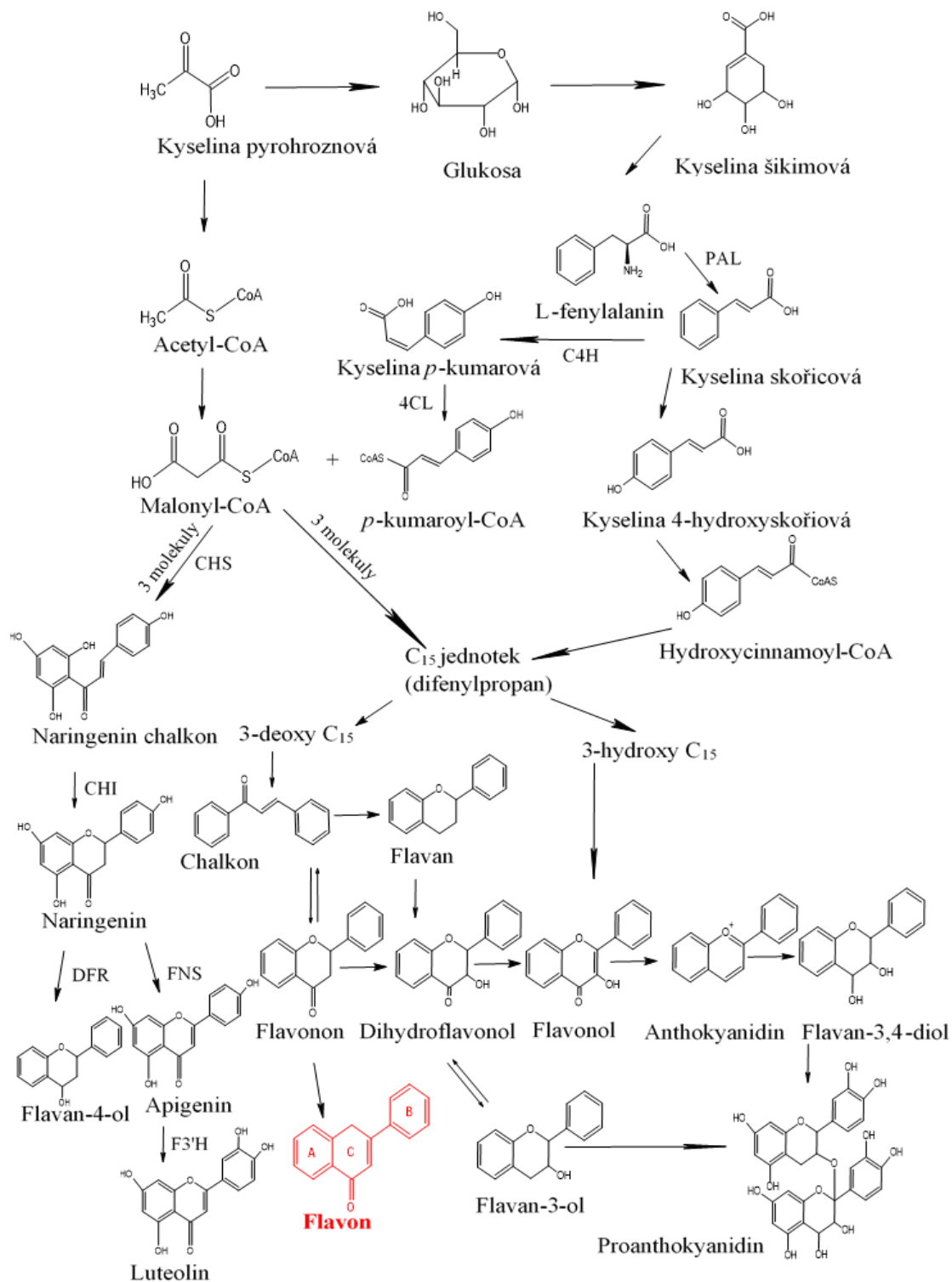
Flavonoidy jsou sekundárními metabolity rostlin. Ty je syntetizují za různými účely. Poskytují rostlině zabarvení, aby přilákala opylovače, chrání ji před UV (ultrafialové) zářením, patogeny a mikroby, protože slouží jako signální molekuly. Kromě toho mají důležitou ekologickou roli, až 20 % uhlíku fixovaného rostlinou k fotosyntéze přechází do fenylypropanoidové dráhy a dochází tak k tvorbě fenolických sloučenin.

Jejich struktura se skládá primárně z 2-fenylchromanů (flavonoidy, včetně flavanonů, flavonů, flavonolů, flavan-3-olů a antokyanidinů) nebo z 3-fenylchromanů (isoflavonoidy) (Dixon a Pasinetti, 2010).

### 2.3.1 Biosyntéza flavonů/flavonoidů

V sóji se flavony vyskytují nejvíce právě v trichomech (Toda *et al.*, 2002). Oba prekurzory jsou odvozeny z metabolismu sacharidů a fenylypropanoidové dráhy (Forkmann a Heller, 1999). A kruh flavonoidů je syntetizován sekundární cestou kondenzací 3 molů malonyl-CoA, který vznikl z glukosy acetátovou cestou (Obr. 3) (Formica et Regelson, 1995). Kruhy C a B mají také jako výchozí látku glukosu, ale vznikají šikimátovou dráhou za vzniku kyseliny skořicové (enzym fenylyalanin amoniak lyasa, PAL), 4-hydroxyskořicové a její redukované formy, kyseliny kumarové (enzym cinamát 4-hydroxylasa (C4H), který je enzymem 4-kumarát-CoA ligasa (4CL) přeměněn na *p*-kumaroyl-CoA. Tyto devítiuhlíkaté produkty kondenzují spolu s produkty malonátu, které mají šestiuhlíkatou kostru, v přítomnosti chalkonsynthasy (CHS), (Obr. 3). Vznikají tak chalkony a po uzavření kruhu a následnou hydratací vznikají 3-hydroxyflavonoidy (katechiny), 3,4-diolové flavonoidy (kvercetin), polyfenolické látky a prokyanidiny (Formica a Regelson, 1995). Isomerizace naringenin chalkonu chalkonisomerasou (CHI) na naringenin vede v přítomnosti *FNS* k tvorbě apigeninu, který může být dále přeměněn pomocí enzymu *F3'H* na luteolin. Dihydroflavonol-4-reduktasa (DFR) katalyzuje jinou dráhu naringenin, a to vznik flavan-4-olu. Tento bezbarvý prekurzor dává polymerací červené zabarvení (vznik flobafenu).





Obr. 3 Biosyntéza flavonoidů a jejich přeměna za přítomnosti enzymů. Kruh A vzniká sekundární cestou, kruhy B a C šikimátovou. Vznik několika derivátů. (Vytvořeno v programu ACD/ChemSketch podle Wollgast a Anklam, 2000)

Enzym CHS katalyzuje první krok biosyntézy flavonoidů, při kterém vzniká chalkon (žluté zbarvení), který je ale uchováván jen v některých rostlinách. Například rajčata uchovávají chalkon ve své slupce (Muir *et al.*, 2001) a astra čínská spolu s hvozdíkem zahradním mají chalkon ve svých okvětních lístcích (Schijlen *et al.*, 2004). CHS má dva typy, z nichž typ enzymu izomerizující 6'-hydroxychalkony a 6'-deoxychalkony je obsažen hlavně v luštěninách (Joung *et al.*, 2003).

Dalšími katalyzovanými reakcemi se tvoří jiné druhy flavonoidů, a to hlavně antokyany, z jejichž meziproductů při jejich tvorbě dále vznikají i aurony, flavony, proantokyanidiny a flavonoly. Metabolická dráha vede z chalkonů přes flavanony až enzymem isoflavonsynthasou (IFS) k isoflavonoidům. Ty se v nezralém embryu sóji akumulují, což je spojeno se zvýšenou expresí genu MYB *GmMYB12B2*, který reguluje expresi chalkonsynthasy 8 (jedna z devíti kopií genu chalkonsynthasy) a podílí se na syntéze isoflavonoidů (Liu *et al.*, 2015). Ty jsou mimo jiné důležité pro symbiotický vztah sóji s diazotrofními aerobními bakteriemi (rody *Rhizobium*, *Sinorhizobium*, *Azorhizobium*, *Bradyrhizobium*) (Cooper, 2004).

4,2',4',6'-tetrahydroxychalkon vzniká kondenzací 3 acetátových jednotek za přítomnosti enzymu CHS. Tato reakce začíná reakcí malonyl-CoA s p-kumaroyl-CoA (Holton a Cornisch, 1995). Následuje hydroxylace flavanonů na uhlíku číslo 3 pomocí enzymu flavanon-3-hydroxylasa za vzniku dihydroflavonolů. Hydroxylací naringeninů vzniká dihydrokaempferol, který je dále hydrolyzován pomocí enzymů, které jsou asociované s P450, v polohách 3' a 5' (tvorba antokyanů odvozených od delphinidinu) nebo v poloze 3' na B-kruhu. Pokud dochází k hydrolyzaci na pozici 3' vznikají pigmenty odvozené od kyanidinu za přítomnosti enzymu F3'H (Toda *et al.*, 2002; Schijlen *et al.*, 2004).

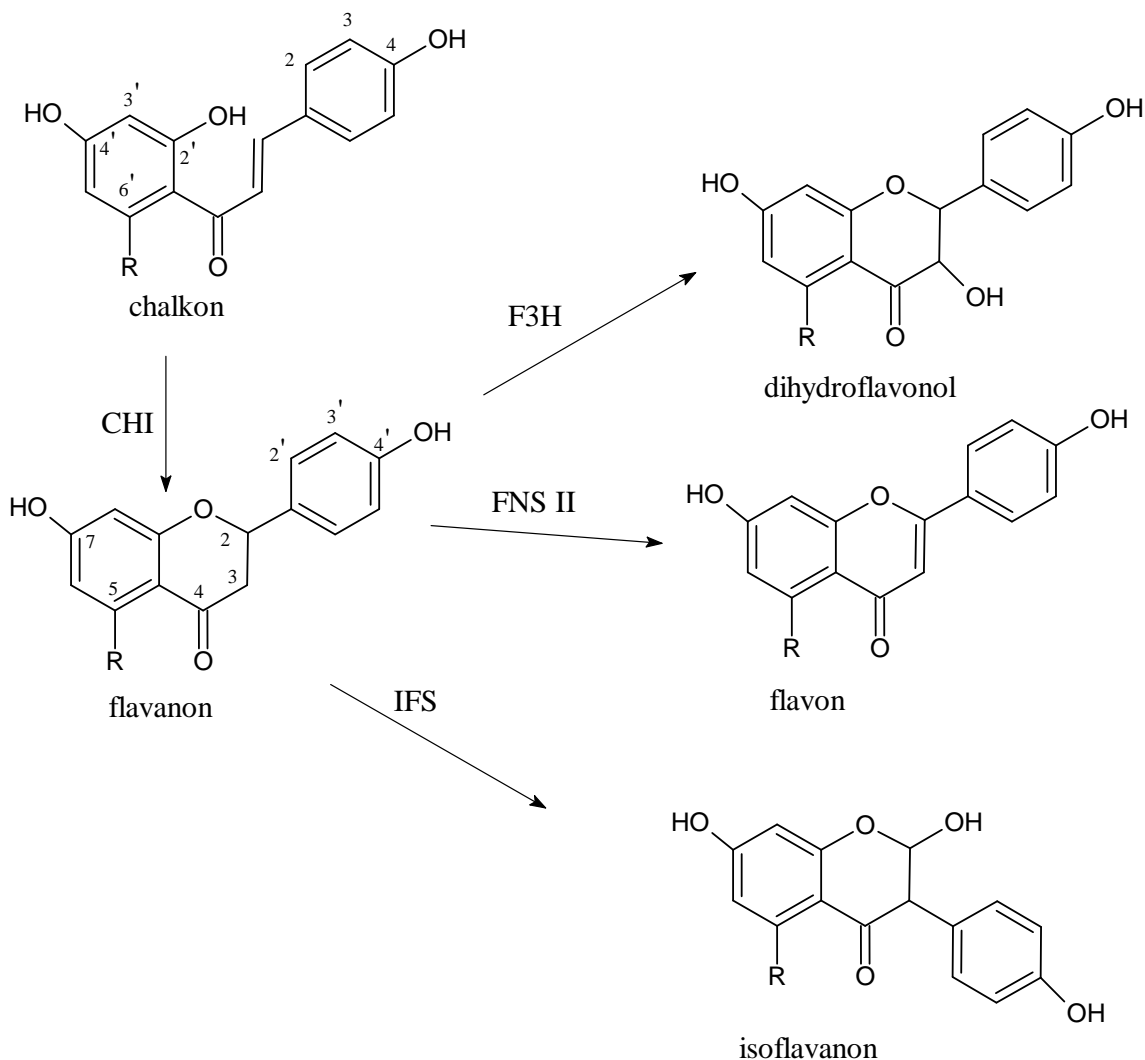
Leukoanthokyanidiny, vznikající redukcí dihydroflavonolů, jsou stejně jako katechiny a proanthokyanidiny prekurzory syntézy antokyaninů za přítomnosti enzymu antokyanidinsyntázy. Anthokyaniny jsou v posledním kroku své syntézy za přítomnosti flavonoid-3-O-glukosyltransferasy stabilizovány přenosem glukosy z UDP-glukosy na volnou hydroxylovou skupinu na uhlíku 3 (Schijlen *et al.*, 2004).

B kruh flavonoidů může být hydroxylován buďto v poloze 3' (pro pigmenty na bázi kyanidinu) nebo na místech 3' a 5' (pigmenty na bázi delphinidinu). Hydroxylace se účastní enzymy F3'H a flavonoid-3',5'-hydroxylasa vyžadující kofaktor NADPH.

Pomocí chromatografie bylo prokázáno, že alely *T* a *W1* kódují oba enzymy a jsou tak zodpovědné za tvorbu flavonoidů (Toda *et al.*, 2002).

### 2.3.2 Degradace chalkonů

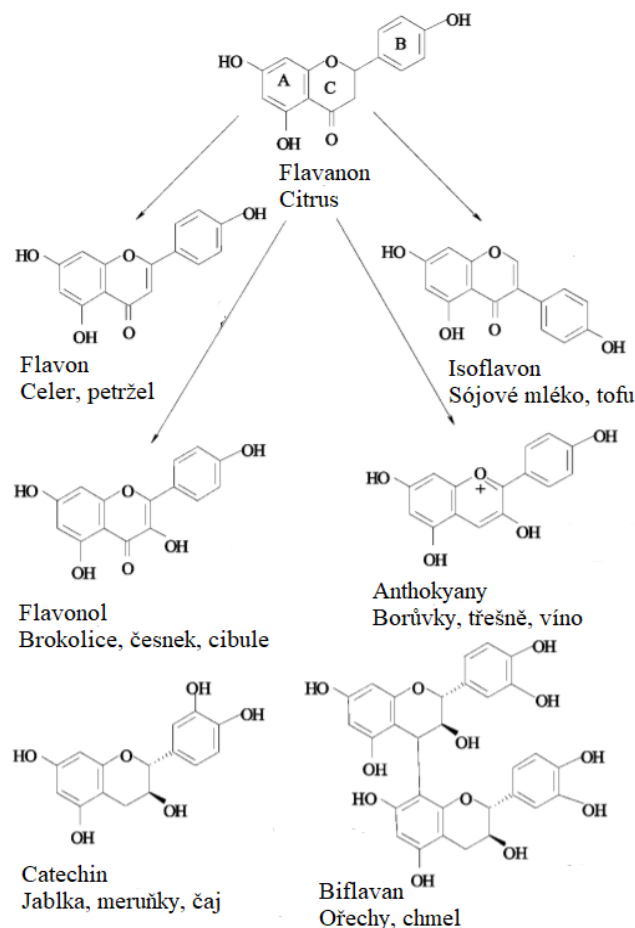
Rostliny se chalkonu zbavují buď spontánně jeho izomerací na naringenin (Holton a Cornisch, 1995) nebo právě pomocí CHI (Obr. 4), která způsobí izomeraci naringenin-chalkonu na flavanon-naringenin (Forkmann a Heller, 1999).



