

## Oponentský posudek

Na disertační práci Ing. Jana Říhy zpracovanou na téma: „**Microsatellite Markers and Genetic Diversity Issues in Cattle**“

*Školitel:* Prof. Ing. Jindřich Čítek, CSc.

*Školitel specialista:* Prof. Ing. Josef Dvořák, CSc., dr. h. c.

*Doktorský studijní program:* Biotechnologie

*Doktorský studijní obor:* Zemědělské biotechnologie

*Oponent:* Doc. Ing. Tomáš Urban, Ph.D.

---

Předložená disertační práce se zabývá matematicko statistickou analýzou markerů typu mikrosatelity pro vyhodnocování genetické diverzity u skotu. Přičemž jsou popsány parametry genetické diverzity pomocí mikrosatelitů u několika plemen skotu v České republice, vyhodnocování testů paternity u skotu a možnosti jejich využití, a aplikované výsledky získávání genotypových dat a vyvinutý software, který je využitelný v laboratorní praxi servisní laboratoře. Posledním záměrem bylo určování jedinců aplikací algoritmů strojového učení na základě genetického profilu mikrosatelitů na úrovni plemene.

Disertační práce vznikla ve spolupráci pracoviště Katedry genetiky, šlechtění a výživy zvířat ZF JU v Českých Budějovicích a pracoviště laboratoře agrogenomiky při Ústavu morfologie, fyziologie a genetiky zvířat AF MENDELU. Práce má rozsah 191 stran bez příloh. Spis je psán v anglickém jazyce s českým souhrnem.

Předložená disertační práce má obvyklé členění. Po úvodu a cíli, následuje přehled literatury, který je členěn do sedmi podkapitol zabývajících se popisem evoluce skotu a jeho genomu, charakterizace plemen, popis parametrů využívaných v genetice populací, metody strojového učení používané v genetice populací, mikrosatelitní markery a jejich využívání v hodnocení diverzity u populací skotu. Disertační práce vychází ze 155 literárních zdrojů, především zahraničních citací.

**Hodnocení formální stránky práce:** Text má kvalitní zpracování z pohledu typografického. Přesto hodnotím negativně množství textu na stránce, i když při počítačovém zpracování textu není možné zcela dodržet normované hodnoty (30 řádků na stránku a 60 úhozů na řádek). Text je průměrně vysázen kolem 48 řádků na stránku a mezi 70 – 80 úhozů na řádku. Textu je prokládán mnoha tabulkami a obrázky, přičemž chybí seznam tabulek a obrázků. Řada tabulek a obrázků bylo možné vložit do příloh. Zpřehlednil by se tím samotný spis.

Každá tabulka má být samovysvětlující a odkazy na vysvětlení zkratk plemen na jinou tabulku není zcela správné (např. Tab. 5.13, str. 76). Bohužel ve spisu chybí seznam používaných zkratk.

Materiál a metodika uvádí konkrétní použité mikrosatelity a popisuje strukturu databází. Z metodiky vyplývá, že pro výpočty standardních parametrů diverzity a odhady a validace testování paternity byly použity běžné softwary (zejména PowerMarker). Zde uváděné vzorce bylo zbytečné uvádět. Dále je zde popsán postup programování databáze využitelnou pro konkrétní referenční laboratoř a posouzení algoritmů strojového učení pro diskriminaci plemene podle genotypových dat.

Kapitola Výsledky a diskuze je uvedena na 95 stranách a je tvořena 4 podkapitolami podle dílčích cílů s vyhodnocenými a diskutovanými výsledky. Souhrnně a pro jednotlivá plemena byla spočítána genetická diverzita a vytvořeny fylogenetické stromy metodami NJ a UPGMA softwarem PowerMarker v3.25. Další částí této kapitoly je popis vyvinutého softwaru (databáze a grafické výstupy) pro referenční laboratoř. Poslední částí jsou výsledky aplikace různých metod strojového učení pro diskriminaci plemen skotu pomocí softwaru Weka 3.6. Velmi zajímavé je porovnání výsledků získaných u skotu s výsledky u koní.

