Aplikace genomické selekce ve šlechtění obilnin

Metodika pro praxi

RNDr. Mgr. Leona Svobodová, Ph.D.
Mgr. Alena Hanzalová, Ph.D.
Ing. Michaela Jungová, Ph.D.
Ing. Tereza Sovová, Ph.D.
Ing. Pavel Svoboda, Ph.D.

Praha 2026

Národní centrum zemědělského a potravinářského výzkumu

Dedikace: Metodika vznikla za finanční podpory MZe ČR a je výstupem řešení projektu NAZV QK22010298: „Aplikace nového přístupu Genomic selection ve šlechtění pšenice“.

Poslání a uživatelé metodiky: Metodika popisuje postup nové šlechtitelské techniky genomické selekce u zemědělsky využívaných samosprašných obilnin. Metodika je určena zejména pro šlechtitelskou praxi.

Metodika byla schválena Ministerstvem zemědělství České republiky dne 10. 7. 2026 pod č. MZE-52451/2026-13123 a doporučena pro využití v praxi.
O uplatnění metodiky byla uzavřena smlouva o využití s firmou SELGEN, a.s. dne 25. 5. 2026 pod č. 221136-2026.

Prohlášení o podílu práce autorů metodiky: Na vypracování metodiky se podílelo celkem 5 spoluautorů, jejich podíly činí: RNDr. Mgr. Leona Svobodová, Ph.D. (20 %), Mgr. Alena Hanzalová, Ph.D. (20 %), Ing. Michaela Jungová, Ph.D. (20 %), Ing. Tereza Sovová, Ph.D. (20 %), Ing. Pavel Svoboda, Ph.D. (20 %)

Národní centrum zemědělského a potravinářského výzkumu

Aplikace genomické selekce ve šlechtění obilnin

Metodika NmetS

Autorský tým: RNDr. Mgr. Leona Svobodová, Ph.D.[a)], Mgr. Alena Hanzalová, Ph.D.[a)], Ing. Michaela Jungová, Ph.D.[a)], Ing. Tereza Sovová, Ph.D.[a)], Ing. Pavel Svoboda, Ph.D.[a)]
[a)] Národní centrum zemědělského a potravinářského výzkumu, v.v.i., Drnovská 507/73, 161 06 Praha 6 – Ruzyně
Korespondenční autor: Leona Svobodová, leona.svobodova@carc.cz

Foto: Archiv autorů (pokud není uvedeno jinak)

Oponenti:
Ing. Lenka Clowezová
prof. Ing. Vladislav Čurn, Ph.D.

Aplikace genomické selekce ve šlechtění obilnin

Anotace Předkládaná metodika představuje postup, jak aplikovat genomickou selekci ve šlechtění obilnin. Tento postup je běžně využíván ve šlechtění hospodářských zvířat a v posledních přibližně deseti letech i u mnoha druhů plodin. U zemědělských plodin v České republice však dosud využit nebyl. Vývoj postupů celogenomového genotypování a jejich finanční zpřístupnění znamená revoluci ve studiu genetické diverzity a jejího využití ve šlechtění. Dále rozvoj hardware a vývoj programů zpracovávajících velké objemy genomických dat umožňuje provázat výsledky genotypové analýzy s výsledky fenotypového hodnocení a navrhovat matematické modely s praktickými výstupy, které umožní výběr genotypů s požadovanými hodnotami znaků ať už jako rodiče do křížení a/nebo pro selekci šlechtitelských linií v raných fázích šlechtitelského procesu.
Uplatněná metodika vznikla za finanční podpory NAZV a je výstupem řešení projektu QK22010298.

Klíčová slova: nové šlechtitelské techniky, genomická selekce, celogenomové genotypování, genetická variabilita, smíšené lineární modely

Application of Genomic selection in cereal breeding

Annotation Presented methodology introduces a procedure, how to apply genomic selection in cereal breeding. This procedure is commonly used in livestock breeding, and in last decade in crop breeding. However, in the Czech Republic, it has not been used so far in crop breeding. The development of whole-genome genotyping and the decrease of its prize lead to a revolution in the study of genetic diversity and its use. Further, the development of hardware and appropriate software handling with a huge genomic data allows tying results of genotype analysis and phenotype evaluation. Suitable mathematical models are proposed with practical outputs leading to the selection of genotypes with required traits as parental lines or breeding lines in the early stages of breeding process.

The applied methodology was created with the financial support of NAZV, project number QK22010298.

Keywords: new breeding techniques, genomic selection, whole-genome genotyping, genetic variability, mixed linear models

ISBN 978-80-7427-466-4

©Národní centrum zemědělského a potravinářského výzkumu, v.v.i., Praha, 2026

Obsah

I. Cíl metodiky 5
II. Vlastní popis metodiky 6
1. Úvod 6
2. Genomická selekce 7
3. Důležité vstupní informace 10
3.1. Pšenice 10
3.2. Rezistence pšenice ke rzi travní a rzi pšeničné 10
4. Trénovací populace 11
4.1. Fenotypová analýza 11
4.2. Genotypová analýza 15
4.3. Navržení a testování modelů genomické selekce 17
4.3.1. Unitrait uspořádání 17
4.3.2. Multitrait uspořádání 20
5. Šlechtitelská populace 22
5.1. Genotypová analýza 22
5.2. Genomická selekce pro jeden znak 23
5.3. Genomická selekce pro více znaků 25
6. Závěr 30
III. Srovnání novosti postupů 31
IV. Popis uplatnění metodiky 31
V. Ekonomické aspekty 32
VI. Seznam použité literatury 33
VII. Seznam publikací, které předcházely metodice 36

I. Cíl metodiky

Uvedená metodika obsahuje optimalizovaný postup genomické selekce u obilnin.

Cílem je:

- Výběr rodičovských komponent na základě genotypových dat a hodnot fenotypových znaků, na které je šlechtitelský program zaměřen
- Efektivní výběr šlechtitelských linií v raných fázích vývoje při udržení optimální míry genetické diverzity daného souboru genotypů

II. Vlastní popis metodiky

1. Úvod

Odrůdová rozmanitost hospodářských rostlin je jedním z nejvýznamnějších faktorů rozvoje zemědělství. Výběr preferovaných vlastností deklarovaný příslušnou odrůdou umožňuje zemědělcům a zpracovatelům dosahovat vyšších hospodářských výsledků, kvalitnějšího zpracování zrna a přípravy hodnotnějších potravin.

Podmínkou pro zachování odrůdové rozmanitosti a výkonnosti nových odrůd je dobře nastavený šlechtitelský proces. Šlechtitelské programy se liší zejména podle toho, na jaké typy znaků se šlechtí. Většina hospodářsky významných znaků je podmíněna mnoha geny, jejichž účinky se sčítají nebo se epistaticky ovlivňují (Boone et al., 2007). Taková místa v genomu rostlin, která obsahují geny, které ovlivňují určitý fenotypový projev, nazýváme anglicky Quantitative Trait Loci (QTLs). S jejich objevováním se začalo v osmdesátých letech, kdy to umožnil rozvoj metod molekulární biologie a identifikace tzv. DNA markerů. Jedním z hlavních využití molekulárních markerů v zemědělském výzkumu byla konstrukce tzv. vazebných map jednotlivých druhů plodin. Tyto mapy byly využity zejména k identifikaci oblastí chromozomů, které obsahují geny, které kontrolují znaky podmíněné jedním či několika málo geny a poté i QTLs (Mohan et al., 1997).

DNA markery, které se ukázaly být v těsné vazbě k hledanému znaku, mohou být využity jako nástroj k selekci šlechtitelských linií. Celý postup se nazývá anglicky Marker Assisted Selection (MAS) – tedy selekce za pomoci DNA markerů (Ribaut and Hoisington, 1998). Tato selekce je efektivní, spolehlivá a levnější ve srovnání s klasickým výběrem (Collard et al., 2005). Používání MAS ve šlechtění tak odstartovalo éru tzv. molekulárního šlechtění (Rafalski and Tingey, 1993).

MAS je velmi výhodné používat v případě znaků podmíněných několika málo geny s velkým účinkem. Avšak u znaků podmíněných kvantitativně, tj. mnoha geny s malými účinky je tato metoda neefektivní a nepraktická (Buhlakoti et al., 2022). Proto byl vyvinut nový postup, který se nazývá genomická selekce (Genomic selection, GS).

2. Genomická selekce

Tento postup umožňuje selekci znaků pomocí genetické hodnoty jedince získané s využitím efektů markerů o dostatečné hustotě rozmístění po celém genomu rostliny (Meuwissen et al., 2001). Nejprve jsou tedy spočteny efekty všech markerů a jejich suma je použita pro výpočet šlechtitelských hodnot (Genomic-Estimated Breeding Values - GEBVs) pro každý genotyp.

Proces genomické selekce pro vybraný znak začíná vždy výběrem tzv. trénovací populace. Je to soubor genotypů, u kterých je k dispozici jak hodnocení fenotypového projevu zkoumaného znaku, tak i genotypová analýza. Tato data slouží k vytvoření modelu zobecněné lineární regrese, kde fenotypové hodnocení je použito jako odpověď a genotypová data jako prediktor. Druhým krokem je výběr šlechtitelské populace, kde jsou k dispozici pouze genotypová data. Na jejich základě s využitím modelu z předchozího kroku je možno předpovědět hodnoty GEBV genotypů šlechtitelské populace.