Obr. 4 Degradace chalkonů. V prvním kroku jejich přeměna na flavanon pomocí CHI a následně tři alternativní cesty odbourání. Buďto vzniká dihydroflavonol, flavon nebo isoflavanon. (Vytvořeno v programu ACD/ChemSketch podle Ralston *et al.*, 2005)

### **2.3.3 Vliv konzumace pigmentů na lidské zdraví**

Isoflavonoidy a flavonoidy mají kromě biologických funkcí u rostlin i důležité účinky na lidské zdraví jako je ochrana enzymů či vitamínů. Jsou to semiesenciální složky potravy. V potravinách se nejčastěji nachází světle žluté flavonoly (kvercetin, kaempferol a myricetin), konkrétně v cibuli, červeném vínu či jablku (Obr. 5) (Yao *et al.*, 2004). Složky černého a zeleného čaje mají schopnost inhibovat proliferaci nádorových buněk nebo zastavit buněčný cyklus apoptózou (Yao *et al.*, 2004). Inhibují bioaktivační enzymy a aktivují detoxikační enzymy. Dokážou inhibovat buněčný cyklus, oxidační stres, apoptózu, a naopak aktivují imunitní systém. Mají pozitivní vliv na metabolismus tuků a snižují počet kardiovaskulárních chorob. Zpomalují menopauzu u žen a zmírňují její následky, u mužů mají pozitivní vliv na vypadávání vlasů (Goluch-Koniuszy, 2016). Zabraňují oxidaci lipoproteinů s nízkou hustotou, tím se snižuje riziko aterosklerózy. Kvercetin inhibuje agregaci trombocytů. Flavonoidy zlepšují paměť, konsolidaci a pozitivně působí na kognitivní funkce i proti Alzheimerově chorobě (Dixon a Pasinetti, 2010).



Obr. 5 Výskyt jednotlivých flavonoidů v potravinách a jejich struktura. (Vytvořeno v programu ACD/ChemSketch podle Yao *et al.*, 2004)

Vykazují antialergickou, antivirovou, antiproliferační aktivitu, antidiabetickou i antineoplastickou aktivitu. Protizánětlivě působení spočívá v inhibici cyklooxygenasy 2, která napomáhá k rozvoji zánětu (Raso *et al.*, 2001). Tyto účinky má biacalin, který zároveň působí proti viru HIV 1 (virus lidské imunitní nedostatečnosti typu) tím, že brání jeho vstupu do cílových buněk. Bezbarvé isoflavony tvořící se v sóji *Glycine max* a dalších luštěninách mohou pomoci k zabránění tvorby hormonálně závislých nádorů, to je ale stále ve fázi výzkumu (Ralston *et al.*, 2005). Zatím bylo prokázáno, že geneticky modifikovaná rajčata s vysokým obsahem antokyanů (*Solanum lycopersicum*) prodlužují životnost myši náchylných k rakovině (Dixon a Pasinetti, 2010). Isoflavony ze sóji (genistein, daidzein a biochanin A) působí proti karcinomu mléčné žlázy.

Flavonoidy výrazně podporují zdraví tím, že působí i jako silné antioxidanty s vysokou antioxidační kapacitou *in vivo* i *in vitro*. Jejich množství ve stravě je tedy

spojeno s doporučeným denním příjmem antioxidantů. Do krevního oběhu se dostávají velmi rychle po požití a dochází ke snížení poškození DNA oxidací (Benzie *et al.*, 1999). Mohou vychytávat volné radikály a různé formy reaktivního kyslíku díky přítomnosti hydroxylové skupiny ve své struktuře. Ta se váže s nepárovým elektronem radikálu a tím snižuje i jeho aktivitu. Vzniklý nepárový elektron původní hydroxylové skupiny se váže s nepárovým elektronem skupiny, která vznikla stejným způsobem. Dalšími mechanismy antioxidačních účinků je chelatace kovů (např. mědi nebo železa) a zhášení singletového kyslíku. Bylo prokázáno, že čím více má flavonoid ve své struktuře hydroxylových skupin a čím méně je glykosylován, tím větší má antioxidační aktivitu (Frankel *et al.*, 1997). Vyšší aktivita je pozorována také spíše v hydrofilním (poměr disociační konstanty kyseliny ku rovnovážné konstantě u naringenu je roven  $2,73 \pm 0,70$ ) než v lipofilním prostředí ( $-1,88 \pm 0,70$ ) (Yao *et al.*, 2004).

Flavonoidy přijaté z rostlinné potravy jsou modifikovány methylovací nebo glykosylací jak u rostlin, tak u živočichů. Ve střevech se nejlépe vstřebávají isoflavony a nejhůře anthokyany. Musí být schopny projít hematoencefalickou bariérou, aby mohly interagovat s mozgovými buňkami a mohla tak dojít ke stimulaci požadované odpovědi nervové soustavy. Aktivní vazebná místa pro flavonoidy v mozku zatím nejsou dostatečně prozkoumána (Dixon a Pasinetti, 2010).

#### **2.3.4 Výhodné znaky spojené s šedými trichomy**

Pokud je sója *Glycine max* [L.] Merr. vystavena během kvetení nízkým teplotám (10-18 °C), což se týká oblastí ve vyšších zeměpisných či nadmořských šířkách, dochází ke zpomalování růstu, ztrátě květů a nezralých lusků, k hnědnutí kolem pupku, k praskání semen a také se snižuje výnos. S tím souvisí právě *T* alela a alely spojené s dobou kvetení a dozrávání (jejich účinky mohou být aditivní) (Sunada a Ito, 1982).

Alela *T* potlačuje vznik pigmentů v oblasti kolem pupku tak, že inhibuje oxidaci fenolických sloučenin. Také potlačuje praskání semen. Alela *I* (zodpovědná za distribuci pigmentů v semenech) nemá tak silný účinek, ale také potlačuje pigmentaci obalu semene i praskání (genotyp *tt*) (Takahashi, 1997). Existuje genetická variace mezi žlutou barvou pupku (alela *I*) a šedou barvou trichomů (*t*) (Takahashi a Abe, 1994).

Na stres chladem jsou obecně náchylnější odrůdy se žlutou barvou pupku, které mají nejčastěji šedé trichomy (Sunada a Ito, 1982), ale náchylnost na nízké teploty souvisí, podle pozorování pěstitelů v Japonsku, spíše jen s barvou trichomů. Takahashi *et al.*,

(2005) provedli studii, která spočívala ve vystavení linií chladu. Vyšší hmotnost byla u linií s *T* alelou, ne s recesivní alelou *t*. Bylo studováno více alel, ale během nízkých teplot může mít příznivý účinek i alela *T*. I z tohoto důvodu je důležitá znalost genetického pozadí variet s šedými trichomy.

Pro další šlechtění sóji, aby byl zvýšen jak výnos, tak i obsah proteinů, je vhodný kultivar TN11-5102, který má šedé trichomy. Rostlina obsahuje 421 g·kg<sup>-1</sup> proteinů a 216 g·kg<sup>-1</sup> oleje suché hmotnosti a je schopná produkovat až 490 g·kg<sup>-1</sup> proteinů v sójové mouce. Vykazuje odolnosti vůči polehávání, rezistenci vůči hád'átku i vůči vzniku nádorů (Pantalone a Smallwood, 2018).

## 2.4 Albedo plodin

Albedo je schopnost povrchu odrážet dopadající záření. Vyjadřuje se jako poměr odraženého elektromagnetického záření k záření dopadajícímu. Čím více záření je odraženo, tím méně energie je pohlceno planetou. Této vlastnosti se dá využít pro snížení teplot, které se neustále zvyšují vlivem globálního oteplování (zvyšování koncentrace CO<sub>2</sub> v atmosféře).

V současné době existuje několik způsobů zvýšení albeda povrchu včetně změn ve výběru typu plodiny, využití orbitálních slunečních clon (Angel, 2006), zvýšení albeda střech a chodníků a konverze jehličnatých lesních plantáží na listnaté. Další strategie, které mohou být vyvinuty v budoucnu by mohly zahrnovat modifikace morfologie rostlin (Doughty *et al.*, 2011).

Na albedo rostlin má vliv i lesk listu či přítomnost trichomů a jejich morfologie. Pěstováním rostlin s vyšším albedem baldachýnu by teoreticky mohlo vést ke snížení teplot v oblasti nad Severní Amerikou a Eurasií (až o 1 °C) během léta, zatímco v nízkých zeměpisných šířkách v oblasti Jihovýchodní Asie je největší ochlazování zaznamenáno v zimě (Ridgwell *et al.*, 2009). Je důležité, aby nebylo dosaženo vyššího albeda baldachýnu na úkor sníženého výnosu plodiny v důsledku nadměrného snížení absorpce fotosynteticky aktivního záření (PAR) (Ridgwell *et al.*, 2009).

Zvýšení albeda v zemědělství může ochladit klima ve vyšších zeměpisných šířkách, ale zvýšení albeda rostlin tak, aby se vyrovnalo oteplení v budoucnosti bude vyžadovat větší změny v rostlinách, než jsou v současné době k dispozici (Doughty *et al.*, 2011).

### 2.4.1 Vliv barvy na albedo

Ze studie od Doughty *et al.*, 2011 vyplývá, že je nevhodnější pěstovat světlejší rostliny, jako je *Encelia farinosa*, která odráží záření přednostněji v NIR (blízká infračervená spektroskopie, ta se nevyužívá při fotosyntéze) (Ehleringer a Bjorkman, 1978). Bylo také zjištěno, že izogenní linie se žlutohnědými trichomy odrážely dopadající záření významněji než ty s šedými trichomy. Linie sóji s kudrnatými a šedými trichomy odrážely dopadající záření ve viditelné oblasti nejvýznamněji (Doughty *et al.*, 2011). Šedé odrůdy pšenice a ječmene pěstovány za podmínek s omezeným množstvím vody vykazují vyšší výnosy zrna ve srovnání s jejich protějšky, což naznačuje zvýšenou efektivitu ve využívání vody (menší zahřívání listů), která převažuje nad snížením fotosyntézy. Mezi mutanty čiroku dvoubarevného byly pozorovány rozdíly v odrazivosti jednotlivých listů s rozdílnou strukturou vosku až o 0,16 a o 0,19. Důležitá je také morfologie vrchlíku, přičemž mezi odrůdami kukuřice se albedo baldachýnu liší až o 0,08 (Ridgwell *et al.*, 2009). Ačkoli barva, orientace a hustota trichomů mají jen nepatrné dopady na velikost albeda listů plodin dá se očekávat, že ve spojitosti s dalšími znaky (jako je intenzita zbarvení listů) mohou albedo sóji zlepšit.

### 2.5 Fenotypová změna jako důsledek přirozených genetických variant

Každý genom je individuální a od ostatních se liší ve smyslu přirozeně se vyskytujících genetických variant. Mezi ty patří SNP, delece, inserce, variabilita počtu kopií segmentů DNA (deoxyribonukleová kyselina) (CNV – copy number variation, variabilita počtu kopií) nebo tautomerní přesmyky (které vznikají přechodem vodíkových atomů z jedné pozice purinu nebo pyrimidinu na jinou a dochází tak ke vzniku bodových mutací). Jakákoli mutace může způsobit změnu ve vlastnostech jedince neboli změnu fenotypu. Podle efektu, které mutace způsobují se dělí na tiché (nemají žádný význam, kódují stejnou aminokyselinu jako před mutací), missence (dochází ke změně aminokyseliny) a na nonsense (dochází ke vzniku stop kodonu a zkrácení polyptidového řetězce).

Ke vzniku stop kodonu došlo například právě v *T* alele v genu pro F3'H na chromosomu 6 výměnou adeninu (A) za thymin (T) nebo právě v *Td* alele v genu *Glyma.03g258700* na chromosomu 3 záměnou cytosinu (C) za T (Tab. 1). K posunu čtecího rámce došlo na chromosomu 3, kdy místo referenční alely (REF) GA (guanin, adenin) je alternativní G. V oblasti kódující sekvence genu *Glyma.03g258700* bylo



nalezeno analýzou SNPviz v2.0 (zobrazovač jednonukleotidových polymorfismů ve verzi 2.0 vyvíjený v rámci spolupráce Skupiny genomiky luštěnin Katedry biochemie s kolegy z University of Missouri, USA: Zeng *et al.*, <http://soykb.org/newSNPViz/> (Zeng *et al.*, 2020; Zeng *et al.*, 2021) rovněž několik tichých mutací, které nevedou k žádné změně aminokyseliny.

Tab. 1 Publikované mutace na chromosomu 6 a 3, pozice, na kterých se nacházejí a popis konkrétní mutace. Gen *Glyma.06g202300* je orientován ve směru, gen *Glyma.03g258700* proti. Zobrazení REF alely i alternativní (ALT) alely.

Chromosom	Gen	Pozice	REF	ALT	Mutace	Orientace genu
6	<i>Glyma.06g202300</i>	18737366	TC	T	posun čtecího rámce, delece C	+
		18737236	A	T	nonsense, Lys344*	
		18737324	C	T	missence, Pro373Leu	
3	<i>Glyma.03g258700</i>	45301305	GA	G	posun čtecího rámce, delece T	-
		45301308	TA	T	posun čtecího rámce, delece T	
		45301350	C	T	nonsense	

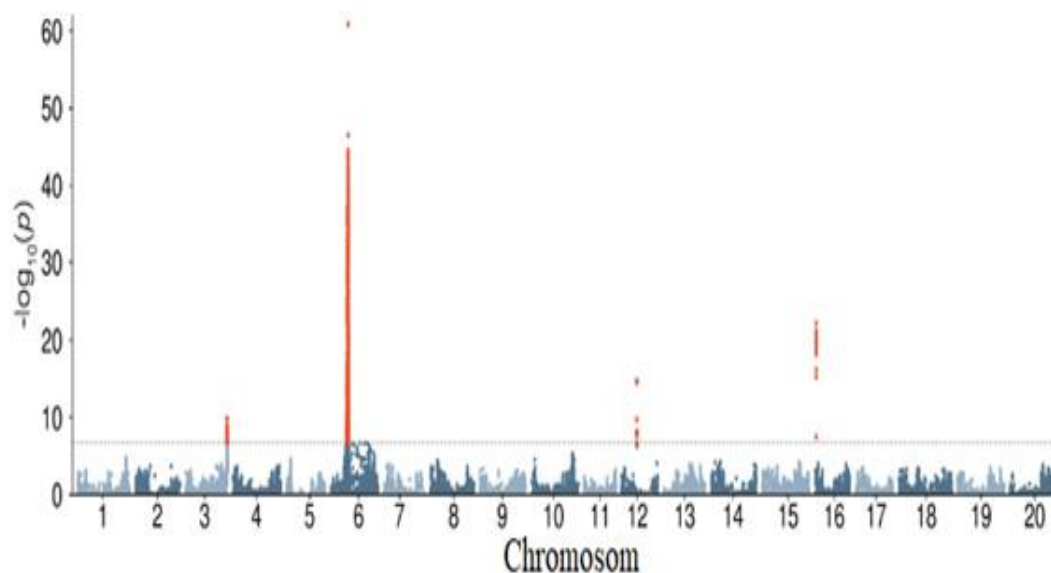
Celogenomové sekvenování je používáno pro určení ideálně celé sekvence genomu DNA. Jde o sekvenování celé DNA v chromosomech, mitochondriích nebo v chloroplastech u rostlin. Nízkohustotní genotypizace je často zaměřena na analýzu specifických signálních drah nebo na soubory biomarkerů.

## 2.6 GWAS

Metoda asociace genotypu na fenotyp (= genome-wide association study) je určena k detekci asociací mezi genetickými variantami a znaky ve vzorcích zkoumané populace. V ideálním případě je takto možno identifikovat mutace přímo v konkrétních jinak neznámých genech. Klasickým výstupem GWAS je vynesení nazývané Manhattan plot znázorňující vynesení  $-\log_{10}$  hodnoty  $p$  na SNP pozicích v jednotlivých chromosomech.

GWAS analýza na celogenomových datech je výpočetně náročná a nedostatečná kvalita informace o insercích a delecích GWAS komplikuje. Proto byla vynalezena celá řada DNA čipů, které jsou omezené pouze na SNPy a to pouze na jejich zlomek z celkového množství SNP v genomu organismu. U sóji je nejoblíbenějším čipem SoySNP50K (okolo 3 %). Použití DNA čipů je však velice oblíbené, protože celý proces analýzy významně zlevňuje a urychluje.