Po souhrnu, který je i v českém jazyce, následují kapitoly Praktické využití výsledků a Teoretické závěry a budoucí práce.

*Dotazy oponenta k předložené disertační práci a náměty do diskuse:*

---

Proč nebyly uvedeny frekvence alel a genotypů ve výsledcích? Jsou to základní údaje, z kterých jsou počítány další parametry.

V metodice chybí popis výpočtu parametru  $F_{IS}$ .

V literárním přehledu se popisují všechny tři F statistiky ( $F_{IS}$ ,  $F_{IT}$  a  $F_{ST}$ ), ve výsledcích byl použit jen parametr  $F_{IS}$ . Proč nebyla hodnocena vnitro i mezipopulační diverzita komplexně všemi F statistikami?

Bylo by vhodnější využít pro konstrukci fylogenetických stromů i jiné metody (např. ML)? Je to proto, že byl použit jeden konkrétní software?

Proč byla sledována vhodnost sady použitých mikrosatelitů pro vyloučení rodičovství, když tento panel mikrosatelitů (dle ISAG) je standardně rutinně využíván a tedy ověřen?

Proč tyto původní výsledky nebyly publikovány v časopisu s impakt faktorem, nebo jsem alespoň publikaci popisující výsledky uváděné v této disertační práci nenašel v seznamu publikací (podle § 47, 111/1998 Sb.)<sup>1</sup>?

Je vytvořený software právně chráněn?

---

<sup>1</sup> . Disertační práce musí obsahovat původní a uveřejněné výsledky nebo výsledky přijaté k uveřejnění (**§ 47, zákona 111/1998 Sb.**).

Musí obsahovat původní výsledky, prokazatelně vyplývající z disertační práce, které byl uveřejněny ve vědeckých časopisech, případně výsledky prokazatelně přijaté k uveřejnění. Student je povinen před obhajobou disertační práce publikovat nejméně jednu práci ve vědeckém časopise s impakt faktorem, která byla prokazatelně zpracována z výsledků získaných při zpracování disertační práce (**bod 9.2 opatření děkana ZF JU v Českých Budějovicích č. 20/2007**).

Materiál a metodika uvádí konkrétní použité mikrosatelity a popisuje strukturu databází. Z metodiky vyplývá, že pro výpočty standardních parametrů diverzity a odhady a validace testování paternity byly použity běžné softwary (zejména PowerMarker). Zde uváděné vzorce bylo zbytečné uvádět. Dále je zde popsán postup programování databáze využitelnou pro konkrétní referenční laboratoř a posouzení algoritmů strojového učení pro diskriminaci plemene podle genotypových dat.

Kapitola Výsledky a diskuze je uvedena na 95 stranách a je tvořena 4 podkapitolami podle dílčích cílů s vyhodnocenými a diskutovanými výsledky. Souhrnně a pro jednotlivá plemena byla spočítána genetická diverzita a vytvořeny fylogenetické stromy metodami NJ a UPGMA softwarem PowerMarker v3.25. Další částí této kapitoly je popis vyvinutého softwaru (databáze a grafické výstupy) pro referenční laboratoř. Poslední částí jsou výsledky aplikace různých metod strojového učení pro diskriminaci plemen skotu pomocí softwaru Weka 3.6. Velmi zajímavé je porovnání výsledků získaných u skotu s výsledky u koní.

Po souhrnu, který je i v českém jazyce, následují kapitoly Praktické využití výsledků a Teoretické závěry a budoucí práce.

*Dotazy oponenta k předložené disertační práci a náměty do diskuse:*

---

Proč nebyly uvedeny frekvence alel a genotypů ve výsledcích? Jsou to základní údaje, z kterých jsou počítány další parametry.

V metodice chybí popis výpočtu parametru  $F_{IS}$ .