Hlavní výhodou genomické selekce je, že umožňuje výrazně zkrátit šlechtitelský proces oproti klasickému výběru linií každé generace a zároveň snižuje náklady na masivní fenotypování (Heffner et al., 2010). Na druhou stranu existuje několik faktorů, které ovlivňují přesnost genomické predikce, která musí být validována na základě fenotypového hodnocení: velikost trénovací a šlechtitelské populace, genetická diverzita šlechtitelské populace, hodnota dědivosti znaku, vliv prostředí na hodnotu znaku, hustota markerů a genetický vztah mezi trénovací a šlechtitelskou populací (Daetwyler et al., 2010; Clark et al., 2011; Howard et al., 2014; Verges et al., 2020). Obecně platí, že čím vyšší je hodnota dědivosti znaku, tím vyšších hodnot přesnosti genomické selekce je možno dosáhnout. Čím je hodnota dědivosti znaku menší, tím více genotypů by měla zahrnovat trénovací populace (Budhlakoti et al., 2022). Šlechtitelská populace by měla vycházet z té trénovací, měla by mít přibližně stejnou úroveň variability. V žádné z obou populací by neměla být výrazná populační struktura, která by zkreslovala hodnoty GEBV. A pokud je hodnota znaku výrazně ovlivňována prostředím, je nutno provést pečlivé hodnocení znaku u trénovací populace na několika lokalitách v průběhu alespoň 3 let. Otázka výběru markerovacího systému a počtu markerů je rovněž důležitá.

Většinou se využívají jednobodové mutace (Single Nucleotide Polymorphisms, SNPs) v uspořádání na čipu nebo v DArT analýze (Diversity Arrays Technology). Minimální počet markerů by měl být alespoň 2000 SNPs (Abed et al., 2018), u druhů jako je pšenice musí být tento počet pro pokrytí diverzity šlechtitelské populace vyšší. Markery by měly pokrývat, pokud možno celý genom a ukazuje se, že je lepší, pokud jsou lokalizovány do mezigenových oblastí, neboť jsou variabilnější a zároveň jsou zde lokalizovány regulační elementy, jako například promotory, zesilovače apod. (Barrett et al., 2012).
Výběr vhodných individuálních genotypů v genomické selekci probíhá pomocí modelu, který vychází ze zobecněné lineární regrese:

$$Y = \ln\mu + X\beta + \varepsilon,$$

kde $\boldsymbol{Y}$ = n × 1 je vektor či matice fenotypových hodnot pro n genotypů a l znaků, $\boldsymbol{\mu}$ je průměr, $\boldsymbol{\beta}$ = p × l je vektor či matice efektů *p* markerů *l* znaků, $\boldsymbol{\varepsilon}$ = n × l je vektor či matice náhodných reziduálních efektů ($\varepsilon \sim N(0,\sigma_e^2)$) a **X** je matice o n × p prvcích, kde řádky reprezentují genotypy n a sloupce markery p. Problém nastává u lineárních modelů všude tam, kde počet proměnných p, tedy v tomto případě počet markerů výrazně převyšuje počet vzorků n, v tomto případě počet genotypů trénovací a šlechtitelské populace (p >> n). Meuwissen et al. (2001) provedl redukci počtu markerů tak, že použil modifikaci lineární regrese s uplatněním metody nejmenších čtverců na každý marker zvlášť. Markery se statisticky významnými efekty byly vybrány pomocí hodnoty logaritmu pravděpodobnosti a dále použity ve výpočtu šlechtitelských hodnot. Nyní existuje více metod, které řeší tento problém „přeparametrizace“, například tzv. Ridge Regression (RR) modely, které řeší zároveň i problém kolinearity markerů, tedy jejich vzájemnou korelaci (Meuwissen et al., 2001). Pomocí nastavitelných penalizačních parametrů snižují efekt korelovaných a statisticky nevýznamných markerů. Dalším modelem je LASSO (Least Absolute Shrinkage and Selection Operator), kde je použit jiný typ penalizačního kritéria výběru markerů (Tibshirani, 1996; Usai et al., 2009). RR model předpokládá, že každý marker přispívá stejnou mírou variance, ale to u všech vlastností neplatí. Pro tyto případy byly vyvinuty modely využívající

podmíněnou pravděpodobnost – tzv. Bayesovské modely (Meuwissen et al., 2001; de los Campos et al., 2009; Habier et al., 2011) a jejich kombinace s výše uvedenými metodami Bayesian LASSO a Bayesian ridge regression (BRR).

Uvedené modely fungují dobře u znaků podmíněných geny s aditivním účinkem, ale selhávají u znaků podmíněných geny s epistatickým působením. V těchto případech je třeba použít neparametrické nebo semiparametrické modely (Gianola et al., 2006; de los Campos et al., 2010) jako je například Reproductive kernel Hilbert Space (RKHS, Gianola et al., 2006), Support vector machine (SVM, Maenhout et al., 2007), Random Forest (RF, Holliday et al., 2012) a mnoho dalších.

Všechny modely, o nichž byla dosud řeč, se týkaly vždy jednoho znaku. Existují však přístupy, které sestavují a optimalizují modely pro více znaků najednou. Pokud jsou sledované znaky navzájem korelované, byť vykazují malé hodnoty dědivosti, přesnost odhadu GEBV se zvyšuje, neboť geny mohou ovlivňovat více znaků zároveň. Bylo takto vytvořeno několik modelů a testována jejich účinnost (Budhlakoti et al., 2019).

Výběr vhodného modelu závisí na tom, zda je genomická selekce prováděna na jeden či více znaků a jaké je chování vybraného znaku, či znaků. Je třeba testovat vždy několik modelů a na základě parametrů: velikosti reziduální variance, hodnoty DIC (Deviance Information Criterion), efektivního počtu parametrů pD a přesnosti odhadu GEBV vybrat ten nejvhodnější.

Postup výběru vhodného predikčního modelu bude dále popsán na základě konkrétního případu aplikace genomické selekce ve šlechtění pšenice na rezistenci ke rzi travní a rzi pšeničné. Aplikace metodiky na jiné druhy obilnin a jiné znaky bude vyžadovat optimalizaci dle výše uvedených návodů a kritérií.

3. Důležité vstupní informace

2.1. Pšenice

Pšenice (*Triticum aestivum* L.) je allohexaploid se třemi genomy (AABBDD) a základním chromozomovým číslem x = 7 (2n = 6x = 42). Celková velikost genomu je 16-17 Gb (Bennett and Smith, 1976). Genom A pochází z *Triticum urartu*, genom B z předchůdce *Aegilops speltoides* a genom D z *Aegilops tauschii*. Genomy tudíž nejsou vzhledem k obsahu genů totožné, vykazují

ale vysokou míru syntenie a jejich odpovídající chromozomy jsou homoeologní. Tento fakt je třeba vzít v úvahu při výběru metody genotypování, aby byl pokryt celý genom tedy všech 7 chromozomů všech tří genomů.

Pro studium diverzity bylo postupně vyvinuto několik metod. Pro účely genomické selekce se využívají výsledky celogenomového sekvenování a identifikace polymorfismů, zejména SNP (Single Nucleotide Polymorphism). Vlastní analýzy se pak provádějí pro desítky tisíc těchto SNP najednou pomocí čipů či technologie DArT.

2.2. Rezistence pšenice ke rzi travní a rzi pšeničné

Rez travní (*Puccinia graminis* f. sp. *tritici*) a rez pšeničná (*Puccinia triticina;* syn. *Puccinia recondita f. sp. tritici*) jsou biotrofní druhy hub parazitující na pšenici. V případě pro jejich růst vhodných podmínek, mohou způsobovat 50–80 procentní ztráty výnosu. Šlechtění pšenice na rezistenci k nim je tedy velmi důležité.

Reakce pšenice ke rzím je buď rasově specifická, nebo rasově nespecifická. Rasově specifická odpověď hostitele je indukována jako specifická odpověď hostitele na konkrétní rasu patogena ve stádiu klíčních rostlin i v dospělosti. Tento typ rezistence je podmíněn geny tzv. velkého účinku. Tyto geny rezistence bývají často překonány nově vznikajícími rasami rzi. Rasy jsou vystaveny silnému selekčnímu tlaku uvnitř populace patogena i selekčnímu tlaku prostředí, tedy i obranou reakcí hostitele, a to eliminuje počet ras patogena na několik převládajících typů. Vznikají a rozšiřují se tak rasy s novými virulencemi, které překonávají geny rezistence pěstovaných

odrůd, což vede ke ztrátě odolnosti založené na genech s velkým účinkem (Babu et al., 2020). U tohoto typu rezistence je rovněž časté epistatické působení, kdy jeden gen může maskovat nebo modifikovat účinek jiných genů. Epistáze se projevuje například tehdy, když přítomnost jednoho genu posílí nebo naopak potlačí účinnost jiného obranného mechanismu. V současné době je známo 60 genů rezistence ke rzi travní (*Sr*) a 80 genů rezistence ke rzi pšeničné (*Lr*).

Rasově nespecifická rezistence ke rzím se projevuje u pšenice ve stádiu dospělosti. Je považována za trvalou, neboť nevytváří tak silný selekční tlak na populaci patogena. Je podmíněna mnoha geny malého účinku, dosud popisovanými jako tzv. QTL (Quantitative Trait Locus). Typické pro tyto geny/ QTL je jejich aditivní působení, což vede k plynulé škále odolnosti a často k odolnosti nejen ke rzím, ale i k dalším druhům patogenů. Na druhou stranu je tento typ rezistence výrazně podmíněn prostředím, a proto je nutné hodnocení provádět po dobu několika let na více lokalitách.

3. Trénovací populace

Na základě předchozích znalostí o rezistenci genotypů ke rzi travní a rzi pšeničné bylo vybráno celkem 186 genotypů pšenice zahrnujících odrůdy a šlechtitelské linie, z toho 62 jarního typu a 124 ozimého typu. Osivo bylo získáno z ÚKZÚZ a ze šlechtitelské stanice Selgen – Stupice. U tohoto souboru byla provedena genotypová a fenotypová analýza.

3.1. Fenotypová analýza

Fenotypová analýza znaků rezistence ke rzi travní a rezistence ke rzi pšeničné byla prováděna u dospělých rostlin v polních podmínkách. Zároveň s klasickým hodnocením byla u všech genotypů trénovací populace provedena rovněž kvantifikace obou patogenů metodou real-time PCR.