Za účelem identifikace konkrétních genů a jejich mutací ve známých alelách lokusů mapovaných na důležité agronomické znaky sóji byla provedena metoda GWAS. Pro 809 druhů sóji a pro 84 agronomických znaků bylo zjištěno 245 lokusů významných pro agronomické znaky sóji (Fang *et al.*, 2017). Kromě 14 alel souvisejících se syntézou oleje byly zjištěny i významné signály pro pigmentaci trichomů na 4 různých chromosomech (Obr. 6). Nejsilnější již publikovaný signál pro barvu trichomů je na chromosomu 6. Ke konci chromosomu 3 je zaznamenán signál odpovídající světle šedé barvě trichomů odpovídající publikované mutaci na pozici 45 301 350. Nově je zaznamenán signál i na chromosomu 12 a na začátku chromosomu 16, kde by mohly být nalezeny další geny s mutacemi, které by byly zodpovědné za pigmentaci trichomů. Velikost prohledávané oblasti spolu s oblastmi s nejvyššími asociacemi jsou uvedeny v tabulce 2. Autoři také predikovali kandidátní gen *Glyma.16g016500*, který ale nebyl anotován s funkcí žádného známého proteinu.



Obr. 6 Manhattan plot pro pigmentaci trichomů. Vynesení  $-\log_{10}(p)$  proti SNP pozicím na 20 chromosomech. Kromě chromosomů 3 a 6, o kterých už víme, že obsahují mutace zodpovědné za barvu trichomů, jde vidět vidíme signál i na chromosomu 12 a 16. Tam by mohly být nalezeny další mutace, které by dovysvětlily barvu u variet bez publikovaných mutací. (Převzato a upraveno podle Fang *et al.*, 2017)

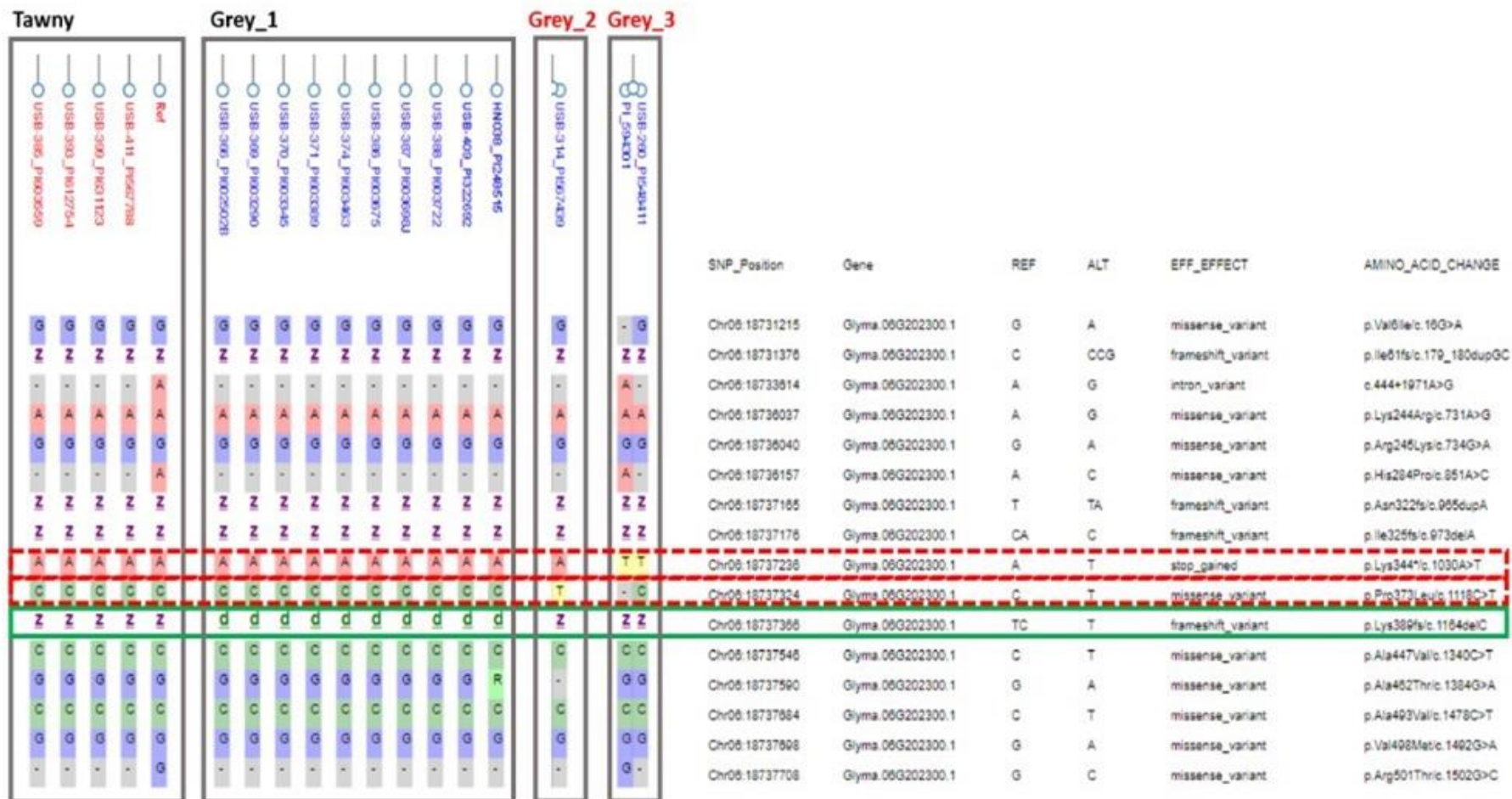
Tab. 2 Oblast hledání v GWAS na chromosomech 3, 6, 12 a 16 společně s pozicí s nejvyšší asociací a hodnotami  $-\log_{10}(p)$ . (Převzato z Fang *et al.*, 2017)

Chromosom	Začátek	Konec	Nejvyšší asociace	$-\log_{10}(p)$
3	45 188 204	45 546 666	45 301 350	9,88
6	16 701 822	19 482 848	18 734 414	60,86
12	16 899 433	16 906 312	16 899 433	14,77
16	1 439 581	1 445 782	1 440 263	22,20

### 2.6.1 Identifikace alternativních mutací v *t* alele

*T* alela je reprezentovaná v genu *Glyma.06g202300* a *Td* alela v *Glyma.03g258700*. *T* alela má na barvu trichomů větší vliv, dominantní alela poskytuje žlutohnědou barvu, recesivní šedou barvu (Iwashina *et al.*, 2006).

Bylo prokázáno, že mutace způsobená delecí C na chromosomu 6 v pozici 18 737 366 *G. max* Williams 82.a2.v1 vede k šedé barvě trichomů. Delece změní následující čtecí rámeček, což způsobí zkrácení polypeptidu, kde následně chybí konsenzuální sekvence GGEK a doména vázající hem (Toda *et al.*, 2002). Později se ale přišlo na to, že tato mutace není u 6,7 % z celkového množství pozorovaných linií s šedými trichomy (Zeng a Škrabišová, 2021), což vede k závěru, že existují další mutace vysvětlující šedou barvu trichomů. Pomocí SNPviz v2.0 byla prozkoumána celá genomová oblast na chromosomu 6 u žlutohnědých a šedých trichomů a byly odhaleny dvě další mutace (Obr. 7). První vede ke změně A za T na místě 18 737 236 a vede k předčasnému stop kodonu Lys344\*. Druhá spočívá v záměně C za T na pozici 18 737 324 a vede ke změně aminokyseliny Pro373Leu (Zeng a Škrabišová, 2021). Tyto mutace vysvětlují jen šedou barvu trichomů, ale ne téměř šedou ani světle šedou.

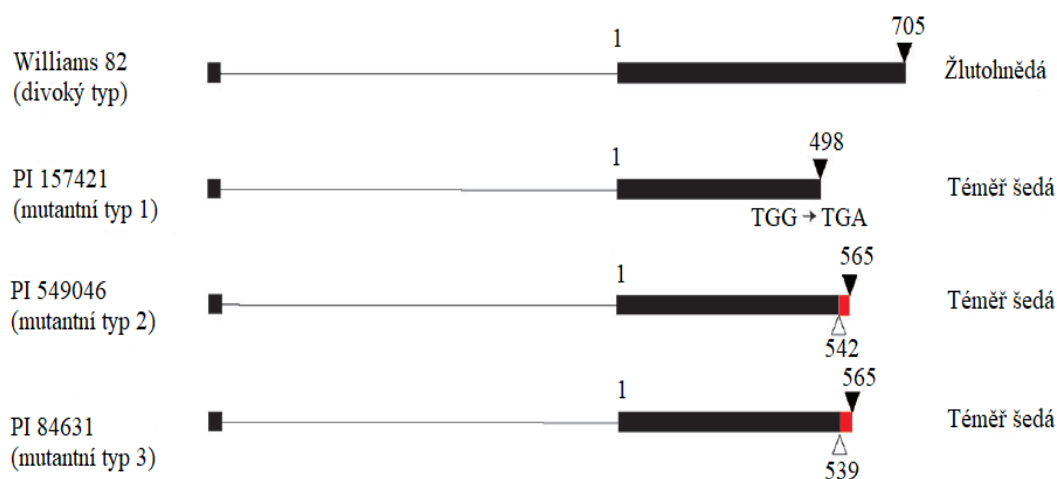


Obr. 7 SNPviz v2.0 vizualizace *T* lokusu haplotypu na chromosomu 6 - *Glyma.06g202300* (pouze kódující oblast genu). Žlutohnědá (červeně) a šedá (modře). Delece C na chromosomu 6 v pozici 18 737 366 *G. max* Williams 82.a2.v1 v *t* allele (zeleně). Alternativní mutace v *t* objevené v tomto případě použitím SNPviz v2.0 jsou označeny červenými čárkovanými čarami. (Převzato z Zeng a Škrabišová, 2021)

## 2.6.2 Mutace v *Td* alele

Yan *et al.*, (2020) zjistili, že dominantní alele *Td* produkovala v trichomech recesivní *td* alela malé množství transkriptů *Glyma.03g258700*, *FNS II-1* a *FNS II-2*. WT proteinu *Glyma.03g258700* se váže na *FNS II* a zvyšuje transkripci, ale zmutované proteiny v *Glyma.03g258700* vedou ke snížení obsahu flavonu a zředění barvy trichomů.

Pro chromosom 3 byly dosud potvrzené 3 mutace (Obr. 8). Srovnáním sekvence genomu divokého typu Williams 82 se žlutohnědými trichomy se sekvencemi rostlin s téměř šedými trichomy (PI 157421, PI 549046 a PI 84631) (PI = plant introduction) byly zjištěny mutace v oblasti *Glyma.03g258700*, což je transkripční faktor MYB. V PI 157421 došlo k substituci G za A (TGG → TGA). V PI 549046 i PI 84631 došlo k delecím, které vedly k předčasnému stop kodonu. *Glyma.03g258700* je oproti divokého typu kratší a nemá počáteční kodon a může odpovídat alele *Td*. Jak autoři práce sami diskutují, je možný další dosud nepopsaný regulační mechanismus pigmentace trichomů.

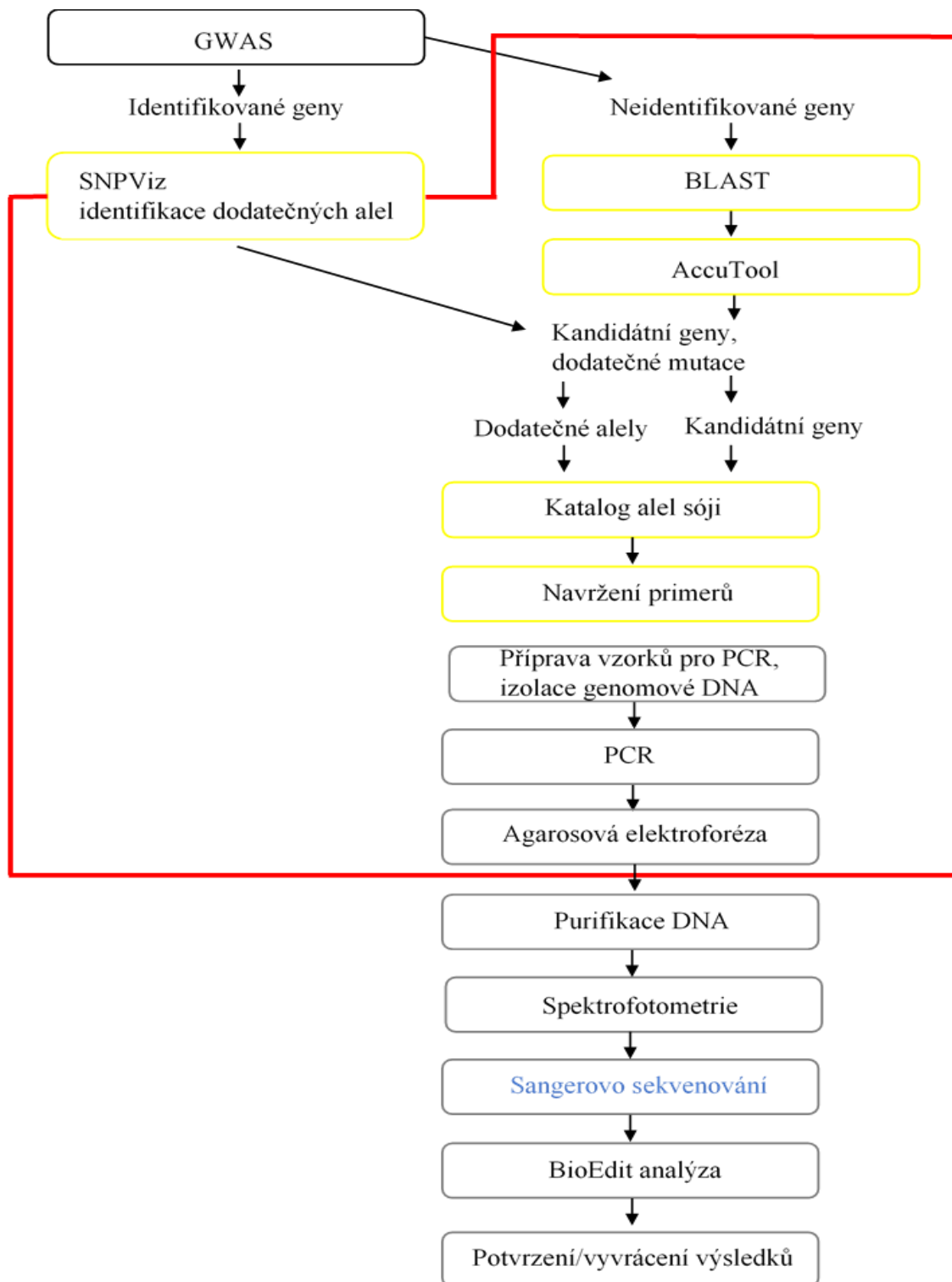


Obr. 8 Struktura genu *Glyma.03g258700* divokého typu Williams 82 se žlutohnědými trichomy a třech mutantních typů s téměř šedými trichomy. Exony jsou zobrazeny rámečky a introny čarou. Stop kodony a delece jsou zobrazeny černými a bílými trojúhelníky. V oblasti zobrazené červeně došlo k posunu čtecího rámce vlivem delece. Čísla znázorňují pozici nukleotidů začínající od druhého exonu. (Převzato a upraveno podle Yan *et al.*, 2020)

## 3 EXPERIMENTÁLNÍ ČÁST

### 3.1 Návrh experimentů

Návaznost experimentů využitých pro tuto práci zobrazuje obrázek 9. Základem byly výsledky získané z GWAS dle Fang *et al.*, 2017. U identifikovaného genu *Glyma.03g258700* jsem prověřila případné existence dalších alel dle vzoru *T* (viz Obr. 7). Následně jsem hledala dodatečné geny potenciálně zodpovědné za změnu pigmentace trichomů u variet s dominantními alelami *T* a *Td* na chromosomu 12 a 16. Při tom jsem vycházela z logiky, že kandidátní geny by mohly být do určité míry homologní na již identifikované geny pro F3'H a/nebo MYB. Pro identifikaci kandidátních genů jsem provedla řadu analýz *in silico* za pomoci uvedených nástrojů/webových aplikací. Rozsah experimentů *in silico* je na obrázku 9 zobrazeno žlutě a laboratorního ověření je zobrazeno šedě. Výjimkou je sekvenování Sangerovou metodou, které je komerčním servisem (modrý text). Většinu laboratorní části nebylo možno absolvovat z důvodu pandemické situace v akademickém roce 2020-2021.



Obr. 9 Schéma experimentů. Bioinformatické nástroje jsou zobrazeny žlutě, zatímco práce prováděná v laboratoři šedě. Červená část dokumentuje mnou provedené experimenty.