V literárním přehledu se popisují všechny tři F statistiky ( $F_{IS}$ ,  $F_{IT}$  a  $F_{ST}$ ), ve výsledcích byl použit jen parametr  $F_{IS}$ . Proč nebyla hodnocena vnitro i mezipopulační diverzita komplexně všemi F statistikami?

Bylo by vhodnější využít pro konstrukci fylogenetických stromů i jiné metody (např. ML)? Je to proto, že byl použit jeden konkrétní software?

Proč byla sledována vhodnost sady použitých mikrosatelitů pro vyloučení rodičovství, když tento panel mikrosatelitů (dle ISAG) je standardně rutinně využíván a tedy ověřen?

Proč tyto původní výsledky nebyly publikovány v časopisu s impakt faktorem, nebo jsem alespoň publikaci popisující výsledky uváděné v této disertační práci nenašel v seznamu publikací (podle § 47, 111/1998 Sb.)<sup>1</sup>?

Je vytvořený software právně chráněn?

---

<sup>1</sup> . Disertační práce musí obsahovat původní a uveřejněné výsledky nebo výsledky přijaté k uveřejnění (**§ 47, zákona 111/1998 Sb.**).

Musí obsahovat původní výsledky, prokazatelně vyplývající z disertační práce, které byl uveřejněny ve vědeckých časopisech, případně výsledky prokazatelně přijaté k uveřejnění. Student je povinen před obhajobou disertační práce publikovat nejméně jednu práci ve vědeckém časopise s impakt faktorem, která byla prokazatelně zpracována z výsledků získaných při zpracování disertační práce (**bod 9.2 opatření děkana ZF JU v Českých Budějovicích č. 20/2007**).

Proč nebyly parametry genetické diverzity u plemen skotu vypočítány ve vytvořeném programu a porovnány se standardními programy používanými v genetice populací (třeba použitým PowerMarker)?

Jaké další typy markerů (spojené s pokroky v genomice) by se v současnosti mohly využívat v genetice populací a bylo by možné u nich využívat metody strojového učení?

Kolik jedinců by bylo potřeba v každé subpopulaci, aby diskriminace plemene byla průkazná?

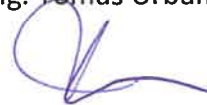
### **Závěrečné hodnocení**

Při celkovém posuzování doktorské disertační práce Ing. Jana Říhy konstatuji, že téma sice není zcela nejaktuálnější z pohledu typu použitého genetického markeru, ale odpovídá současnému stavu rutinního testování v akreditované laboratoři. Zpracování tématu je náročné z hlediska analytických postupů a interpretace. Výsledky mají teoretický i praktický význam pro další vědeckou činnost, dokonce byly již realizovány. Práce splnila svůj cíl i věcné a formální podmínky.

Doporučuji disertaci k obhajobě před oborovou komisí a po jejím úspěšném obhájení doporučuji jmenovanému udělit akademický titul „doktor“, ve zkratce Ph.D.

V Brně dne 5. 9. 2012

doc. Ing. Tomáš Urban, Ph.D.



OPONENTSKÝ POSUDEK  
na doktorskou disertační práci Mgr. Jana Říhy  
„**Microsatellite Markers and Genetic Diversity Issues in Cattle**”

Školitel: prof. ing. Jindřich Čítek, CSc. (JČU České Budějovice)

Školitel specialista: Prof. Ing. Josef Dvořák, CSc., dr.h.c. (MU Brno)

Předložená práce řeší aktuální téma všestranného využití výsledků a aplikace molekulárně genetických studií mikrosatelitních markerů u zvířat, v tomto konkrétním případě u skotu.

Disertační práce je sepsána v anglickém jazyce na 191 stránkách členěných do obvyklých 10 kapitol, práce obsahuje 51 tabulek a 65 obrázků v textu práce. V seznamu použité literatury je uvedeno 155 aktuálních pramenů dané a související literatury zásadně zahraniční provenience. V příloze je uveden životopis a seznam publikací doktoranda.