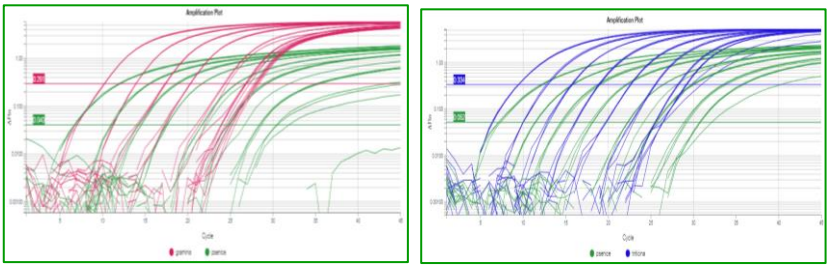
Polní pokusy byly prováděny po dobu tří let na dvou lokalitách. Výsevy polních pokusů proběhly dle standardní metodiky, tj. velikost parcel byla 0,25 m^2, umístění jednotlivých genotypů bylo znáhodněno. Soubory odrůd

byly testovány na odolnost ke rzi travní a rzi pšeničné při umělé infekci oběma patogeny. Pro každou ze zkoušených chorob byl pokus vyset ve dvou opakováních. Lokality pro zkoušení každé z chorob byly od sebe vzdáleny a izolovány obsevem tak, aby nedošlo ke kontaminaci.
Inokulace byla prováděna formou nástřiku i formou infekční aplikace do stébel. Jako inokulum byla použita směs urediospor odlišných fyziologických ras s definovanou virulencí/avirulencí. Virulence inokula byla determinována na téměř izogenních liniích (NILs (*Lr*) a NILS (*Sr*). Zdrojem infekčního materiálu byly listy pšenice napadené rzemi ze sběrů z celé ČR, a to z materiálu získaného z předchozí vegetační sezóny. Determinované inokulum jednotlivých patogenů bylo rozmnoženo

ve skleníkových podmínkách na klíčních rostlinách náchylné odrůdy (Michigan Amber). Termín umělé inokulace se pro jednotlivé druhy rzí lišil vzhledem k jejich různým teplotním nárokům při vzniku infekce, inkubaci a dalším šíření choroby.
Inokulace rostlin v pokusu se rzí pšeničnou byla provedena na rostliny ve fázi BBCH 29–31 postřikem urediospor ve vodní suspenzi. Pro zajištění vysoké vzdušné vlhkosti byly inokulované rostliny pět dní zakryty fólií. Inokulovány byly odrůdy použité v prosevném řádku jako spreader (Michigan Amber, Baletka, Victo). Ze spreaderu se pak urediospory rzí rovnoměrně rozšířily na přilehlé parcely se zkoušenými odrůdami. Opakovaná inokulace byla provedena počátkem sloupkování injekcí vodní suspenze urediospor do stébel „spreadera“, a to v části pokusu, kde byla zkoušena odolnost ke rzi pšeničné i v části pokusu se rzí travní. Rzí pšeničnou byly rostliny inokulovány injekčně na konci dubna a rzí travní v prvním květnovém týdnu.
Hodnocení obou částí pokusu proběhlo ve třech termínech podle rozvoje jednotlivých chorob s použitím stupnice ÚKZÚZ (1–9, 1 – náchylná, 9 – odolná). Průměry opakovaných hodnocení pro každý genotyp, rok a lokalitu byly zaznamenány pro názvem proměnné ***hodnocení Pg***, ***hodnocení Pt***.
Pro kvantitativní stanovení obsahu patogena ve vzorku byly nejprve na základě v databázích dostupných sekvencí genu pro β-tubulin u rzi travní, rzi pšeničné a u pšenice navrženy primery a TaqMan sondy. Byla provedena optimalizace reakcí v duplexovém uspořádání patogen-hostitel

s tím, že TaqMan sonda pro patogena byla fluorescenčně značena 6-FAM a TaqMan sonda pro pšenici byla značena VIC. Jako standardy pro kvantifikaci byly připraveny ředicí řady směsi vzorku DNA hub (*P. triticina* nebo *P. graminis*) a DNA pšenice.

Obrázek 1 – Průběhy real-time TaqMan PCR pro ředicí řady DNA *P. triticina*/ *P. graminis* a pšenice v duplexovém uspořádání

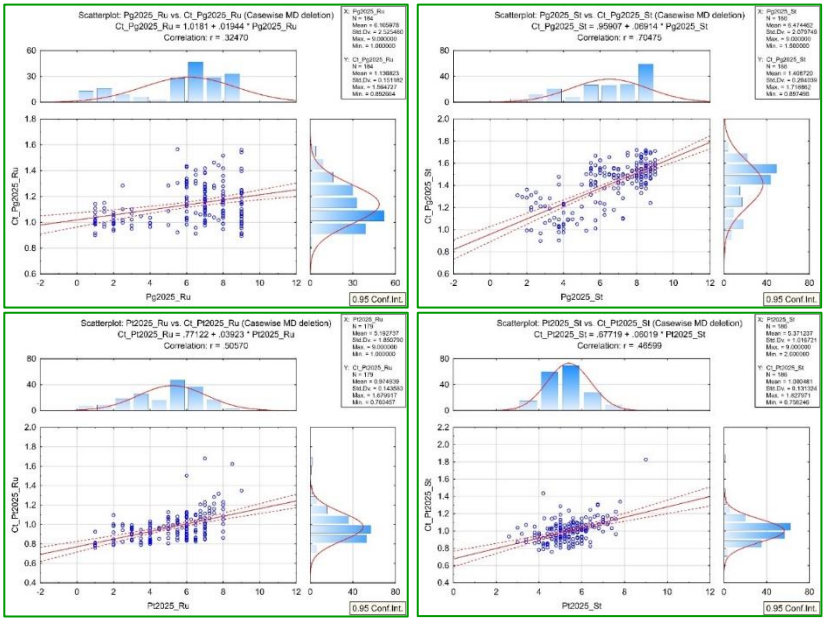


Každý rok na obou lokalitách byly zároveň s posledním hodnocením odebrány vzorky. Z každé parcely bylo odebráno náhodně min. 5 celých odnoží z krajů a prostředku parcely. Vzorky byly uchovávány v papírových pytlích. DNA byla extrahována pomocí CTAB podle protokolu optimalizovaného pro extrakci DNA ze suchých rostlin a herbářových sbírek. Kvalita a koncentrace extrahované DNA byla ověřena elektroforeticky v 0,7 % agarózovém gelu a spektrofotometricky na Nano-photometru Implen.

Real-time PCR byly prováděny v přístrojích ABI790 nebo Quant Studio 6 (Applied Biosystems) v optických PCR destičkách krytých optickými víčky (MicroAmp Optical 96-well reaction plate, MicroAmp Optical 8-Cap Strip) v reakčním objemu 20 µl obsahujícím 0,25 µM páru primerů pro rez travní nebo rez pšeničnou, 0,25 µM páru primerů pro pšenici, 0,25 µM TaqMan sondy pro rez travní nebo rez pšeničnou, 0,25 µM TaqMan sondy pro pšenici, 10 µl 2xTaqMan Universal Master Mix II (Applied Biosystems), 5 µl H_2O a 2 µl vzorku DNA. Všechny vzorky byly hodnoceny ve třech opakováních. Z hodnot Ct houbového patogena a Ct pšenice v daném vzorku byl vypočten převrácený poměr patogena k hostiteli, tedy $Ct_{pšenice}$ / $Ct_{patogen}$, aby výsledná hodnota pozitivně korelovala se škálou hodnocení. Tyto výsledné hodnoty pro každý genotyp, rok a lokalitu byly zaznamenány do tabulky jako proměnné ***stanovení Pg***, ***stanovení Pt***.

Každoročně byly vypočteny korelace mezi hodnocením a stanovením jako kontrola koherence výsledků (Obrázek 2).

Obrázek 2 – Příklad výsledku korelační analýzy mezi vizuálním hodnocením a stanovením *P. graminis* a *P. triticina* v pletivech hostitele



Po skončení celého tříletého pokusu byly z hodnot všech čtyř proměnných spočteny tzv. BLUP (Best Linear Unbiased Prediction) hodnoty pomocí lineárního smíšeného modelu za použití balíčku ASReml v. 4.1.0.149 v programu R. Zároveň byly pomocí uvedeného programového balíčku spočteny hodnoty dědivosti dle vzorce:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GE}^2}{e} + \frac{\sigma_e^2}{e}},$$

kde σ_G^2 je genetická variance (efekt kultivaru), σ_{GE}^2 je celková variance interakce genotyp × prostředí (suma komponentů kultivar × lokalita

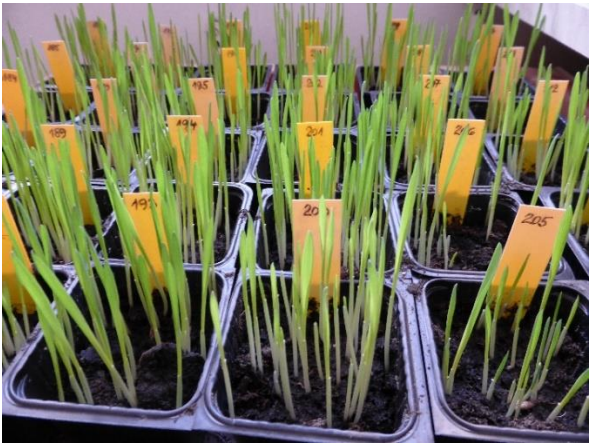
a kultivar × rok), σ_e^2 je reziduální variance a *e* je počet „prostředí“ (kombinace rok × lokalita).

V našem případě byly hodnoty dědivosti H^2 = 0,95 pro *hodnocení Pg*, H^2 = 0,80 pro *hodnocení Pt*, H^2 = 0,79 pro *stanovení Pg* a H^2 = 0,58 pro *stanovení Pt*. Tyto hodnoty jsou potřebné pro výpočet přesnosti modelu genomické selekce.