## 3.2 Provedené experimenty

### 3.2.1 SNPviz v2.0

SNPviz v2.0 je webový nástroj pro zobrazování a analýzu haplotypů. Haplotyp je skupina genů v organismu, která zároveň označuje skupinu mutací děděných společně.

Na rozdíl od ostatních dostupných zobrazovačů haplotypů (HaploPainter – Peter a Nürnberg, 2004; Haploview – Barrett *et al.*, 2004; Haplotrips – Marnetto a Huerta – Sánchez, 2017) je tento obohacen o funkcionality, které kromě jednoduchého zobrazení zvolené oblasti genomu umožňují analyzovat asociaci fenotypu a pravděpodobnost, s jakou konkrétní mutace vysvětluje studovaný fenotyp. Dokáže vizualizovat velkokapacitní bloky haplotypů. Seřadí je podle jejich pozici v chromosomu s ohledem na přesnost fenotypu na genotyp, proteinovou rodinu nebo jejich funkci. Pro analýzu genomu sóji pomocí SNPviz v2.0 je možné využít veřejně dostupných data setů obsahujících soubory varet s resekvenovaným genomem (dohromady 775 varet, Škrabišová *et al.*, 2021a).

Na stránkách soykb.org v sekci nástroje jsem vybrala nástroj SNPviz v2.0. Zvolila jsem verzi genomu Wm82.a2.v1, ta je obecně vnímána jako nejpřesnější. Pro zadání konkrétní oblasti pro vizualizaci jsem použila koordináty genu *Td* alely *Glyma.03g258700* (podle databáze anotovaných genů SoyBase, <https://www.soybase.org/>). Startovní pozice je chr03: 45 300 865 a konečná je chr03: 45 302 850. Zvolila jsem možnost zobrazit všechny SNP a indely (inzerce a delece) na chromosomu 3. Jako data set jsem zvolila Soy775 (Škrabišová *et al.*, 2021a), u kterého byly zobrazované oblasti omezené pouze na kódující oblast genu (zvolením možnosti varianty tohoto data setu “Modifying Soy775 with 50K positions”). Nakonec jsem vložila fenotypový soubor (pro barvu trichomů). Jako WT je číslem 1 označen žlutohnědý, jako mutantní je číslem 2 označen šedý. Téměř žlutohnědý fenotyp je kódován číslem 3 a světle žlutohnědý 4. Klastrováním byly získány výsledky, které po úpravě dokumentuje Obr. 10 (v kapitole Výsledky a diskuze).

### 3.2.2 AccuTool

AccuTool (Accuracy Tool) je webový nástroj (<http://soykb.org/AccuTool/index.php>) pro post-GWAS analýzu míry asociace fenotypu s každou jednotlivou mutací v oblasti asociace genomu v definované oblasti chromosomu. Nástroj používá pro vyjádření této míry asociace tzv. přesnost (průměrnou přesnost neboli Avg a pesimistickou kombinovanou přesnost Com) a to pro soubor 775 variet sóji (<https://github.com/nad7wf/AccuTool>, Škrabišová *et al.*, 2021a). Nástroj používá jako referenční genotyp varietu Williams 82.

Nástroj jsem použila pro chromosom 12 i 16, ve kterých byly Fang *et al.*, 2017 metodou GWAS pro pigmentaci trichomů nalezeny statisticky významné signály (Obr. 6). Prohledávala jsem vždy oblast velkou 2 Mbp. U chromosomu 12 to byla konkrétně oblast 15 899 433-17 899 433 a u chromosomu 16 to byla oblast 440 263-2 440 263. Oblasti chr16:899 433 a chr12:1 440 263 byly ve Fang *et al.*, 2017 publikovány jako ty s nejvyšší asociací. Jako referenční fenotyp jsem vždy vybrala WT, to znamená v tomto případě žlutohnědé referenční variety. Nakonec jsem nahrála fenotypový soubor, který přiřadil reálný fenotyp pigmentace trichomů všem 775 varietám. V AccuTool je možno použít pouze binárního rozdělení fenotypů, WT (žlutohnědý) = 1, mutantní (šedý) = 2. Prohledávanou oblast jsem analyzovala dvakrát, jednou celou genomovou oblast a podruhé oblast omezenou na kódující oblasti genu. Mutace v kódujících oblastech již anotovaných genů se biochemicky vysvětlují snadněji než mutace, které se vyskytují v oblastech mimo geny.

Kandidátní geny jsem získala filtrováním výsledků, které mají přesnost pro WT fenotyp větší nebo rovnu 90 %, a to proto, že se předpokládá, že genetické pozadí původního žlutohnědého fenotypu všech variet je stejné.

Při hledání jsem ověřila i případné homologní geny pro F3'H a MYB. Seřazením výsledků přesnosti od největší přesnosti po nejmenší jsem postupně prohledala všechny geny, dokud jsem nenalezla takové, které by svou anotací naznačovali s plněním své funkce v metabolismu a regulace flavonoidů. Kandidátní geny s nejvyššími výsledky přesnosti jsem pro ověření srovnala pomocí nástroje BLAST s chromosomy 6 i 3.

### 3.2.3 BLAST

BLAST (basic local alignment search tool) je algoritmus využívaný v bioinformatice pro srovnání informací z primární sekvence proteinů se sekvencemi z databáze nebo srovnání na základě jejich domén (<https://www.soybase.org/GlycineBlastPages/index.php>). BLAST dokáže najít domény i kratší úseky sekvence, které jsou nad definovanou hranicí podobnosti (Altschul *et al.*, 1990). Analogickou variantou pro nukleové sekvence je BLASTn a pro proteinové BLASTp.

Zadáním ID genu *Glyma.03g258700* nebo *Glyma.06g202300* s výběrem modelu genů Glyma2.0 byly načteny informace o daném genu. V sekci “predicted protein sequences” (předpovězení proteinové sekvence) jsem zvolila BLAST “Sequences at SoyBase” (provést BLAST metodu na stránkách SoyBase). Jako program jsem zvolila “protein => protein”, hledání v databázi Williams 82 Assembly 2 Annotation 1 (Wm82.a2.v1) Protein Sequences. Snahou bylo ověřit hypotézu, že na chromosomech 12 a 16 se nacházejí geny pro proteiny se stejnou funkcí jako ty na chromosomech 3 a 6 (F3'H a MYB transkripční faktor).

### 3.2.4 Katalog alel sóji

Soybean Allele Catalog je webový nástroj na stránkách SoyKB (<http://soykb.org/GenescapeAnalysis/search.php>). Pomáhá při hledání nových alel a pro identifikaci SNP. Zobrazuje genetickou variaci napříč varietami divoké sóji, domestikovaných variet, a elitních kultivarů ze Severní Ameriky (agronomicky významné kultivary) a také uvádí počet všech variet s danou mutací pro předem zvolený gen. Počet variet je zatím 893, ale postupně se k nim přidávají další, které mají resekvenovaný genom. Podobně jako v nástroji SNPviz v2.0 tento zobrazuje pozice, na kterých se mutace nachází. Omezení nástroje spočívá v zobrazování kódujících oblastí regionů genů, ale nezobrazí například promotory nebo introny.

Vyhledávání v katalogu je dvojího typu. Dá se hledat jen podle ID (identifikace) genu, nebo se dají vyhledat konkrétní linie díky jejich PI nebo ID. Z nástroje SNPviz v2.0 jsem vybrala názvy všech variet s šedými trichomy, které nemají na chromosomu 6 žádnou z publikovaných mutací, která by vysvětlovala jejich barvu. Celkem jsem takto

identifikovala 7 variet. Vložila jsem jejich identifikační čísla do kolonky “accession” a název genu *Glyma.06g202300*.

Druhým způsobem hledání byl vyhledán homologní gen *Glyma.12g137000*. Zde neměla většina variet dostupných v tomto katalogu žádnou mutaci, což svědčí o pravděpodobné důležitosti konzervované sekvence genu. Mutace na pozici 16 285 557 je u HN040\_PI378702, která má ale žlutohnědé trichomy. Stejným způsobem jsem prohledala i homologní geny *Glyma.16g021200* a *Glyma.16g017500*.

### **3.2.5 Analýza zastoupení jednotlivých fenotypů pigmentace trichomů**

Pro grafické zobrazení zastoupení vysvětlených a nevysvětlených mutací (Obr. 11 a Obr. 12) jsem prohledala gen *Glyma.06g202300* (pro šedou) a *Glyma.03g258700* (pro světle žlutohnědou). Získaný počet vysvětlených mutací jsem odečetla od celkového počtu variet s daným fenotypem. Pro přesnější vyhodnocení počtu u šedých trichomů na chromosomu 6 jsem pomocí AccuTool našla marker ss715593787 s průměrnou přesností 91,3 %, kombinovanou přesností 59,9 %, přesností pro WT 97,3 % a pro mutantní 85,4 %. Díky němu mohu na 97 % říct, které trichomy jsou žlutohnědé.

### **3.3 Plánované experimenty**

#### **3.3.1 Potřebné chemikálie**

- Ethylendiamintetraoctová kyselina (Fluka, Česká republika)
- Chlorid sodný (Sigma-Aldrich, Německo)
- Kyselina chlorovodíková (Lach-Ner, Česká republika)
- Tris(hydroxymethyl)aminomethan (Sigma-Aldrich, Německo)
- Triton X-100 (Sigma-Aldrich, Německo)

#### **3.3.2 Přístroje a vybavení**

- Automatické pipety (Eppendorf, Německo)
- Analytické váhy (Sartorius, Německo)
- Digitální pH metr (InoLab, Německo)
- T100<sup>TM</sup> Thermal cycler (BioRad, USA)
- Elektroforetická cela (Bio-Rad USA)
- Elektroforetická cela (Bio-Rad USA)
- Elektromagnetická míchačka (Biosan, Lotyšsko)
- Gel Doc EZ systém (BioRad, Německo)
- Předvážky (KERN, Německo)
- Vortex V-1 plus (Biosan, Lotyšsko)

#### **3.3.3 Roztoky**

- TAE pufr (Tris-acetát-EDTA, 0,04 M Tris-kyselina octová, 0,001 M EDTA, pH 8,3)
- TENT pufr (10 mM Tris-Cl, 1 mM EDTA, 12 mM NaCl, 2,5% Triton X-100, pH 8)
- TE pufr (10 mM Tris-Cl, 0,1 mM EDTA, pH 8)

#### **3.3.4 Reagencie pro molekulární biologii**

- 5mM dNTP (Thermo Scientific, USA)
- Barvivo GelRed® (Biotinum, USA)
- Hmotnostní standard GeneRuler 1kbp DNA ladder (Thermo Scientific, USA)
- iProof<sup>TM</sup> High-Fidelity DNA polymerasa (BioRad, Německo)

- iProof™ High-Fidelity pufr (BioRad, Německo)
- Nanášecí pufr – 6x DNA Gel Loading Dye (Thermo Scientific, USA)
- NucleoSpin Gel and PCR Clean-Up (Macherey-Nagel, Německo)
- Primery *Gm09g278000fw* a *Gm09g278000rev* (Sigma-Aldrich, USA)
- Whatman® FTA® Cards (Sigma-Aldrich, USA)
- Wizard SV Gel and PCR Clean-Up Systém (Promega, USA)

### 3.3.5 Biologický materiál

Pro ověření nalezených alternativních mutací v genu *t* alely *Glyma.06g202300* by byly použity variety z databáze GRIN (USDA Germplasm Resources Information Network, <https://soybase.org/grindata/>), tabulka 3.

Tab. 3 Variety s šedými, téměř šedými, žlutohnědými a světle žlutohnědými trichomy, které by byly použity v této práci.

PI	USB ID	Barva trichomů
PI 497967	USB-226	
PI 166105	USB-004	
PI 549018	USB-035	
PI 379618	USB-143	
PI 548619	USB-271	
PI 504288	USB-227	
PI 548193	USB-022	
PI 567231	USB-295	
PI 417215	USB-154	
PI 417500	USB-159	
PI 567788	USB-411	
PI 165675	USB-108	
PI 404187	USB-148	
PI 548473	USB-262	žlutohnědá
PI 507180	USB-017	
PI 578495	USB-044	
PI 361093	USB-011	
PI 518668	USB-018	
PI 567525	USB-317	
PI 360957	USB-009	
PI 567426	USB-040	
PI 548656	USB-001	
PI 548978	USB-276	
PI 518727	USB-241	
PI 468908	USB-014	
PI 549028	USB-279	

Tab. 3 Variety s šedými, téměř šedými, žlutohnědými a světle žlutohnědými trichomy, které by byly použity v této práci (*pokračování*).

PI	USB ID	Barva trichomů
PI 342434	USB-128	
PI 548520	USB-032	
PI 549041 A	USB-037	
PI 592523	USB-045	
PI 123440	USB-101	
PI 438335	USB-013	
PI 437662	USB-182	
PI 594880	USB-357	
PI 603397	USB-372	
PI 603442	USB-050	
PI 592940	USB-346	žlutohnědá
PI 592954	USB-046	
PI 180501	USB-005	
PI 548325	USB-024	
PI 578309	USB-043	
PI 548256	USB-254	
PI 86904	USB-084	
PI 81785	USB-003	
PI 547481	(pa1pa2 NIL L67-497)	
PI 547528	(pa2 NIL L70-4558)	
PI 547634	(pa1 NIL L76-1291)	
PI 507471	USB-236	
PI 567407	USB- 306	
PI 603290	USB-369	
PI 567428	USB-312	
PI 567576	USB-320	
PI 548561	USB-033	
PI 507088	USB-232	
PI 603722	USB-388	
PI 567746	USB-326	
PI 549017	USB-277	
PI 417479	USB-158	šedá
PI 567558	USB-041	
PI 430595	USB-167	
PI 548447	USB-028	
PI 587804	USB-342	
PI 549026	USB-036	
PI 598358	USB-364	
PI 597476	USB-048	
PI 561387	USB-285	
PI 594307	USB-047	
PI 416838	USB-153	

Tab. 3 Variety s šedými, téměř šedými, žlutohnědými a světle žlutohnědými trichomy, které by byly použity v této práci (*pokračování*).

PI	USB ID	Barva trichomů
PI 407742	USB-151	
PI 324924	USB-127	
PI 495020	USB-223	
PI 612730	USB-392	
PI 548364	USB-026	
PI 602993	USB-367	
PI 603675	USB-386	
PI 507467	USB-235	
PI 361087	USB-010	šedá
PI 567675	USB-322	
PI 578504	USB-336	
PI 592952	USB-347	
FC 33243	USB-002	
PI 62203	USB-066	
PI 95860	USB-098	
PI 547742	(pa1pa2 NIL L69-6095)	
PI 547744	(pa2 NIL L70-4112)	
PI 547733	(pa1 NIL L67-271)	
PI 548490	USB-031	
PI 548452	USB-029	
PI 597464	USB-361	světle žlutohnědá
PI 548360	USB-025	
PI 090763	HN-004	

### 3.3.6 Navržení primerů

S pomocí programu Primer-BLAST ([https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK\\_LOC=BlastHome](https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome)) jsem navrhla primery pro gen *Glyma.16g16500*. Jeho nukleotidovou sekvenci jsem získala na stránkách SoyBase zadáním ID genu. Sekvenci jsem zkopírovala do sekce PCR template a jako organismus jsem zvolila *Glycine max* (taxid:3847). Program generuje primery automaticky na základě předvolených obecných kritérií a standardních parametrů pro PCR primery (Tab. 4).



Tab. 4 Výsledky navržení primerů pro *Glyma.16g16500* získané z NCBI Primer-BLAST. Navržení 10 vhodných párů primerů a jejich specifikace (T<sub>m</sub> je teplota tání).