Po prostudování práce je možno konstatovat, že předložená práce naplňuje cíle práce jak jsou uvedeny na str. 12., a to ve velmi vysoké kvalitě.

Literární přehled je zpracován v rozsahu 39 stran - první část charakterizuje stručně plemena skotu zahrnutá do šetření, druhá se zabývá široce teorií populačně-genetických biometrických postupů použitelných pro zpracování souborů mikrosatelitních profilů populací plemen skotu vybraných k šetření pro popis genetické diverzity, modelovým studiím software pro zpracovávání datových souborů mikrosatelitních markerů vycházejícími z rutinního provozu genetické laboratoře. Tato kapitola poskytuje ucelenou informaci o studované problematice. Velmi dobře a přehledně je zpracováno pojednání algoritmech pro plemennou diskriminaci na základě sady mikrosatelitů – tato část ukazuje na velmi dobré metodologické znalosti autora v informatice využitelné v genetických analýzách.

Kapitola Materiál a metodika ve vyčerpávajícím rozsahu uvádí použité soubory zvířat 8 plemen a skupiny kříženců, soupravu využitých mikrosatelitů, pracovní postupy a metody pro genetickou analýzu, validitu testování paternity, tvorbu software pro zpracování datových souborů markerů, algoritmy metod strojového učení – jsou zvoleny jako vhodné a adekvátní, umožňující vyčerpávající vyhodnocení získaných dat.

Kapitola Výsledky a diskuze charakterizuje výsledky vlastní práce v logickém sledu, tvoří největší část práce.

Analýza ISAG sady mikrosatelitů určené k testaci paternity využitá k posouzení genetické diverzity potvrzuje poznatky o diverzitě a genetických vzdálenostech studovaných plemen – kladné je zjištění, že vhodná šlechtitelská strategie dokáže udržet (zlepšit) potřebnou genetickou variabilitu i v intenzivně šlechtěné populaci.

U daných souborů 8 plemen bylo ověřována - potvrzována vhodnost využití ISAG sady mikrosatelitů k testaci paternity, výsledky studie použitelnost této sady potvrzují. Dále jsou uvedeny návrhy softwarové aplikace pro práci s rozsáhlými soubory markerů produkovaných laboratoří – funkčnost systému potvrzuje jeho využívání akreditovanou laboratoří.

Byla ověřována možnost plemenné diskriminace na základě mikrosatelitních markerů

využitím 10 algoritmů metod strojového učení. Rozsáhlá studie napovídá možnost úspěšného řešení problému – studie ukázala nejlepší výsledek u algoritmu Bayesovských sítí – zde 84 % správně určených případů. Je konstatována možnost určit plemennou příslušnost skotu na základě mikrosatelitních markerů.

Vzhledem k faktu, že poznatky o využití mikrosatelitů doktorand publikoval též v impaktovaných časopisech (viz str. 180 -191 disertace), byly již uváděné výsledky podrobeny oponentnímu řízení.

S ohledem na získané poznatky jsou dosažené výsledky práce přínosem jak po stránce teoretické i metodické pro další výzkum, tak i pro aplikaci v praxi.

Předložená disertační práce taktéž po stránce formální splňuje plně požadavky kladené na tento druh prací, autor se však nevyvaroval chyb v textu způsobených překlepy.

K autorovi práce mám následující dotazy:

- Jaký počet mikrosatelitů sady považuje za dostatečný pro spolehlivé určení plemenné příslušnosti?
- Zvýší se diskriminační schopnost plemenné příslušnosti u metody zařazením jiných markerů kromě mikrosatelitů?