3.2. Genotypová analýza

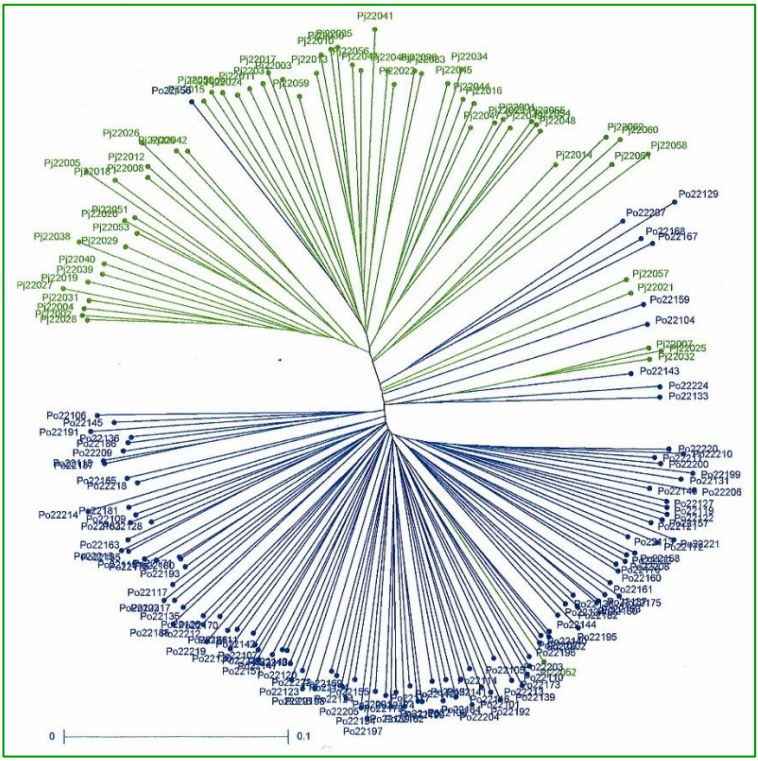
U všech 186 genotypů trénovací populace pšenice byl proveden výběr tzv. SSD (Single Seed Descent) linií a jejich osivo namnoženo. Byly odebrány vzorky mladých listových pletiv (Obrázek 3) a uchovávány při teplotě -80°C. DNA byla extrahována pomocí CTAB podle optimalizovaného protokolu. Kvalita a koncentrace byla ověřena elektroforeticky a spektrofotometricky pomocí fluorometru Qubit4 (Invitrogen). Všech 186 vzorků bylo odesláno na analýzu DArT Seq do Austrálie. Jako referenční model byla zvolena odrůda pšenice seté Chinese Spring 21 a konsensuální mapa: Wheat_Consensus Map_version_4.

Obrázek 3 – Ilustrační obrázek – příklad rostlin pro odběr vzorků na extrakci DNA pro genotypování



Výsledky analýzy byly staženy, převedeny z formátu *.csv do formátu *.xlsx a uloženy. Pro další zpracování byl vybrán soubor Report_*_SNP_mapping_2.csv, kde hodnota „0“ značí alelu reference, „1“ alternativní alelu a „2“ heterozygotní sestavu alel. Nejprve bylo provedeno filtrování dat - tj. v souboru byla ponechána pouze data s hodnotou PIC ≥ 0,05 a chybějícími daty ≤ 10 %. Z původních 103307 markerů jich v souboru zůstalo 13275. Pak byla provedena transpozice dat, tj. výměna řádků za sloupce a opačně. Zároveň byla identifikována neúplnost dat u jednoho genotypu. Finální datová matice tedy obsahovala 185 řádků (genotypů) a 13275 sloupců (markerů).

Obrázek 4 – Dendrogram sestrojený metodou nejbližších sousedů na základě matice genetických vzdáleností vypočtené z dat DArT analýzy u 185 genotypů pšenice; zeleně jsou označeny jařiny a modře ozimy



Pomocí programu TASSEL v. 5.2.93 (Bradbury et al., 2007) byla vypočtena matice genetických vzdáleností, převedena do programu Darwin (Perrier and Jacquemoud-Collet, 2006), kde byl sestrojen dendrogram metodou nejbližších sousedů (Obrázek 4). Z dendrogramu jsou patrné dvě hlavní skupiny, které odpovídají jarnímu a ozimému typu pšenice, přičemž u malé části genotypů na rozhraní těchto dvou shluků dochází k mírnému překryvu. Další výrazná struktura v dendrogramu není vidět, což je pro dosažení dobré přesnosti genomické selekce nutné. Ověření genetické struktury bylo také ověřeno pomocí analýzy MDS (Multi Dimensional Scaling) a pomocí Bayesovského přístupu v programu Structure.

3.3. Navržení a testování modelů genomické selekce

Pro navržení modelu byla využita matice genotypových dat o 185 řádcích (genotypy pšenice) a 13275 sloupcích (SNP markery z DArT analýzy) a matice fenotypových dat o 185 řádcích (genotypy pšenice) a 4 sloupcích (hodnoty BLUP pro proměnné *hodnocení Pg*, *hodnocení Pt*, *stanovení Pg* a *stanovení Pt*). Výběr modelu se lišil pro unitrait (pro 1 znak) a multitrait (pro více znaků) uspořádání.

3.3.1. Unitrait uspořádání

Jelikož jsme se zaměřili na rasově nespecifickou rezistenci ke rzi travní a rzi pšeničné a hodnocení jsme prováděli v polních podmínkách, mohli jsme předpokládat aditivní účinky genů rezistence. Rovněž víme, že každý marker nepřispívá k celkové varianci stejně. Proto jsme zvolili parametrické modely genomické selekce Bayesovského typu: Bayes A, Bayes B, Bayes C, Flat, Bayes LASSO a BRR. S těmito modely jsme pracovali v programovém balíčku BGLR v R-studiu (Pérez and de los Campos, 2014). Program je volně dostupný na webu: https://CRAN.R-project.org/package=BGLR. Pro úvodní testování byla použita data pro všech 185 genotypů pšenice, počet iterací 4500 po zahřívacích 500 iteracích. Výsledky testování jsou uvedeny v Tabulce 1.

Tabulka 1 – Výběrové parametry testovaných modelů

Znak	Model*	varE	pD	DIC	r(y,yHat)
Hodnocení Pg	FX	4,050 (4,378)	129,1	857,2	0,999
	BRR	0,509 (0,110)	122,8	500,6	0,991
	BL	0,496 (0,099)	123,6	497,7	0,990
	BA	0,455 (0,097)	122,8	478,1	0,989
	BB	0,466 (0,096)	112,0	474,6	0,984
	BC	0,473 (0,105)	116,7	478,2	0,989
Stanovení Pg	FX	0,010 (0,007)	152,7	-199,7	0,994
	BRR	0,003 (0,001)	110,6	-466,3	0,980
	BL	0,004 (0,001)	89,5	-436,9	0,963
	BA	0,003 (0,001)	109,3	-476,7	0,993
	BB	0,003 (0,001)	99,5	-466,1	0,963
	BC	0,003 (0,001)	114,1	-474,4	0,980
Hodnocení Pt	FX	0,521 (0,582)	133,8	487,2	0,999
	BRR	0,198 (0,066)	95,8	305,2	0,979
	BL	0,208 (0,051)	98,9	321,2	0,977
	BA	0,138 (0,039)	126,5	269,0	0,992
	BB	0,172 (0,051)	108,4	295,4	0,984
	BC	0,159 (0,048)	115,9	284,5	0,988
Stanovení Pt	FX	0,004 (0,004)	137,5	-401,7	0,999
	BRR	0,002 (0,001)	51,5	-551,4	0,878
	BL	0,002 (0,0002)	68,8	-555,2	0,911
	BA	0,001 (0,0003)	112,0	-608,2	0,979
	BB	0,002 (0,0005)	83,4	-583,4	0,953
	BC	0,002 (0,0004)	78,3	-568,4	0,939

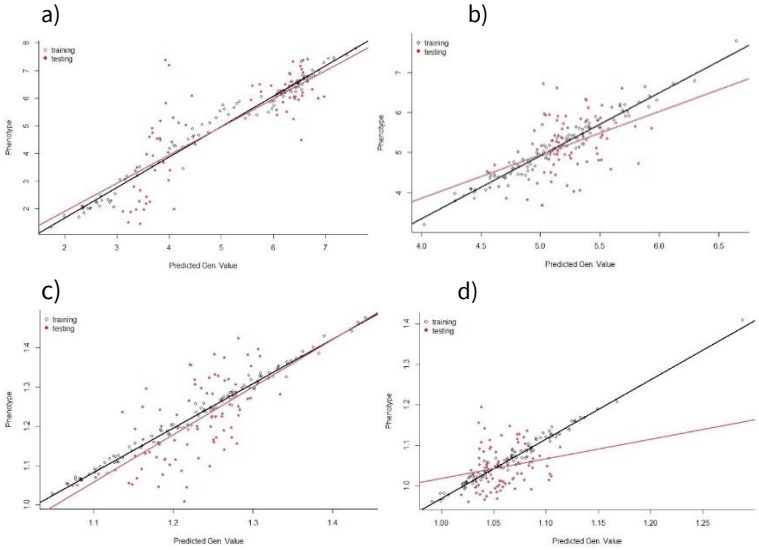
*FX je model Flat; BRR je Bayesian Ridge Regression model; BL je Bayesian LASSO model; BA je Bayes A model; BB je Bayes B model; BC je Bayes C model

Na základě nejnižších hodnot parametru DIC a zároveň nejnižších hodnot reziduální variance ***var E*** a zároveň nejvyšších hodnot korelace mezi původní hodnotou znaku y a vypočtenou hodnotou znaku yHat a zároveň relativně vysoké hodnoty komplexity modelu pD byl pro další analýzy zvolen model Bayes A.