Pár	Typ primeru	Sekvence od 5' konce	T <sub>m</sub> (°C)	Délka (NTP)	Obsah GC (%)	Délka ampliconu
1	přímý	GTTCCATGGACGGGGACAAA	60,25	20	55,00	142
	zpětný	ATGGCTTGGAACCCTGTCTG	59,96	20	55,00	
2	přímý	ACGGGGACAAAATGATGGGG	60,3	20	55,00	134
	zpětný	AATGGCTTGGAACCCTGTCT	59,22	20	50,00	
3	přímý	ATGGGGCAGAGACCTCATCA	60,33	20	55,00	122
	zpětný	CAAATGGCTTGGAACCCTGTG	59,73	21	52,38	
4	přímý	TGATGGGGCAGAGACCTCAT	60,33	20	55,00	124
	zpětný	CAAATGGCTTGGAACCCTGT	58,66	20	50,00	
5	přímý	ATGATGGGGCAGAGACCTCA	60,33	20	55,00	126
	zpětný	ACAAATGGCTTGGAACCCTG	58,66	20	50,00	
6	přímý	TTCCATGGACGGGGACAAA	59,52	20	50,00	140
	zpětný	TGGCTTGGAACCCTGTCTG	59,54	19	57,89	
7	přímý	GGGGACAAAATGATGGGGCA	60,62	20	55,00	132
	zpětný	AATGGCTTGGAACCCTGTCTG	60,55	21	52,38	
8	přímý	CAAGTTCCATGGACGGGGAC	60,68	20	60,00	147
	zpětný	AAATGGCTTGGAACCCTGTG	58,37	20	50,00	
9	přímý	GGGACAAAATGATGGGGCAG	59,17	20	55,00	135
	zpětný	TACAAATGGTTGGAACCCTG	58,48	21	47,62	
10	přímý	TTGCCAACTTCCATGGACGG	60,89	20	55,00	151
	zpětný	AAATGGCTTGGAACCCTGTCTG	61,08	22	50,00	

### 3.3.7 Karty typu FTA®, Whatman™

Karty byly vyvinuty pro sběr, transport, izolaci a archivaci nukleových kyselin při pokojové teplotě po několik let. Celá příprava trvá méně než 30 minut. Výhodou je, že při skladování nezaberou moc místa a dají se poslat poštou bez nutnosti zvláštní manipulace, což z nich činí žádaný a dostupný artikl. Rozsah jejich použití je široký. Využívají se v genomice, forenzní medicíně nebo při objevu nových léků. Karty FTA® (Flinders Technology Associates) byly testovány v celé škále vzorků, ať už na vzorcích mikroorganismů, krve nebo vzorcích rostlinných materiálů.

Karty jsou napuštěny pohlcovači volných radikálů a chemikáliemi, aby došlo k lýzi buněk, denaturaci proteinů a díky tomu jsou také rychle inaktivovány patogeny. Chemikálie zabraňují degradaci DNA, chrání vzorek proti oxidaci, napadení mikroorganismy, nukleasami a UV zářením. Po aplikaci vzorku na kartu se změní její zabarvení, aby bylo jasné, na kterém místě se nachází nanesený vzorek. Před analýzou je DNA třeba z nosiče nejprve uvolnit a purifikovat, přičemž ne vždy se podaří získat

dostatek použitelné DNA, případně může dojít ke kontaminaci vzorku. Po vyříznutí vzorku z karty je v mikrozkuhavce promyt dvěma pufry: čistícím činidlem a TE pufrem (10 mM Tris-Cl, 0,1mM EDTA, pH = 8) (de Vargas Wolfgramm *et al.*, 2008; Thompson a Hrabak, 2018).

Existuje více typů FTA® karet. Klasická karta má čtyři vzorkovací plochy pro 500 µl krve nebo 100 µl rostlinného homogenátu. Mini karta má o polovinu menší plochu a kapacitu pro vzorky. Karta pro nejmenší obsah vzorků je mikro karta s jednou vzorkovací plochou. Dále existují genetické, identifikační a testovací karty typu FTA®.

### 3.3.8 Příprava vzorku pro PCR

FTA karty obsahují DNA kultivarů sóji, která je na ně přenášena přitlačením listu, kdy jeho spodní strana směřovala směrem vzhůru. Použitý list je odebrán a karta se vzorkem tak může být skladována při laboratorní teplotě pro další použití. Je nutné pracovat sterilně pro zamezení kontaminace.

Pro izolaci DNA z FTA karet jsem sterilním skalpelem vyřízla z vhodného místa karty disk o průměru asi 1 mm. Ten jsem vložila do vhodně popsané mikrozkuhavky a pomocí automatické pipety jsem ho promyla 3x 100 µl TENT pufrem a 2x 100 µl TE pufrem. Disk jsem nechala vysušit v digestoři na půl hodiny.

### 3.3.9 PCR

Polymerázová řetězová reakce (= polymerase chain reaction) je klonovací technika *in vitro* nevyužívající mikroorganismy a sloužící primárně k amplifikaci požadovaného úseku DNA. Proces probíhá v cyklech (asi 25-40 opakování). Generuje velké množství kopií DNA sekvence i z velmi malého množství původního materiálu. Umožňuje rychlou selekci, izolaci a amplifikaci požadovaných úseků DNA a také můžeme díky ní zkoumat nukleotidové sekvence ze starodávných dochovaných vzorků, které byly vysušeny, zmrazeny, ukryty v anaerobních sedimentech nebo namočeny v alkoholu nebo formalinu (Simon *et al.*, 1991).

V roce 1971 bylo teoreticky předvedeno, jak by se tato metoda mohla provádět, ale vše bylo reálně provedeno až za dalších 14 let Karym Mullisem (1986). Velký význam ve vývoji má bakterie *Thermus aquaticus*, ze které byla izolována DNA polymerasa

(tzv. *Taq* polymerasa), která je tepelně stabilní do 95 °C (Lorenz, 2012). Dnes se používá termostabilnější *Pyrococcus furiosus*. Když PCR selže, může to vést k mnoha nespecifickým produktům DNA různých velikostí, které se projevují rozmazáním pásů na agarosovém gelu, nebo se nemusí tvořit žádné kopie původních sekvencí DNA (Lorenz, 2012).

PCR reakce probíhá ve třech krocích: denaturace templátu, annealing a fáze prodlužování neboli elongace. V prvním kroku dochází k rozpletení dvouvláknové templátové DNA tak, že teplota rychle vzroste na 92-96 °C, kde se udržuje 15 až 60 sekund. Poté teplota rychle klesne na přibližně 50 °C. To umožňuje navázání primerů, které jsou v přebytku. Primery směřují DNA polymerasu syntetizovat komplementární vlákna k původní DNA. Ve třetím kroku je teplota 72-74 °C pro co největší využití *Taq* polymerasy. Bylo zjištěno, že při této teplotě *Taq* přidává 35 až 100 nukleotidů za sekundu (Simon *et al.*, 1991). Reakční směs obsahuje templátovou DNA, dva oligonukleotidové primery (každý komplementární k vláknům DNA, které se amplifikují, velikost 18-25 bp-párů bází), pufr, čtyři nukleotidy dNTP ve stejných poměrech a *Taq* polymerasu. Cyklus končí, když se teplota opět zvýší na 92 °C a vlákna se vzdálí.

K vysušenému disku FTA karty jsem ve sterilním boxu vysvíceném UV na 20 minut podle tabulky 5 přidávala roztoky pro amplifikaci DNA. Takto připravené vzorky o objemu 50 µl jsem vložila asi na hodinu a půl do přístroje Thermal Cycler, který má předem nastavené podmínky (Tab. 6).

Tab. 5 Objemy roztoků potřebných pro přípravu reakční směsi do PCR a výsledné koncentrace.

Roztok	Objem (µl)	Výsledná koncentrace
iProof High-Fidelity pufr	10	1x
Mix dNTP	2	200 µM (každý dNTP)
Fw primer	1	0,5 µM
Rev primer	1	0,5 µM
Sterilní diH <sub>2</sub> O	35,75	
iProof High-Fidelity DNA polymerasa	0,25	0,02 U/µl
Celkový objem	50	

Tab. 6 Jednotlivé fáze PCR reakce a podmínky, za kterých probíhají (teplota, čas a počet cyklů).

Fáze	Teplota (°C)	Čas (s)	Počet cyklů
Iničiační denaturace	98	30	1
Denaturace	98	10	
Annealing	67	20	30
Elongace	72	30	
Finální elongace	72	300	1

### 3.3.10 Elektroforéza v agarosovém gelu

Gelová elektroforéza je separační metoda, sloužící k oddělení fragmentů nukleových kyselin vlivem odlišné pohyblivosti ve stejnosměrném elektrickém poli.

Agarosa, produkt extrahovaný z mořských řas, je lineární polymer derivátů galaktopyranosy. Gely se připravují rozpuštěním agarosy v teplém elektroforetickém pufru. Po ochlazení gelové směsi na 50 °C je agarosový roztok nalit mezi skleněné desky. Gely s obsahem agarosy méně než 0,5 % jsou poměrně křehké a musí být použity v horizontálním uspořádání (Boyer, 2012).

Vzorek, který se má separovat, se napipetuje do jamky vyrobené pomocí hřebínku a napětí je přiváděno, dokud nedojde k úplné separaci. Nukleové kyseliny lze na gelu vizualizovat nebo lze sledovat během elektroforézy barvivy jako například ethidiumbromidem nebo barvivy GelRed a GelGreen. Rychlost migrace nukleových kyselin v gelu závisí na procentuálním obsahu agarosy, přítomnosti barviv, velikosti napětí a na velikosti a konformaci fragmentu DNA. Rychlost migrace molekul je nepřímo úměrná logaritmu jejich molekulových hmotností (Boyer, 2012).

Pro elektroforézu jsem nejdříve připravila 0,8% agarosový gel rozpuštěním 0,8 g agarosy ve 100 ml 1x TAE pufru. Do gelu jsem vložila z obou stran hřebínek pro vytvoření jamek. Když gel ztuhl, přenesla jsem ho do vodivostní cely a zalila TAE pufrem.

Do první jamky jsem pomocí pipety nanasla hmotnostní standard GeneRuler 1kb DNA ladder a do zbylých jamek vždy po 5 µl vzorku smíchaného s 1 µl nanášecího pufru Gel Loading Dye. Elektroforéza probíhala při napětí 90 V po dobu jedné hodiny.

Po hodině jsem vodivostní celu odpojila od zdroje napětí a gel jsem barvila barvivem GelRed® na třepačce. Vizualizaci gelu jsem provedla pomocí programu Image Lab Software v přístroji Gel Doc™ EZ Imager (BioRad).

### **3.3.11 Přečištění produktu PCR**

Pro přečištění DNA se používají dva sety. Set Wizard SV Gel and PCR Clean-Up System se používá jako první. Ke vzorkům se přidá roztok Membrane Binding Solution (MBS) v poměru 1:1. Do mikrozkušavky s kolonkou se silikátovým nosičem byl napipetován roztok amplifikované DNA s MBS roztokem. Směs se v mikrozkušavce inkubuje asi 1 minutu a poté je 1 minutu centrifugována při 16 000 g při teplotě -4°C. Po odlití filtrátu se k DNA navázané na nosič přidává 700 µl promývacího roztoku Membrane Wash Solution (MWS) a vše je centrifugováno za stejných podmínek jako poprvé. Filtrát se odlije, přidá se 500 µl MWS a centrifuguje se 5 minut při 16 000 g. Odlitím filtrátu a centrifugováním byly odstraněny zbytky promývacího roztoku, kolonka je přenesena do mikrozkušavky o objemu 1,5 ml a je k ní přidáno 50 µl vody. Po jedné minutě je provedena centrifugace po dobu 2 minut při 16 000 g. Přečištěný produkt je uchován v mikrozkušavce.

Po spotřebě prvního setu je použit druhý, Macherey-Nagel NucleoSpin Gel and PCR Clean-Up. Ke vzorkům se přidá roztok NT1 ve dvojnásobném množství a směs je pipetována do kolonky v mikrozkušavce. Po centrifugaci, která trvá 30 s při 11 000 g se odlije filtrát a přidá se 700 µl roztoku NT3. Vše je centrifugováno za stejných podmínek. Po odlití druhého filtrátu je 1 minutu prováděna centrifugace pro odstranění NT3. Ke kolonce v čisté mikrozkušavce se napipetuje 15 µl pufru NE. Po inkubaci trvající 1 minutu při laboratorní teplotě se provede poslední centrifugace při 11 000 g po dobu 1 minutu. Přečištěný PCR produkt byl uchován v mikrozkušavce, kolonka byla odstraněna.

### **3.3.12 Spektrofotometrické stanovení koncentrace DNA**

Pro sekvenaci je nutné nejdříve spektrofotometricky stanovit koncentraci DNA. K tomu se používá mikrodestičkový reader Synergy H1 a software Gen5 3.06 (Biotek, USA). Jako program je nejvhodnější program kvantifikace nukleových kyselin pro dvouvláknovou DNA. Měří se vzorek (1 µl) proti blanku (destilovaná voda).

### 3.3.13 Sekvenování Sangerovou metodou

Metoda sloužící ke zjištění pořadí nukleotidů v DNA nebo v RNA. Pro sekvenování je důležitá přítomnost primeru, DNA polymerasy, všech čtyřech dNTP (deoxyribonukleotidtrifosfátů) a ddNTP (dideoxyribonukleotidtrifosfát). Když je na nově syntetizované vlákno DNA přidán ddNTP, dochází k terminaci řetězce, protože ddNTP neobsahuje žádnou OH<sup>-</sup> skupinu, přes kterou by se mohly navázat další nukleotidy (Boyer, 2012). Je možné separovat takové fragmenty DNA, které se liší jen jedním nukleotidem. Detekce fragmentů se provádí chemiluminiscenčním značením nebo značením <sup>32</sup>P (Sanger *et al.*, 1977; Dovichi, 1997).

Vzorky pro sekvenování se připravují smícháním 5 μl vzorku s 5 μl 5 μM roztoku forward primeru. Pokud není množství PCR produktu v 5 μl vzorku kolem 200 ng, musí se upravit poměr vzorku ku primeru. Sekvenování vzorků známé koncentrace bylo provedeno jako komerční servis u firmy SEQme (SEQme s.r.o., Dobříš, Česká republika).

### 3.3.14 Analýza výsledků pomocí programu BioEdit

BioEdit sice neslouží přímo pro sekvenční analýzu, ale pomáhá s úpravou, manipulací a analýzou sekvencí, jejichž délka i počet je omezený. Program fungující ve verzích Windows 95/98/NT/2000/XP je volně přístupný a přizpůsobený běžným uživatelům (Hall, 2004).

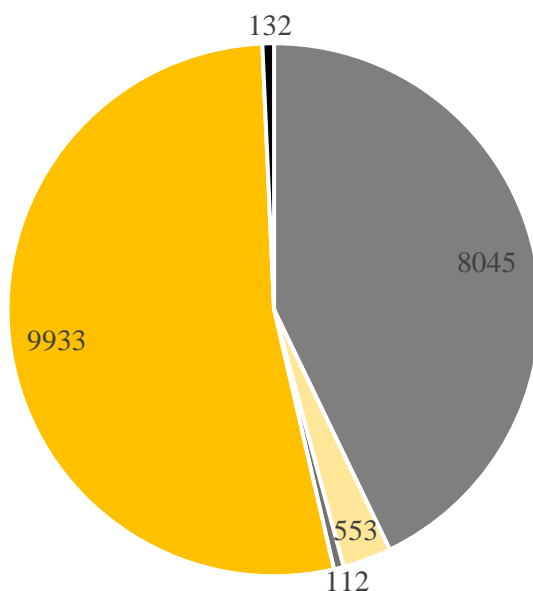
## 4 VÝSLEDKY A DISKUZE

### 4.1 Analýza zastoupení jednotlivých fenotypů pigmentace trichomů variet sóji z databáze GRIN

Databáze GRIN obsahuje kolem 20 tisíc variet sóji (*Glycine max* i *Glycine soja*), přičemž fenotyp pro pigmentaci trichomů byl stanoven u 18643 z nich, původní žlutohnědá barva byla pozorována pouze u 53,28 % variet, kdežto alternativní šedá u téměř 43,15 % variet. To naznačuje, že šedé trichomy jsou alternativním fenotypem s vysokou frekvencí výskytu. Na obrázku 10 můžeme vidět kromě největšího zastoupení variet s žlutohnědými a šedými trichomy 132 variet je bez známého fenotypu trichomů a dále 112 variet s téměř šedými a 553 se světle žlutohnědými trichomy. Šedé i žlutohnědé zbarvení je vysvětleno mutacemi na chromosomech 6 a 3. Světle žlutohnědé i šedé trichomy nemají dosud známé specifické mutace. Celkem zbývá popsat mutace u 685 (3,67 %) variet se známým fenotypem a u 132 chybí fenotyp popsat.