#### **Závěr:**

Předložená kvalitní disertační práce prokazuje, že Mgr. Jan Říha má velmi dobré znalosti v oboru informatiky a její aplikace v populační genetice, zvládá práci s vědeckou literaturou a je schopen realizovat a vyhodnotit náročné vědecké experimenty. Výsledky jeho práce rozšiřují možnosti aplikace poznatků molekulární genetiky ve šlechtění zvířat, mají význam pro rozvoj oboru a jsou následně východiskem pro další experimentální práci. Autor výsledky své práce publikoval v impaktovaném tisku. Kvalita doktorské disertační práce Mgr. Říhy též svědčí o velmi dobrých studijních a pracovních podmínkách na obou pracovištích a zároveň se projevuje kvalitní vedení doktoranda oběma školiteli, prof. Čítkem a prof. Dvořákem. Na základě všech výše uvedených skutečností souhlasím, aby disertační práce Mgr. Jana Říhy byla přijata k obhajobě a po jejím úspěšném obhájení byl Mgr. Říhovi udělen akademický titul „Doktor“ (PhD).

V Praze 4.9.2012

  
doc. Ing. Ivan Majzlík, CSc  
ČZU v Praze

## Oponentský posudek doktorské disertační práce

**Mgr. Jana Ř Í H Y,**

studenta doktorského studijního programu Biotechnologie, oboru Zemědělské biotechnologie, na Zemědělské fakultě Jihočeské univerzity v Českých Budějovicích, s názvem:

### **Microsatellite Markers and Genetic Diversity Issues in Cattle**

**(Studium mikrosatelitních markerů v úlohách genetické diverzity skotu)**

Oponent: prof. Ing. Jiří Kuciel, CSc.,  
profesor oboru: Genetika živočichů.

#### **Aktuálnost zadaného tématu disertační práce**

Současný intenzivní rozvoj metod molekulární genetiky a možnosti genotypování jednotlivých zvířat přináší nutnost využití aplikace nových přístupů s manipulací a ve vyhodnocení získaných dat. Pro jejich efektivní použití je nutná tvorba nové softwarové aplikace na základě algoritmů tzv. strojového učení. Těmito otázkami se zabývá předložená disertační práce.

Poznatky a určení genetické dynamiky různých genotypů u různých plemen jsou nezbytné pro zamezení nevhodných změn ve snížení genetické variability různých populací zvláště u skotu. Tyto otázky jsou předmětem výzkumu mnoha vědeckých týmů v různých zemích a je nutno je vyhodnotit i v našich chovatelských podmínkách. Studium genetické rozdílnosti skotu je významným tématem vědeckého výzkumu v důsledku širokého používání inseminace ve všech chovatelsky vyspělých státech, v možnosti nakupování a v rychlém transportu inseminačních dávek po celém světě, což do značné míry ovlivňuje genetické složení různých populací skotu. Poznatky a určení genetické dynamiky pomocí mikrosatelitních lokusů různých plemen jsou nutné pro zamezení nevhodných změn ve snížení genetické vzdálenosti různých populací.

Zadané téma disertační práce považují za aktuální a vhodné k řešení jako podklad k disertačnímu řízení.

## **Splnění sledovaných cílů a zpracování disertační práce**

Cíle disertační práce autor popsal v úvodu a byly zaměřeny na:

- Popis genetické diverzity a charakteristiky vybraných subpopulací různých plemen skotu chovaných v České republice
- Ocenění platnosti testování paternity skotu podle mikrosatelitních lokusů u vybraných plemen skotu
- Získání genotypových dat a tvorba softwarové aplikace k možnosti jejího využití ve specifických podmínkách
- Rozlišení jedinců podle genetického profilu mikrosatelitů na úrovni plemene, za pomoci algoritmů tzv. strojového učení.

Autor k přípravě a zpracování disertační práce přistoupil na základě studia a citace 155 většinou zahraničních literárních publikací. Všechny práce zahrnuté do seznamu použité literatury, převážně citované dle požadované normy plně využil v přehledu dosavadních poznatků a v diskusi k vlastním výsledkům. Metodické postupy odpovídaly vytýčeným cílům. Výsledky uvedl v tabulkách a v přehledném grafickém vyjádření, které diskutoval s dosud známými poznatky. Závěr práce byl napsán nezvykle příliš rozsáhle na čtyřech stranách a měl být konkrétní. Text disertace byl doplněn o praktické využití výsledků a o teoretická východiska dalšího výzkumu, bez uplatnění diskuse.