V dalším kroku byla stanovena přesnost modelu. Program byl spuštěn tak, aby provedl 100 opakování výběru 80 genotypů, u kterých nastaví hodnotu znaku jako neznámou (NA) a spočte příslušné hodnoty znaku (GEBV, v programu má tato proměnná název *yHat*). Počet iterací byl nastaven na 5000 po 1000 zahřívacích iteracích. Korelace mezi naměřenou hodnotou znaku (BLUP, v programu proměnná *y*) a vypočtenou hodnotou GEBV (*yHat*) je uvedena v Tabulce 2 a na Obrázku 5.

Přesnost selekce, která vyjadřuje vztah mezi skutečnou šlechtitelskou hodnotou genotypu a znakem, podle kterého genotyp vybíráme se pak vypočte jako podíl uvedených hodnot korelace a hodnoty dědivosti znaku (H^2).

Obrázek 5 – Grafy korelace mezi vypočtenou a naměřenou hodnotou znaku a) *hodnocení Pg*; b) *hodnocení Pt*; c) *stanovení Pg*; d) *stanovení Pt* pro model genomické selekce Bayes A.



Tabulka 2 – Výpočet hodnot přesnosti pro model Bayes A pro 4 znaky u pšenice

Znak	r(y,yHat)	H^2	Přesnost
Hodnocení Pg	0.7799613	0.94973	0.82
Stanovení Pg	0.6650436	0.80608	0.83
Hodnocení Pt	0.3731096	0.74545	0.50
Stanovení Pt	0.2891272	0.58801	0.49

Z tabulky vyplývá, že u rzi travní je přesnost genomické selekce vysoká, u rzi pšeničné je nižší, ale ve srovnání s jinými znaky u pšenice ji lze považovat za střední. Model zahrnuje regresní koeficienty – efekty markerů, penalizační kritéria markerů a prvky kovarianční matice náhodných efektů. Na základě tohoto modelu jsou vypočteny tzv. GEBV hodnoty (Genomic Estimated Breeding Values) u genotypů šlechtitelské populace.

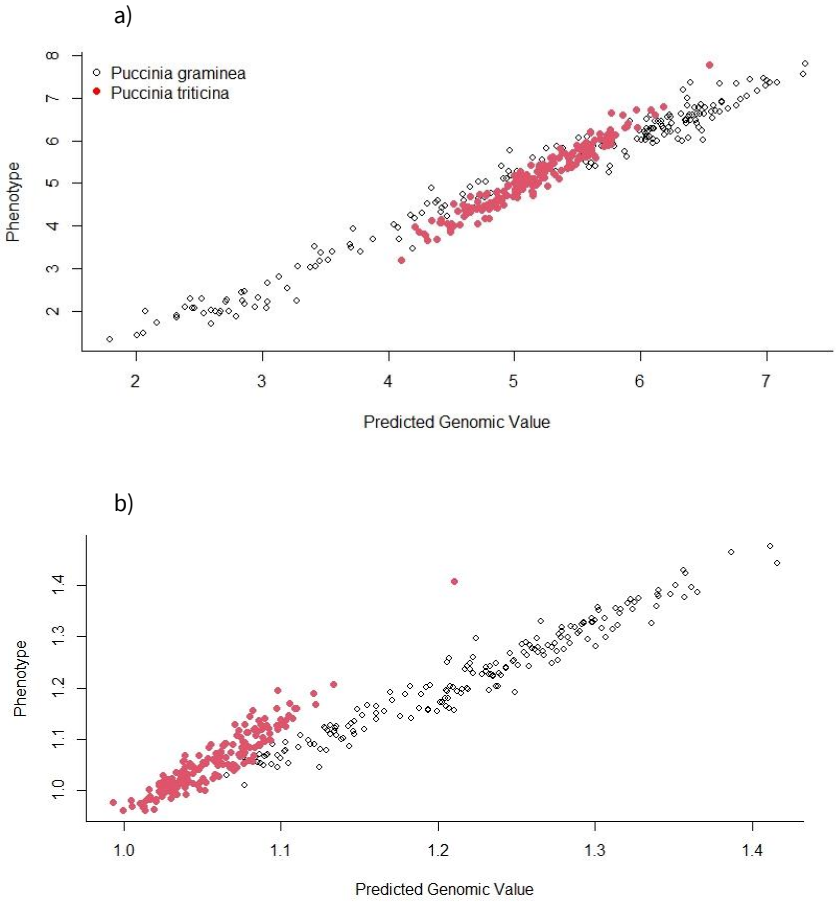
3.3.2. Multitrait uspořádání

Bylo testováno několik modelů počítajících odhady genetických hodnot znaků v případě, že do analýzy vstupují všechny čtyři znaky najednou. V programovém balíčku BGLR byly testovány dva modely: SpikeSlab a BRR a v programovém balíčku MTGS model mLASSO (Budhlakoti et al., 2019b). Program MTGS je volně dostupný na webu: https://CRAN.R-project.org/package=MTGS. Parametr DIC vykazoval nejnižší hodnotu pro model BRR, a to i v případech různých kombinací znaků i pro všechny 4 znaky najednou. Na Obrázku 6 je uveden příklad korelace mezi vypočtenou a naměřenou hodnotou znaku zvlášť pro hodnocení a zvlášť pro stanovení, protože se tyto dvě kategorie proměnných lišily v rozsahu hodnot.

Jako v předchozím případě, model zahrnuje regresní koeficienty – efekty markerů, penalizační kritéria markerů a prvky kovarianční matice náhodných efektů. Na základě tohoto modelu jsou vypočteny tzv. GEBV hodnoty u genotypů šlechtitelské populace.

Výhodou tohoto uspořádání je rychlejší statistické zpracování dat, neboť se hodnoty GEBV počítají pro všechny znaky najednou. Navíc je umožněna návazná analýza funkce očekávaných ztrát – viz. dále.

Obrázek 6 - Grafy korelace mezi vypočtenou a naměřenou hodnotou znaku a) *hodnocení Pg* a *hodnocení Pt*; b) *stanovení Pg* a *stanovení Pt* pro model genomické selekce BRR



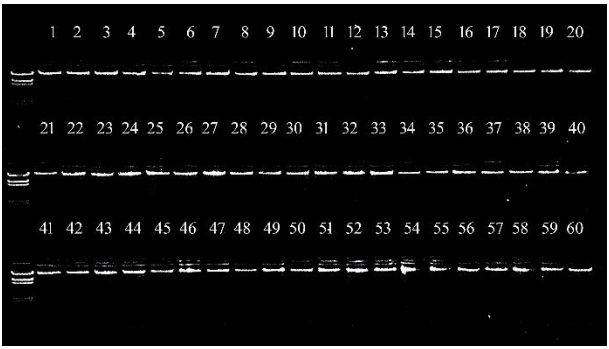
4. Šlechtitelská populace

Do šlechtitelské populace bylo vybráno 103 genotypů pšenice, které zahrnují pokročilé linie šlechtitelského programu a vycházejí z odrůd a linií pšenice vybraných do trénovací populace. Vzorky listových pletiv ve fázi BBCH13 byly odebrány přímo z pokusných parcelek – řádkových výsevů – a uchovávány při teplotě -80°C. Bylo vždy odebráno 5 listů z 5 rostlin.

4.1. Genotypová analýza

Z odebraných vzorků byla extrahována DNA pomocí CTAB podle optimalizovaného protokolu. Kvalita a koncentrace byla ověřena elektroforeticky a spektrofotometricky pomocí fluorometru Qubit4 (Invitrogen). Roztoky vzorků DNA byly ředěny na koncentraci 100 ng/µl (Obrázek 7). Všech 103 vzorků bylo odesláno na analýzu DArT Seq do Austrálie.

Obrázek 7 – Příklad elektroforetogramu vzorků DNA ředěných na koncentraci 100 ng/µl



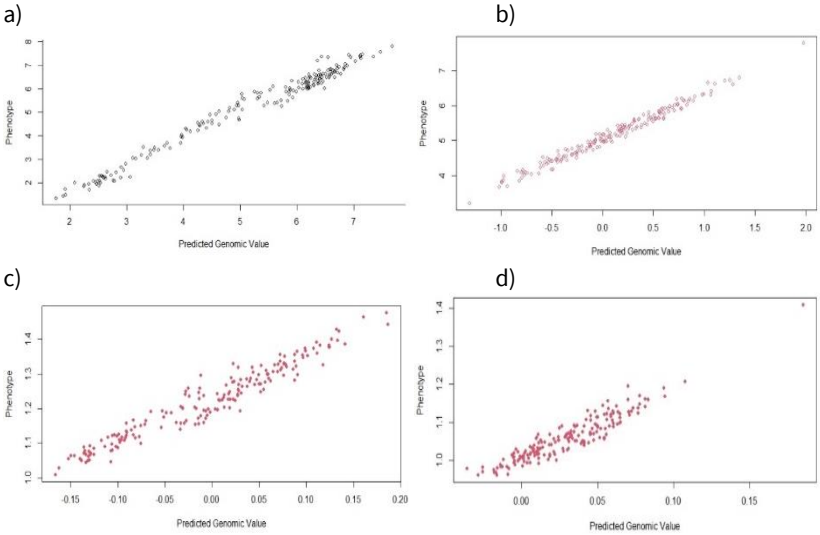
Jako u trénovací populace byla zvolena odrůda pšenice seté Chinese Spring 21 a konsensuální mapa: Wheat_ConsensusMap_version_4 jako referenční. Výsledky analýzy byly staženy, převedeny z formátu *.csv do formátu *.xlsx a uloženy. Pro další zpracování byl vybrán soubor Report_*_SNP_mapping_2.csv, kde hodnota „0“ značí alelu reference, „1“ alternativní alelu a „2“ heterozygotní sestavu alel. Pro zachování koherence výsledků je třeba sjednotit obě matice dat pro trénovací i šlechtitelskou populaci tak, aby byla identická sada markerů a jednotné kódování alel. Proto byla nejprve

použita funkce inner_join v balíčku dplyr v prostředí programu R (je to jedna z několika možností), která sjednotila obě tabulky průnikem přes společné markery. Průnikem obou tabulek s 103308 a 55613 markery byla matice dat s 288 genotypy pšenice a 21883 markery. Poté bylo provedeno filtrování dat - tj. v souboru byla ponechána pouze data s hodnotou PIC ≥ 0,05 a chybějícími daty ≤ 10 %. Pak byla provedena transpozice dat, tj. výměna řádků za sloupce a opačně a dopočet dat na základě pravděpodobnosti. Finální datová matice tedy obsahovala 288 řádků (genotypů pšenice) a 10092 sloupců (markerů).

