



Analýza stavu biodiverzity živočíšnych genetických zdrojov pomocou genomických údajov

Modul no. 2: Conservation and Sustainable Use of Animal Genetic Resources

Nina Moravčíková

Slovenská poľnohospodárska univerzita v Nitre

Fakulta agrobiológie a potravinových zdrojov

Nástroje pre analýzu variability genómu

Genóm:

- haploidná chromozómová sada/kompletná sekvencia DNA

Analýza DNA:

- genetické markery
- celogenómové sekvenovanie
- genotypizačné čipy



Co-funded by
the European Union

Nástroje pre analýzu variability genómu

Genetický marker

- akýkoľvek charakteristický znak alebo prejav organizmu, ktorý môže byť použitý na určenie špecifického chromozómu, bunky alebo jedinca
- gény, krátke DNA úseky, heterochromatínové oblasti chromozómov, prípadne iné prejavy genotypu, chromozómov alebo karyotypu
- polymorfný znak (variant), ktorý vykazuje mendelistickú dedičnosť a je vo vzťahu s genetickou variabilitou znaku (vlastnosti), ktorá je zo šľachtiteľského hľadiska významná



Co-funded by
the European Union

Nástroje pre analýzu variability genómu

Perspektívne markery produkčných vlastností organizmov

- kandidátske gény - ich rôzne alely a genotypy ovplyvňujú kvantitatívne znaky
- QTL - lokusy pre kvantitatívne vlastnosti



Co-funded by
the European Union

Nástroje pre analýzu variability genómu

DNA markery

- priama schopnosť detekcie alel v sekvenciách nukleotidov
- ↑ polymorfizmus
- dominantná/kodominantná dedičnosť
- častý výskyt v genóme
- ľahké a rýchle testovanie
- vysoká reprodukovateľnosť



Co-funded by
the European Union

Nástroje pre analýzu variability genómu

SNP – jednonukleotidový polymorfizmus

- vzniká zámenou nukleotidu v DNA v určitom mieste (bodová mutácia)
- výskyt v genóme: 100-300 bp →
najčastejší typ polymorfizmov DNA
- vyskytuje sa v populácii s frekvenciou $> 1\%$
- bialelický marker
- informujú o variabilite génov v rodinách, líniiach a populáciách



Co-funded by
the European Union

Nástroje pre analýzu variability genómu

Celogenómové sekvenovanie

- sekvenovanie DNA je stanovenie presného poradia jednotlivých nukleotidov v reťazci molekuly DNA, teda určenie primárnej štruktúry DNA
 - klasické metódy sekvenovania: Maxam-Gilbertová a Sangerová metóda sekvenovania
 - nové sekvenačné technológie – sekvenovanie novej generácie (NGS): Jedná sa o viacero platforiem (Illumina, Ion Torrent, 454 a ďalšie), ktoré sa líšia v technickom prístupe k sekvenovaniu, ale prinášajú porovnateľné výstupy



Co-funded by
the European Union

Nástroje pre analýzu variability genómu

SNP čipy

- umožňujú analyzovať súčasne niekoľko tisíc SNP markerov rovnomerne distribuovaných v genóme
- dostupné pre väčšinu druhov hospodárskych a spoločenských zvierat (hovädzí dobytok, kone, ovce, kozy, ošípané, hydina, psy alebo mačky)
- v súčasnosti celosvetovo najpopulárnejší nástroj pre testovanie napr.:
 - analýza rodičovstva
 - stavu genomickej diverzity
 - predispozície jedincov a populácií
 - celogenómové asociačné štúdie
 - odhad genomických plemenných hodnôt
 -



Co-funded by
the European Union

Homozygotnosť genómu a genomický inbríding

- **autozygotnosť genómu** – všetky alely alebo chromozomálne úseky DNA, ktoré sú identické podľa pôvodu (IBD) → odvodené od spoločného predka
- **homozygotné úseky (ROH)** – genomické oblasti so špecifickým počtom za sebou idúcich homozygotných SNP markerov
- vplyv prírodnej a zámernej selekcie, rekombinácie, väzbovej nerovnováhy, mutácií....
- dĺžka ROH úsekov v genóme korešponduje so vzdialenosťou predkov v rodokmeni jedinca