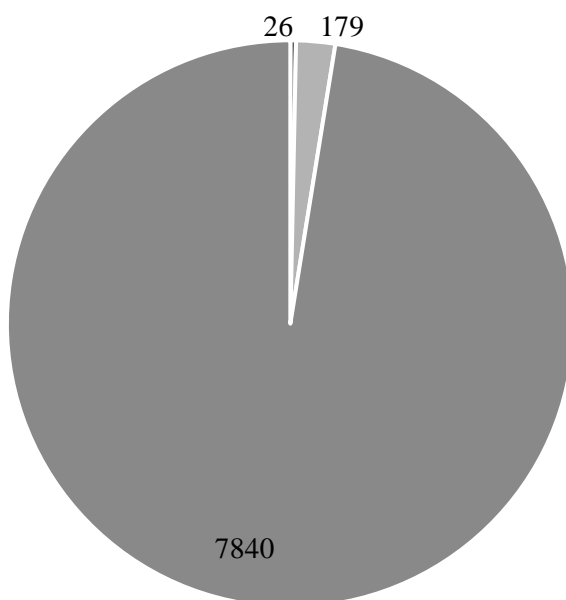
Zastoupení variet s šedými trichomy s vysvětleným a nevysvětleným fenotypem zobrazuje graf na obrázku 11, kde vidíme, že počet neosekvenovaných variet je 7840 a počet variet s resekvenovaným genomem je 205. 179 s publikovanými mutacemi *t* alely a 26 s funkčními (dominantními) *T* alelami (jen 0,32 %).

V grafech na obrázcích 12 a 13 jsou znázorněny variety se světle žlutohnědými a téměř šedými trichomy, jejichž fenotyp je vysvětlený publikovanými mutacemi v *Td* v genu *Glyma.03g258700* a světle žlutohnědé a téměř šedé, které zatím nemají osekvenovaný celý genom, a tedy nemají vysvětlený fenotyp. Z grafu na obrázku 13 plyne, že většina variet se světle žlutohnědými trichomy obsahuje alternativní alelu *td*, stejně tak jako variety s téměř šedými trichomy. Z nich pouze jedna má REF alelu *Td*.



■ Šedé ■ Světle žlutohnědé ■ Téměř šedé ■ Žlutohnědé ■ S neznámým fenotypem

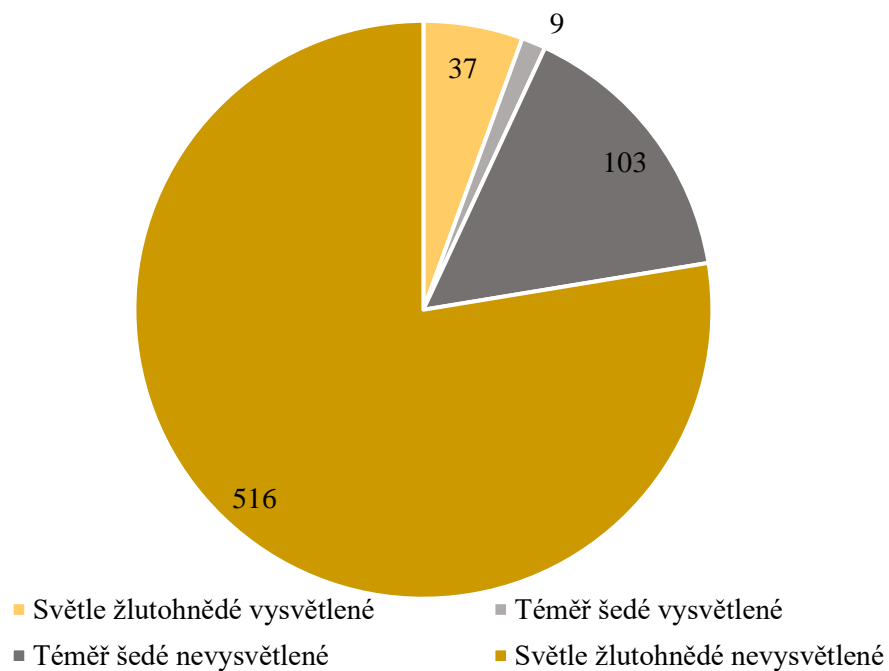
Obr. 10 Koláčový graf zobrazující zastoupení jednotlivých druhů pigmentace trichomů sóji všech variet z databáze GRIN. Největší zastoupení variet je se žlutohnědými a šedými trichomy. Zbývá popsat mutace u 112 téměř šedých a 553 světle žlutohnědých trichomů. 132 variet nemá fenotyp popsán.



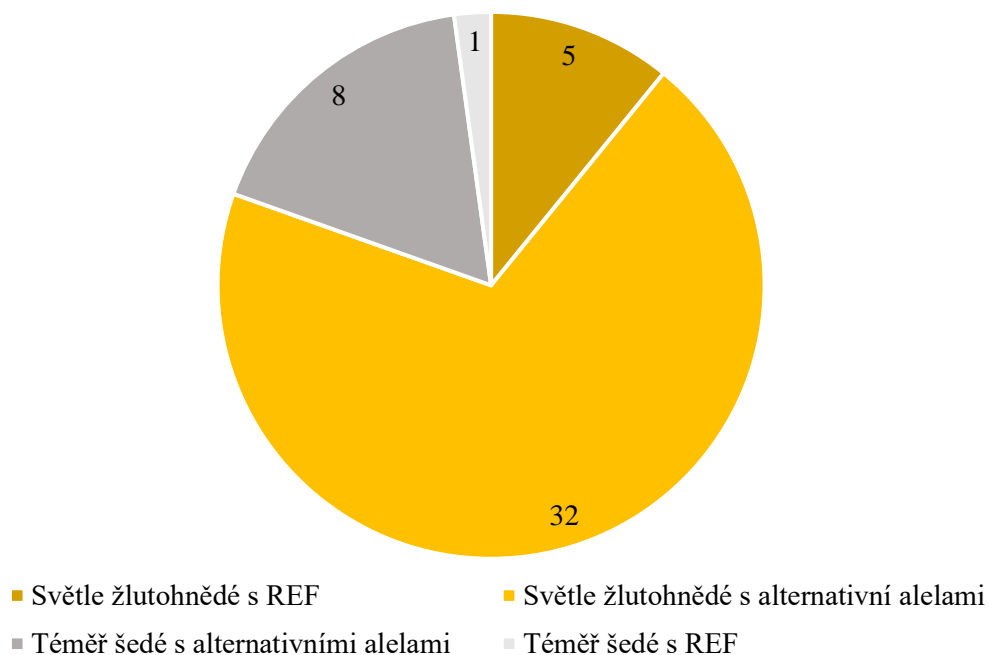
■ REF ■ Vysvětleno alternativními alelami mutacemi ■ Neosekvenovaný celý genom

Obr. 11 Koláčový graf zobrazující zastoupení variet s šedými trichomy s fenotypem vysvětleným alternativními alelami (*t* alelu má 179 variet), vysvětleným REF *T* alelou (26 variet) a s neosekvenovaným celým genomem (celkem 7840 variet).





Obr. 12 Koláčový graf variet s původní žlutohnědou a s téměř šedou pigmentací trichomů s resekvenovaným (obsahující *Td* alelu nebo alternativní alelu na genu *Glyma.06g258700*) a neosekvenovaným genomem.



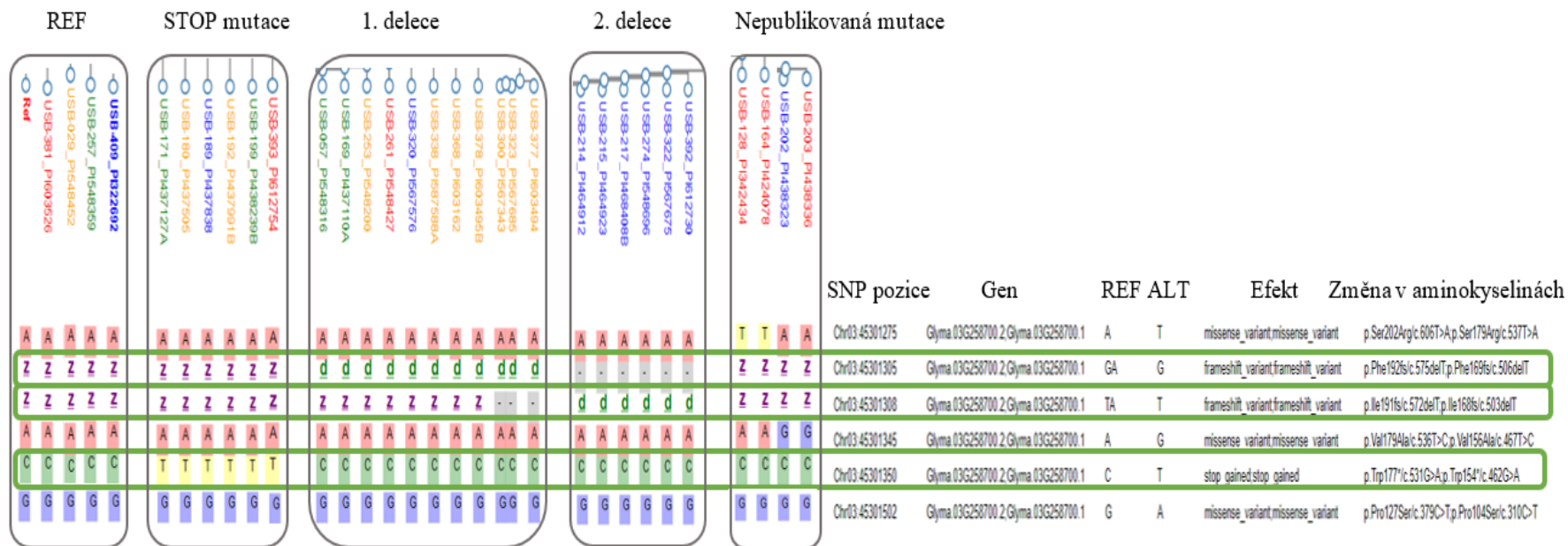
Obr. 13 Koláčový graf variet s resekvenovaným genomem se světle žlutohnědými a téměř šedými trichomy. Většina variet se světle žlutohnědými (32) a téměř šedými (8) trichomy je vysvětlena alternativní alelou *td*. 1 varieta s téměř šedými a 5 variet se světle žlutohnědými trichomy jsou vysvětleny REF alelou *Td* (gen *Glyma.03g258700*).

## 4.2 Ověření případných alternativních *td* alel pomocí SNPviz v2.0

Nástroj SNPviz v2.0 pro *Glyma.03g258700* byl využit pro nalezení nových mutací, které by dovysvětlily fenotyp u variet neobsahujících dosud publikované mutace analogicky s nově identifikovanými mutacemi v alele *t*. Nástroj ukázal mutace publikované Yan *et al.* (2020), které by podle autorů měly být specificky přítomny u variet s téměř šedými trichomy (Obr. 8). Výsledky můžeme vidět na obrázku 14. Publikované mutace jsou v zeleném rámečku. Jedná se o dvě delece a mutaci vedoucí ke vzniku stop kodonu. 1. delece na pozici chr03:45 301 305 ukazuje, že se nachází nejčastěji u trichomů se světle žlutohnědou pigmentací. Je ale možné, že se mutace podílí na více typech zbarvení. Naopak 2. delece na pozici 45 301 308 se s vysokou pravděpodobností vyskytuje jen u šedých trichomů, není to mutace specifická pro zředění šedé barvy. Existuje ještě třetí publikovaná mutace, která má oproti REF alele (Obr. 14) místo C na pozici chr03: 45301350 T a vede tak k předčasnému stop kodonu. Pomocí SNPviz v2.0 (Obr. 14) se ukázalo, že je přítomna u všech studovaných typů pigmentace. V posledním sektoru na obrázku 6 jsou první dva sloupce s mutací, která nezpůsobuje v konečném důsledku změnu aminokyseliny, stejně tak jako mutace v dalších dvou sloupcích, kde je zařazen G.

Na chromosomu 3 se mi podařilo najít všechny tři publikované mutace, ale žádnou novou, která by způsobovala změnu aminokyseliny a byla by tak zodpovědná za fenotyp téměř žlutohnědých, světle šedých a šedých variet s dominantními *T* a/nebo *Td* alelami. Pozorovala jsem celou řadu dalších mutací, které se nejeví jako specifické vůči sledovanému fenotypu, ale naznačují, že gen *Glyma.06g258700* nemá konzervovanou sekvenci.

Manhattan plot pro barvu trichomů z Fang *et al.*, 2017 (Obr. 6) ukázal signály na chromosomech 3, 6, 12 a 16, kdy na chromosomu 6 byl nejsilnější, a kromě publikovaných mutací byly pomocí SNPviz v2.0 nalezeny i dvě nové, které by mohly být zodpovědné za barvu trichomů (Obr. 7) (Zeng a Škrabišová, 2021). Zbylé linie neobsahujících ani jednu z již publikovaných mutací by tedy mohly mít mutace na chromosomu 12 nebo 16. Na obou je dokonce signál z GWAS (Obr. 6) vyšší než na chromosomu 3. Mohly by být nalezeny nové mutace na jednom nebo druhém chromosomu, které by vysvětlovaly zbarvení trichomů.



Obr. 14 SNPviz v2.0 vizualizace *Td/tr* alel v *Glyma.03g258700* (pouze kódující oblast genu). Červeně je zobrazený žlutohnědý, modře šedý, zeleně téměř šedý a žlutě světle žlutohnědý fenotyp trichomů. Zelený rámeček znázorňuje tři dosud publikované mutace. Nově nalezené mutace nejsou specifické pro pigmentaci trichomů. (Upraveno podle SoyKB: Soybean Knowledge Base)

### 4.3 Analýza přesnosti asociace fenotypu pro mutace v kandidátních genech pigmentace trichomů pomocí AccuTool

Sestupně seřazené výsledky průměrných přesností u chromosomů 12 a 16 byly prohledány pomocí odkazů na anotace pro daný gen v SoyBase. Tímto byl získán jeden kandidátní gen na chromosomu 12, *Glyma.12g137000* anotovaný jako flavonol reduktasa/cinnamoyl-CoA reduktasa/dihydroflavonol 4-reduktasa, která by se dle anotované biochemické funkce mohla podílet na biosyntéze flavonoidů a tedy i pigmentů trichomů sóji. Dva kandidátní geny byly nalezeny na chromosomu 16, a to *Glyma.16g021200* a *Glyma.16g017500*. První z nich obsahuje cytochrom P450, má monooxygenasovou/flavonoid-3'-monooxygenasovou aktivitu a podílí se na biosyntéze luteolinu a flavonolu. Druhý nalezený gen na chromosomu 16 se podílí na biosyntéze flavonoidů a kóduje mimo jiné flavanon-3-dioxygenasovou aktivitu.

Výsledky jsou shrnuty v tabulkách 7, 8, 9 a 10, kde jsou ukázány jen reprezentativní výsledky. Čísla znázorňující různé druhy filtrování mají tento význam: 1 = varianty seřazené sestupně podle průměrné přesnosti, 2 = varianty seřazené podle kombinované pesimistické přesnosti, 3 = filtrované varianty podle přesnosti pro WT = 100 % a zároveň seřazené podle průměrné přesnosti, 4 = filtrování podle nejvyšší přesnosti pro WT a zároveň seřazené podle nejvyšší průměrné přesnosti, 5 = filtrování pro mutantní přesnost  $\leq 90$  % seřazeno sestupně podle kombinované pesimistické přesnosti, 6 = homologní geny filtrovány pro mutantní přesnost  $\leq 90$  % seřazeno sestupně podle průměrné přesnosti, 7 = varianty bez modifikací filtrovány pro mutantní přesnost  $\leq 90$  % seřazeno sestupně podle kombinované pesimistické přesnosti.

U genů v tabulce 5 filtrovaných pro mutace nacházející se v kódujících oblastech genů je vidět, že daná oblast je zatížena nedostatečnou kvalitou sekvenování, která se odráží v množství variet s chybějící informací o přítomnosti referenční nebo alternativní alel. Bylo potvrzeno, že neexistují žádné popsané geny (u variet s šedými trichomy), které by neměly homologní pozadí u *T* fenotypu.

Z filtrování podle 6 plyne, že neexistuje 100% přesnost pro žlutohnědý fenotyp (WT), což je ale v rozporu s hypotézou, že všechny variety se žlutohnědými trichomy mají homogenní pozadí. Výsledky filtrováním podle 7 (kandidátní gen) říkají, že mutace zodpovědné za šedý fenotyp by mohly vzniknout v rámci jedné evoluční příhody/mutace, pokud by všechny variety s šedými trichomy nevysvětlené mutacemi

na chromosomu 6 byly vysvětleny mutacemi na chromosomu 12. Získané výsledky jsou ale výsledky jen z kódujících oblastí genů. Ve filtrování podle 7 jsou seřazena data sestupně podle kombinované pesimistické přesnosti, a ne podle průměrné přesnosti, protože tím bychom se nezbavili velkého procenta chybějících dat.