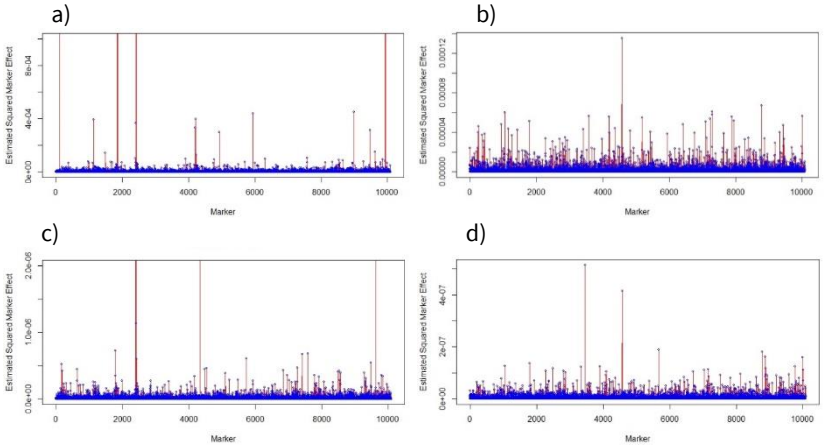
5.2. Genomická selekce pro jeden znak

Pro výpočet hodnot GEBV (Genomic estimated breeding value) pro každou ze 4 proměnných (*hodnocení Pg*, *hodnocení Pt*, *stanovení Pg*, *stanovení Pt*) bylo použit model Bayes A vybraný v předchozím kroku. Počet iterací byl nastaven na 100000 po zahřívacích 10000 iteracích. Výstupem byla tabulka vypočtených hodnot GEBV a jejich standardních odchylek (SD), hodnoty reziduální variance a hodnoty efektů jednotlivých markerů. Grafickými výstupy jsou Obrázky 8, 9 a 10.

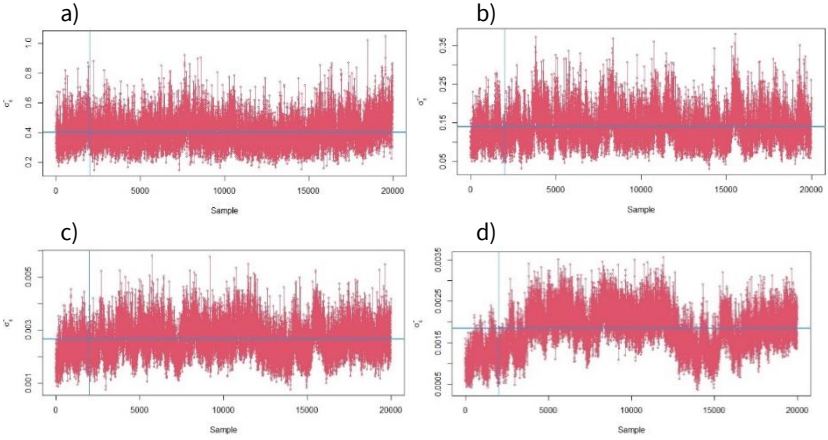
Obrázek 8 – Grafy korelace mezi vypočtenými hodnotami znaku a odpovídajícími naměřenými hodnotami pro genotypy trénovací populace (viz. kap. 4.1.) a) *hodnocení Pg*; b) *hodnocení Pt*; c) *stanovení Pg*; d) *stanovení Pt*.



Obrázek 9 – Grafy efektů markerů pro znaky a) *hodnocení Pg*; b) *hodnocení Pt*; c) *stanovení Pg*; d) *stanovení P*t.



Obrázek 10 – Grafy reziduální variance v prvních 20000 iteracích pro znaky a) *hodnocení Pg*; b) *hodnocení Pt*; c) *stanovení Pg*; d) *stanovení Pt*.



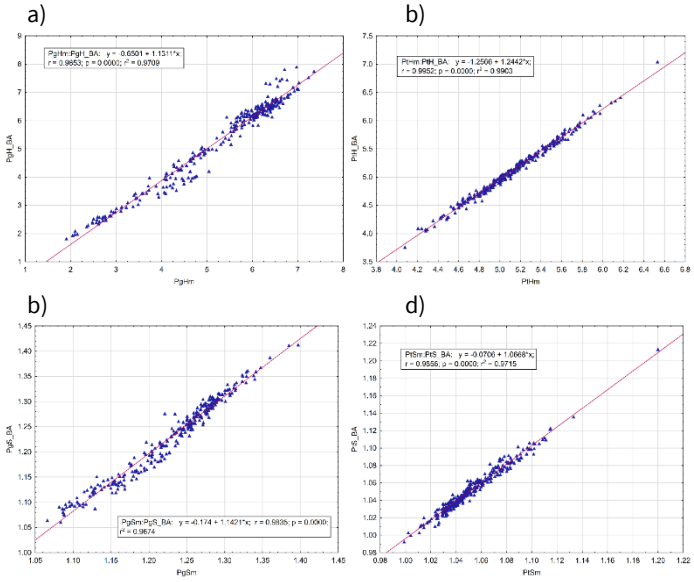
Na obrázku 9 je možno vidět, že některé markery mají významnější vliv na daný znak a mohou být potencionálně využity jako molekulární markery ve šlechtění na rezistenci. U rzi travní bylo takto identifikováno 7 markerů, u rzi pšeničné 2 markery. Jak ukazuje modrá vodorovná linie v grafech obrázku

10, model u jednotlivých znaků vysvětluje 60 % variance u *hodnocení Pg*, 85 % u *hodnocení Pt*, 99,7 % u *stanovení Pg* a 99,8 % u *stanovení Pt*. Mísení (mixing) reziduální variance v jednotlivých krocích počátečních 20000 iterací je regulérní, celkové nastavení na 100000 iterací je tedy postačující (Obrázek 10).

5.3. Genomická selekce pro více znaků

Funkce Multitrait, která je součástí balíčku BGLR pro program R byla použita pro výpočet šlechtitelských hodnot GEBV pro všechny čtyři proměnné najednou. Jako model byl již v předchozím kroku vybrán model BRR. Důležité je nastavení parametrů uložení efektů markerů v každém kroku z nastavených 100000 iterací a nastavení reziduální kovariance (resCov=list(type="UN",saveEffects=TRUE)).
Hodnoty GEBV vypočtené pomocí multitrait analýzy pro všechny 4 znaky korelovaly s hodnotami GEBV vypočtenými pro každý znak zvlášť (r = 0,98 - 1,00; Obrázek 11).

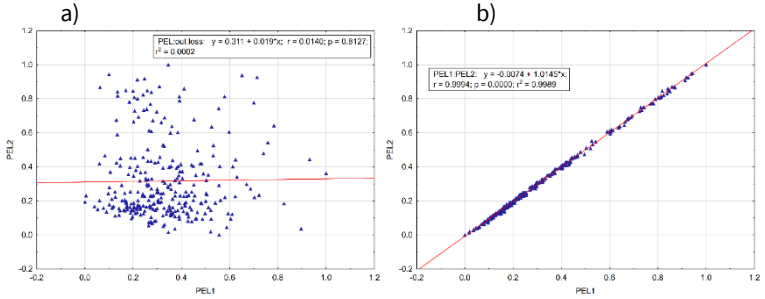
Obrázek 11 – Grafy korelace mezi hodnotami znaku (GEBV) vypočtenými pomocí modelu BRR v multitrait uspořádání (na ose x) a pomocí modelu Bayes A pro každý znak zvlášť: a) *hodnocení Pg*; b) *hodnocení Pt*; c) *stanovení Pg*; d) *stanovení Pt*.



Aby bylo možno vybírat genotypy s kombinovanou rezistencí k oběma patogenům, byl optimalizován výpočet tzv. funkce očekávaných ztrát (Expected loss), což je typ funkce pro vyhodnocování rizik. Používá se zejména v bankovnictví jako klíčový nástroj pro odhad možných finančních ztrát. Ve šlechtění jsou pomocí této funkce vybírány genotypy s co největším ziskem, v našem případě s co největšími hodnotami GEBV požadovaných znaků. K výpočtu byl použit programový balíček MPS v prostředí R, který je volně dostupný na webu: https://github.com/bjesusvh/MPS. Jednotlivé funkce a příklady jsou uvedeny v publikaci Villar-Hernández et al. (2023). Nejprve byly z výstupních dat předchozího kroku vypočteny hodnoty proměnných *B0* (matice průměrů hodnot znaků), *yHat* (matice vypočtených GEBV) a *R* (matice reziduální kovariance). Pak byla použita funkce ApproxMPS, která na rozdíl od přednastavené funkce FastMPS používá souhrnné výsledky po 100000 iteracích jako vstupní hodnoty proměnných. Funkce FastMPS používá jako vstupní hodnoty proměnných matice z hodnot z jednotlivých iterací, z prvních 10000. Hodnoty vypočtené funkce ztrát pak nekorelují mezi nezávislými analýzami (Obrázek 12). V případě, že je třeba použít vyšší počty iterací pro vyrovnání reziduální variance, doporučujeme použít funkci ApproxMPS. Přednastavená metoda výpočtu funkce ztrát „kl“ (Kullback-Leibler) byla ponechána.