Po prostudování anglického rukopisu publikace mohu konstatovat, že disertant **s p l n i l** zadání a vytýčené cíle disertační práce.

## **Výsledky a nové poznatky pro rozvoj vědního poznání**

Autor analyzoval genetickou variabilitu na základě 10-ti mikrosatelitních markerů, které doporučuje ISAG/FAO pro určení genetického typu skotu. Genetickou diverzitu popsal u deseti subpopulací plemen skotu schovaných v našich podmínkách. Určil osum genetických distancí založených na různých teoretických modelech algoritmů a genetických stromů. Pomocí kombinované pravděpodobnosti a hodnoty informačního obsahu vyhodnotil využití sledované sady mikrosatelitů při vylučování paternity u jednotlivých plemen. Navrhl algoritmy a způsob implementace softwarové aplikace pro hodnocení velkých datových sad, pro účely jejich využití. Pro hodnocení plemenného rozlišení využil deset algoritmů strojového učení a diskutoval jejich použití.

## **Připomínky k práci a náměty do diskuse při obhajobě**

### **Formální připomínky k textu:**

- Zjištěné nepřesnosti jsem v textu práce vyznačil a autor by je měl napravit v opravném listu, vloženém do disertační práce před obhajobou.
- V seznamu použité literatury jsou některé citace použitých publikací uvedeny neúplně a je nutné je upřesnit v opravném listu!
- Je správný překlad anglického názvu „machine learning“ do českých pojmů „strojové učení“?



- Český název disertační práce „Studium mikrosatelitních markerů v úlohách genetické diverzity skotu“ není zcela přesný a pro lepší informovanost navrhuji změnu českého názvu na „Studium mikrosatelitních markerů ve výstupech genetické diverzity skotu“.
- Příložený seznam publikací autora měl obsahovat práce týkající se jen řešeného tématu disertačního spisu a navíc není zpracován dle požadovaných norem citace. Rovněž seznam řešených výzkumných úkolů nesouvisí s tématem zadané disertační práce a nebyl důvod jej zde uvádět.

### Náměty do diskuse při obhajobě

- Neovlivnily různé počty hodnocených jedinců v subpopulacích skotu zjištěné výsledky genetické variability mikrosatelitů?
- Byly některé výsledky uvedené v disertační práci publikované ve vědeckém časopisu s impaktním faktorem?
- Na základě kterých principů vznikají rozdílné hodnoty v genetické diverzitě v jednotlivých lokusech mikrosatelitů?
- Byly teoretické modely zjišťovaných distancí porovnány s běžně používanými postupy?
- Pro jaké další výzkumné zaměření v genetice populací lze využít algoritmy tzv. strojového učení v analýze haplotypů a polygenní genetické informace?
- Při obhajobě by měl doktorand diskutovat a upřesnit využití výsledků své práce a zaměření dalšího výzkumu v této oblasti.

### Z á v ě r

Doktorská disertační práce **Mgr. Jan Ř í h y** byla zpracována na základě studia převážně zahraničních publikací. Předložená práce v anglickém jazyku byla vypracována a popsána velmi obsáhle a tím je často méně přehledná. Výsledky práce byly zjištěny za pomoci objektivních genetických metod biometrického vyhodnocení, ale diskuse k nim měla být konkrétní. Práce splňuje požadavky kladené na spis tohoto typu a je dokladem o schopnosti autora samostatně vědecky pracovat.

D o p o r u č u j i přijetí doktorské disertační práce k obhajobě a po úspěšné diskusi udělit Mgr. Janu Ř í h o v i akademický titul

„Doktor“ (Ph.D.).



prof. Ing. Jiří Kuciel, CSc.

V Brně dne 31. 8. 2012