Tab. 7 Výňatek výsledků z AccuTool pro chromosom 12 zahrnující jen modifikace. Obsahuje ID genu, průměrnou přesnost (Avg), kombinovanou pesimistickou přesnost (Comb), přesnost pro WT a pro mutantní typ (MUT). Rozdělení výsledků podle zobrazení všech SNP nebo SNP způsobujících modifikaci. Filtrování podle sedmi různých kritérií. Rozdělení podle filtrování. 1 je seřazení podle Avg, 2 podle Com, 3 podle WT = 100 % a Avg a 4 podle Avg a WT. Zobrazení REF a ALT alel a výsledného efektu (posun čtecího rámce nebo změna aminokyseliny – AA).

Filtr	Gen	Pozice	REF	ALT	Efekt	Avg (%)	Com (%)	WT (%)	MUT (%)
1	<i>Glyma.12g138800</i>	16 740 908	G	T	C335F	64,4	23,4	34,7	94,1
2	<i>Glyma.12g140600</i>	16 241 515	G	T	C855F	53,9	40,8	99,7	8,1
	<i>Glyma.12g137700</i>	16 411 095	C	T	A208V	53,7	40,8	99,7	7,7
3	<i>Glyma.12g138800</i>	16 736 080	CAA GT	C	posun	51,9	8,6	100,0	3,7
4	<i>Glyma.12g140600</i>	17 394 105	CAA	C	posun	53,9	40,8	99,7	8,1

Tab. 8 Výňatek výsledků z AccuTool pro chromosom 12 zahrnující i oblasti bez modifikací. Obsahuje ID genu, průměrnou přesnost (Avg), kombinovanou pesimistickou přesnost (Comb), přesnost pro WT a pro mutantní typ (MUT). Rozdělení výsledků podle zobrazení všech SNP nebo SNP způsobujících modifikaci. Filtrování podle sedmi různých kritérií. Rozdělení podle filtrování. 1 je seřazení podle Avg, 2 podle Com, 3 podle WT = 100 % a Avg, 6 pro MUT ≤ 90 % podle Avg a filtrování 7 je pro MUT ≤ 90 % seřazeno podle Com. Zobrazení REF a ALT alel a výsledného efektu (posun čtecího rámce nebo změna aminokyseliny – AA).

Filtr	Gen	Pozice	REF	ALT	Efekt	Avg (%)	Com (%)	WT (%)	MUT (%)
5	<i>Glyma.12g140800</i>	17 528 650	AA	A	downstream	54,0	29,5	9,0	99,0
			CT						
			TG						
			CT						
			AA						
CT									
1	-	15 906 349	G	A	intergenic	100,0	0,3	100,0	99,7
			C	A	intergenic	79,0	48,5	59,4	98,6
3	<i>Glyma.12g140900</i>	17 604 647	A	G	upstream	100,0	0,8	100,0	100,0
			-	T	intergenic	100,0	0,8	100,0	100,0
			-	AG	intergenic	100,0	0,8	100,0	100,0
6	<i>Glyma.12g137000</i>	16 287 895	G	A	upstream	59,5	39,9	76,3	42,6
			C	T	downstream	51,5	39,1	96,7	6,2
7	-	16 901 926	C	A	intergenic	79,0	48,5	59,4	98,6

Tab. 9 Výňatek výsledků z AccuTool pro chromosom 16 zahrnující jen modifikace. Obsahuje ID genu, průměrnou přesnost (Avg), kombinovanou pesimistickou přesnost (Comb), přesnost pro WT a pro mutantní typ (MUT). Rozdělení výsledků podle zobrazení všech SNP nebo SNP způsobujících modifikaci. Filtrování podle sedmi různých kritérií. Rozdělení podle filtrování. 1 je seřazení podle Avg, 2 podle Com, 3 podle WT = 100 % a Avg, 4 podle Avg a WT a 5 pro MUT ≤ 90 % seřazeno podle Com. Zobrazení REF a ALT alel a výsledného efektu (posun čtecího rámce nebo změna aminokyseliny – AA).

Filtr	Gen	Pozice	REF	ALT	Efekt	Avg (%)	Com (%)	WT (%)	MUT (%)
1	<i>Glyma.16g018400</i>	1 629 041	C	T	D1502N	65,4	39,7	42,9	87,9
2	<i>Glyma.16g024800</i>	2 407 980	C	T	D383N	61,0	42,3	84,8	37,3
	<i>Glyma.16g024700</i>	2 405 127	A	G	I91V	61,5	41,9	83,9	39,0
3	<i>Glyma.16g008900</i>	775 364	G	A	H717Y	51,5	5,7	100,0	3,0
4	<i>Glyma.16g012800</i>	1 124 593	T	A	K245M	57,2	41,3	92,3	22,1
5	<i>Glyma.16g019200</i>	1 693 169	TG	T	posun	64,2	38,6	37,4	90,9
	<i>Glyma.16g019300</i>	1 699 364	A	T	S442C	64,0	38,3	36,7	91,3

Tab. 10 Výňatek výsledků z AccuTool pro chromosom 16 zahrnující jen modifikace. Obsahuje ID genu, průměrnou přesnost (Avg), kombinovanou pesimistickou přesnost (Comb), přesnost pro WT a pro mutantní typ (MUT). Rozdělení výsledků podle zobrazení všech SNP nebo SNP způsobujících modifikaci. Filtrování podle sedmi různých kritérií. Rozdělení podle filtrování. 1 je seřazení podle Avg, 2 podle Com, 3 podle WT = 100 % a Avg a filtrování 7 je pro MUT ≤ 90 % seřazeno podle Com. Zobrazení REF a ALT alel a výsledného efektu (posun čtecího rámce nebo změna aminokyseliny – AA). Gen označený \* je kandidátní gen navržený Fang *et al.*, 2017.

Filtr	Gen	Pozice	REF	ALT	Efekt	Avg (%)	Com (%)	WT (%)	MUT (%)
1	* <i>Glyma.16g016500</i>	1 443 234	TG GA	T	downstream	100,0	11,1	100,0	100,0
2	<i>Glyma.16g016500</i>	1 440 253	G	A	upstream	78,9	49,2	65,6	92,2
	-	1 446 182	A	G	intergenic	83,3	4,1	100,0	66,7
3	<i>Glyma.16g016500</i>	1 441 937	T	G	downstream	75,0	3,5	100,0	50,0
	<i>Glyma.16g020200</i>	1 832 717	AT	A	intron	71,5	1,0	100,0	42,9
			AT						
AT									
7	<i>Glyma.16g016500</i>	1 440 253	G	A	upstream	78,9	49,2	65,6	92,2
	<i>Glyma.16g016500</i>	1 442 501	C	T	downstream	79,6	47,6	66,7	92,4

#### 4.4 Identifikace genů kódujících tvorbu pigmentů v sóji pomocí Katalogu alel sóji

Cílem této analýzy bylo zjistit, zda-li se na chromosomech 12 a 16 nacházejí případné paralogy genů F3'H lokusu *T* a MYB transkripčního faktoru lokusu *Td*. Na základě podobnosti s proteinovými sekvencemi F3'H a MYB nebyly signifikantně nalezeny shodné sekvence na chromosomech 12 ani 16. Kandidátní geny nalezené pomocí nástroje AccuTool si se sekvencemi F3'H a MYB podobné nejsou.

Pomocí nástroje Katalog alel sóji jsem hledala, zda variety s šedými trichomy, které nemají publikované mutace na chromosomu *Glyma.06g202300* nemají mutace na mnou nalezených kandidátních genech *Glyma.12g137000*, *Glyma.16g021200* a *Glyma.16g017500*.

Prohledáváním v genu *Glyma.06g202300* byly nalezeny dvě variety s REF alelami (HN038\_PI248515, PI\_546044) a dalších pět variet bez publikovaných specifických mutací vysvětlujících jejich šedou barvu trichomů. Prohledáváním homologního genu *Glyma.12g137000* nebyla nalezena žádná z variet, která na chromosomu 6 nemá mutaci zodpovědnou za její fenotyp. Na chromosomu 12 nebyla nalezena žádná specifická mutace pro námi hledaný fenotyp. Pro gen *Glyma.16g021200* nebyla nalezena žádná mutace. U druhého homologního genu *Glyma.16g017500* sice byla nalezena mutace, ta ale není u žádné z námi hledaných variet. Žádný z nalezených homologů nemá specifickou mutaci, která by dovysvětlila zbylé linie. Nástroj ale nezobrazuje například promotory, proto bylo nutné znovu použít nástroj SNPviz 2.0, kde jsem se podívala jen na variety s nevysvětleným fenotypem. Analýza, kterou jsem provedla, prohloubila vědění v této oblasti.

#### 4.4.1 Ověření kandidátního genu *Glyma.16g016500* pro tvorbu pigmentů v sóji

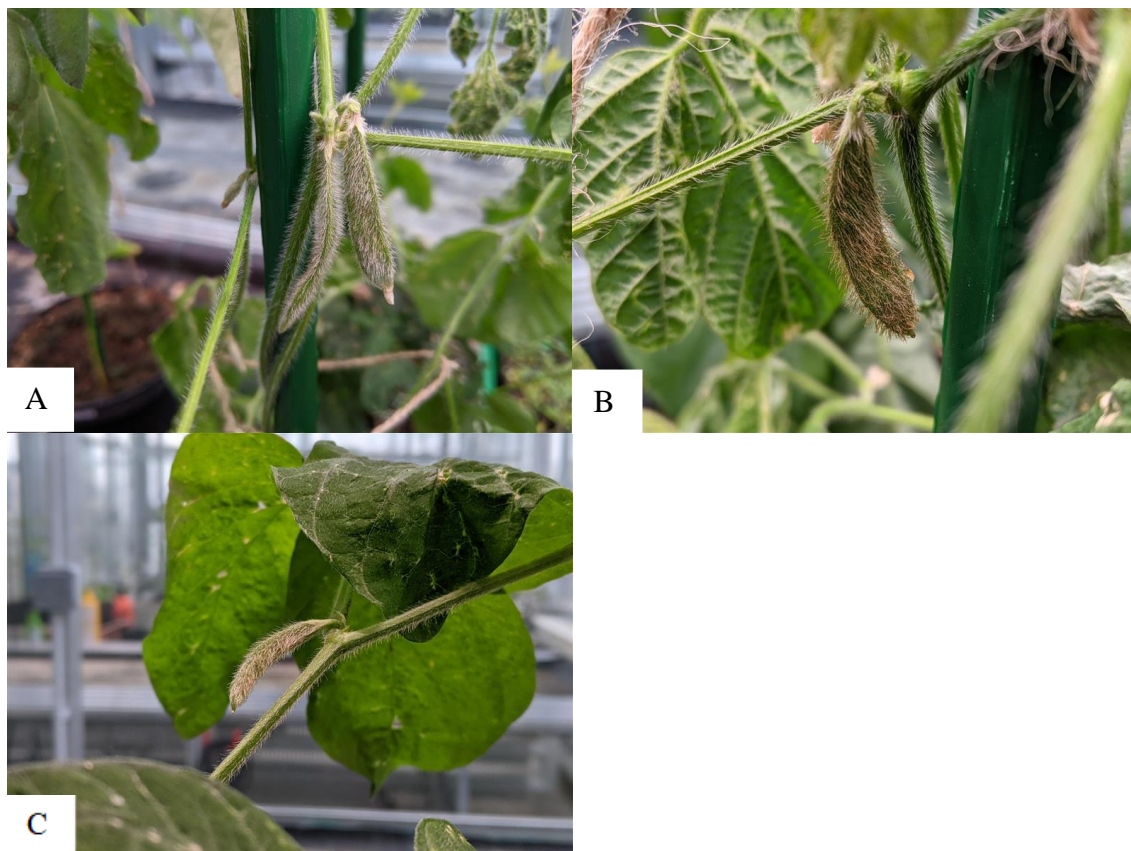
Z dat od Fang *et al.*, 2017 získaných z nástroje AccuTool pro nalezení kandidátních genů na chromosomech 12 a 16 (Tab. 11) byla pro gen *Glyma.16g016500* provedena analýza haplotypu pomocí zobrazovače SNPviz v2.0 a analýza vyhledáváním přirozené genetické variability v konkrétních genech a varietách pomocí nástroje Katalog alel sóji (Tab. 12). Byla nalezena mutace vedoucí k předčasnému stop kodonu u jedné z variet, ale 889 variet má REF alelu. Kandidátní gen *Glyma.16g016500* bych potvrdzovala u reprezentativní variety se světle žlutohnědým fenotypem trichomů (Obr. 15).

Tab. 11 Výsledky hledání homologního genu na chromosomu 12 a 16 získaných z AccuTool.

Chromosom	Gen	Pozice	REF	ALT	Avg (%)	Comb (%)	WT (%)	MUT (%)
12	-	16 899 433	G	T	46,6	31,2	83,9	9,2
16	<i>Glyma.16g016500</i>	1 440 263	T	C	63,7	37,2	34,1	100,0

Tab. 12 Výsledek z Katalogu alel sóji pro homologní gen *Glyma.16g016500*. U jedné variety byl na pozici 1441051 nalezen stop kodon. Většina variet má REF alely.

Chromosom	Pozice	Počet variet	Mutace
16	1 441 136	2	C V79A
	1 441 051	1	stop kodon
	1 440 852	1	T V3L
	-	889	REF alela



Obr. 15 Potvrzení všech kandidátních mutací navržených v této práci. A (USB-320, šedá, PI 567576), B (USB-017, PI507180, žlutohnědá), C (USB-029, světle žlutohnědá, PI 548452) - reprezentativní varieta pro potvrzení identifikovaných kandidátních genů Fang *et al.*, 2017.



## 5 ZÁVĚR

Byla vypracována literární rešerše na téma pigmentace trichomů, rozdílná akumulace pigmentů, biosyntéza flavonoidů a regulace těchto drah. Dále byl popsán vliv trichomů na albedo a jejich účinky na lidské zdraví. Věnovala jsem se také známým alelám a genům, které se týkají zbarvení trichomů sóji luštinaté.

Díky databázi GRIN a Katalogu alel sóji byly vytvořeny grafy pro znázornění zastoupení jednotlivých barev trichomů, přičemž je zde i uvedeno, jaká část variet již byla resekvenována a/nebo je známá mutace způsobující jiný než žlutohnědý fenotyp trichomů.

Kandidátní geny byly identifikovány a ověřeny pomocí bioinformatických nástrojů SNPviz v2.0, AccuTool, BLAST a pomocí Katalogu alel sóji. Analýzou celého genomu pomocí nástroje SNPviz v2.0 byly potvrzeny tři publikované mutace v genu *Glyma.03g258700*, ale nebyly však pozorovány žádné dodatečné mutace, které by byly specifické pro variety s šedými trichomy. Byly identifikovány kandidátní geny na chromosomu 12 a 16. Pomocí nástroje AccuTool byly po sestupném seřazení výsledků podle průměrné přesnosti identifikovány díky anotacím na SoyBase kandidátní geny *Glyma.12g137000*, *Glyma.16g021200* a *Glyma.16g017500*. Byl také nalezen gen *Glyma.16g016500*, který predikovali autoři Fang *et al.*, (2017) Jeho průměrná přesnost byla ale jen 63,7 % a jeho funkce nebyla potvrzena anotací SoyBase jako relevantní k biosyntetické dráze spojené s tvorbou pigmentů v trichomech. Ověřením těchto genů pomocí BLAST nebyla ani v jednom případě nalezena podobnost jejich sekvencí se sekvencemi genů *Glyma.03g258700* a *Glyma.06g202300*. V katalogu alel pro geny *Glyma.12g137000*, *Glyma.16g021200* a *Glyma.16g017500* nebyly nalezeny žádné mutace, které by odpovídaly varietám s nevysvětleným fenotypem v genech a mohly by dovysvětlit fenotyp. Analýzou genu *Glyma.16g016500* bylo zjištěno, že žádná z jeho nalezených mutací není specifická vůči fenotypu šedých trichomů, a proto na základě analýz v této práci může být tento kandidátní gen vyloučen.