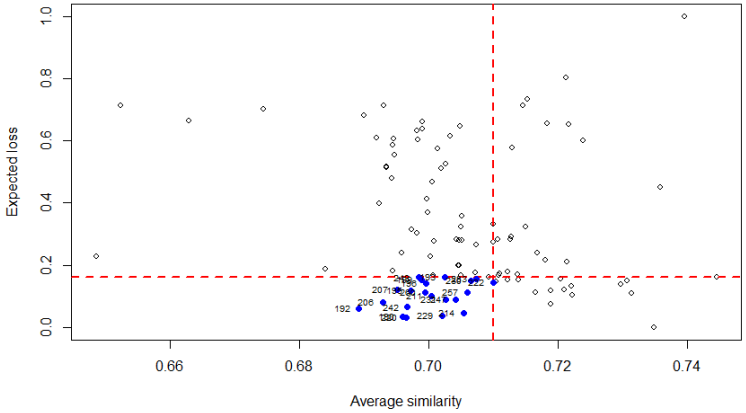
Na závěr byly vypočteny hodnoty podobnosti mezi všemi genotypy pšenice šlechtitelské populace pomocí koeficientů jednoduché shody (Simple Matching coefficients) s využitím funkce SimMatrix v programovém balíčku MPS. Výsledkem je matice n × n, kde n je počet genotypů. Z této matice byla vypočtena průměrná genetická podobnost pro každý genotyp zvlášť a uložena v proměnné *AveSim*.

Obrázek 12 – Graf korelace hodnot funkce ztrát vypočtených v nezávislých analýzách za použití funkce a) FastMPS a b) ApproxMPS



Na závěr byl sestrojen graf, kde na ose x je průměrná genetická podobnost genotypu a na ose y hodnota funkce očekávaných ztrát pro daný genotyp (Obrázek 13). Linie odpovídají kvantilům – 0,35 u funkce očekávaných ztrát a 0,65 u genetické podobnosti genotypů. Tyto hodnoty je možno nastavit libovolně podle požadavků šlechtitelů na počet vybíraných linií. V tomto konkrétním případě bylo vybráno 21 linií pšenice – v grafu jsou reprezentovány označenými zvýrazněnými body v levém dolním kvadrantu (Obrázek 13).

Obrázek 13 – Graf umístění genotypů testovací populace pšenice v 2D prostoru hodnot genetické podobnosti a předpokládaných ztrát; červené čáry reprezentují kvantily 0,35 u hodnot funkce očekávaných ztrát (***expected loss***) a 0,65 u hodnot průměrné podobnosti (***average similarity***)



Tabulka 3 prezentuje vybrané genotypy. Průměrná hodnota změny znaku u selektovaných linií se počítá jako procentuální vyjádření průměrné hodnoty znaku u vybraného souboru genotypů ku průměrné hodnotě znaku nevybraných genotypů. V tomto případě genetický zisk dosáhl 16,1 % u *hodnocení Pg*, 4,9 % u *hodnocení Pt*, 3,1 % u *stanovení Pg* a 0,5 % u *stanovení Pt*. Tyto hodnoty jsou ovlivněny strategií výběru, a tedy preferencemi šlechtitelů. Proto jsou uvedené hodnoty pouze příkladem konkrétního nastavení.

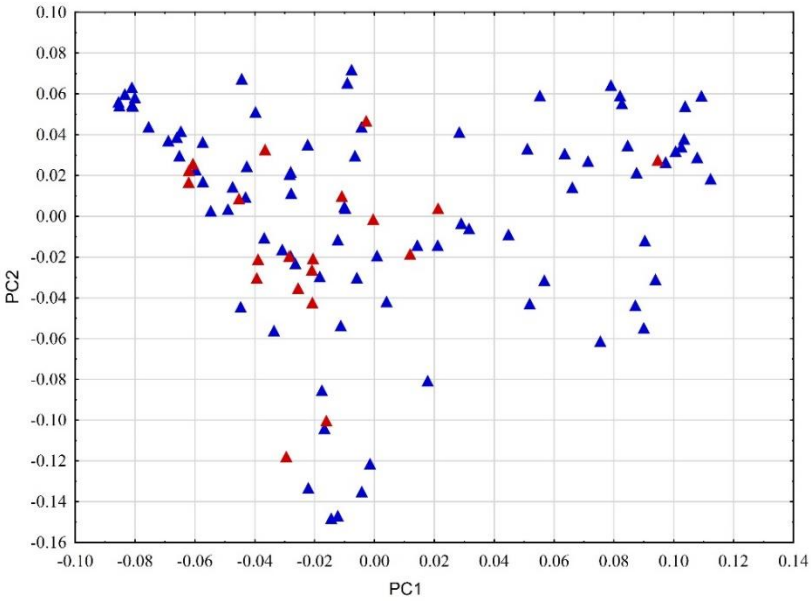
Tabulka 3 – Modelem selektované genotypy pšenice z hlediska vyšší míry kombinované rezistence ke rzi travní a rzi pšeničné a menší genetické podobnosti

ID	Název	PgH	PtH	PgS	PtS	AS
Ta25005	SG-S917-23	6.223±1.407	5.561±0.651	1.254±0.084	1.076±0.046	0.696
Ta25007	SG-S2042-23	6.213±1.224	5.467±0.563	1.268±0.074	1.056±0.04	0.689
Ta25008	SG-S2172-23	5.851±1.231	5.032±0.568	1.281±0.074	1.055±0.04	0.702
Ta25010	SG-S874-23	6.212±1.335	5.187±0.607	1.271±0.08	1.047±0.043	0.697
Ta25011	SG-S1185-23	6.556±1.303	4.998±0.594	1.291±0.078	1.033±0.042	0.700
Ta25014	SG-S2156-23	6.297±1.178	5.053±0.54	1.257±0.07	1.058±0.038	0.699
Ta25021	SG-S1777-24	6.082±1.265	5.193±0.589	1.293±0.076	1.055±0.041	0.693
Ta25022	SG-S2057-24	6.503±1.079	5.2±0.493	1.252±0.064	1.059±0.035	0.695
Ta25026	SG-S1007-24	6.319±1.093	5.134±0.514	1.279±0.066	1.05±0.036	0.700
Ta25029	SG-S1380-24	6.762±1.248	5.347±0.568	1.29±0.073	1.041±0.04	0.705
Ta25035	SG-S1707-24	6.341±1.243	5.488±0.579	1.278±0.075	1.055±0.041	0.696
Ta25037	SG-S1783-24	6.125±1.309	5.103±0.594	1.291±0.078	1.032±0.042	0.710
Ta25044	SG-S1789-24	6.338±1.263	5.438±0.592	1.28±0.075	1.054±0.041	0.702
Ta25045	SG-S1972-24	5.942±1.118	5.256±0.519	1.259±0.068	1.047±0.037	0.707
Ta25046	SG-S2012-24	6.034±1.079	5.435±0.504	1.262±0.064	1.057±0.035	0.703
Ta25057	ST822/25	6.761±1.108	5.413±0.516	1.267±0.066	1.05±0.037	0.697
Ta25058	ST825/25	6.648±1.17	4.95±0.531	1.288±0.07	1.032±0.038	0.698
Ta25062	ST839/25	6.121±1.092	5.372±0.51	1.284±0.066	1.041±0.036	0.704
Ta25072	ST892/25	6.221±1.074	5.187±0.493	1.261±0.064	1.06±0.035	0.706
Ta25075	ST904/25	5.976±1.096	5.216±0.499	1.285±0.066	1.048±0.036	0.699
Ta25098	ST965/25	6.444±1.12	4.942±0.521	1.275±0.067	1.049±0.037	0.707

AS – Average Similarity

Míra diverzity vybraných genotypů byla kromě modelem vypočtené průměrné genetické podobnosti (Tabulka 3) ověřena ještě nezávisle pomocí MDS analýzy genotypových dat šlechtitelské populace pšenice (Obrázek 14).

Obrázek 14 – Graf prvních dvou latentních proměnných MDS analýzy genotypů pšenice šlechtitelské populace; červeně jsou označeny genotypy vybrané modelem



5. Závěr

Genomická selekce je na rozdíl od mapování genů a asociačních studií zaměřena na genotyp jako celek. Vychází z existující genetické diverzity a díky právě ukázanému postupu se snaží co nejvíce této diverzity zachovat při výběru linií s lepšími hodnotami testovaných znaků, než byly na počátku šlechtitelského procesu. Tento postup slouží k výběru genotypů, a tedy jak jako rodičů do dalšího procesu šlechtění, tak i k výběru šlechtitelských linií v raných generacích.

Výsledky rovněž ukázaly, že je možno využít genomickou selekci pro několik znaků najednou. Pro většinu hospodářských znaků, které jsou podmíněny polygenně je genomická selekce dobrou cestou pro šlechtění odrůd s lepšími hodnotami znaků při současném uchování co největší míry genetické diverzity.

III. Srovnání novosti postupů

Dle našich znalostí se jedná o první aplikaci postupu genomické selekce ve šlechtění zemědělských plodin v České republice. Na základě výsledků řešení projektu můžeme tento postup doporučit u všech znaků, jejichž dědičnost je komplexní, je podmíněna polygenně, s aditivními či epistatickými účinky genů či složitým systémem regulací jejich exprese. Je možno tímto způsobem vybírat rodiče nebo selektovat šlechtitelské linie na více znaků najednou, což může zrychlit a zlevnit šlechtitelský proces.
Další pozitivum výstupu projektu je zohlednění genetické variability při výběru linií. Doporučením je snažit se vybírat vstupní soubory genotypů s co největší mírou genetické variability, a i při selektování linií vybírat materiál s co nejnižšími hodnotami genetické podobnosti. Tímto lze docílit udržení efektivní velikosti populace, a tedy optimalizace zisku ze selekce oproti nárůstu inbreedingu. Obecně pro zachování genetické diverzity lze doporučit tzv. prebreeding s využitím planých příbuzných druhů nebo starých krajových odrůd.
Zavedení tohoto přístupu do šlechtitelské praxe je žádoucí a znamená rozšíření portfolia nových šlechtitelských technik.

IV. Popis uplatnění metodiky

Šlechtění obilnin je náročný proces, zvlášť když téměř všechny hospodářsky důležité znaky jsou kvantitativní povahy. Navíc v průběhu posledních více než 100 let šlechtění došlo k výrazné ztrátě genetické diverzity. Tato metodika uvádí odbornou veřejnost do problematiky jedné z nových technik šlechtění – genomické selekce, která představuje odlišný přístup ke šlechtění a zohledňuje i zachování co nevětší míry genetické diverzity, která se ve vstupních populacích nachází.
Uplatnění metodiky v praxi znamená posun know-how šlechtění obilnin v České republice. Proto je metodika předána přední šlechtitelské firmě ČR. Metodika bude dále prezentována odborné veřejnosti formou přednášek, posterů a článků v odborných časopisech tak, aby se povědomí o této nové šlechtitelské technice co nejvíce rozšířilo.

Problematika genomické selekce bude dále rozvíjena i na pracovišti autorů metodiky a optimalizována pro další zemědělské plodiny a důležité hospodářské znaky.

V. Ekonomické aspekty

Šlechtění nové odrůdy pšenice je komplexní, dlouhodobý a finančně velmi náročný proces, který trvá přibližně 10–12 let. Celkové náklady na vyšlechtění jedné nové odrůdy, která se dostane až na trh, se pohybují v řádech desítek milionů korun. Klíčovými faktory nákladů jsou zejména dlouhá doba vývoje odrůdy: 10–12 let intenzivní práce a malá úspěšnost, kdy z tisíců křížení je do registrace úspěšně uvedena jen zanedbatelná část. Náklady zahrnují práci šlechtitelů, laboratorní testy, polní pokusy a registrační řízení. Ve výsledku tak vývoj jedné odrůdy přesahuje 20 až 50 milionů Kč (přepočteno z mezinárodních standardů šlechtění, jelikož přímé fixní náklady v Kč za 1 odrůdu nejsou veřejně uváděny).
Zavedením genomické selekce jako nástroje zejména k výběru kříženců v raných generacích dojde ke zvýšení efektivity šlechtitelského procesu min o 10 %, tedy o 2 až 5 milionů Kč u vývoje jedné odrůdy pšenice. U ostatních obilnin je situace obdobná.

VI. Seznam použité literatury

Abed A., Pérez-Rodriguez P., Crossa J., Belzile F. (2018): When less can be better: How can we make genomic selection more cost-effective and accurate in barley? Theoretical and Applied Genetics 131: 1873-1890.

Babu P., Baranwai D.K., Harikrishna, Pal D., Bharti H., Joshi P., Thiyagarajan B., Gaikwad K.B., Bhardwaj S.C., Singh G.P., Singh A. (2020): Application of genomic tools in wheat breeding to attain durable rust resistance. Frontiers in Plant Science 11: 567147.

Barrett L.W., Fletcher S., Wilton S.D. (2012): Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. Cellular and Molecular Life Sciences 69: 3613-3634.

Bennett M.D., Smith J.B. (1976): Nuclear DNA amounts in angiosperms. Philos. Trans. R. Soc. Lond. Biol. Sci. 274(933): 227-274.

Boone C., Bussey H., Andrews B.J. (2007): Exploring genetic interactions and networks with yeast. Nat Rev Genet. 8(6): 437–449.

Bradbury P.J., Thang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. (2007): TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633-2635. Doi:10.1093/bioinformatics/btm308

Budhlakoti N., Mishra D.C., Rai A., Lal S.B., Chaturvedi K.K., Kumar R.R. (2019): A comparative study of single-trait and multi-trait genomic selection. Journal of Computational Biology 26: 1100-1112.

Budhlakoti N., Mishra D.C., Rai A. (2019b): Package MTGS. 1-6.

Budhlakoti N., Kushwaha A.K., Rai A., Chaturvedi K.K., Kumar A., Pradhan A.K., Kumar U., Kumar R.R., Juliana P., Mishra D.C., Kumar S. (2022): Genomic selection: A tool for accelerating the efficiency of molecular breeding for development of climate-resilient crops. Frontiers in Genetics 13: 832153.

Clark S.A., Hickey J.M., Van der Werf J.H. (2011): Different models of genetic variation and their effect on genomic evaluation. Genetics, Selection and Evolution 43: 18.

Collard B.C.Y., Jahufer M.Z.Z., Brouwer J.B., Pang E.C.K. (2005): An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted for crop improvement: the basic concepts. Euphytica 142: 169-196.

Daetwyler H.D., Hickey J.M., Henshall J.M., Dominik S., Gredler B., van der Werf J.H.J. (2010): Accuracy of estimated genomic breeding values for wool and meat in a multi-breed sheep population. Animal Production Science 50: 1004-1010.

de los Campos G., Naya H., Gianola D., Crossa J., Legarra A., Manfredi E., Weigel K., Cotes J.M. (2009): Predicting quantitative traits with regression models for dense molecular merkers and pedigree. Genetics 182: 375-385.

de los Campos G., Gianola D., Rosa G.J.M., Weigel K.A., Crossa J. (2010): Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. Genetic Research 92: 295-308.

Gianola D., Fernando R.L., Stella A. (2006): Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics 173: 1751-1776.

Habier D., Fernando R.L., Kizilkaya K., Garrick D.J. (2011): Extension of the Bayesian alphabet for genomic selection. BMC Bioinformatics 12: 186-197.

Heffner E.L., Lorenz A.J., Jannink J.L., Sorrels M.E. (2010): Plant breeding with genomic selection: gain per unit time and cost. Crop Science 50: 1651-1690.

Holliday J.A., Wang T., Aitken S. (2012): Predicting adaptive phenotypes from multilocus genotypes in Sitka spruce (*Picea sitchensis*) using random forest. Genes, Genomes and Genetics 2: 1085-1093.

Howard R., Carriquiry A.L., Beavis W.D. (2014): Parametric and nonparametric statistical methods for genomic selection of traits with additive and epistatic genetic architectures. Genes, Genomes, Genetics 4: 1027-1046.

Maenhout S., De Baets B., Haesaert G., Van Bockstaele E. (2007): Support vector machine regression for the prediction of maize hybrid performance. Theoretical and Applied Genetics 115: 1003-1013.

Meuwissen T.H.E., Hayes B.J., Goddard M.E. (2001): Prediction of total genetic values using genome-wide dense marker maps. Genetics 157: 1819-1829.

Mohan M.S., Nair A., Bhagwat T.G., Krishna M., Yano M., Bhatia C.R. Sasaki T. (1997): Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding 3: 87-103.

Pérez P., de los Campos G. (2014): Genome-wide regression and prediction with the BGLR statistical package. Genetics 198: 483-495.

Perrier X, Jacquemoud-Collet JP (2006) DARwin software http://darwin.cirad.fr/darwin, the latest accession: 2017-10-02.

Rafalski J., Tingey S. (1993): Genetic diagnostics in plant breed: RAPDs, microsatellites and machines. Trends in Genetics 9: 275-280.

Ribaut J.M., Hoisington D. (1999): Marker-assisted selection: New tools and strategies. Trends in Plant Science 3: 236-239.

Tibshirani R. (1996): Regression shrinkage and selection via the Lasso. Journal of Royal Statistical Society Series B 58: 267-288.

Usai M.G., Goddard M.E., Hayes B.J. (2009): LASSO with cross-validation for genomic selection. Genetic Resources 91: 427-436.

Verges V.L., Lyerly J., Dong Y., Van Sanford D.A. (2020): Training population design with the use of regional Fusarium head blight nurseries to predict independent breeding lines for FHB traits. Frontiers in Plant Science 11: 1083.

Villar-Hernández B.d.J., Dreisigacker S., Crespo L., Pérez-Rodríguez P., Pérez-Elizalde S., Toledo F., Crossa J. (2023): A Bayesian optimization R package for multitrait parental selection. The Plant Genome 17: e20433.

VII. Seznam publikací, které předcházely metodice

Leišová-Svobodová L., Psota V., Zavřelová M., Kříž M., Mařík P., Nesvadba Z. (2024): Malting quality molecular markers for barley breeding. Czech Journal of Genetics and Plant Breeding 60(2): 70-78.

Jungová M., Leišová-Svobodová L., Hanzalová A., Zelba O., Svoboda P. (2022): Detection and quantification of *Puccinia graminis* and *Puccinia triticina* in host plant tissue by real-timePCR. Kostelecké inspirování, ČZU, Praha 2022: 42-43.

Hanzalová A. (2023): Hodnocení odolnosti jarních odrůd pšenice ke rzím. Úroda 2: 27-28.

Ondreičková K., Hanzalová A., Klčová L., Hudcovicová M., Hrdlicová M., Gregová E., Šliková S. (2025): High genetic variability of *Puccinia triticina* at a specific location in Slovakia: a case study. Cereal Research Communications 53: 1505-1517.

Leišová-Svobodová L., Ježek S., Veškrna O., Hanzalová A., Jungová M., Sovová T., Svoboda P. (2026): Genetic variability of wheat collection selected with respect to the resistance to leaf and stem rusts. Agronomy 16(14): 1322. DOI: 10.3390/agronomy16141322.

Leišová-Svobodová L., Hanzalová A., Veškrna O., Jungová M., Sovová T., Svoboda P. (2026): Multi-trait genomic selection for combined leaf and stem rust resistance in wheat breeding. (*in review*)

Název: Aplikace genomické selekce ve šlechtění obilnin

Autoři: RNDr. Mgr. Leona Svobodová, Ph.D., Mgr. Alena Hanzalová, Ph.D., Ing. Michaela Jungová, Ph.D., Ing. Tereza Sovová, Ph.D., Ing. Pavel Svoboda, Ph.D.

Vydalo: Národní centrum zemědělského a potravinářského výzkumu, v.v.i., Drnovská 507, 161 00 Praha 6 – Ruzyně

Počet stran: 25

Rok vydání: 2026

ISBN: 978-80-7427-466-4