Nalezené kandidátní geny neobsahují specifické mutace, které by dovysvětlily fenotyp pigmentace trichomů. Vyvrátila jsem Fang *et al.*, 2017 predikovaný kandidátní gen *Glyma.16g016500*, který je konzervovaný - neobsahuje žádné potenciálně významné mutace.

## 6 LITERATURA

- Altschul S. F., Gish W., Miller W., Eugene W., Myers E. W., Lipman D. J. (1990): Basic Local Alignment Search Tool. *Journal of Molecular Biology* **215**, 403-410.
- Angel R. (2006): Feasibility of cooling the Earth with a cloud of small spacecraft near the inner Lagrange point (L1), *Proceedings of the National Academy of Sciences* **103**, 17184-17189.
- Bandillo N. B., Lorenz A. J., Graef G. L., Jarquin D., Hyten D. L., Nelson R. L., Specht J. E. (2017): Genome-wide Association Mapping of Qualitatively Inherited Traits in a Germplasm Collection. *The Plant Genome* **10**, 1-18.
- Barrett J. C., Fry B., Maller J., Daly M. J. (2004): Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265.
- Benzie I. F., Szeto Y. T., Strain J. J., Tomlinson B. (1999): Consumption of green tea causes rapid increase in plasma antioxidant power in humans. *Nutrition and cancer* **34**, 83-87.
- Bernard R. L., Singh B. B. (1969): Inheritance of Pubescence Type in Soybeans: Glabrous, Curly, Dense, Sparse, and Puberulent1. *Crop Science* **9**, 192-197.
- Bernard R.L. (1975a): The inheritance of appressed pubescence. *Soybean Genetics Newsletter* **2**, 34-36.
- Boyer R. F. (2012): *Biochemistry laboratory: modern theory and techniques*. 2nd ed., Pearson Education, Boston, 362 stran.
- Brugliera F., Barri-Rewell G., Holton T. A., Mason, J.G. (2002): Isolation and characterization of a flavonoid 3'-hydroxylase cDNA clone corresponding to the Ht1 locus of *Petunia hybrida*. *The Plant Journal* **19**, 441-451.
- Brugliera, F., Barri-Rewell G., Holton T. A., Mason, J.G. (1999): Isolation and characterization of a flavonoid 3-hydroxylase cDNA clone corresponding to the HT1 locus of *Petunia hybrida*. *Plant Journal* **19**, 441-451.
- Bu T., Lu S., Wang K., Dong L., Li S., Xie Q., Xu X., Cheng Q., Chen L., Fang C., Li H., Liu B., Weller J. L., Kong F. (2021): A critical role of the soybean evening complex in the control of photoperiod sensitivity and adaptation. *National Academy of Sciences* **118** <https://doi.org/10.1073/pnas.2010241118>.
- Clarke E., Wiseman J. (2005): Effects of variability in trypsin inhibitor content of soya bean meals on true and apparent ileal digestibility of amino acids and pancreas size in broiler chicks, *Animal Feed Science and Technology* **121**, 125-138.
- Doughty C. E., Field C. B., McMillan A. M. S. (2011): Can crop albedo be increased through the modification of leaf trichomes, and could this cool regional climate?: A letter. *Climatic Change* **104**, 379-387.
- Dovich N. J. (1997): DNA sequencing by capillary electrophoresis. *Electrophoresis* **18**, 2393-2399.
- Ehleringer J. R., Björkman O. (1978): A Comparison of Photosynthetic Characteristics of Species Possessing Glabrous and Pubescent Leaves, *Plant Physiology*, **62**, 185-190.
- Forkmann G., Heller Werner (1999): Biosynthesis of Flavonoids, *Comprehensive Natural Products Chemistry* **1**, 713-748.
- Forkmann, G. (1991): Flavonoids as flower pigments: the formation of the natural spectrum and its extension by genetic engineering. *Plant Breed* **106**, 1-26.
- Formica J. V., Regelson W. (1995): Review of the biology of quercetin and related bioflavonoids, *Food and Chemical Toxicology* **33**, 1061-1080.
- Frankel S., Robinson G. E., Berenbaum M. R. (1997): Antioxidant capacity and correlated characteristics of 14 unifloral honeys. *Journal of Apicultural Research*, **37**, 27-31.
- Fritsch H., Grisebach H. (1975): Biosynthesis of cyanidin in cell cultures of *Haplopappus gracillus*. *Phytochemistry* **14**, 2437-2442.
- Gillman J. D., Tetlow A., Lee J.-D., Shannon J. G., Bilyeu K. (2011): Loss-of-function mutations affecting a specific *Glycine max* R2R3 MYB transcription factor result in brown hilum and brown seed coats. *BMC Plant Biology* **11**, 155.

- Goluch-Koniuszy Z. S. (2016): Nutrition of women with hair loss problem during the period of menopause. *Przegląd menopauzalny = Menopause review* **15**, 56-61.
- Hall T. (2004): BioEdit version 7.0.0. *Distributed by the author, website: www.mbio.ncsu.edu/BioEdit/bioedit.html*.
- Herman C., Adlercreutz T., Goldin B. R., Gorbach S. L., Höckerstedt K. A. V., Watanabe S., Hämäläinen E. K., Markkanen M. H., Mäkelä T. H., Wähälä K. T., Hase T., A., Fotsis T. (1995): Soybean Phytoestrogen Intake and Cancer Risk. *The Journal of Nutrition* **125**, 757S-770S.
- Choung M., Baek I., Kang S., Han W., Shin D., Moon H., Kang K. (2001): Isolation and determination of anthocyanins in seed coats of black soybean (*Glycine max* (L.) Merr.). *Journal of Agricultural Food Chemistry* **49**, 5848–5851.
- Iwashina T., Benitez E. R., Takahashi R. (2006): Analysis of Flavonoids in Pubescence of Soybean Near-isogenic Lines for Pubescence Color Loci. *Journal of Heredity* **97**, 438-443.
- Iwashina T., Githiri S. M., Benitez E. R., Takemura T., Kitajima J., Takahashi R. (2007): Analysis of Flavonoids in Flower Petals of Soybean Near-isogenic Lines for Flower and Pubescence Color Genes. *Journal of Heredity* **98**, 250-257.
- Joung J., Kasthuri G. M., Park J., Kang W., Kim H., Yoon B., Joung H., Jeon J. (2003): An overexpression of chalcone reductase of *Pueraria montana* var. *lobata* alters biosynthesis of anthocyanin and 5'-deoxyflavonoids in transgenic tobacco, *Biochemical and Biophysical Research Communications* **303**, 326-331.
- Kovinich N., Arnason J. T., De Luca V., Miki B. (2010): Coloring Soybeans with Anthocyanins? *The Biological Activity of Phytochemicals* **41**, 47-57.
- Lee J. M., Bush A. L., Specht J. E., Shoemaker R. C. (1999): Mapping of duplicate genes in soybean. *Genome* **42**, 829-836.
- Liu J., Osbourn A., Ma Pengda (2015): MYB Transcription Factors as Regulators of Phenylpropanoid Metabolism in Plants, *Molecular Plant*, **8**, 689-708.
- Liu J., Osbourn A., Ma, P. (2015): MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Molecular Plant* **8**, 689–708.
- Lorenz T. C. (2012): Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. *Journal of Visualized Experiments* **63**, 3998.
- Marnetto D, Huerta-Sánchez E. (2017): Haplotrips: revealing population structure through haplotype visualization. *British Ecological Society* **8**, 1389-1392.
- Morse W. J., Cartter J. L., Williams L. F. (1949): *Soybeans: culture and varieties*. U.S. Dept. of Agriculture, Washington D.C., 37 stran.
- Muir S. R., Collins G. J., Robinson S., Hughes S., Bovy A., de Vos C. H. R., van Tunen A. J., Verhoeven M. E. (2001): Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols, *Nature Biotechnology* **19**, 470-474.
- Mullis K., Faloona F., Scharf S., Saiki R., Horn G., Erlich H. (1986): Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. *Cold Spring Harbor Symposia on Quantitative Biology* **51**, 263-273.
- Pantalone V., Smallwood Ch. (2018): Registration of ‘TN11-5102’ Soybean Cultivar with High Yield and High Protein Meal. *Journal of Plant Registrations* **12**, 304308.
- Ralston L., Subramanian S. Matsuno M., Yu Oliver (2005): Partial Reconstruction of Flavonoid and Isoflavonoid Biosynthesis in Yeast Using Soybean Type I and Type II Chalcone Isomerases, *American Society of Plant Biologists* **137**, 1375-1388.
- Raso G. M., Mèli R., Di Carlo G., Pacilio M., Di Carlo R. (2001): Inhibition of inducible nitric oxide synthase and cyclooxygenase-2 expression by flavonoids in macrophage J774A.1, *Life Sciences* **68**, 921-931.
- Ridgwell A., Singarayer J. S., Hetherington A. M., Valdes P. J. (2009): Tackling Regional Climate Change By Leaf Albedo Bio-geoengineering. *Current Biology* **19**, 146-150.

- Rise-Evans C. A., Miller N. J., Paganga G. (1996): Structure-antioxidant activity relationships of flavonoids and phenolic acids, *Free Radical Biology and Medicine* **20**, 933-956.
- Sanger F., Nicklen S., Coulson A. R. (1977): DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**, 5463-5467.
- Schijlen E. G. W. M., de Vos C. H. R., van Tunen A. J., Bovy A. G. (2004): Modification of flavonoid biosynthesis in crop plants, *Phytochemistry* **65**, 2631-2648.
- Simon C., Franke A., Martin A. (1991): The Polymerase Chain Reaction: DNA Extraction and Amplification. In: *Molecular Techniques in Taxonomy*. (Hewitt G.M., Johnston A.W.B., Young J.P.W. eds.), Springer, Berlin, 329-355.
- Song Q., Hyten D. L., Jia G., Quigley Ch. V., Fickus E. W., Nelson R. L., Cregan P. B. (2013): Development and Evaluation of SoySNP50K, a High-Density Genotyping Array for Soybean, *PLoS ONE* **8**:e54985 <https://doi.org/10.1371/journal.pone.0054985>.
- SoyKB: Soybean Knowledge Base: <http://soykb.org/AccuTool/index.php> (9. 3. 2021).
- SoyKB: Soybean Knowledge Base: <http://soykb.org/GenescapeAnalysis/search.php> (20. 3. 2021).
- SoyKB: Soybean Knowledge Base: <http://soykb.org/newSNPViz/> (28. 2. 2021).
- Sunada K., Ito T. (1982): Soybean grain quality as affected by low temperature treatments in plants: color of hilum, seed coat cracking, *Hokkaido Branch, the Japanese Society of Breeding and Hokkaido Branch, the Crop Science Society of Japan* **22**, 22-34.
- Škrabišová M., Dietz N, Zeng S, Chan YO, Wang J, Liu Y, Biová J, Joshi T, Bilyeu KD (2021): An Inverse GWAS approach reveals the landscape of association for genomic variants and phenotypes. *Nature Scientific Reports* (under review)
- Takahashi R. (1997): Association of Soybean Genes I and T with Low-Temperature Induced Seed Coat Deterioration. *Crop Breeding, Genetics & Cytology* **37**, 1755-1759.
- Takahashi R., Abe J. (1994): Genetic and Linkage Analysis of Low Temperature-Induced Browning in Soybean Seed Coats. *Journal of Heredity* **85**, 447-450.
- Takahashi R., Ohmori R., Kiyose Ch., Momiyama Y., Ohsuzu F., Kondo K. (2005): Antioxidant Activities of Black and Yellow Soybeans against Low Density Lipoprotein Oxidation. *Journal of Agricultural and Food Chemistry* **53**, 4578-4582.
- Thiele H., Nürnberg P. (2004) HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics* **21**, 1730-1732.
- Thompson M. M., Hrabak E. M. (2018): Capture and storage of plant genomic DNA on a readily available cellulose matrix. *BioTechniques* **65**, 285-287.
- Toda K., Yang D., Yamanaka N., Watanabe S., Harada K., Takahashi R. (2002): A single-base deletion in soybean flavonoid 3'-hydroxylase gene is associated with gray pubescence color. *Plant Molecular Biology* **50**, 187-196.
- Tvrzická N. (2016): *Sója a její využití*. Diplomová práce, Jihočeská univerzita v Českých Budějovicích, Česká republika.
- U.S. National Plant Germplasm System: <https://npgsweb.ars-grin.gov/gringlobal/descriptordetail?id=51057> (28. 2. 2021).
- ÚKZÚZ (2019): Seznam doporučených odrůd 2019. Ústřední kontrolní a zkušební ústav zemědělský a komise pro doporučení odrůd hrachu polního a sóje, Česká republika, 2 strany.
- Visscher P. M., Wray N. R., Zhang Q., Sklar P., McCarthy M. I., Brown M. A., Yang J. (2017): 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics* **101**, 5-22.
- Wen Z., Boyse J. F., Song Q., Cregan P. B., Wang D. (2015): Genomic consequences of selection and genome-wide association mapping in soybean. *BMC Genomics* **16**, 671-684.
- Wollgast J., Anklam E. (2000): Review on polyphenols in Theobroma cacao: changes in composition during the manufacture of chocolate and methodology for identification and quantification, *Food Research International* **33**, 423-447.

- Yao L. H., Jiang Y. M., Tomás-Barberán F. A., Datta N., Singanusong R., Chen S. S. (2004): Flavonoids in Food and Their Health Benefits, *Plant Foods for Human Nutrition* **59**, 113-122.
- Zeng S., Škrabišová M., Lyu Z., Chan Y. O., Bilyeu K., Joshi T. (2020): SNPviz v2.0: A web-based tool for enhanced haplotype analysis using large scale resequencing datasets and discovery of phenotypes causative gene using allelic variations. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (IEEE), 1408-1415.
- Zeng S., Škrabišová M., Lyu Z., Chan Y. O., Dietz N., Bilyeu K., Joshi T. (2021): Application of SNPviz v2.0 utilizing next-generation sequencing datasets in the discovery of potential causative mutations in candidate genes associated with phenotypes. *International Journal of Data Mining and Bioinformatics* (under review).

## 7 SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

°C	stupně Celsia
4CL	4-kumarát-CoA ligasa
A	adenin
AA	aminokyselina
AccuTool	Accuracy Tool
ALT	alternativní alela
Avg	průměrná přesnost
BLAST	Basic Local Alignment Search Tool
bp	páry bází
C	cytosin
C4H	cinamát 4-hydroxylasa
cM	centimorgany
cm	centimetr
CNV	variabilita počtu kopií
CO <sub>2</sub>	oxid uhličitý
Com	pesimistická kombinovaná přesnost
ddNTP	dideoxynukleosidtrifosfáty
DFR	dihydroflavonol-4-reduktasa
diH <sub>2</sub> O	deionizovaná voda
DNA	deoxyribonukleová kyselina
dNTP	2'-deoxynukleosid 5'-trifosfát
EDTA	kyselina ethylendiaminotetraoctová
F3'H	flavonoid-3'-hydroxylasa
<i>FNS</i>	flavonsynthasa
FTA	Flinders Technology Associates
G	guanin
g·kg <sup>-1</sup>	gramů na kilogram
GRIN	Germplasm Repository Information Network

GWAS	genome-wide association study
HIV	virus lidské imunitní nedostatečnosti
CHI	chalkonisomerasa
CHS	chalkonsynthasa
ID	identifikace
IFS	isoflavonsynthasa
kbp	tisíc párů bází
M	molární
m	metry
Mbp	milion párů bází
mM	milimolární
MUT	mutantní typ
MYB	rodina transkripčních faktorů
NaCl	chlorid sodný
NADPH	nikotinamidadeninukleotidfosfát
ng	nanogramy
NIR	blízké infračervené záření
NTP	nukleosidtrifosfát
OH <sup>-</sup>	hydroxylový anion
PAL	fenylalanin amoniak lyasa
PAR	fotosynteticky aktivní záření
PI	plant introduction (uvedení rostliny)
REF	referenční
RNA	ribonukleová kyselina
s	sekunda
SNP	jednonukleotidový polymorfismus
SNPViz	vizualizace jednonukleotidových polymorfismů
T	thymin
TAE	tris-acetát-EDTA pufr

TE	tris-EDTA pufr
TENT	tris-EDTA-NaCl-Triton X-100 pufr
T <sub>m</sub>	teplota tání
UV	ultrafialové
v2.0	verze 2.0
v2.0	verze 2.0
WT	divoký typ
μl	mikrolitr