Co-funded by
the European Union

Homozygotnosť genómu a genomický inbríding

Využitie skrínungu ROH úsekov:

odhad genomického inbrídingu na základe podielu homozygotných úsekov v autozomálnom genóme

testovanie vplyvu sekcie na genóm

identifikácia kauzálnych mutácií podieľajúcich sa na kontrole preferovaných znakov a vlastností



Co-funded by
the European Union

Homozygotnosť genómu a genomický inbríding

Odhad koeficienta inbrídingu (McQuillan et al., 2008):

$$F_{ROH} = \frac{\sum L_{ROH}}{L_{autosome}}$$

kde $\sum L_{ROH}$ vyjadruje celkovú dĺžku ROH úsekov v genóme jedinca obsahujúceho špecifický počet za sebou idúcich homozygotných SNP markerov a $L_{autosome}$ je dĺžka autozomálneho genómu odvodená na základe fyzickej pozície testovaných SNP markerov

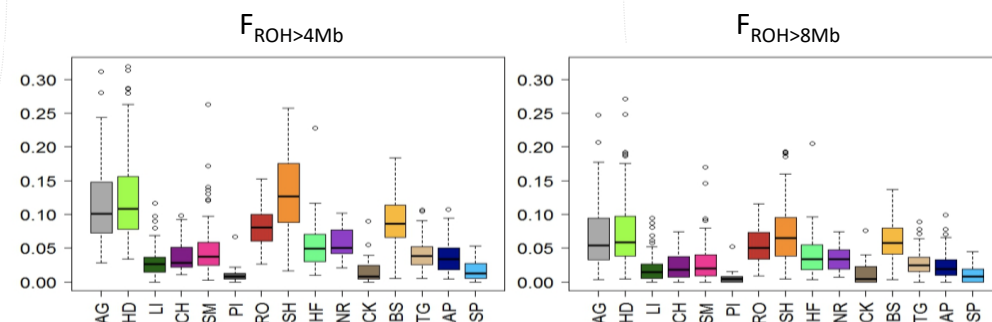
- F_{ROH} : od 0 po 1 (0 po 100%)
- v malých populáciách $\Delta F > 1\%$ (4 % vo veľkých populáciách) môže viesť k signifikantnej strate diverzity a môže ovplyvniť ich dlhodobú prežiteľnosť

Najčastejšie využívané programy:

PLINK v1.9

detectRUNS (program R)

cgaTOH



Genomický inbríding odvodený od homozygotných úsekov na základe počtu generácií od spoločného predka ($F_{ROH>4Mb}$ – 12-13 generácií a $F_{ROH>8Mb}$ – 6-7 generácií) u 15 európskych plemien dobytky (Kukučková et al., 2017)

McQUILLAN, R. et al. 2008. Runs of homozygosity in European populations. In *Am J Hum Genet*, vol. 83, pp. 359-372.

KUKUČKOVÁ, Veronika - MORAVČÍKOVÁ, Nina - FERENČAKOVIČ, Maja - SIMČIČ, Mojca - MÉSZÁROS, Gábor - SÖLKNER, Johann - TRAKOVICKÁ, Anna - KADLEČÍK, Ondrej - CURIK, Ino - KASARDA, Radovan. Genomic characterization of Pinzgau cattle: genetic conservation and breeding perspectives. In *Conservation Genetics*. ISSN 1566-0621, 2017, vol. 18, no. 4, s. 893-910.

Väzbová nerovnováha a efektívna veľkosť populácií ŽGZ

Väzbová nerovnováha (LD)

- nenáhodná asociácia alel v rôznych lokusoch v danej populácii
-
- vplyv rôznych faktorov: selekcia, rekombinácia, mutácie, genetický drift, systém priparovania
- využitie LD: analýza evolučnej histórie populácií, vplyv dopadu selekcie na genóm, odhad efektívnej veľkosti populácií



Co-funded by
the European Union

Väzbová nerovnováha a efektívna veľkosť populácií ŽGZ

Väzbová nerovnováha (LD)

- najčastejšie sa vyjadruje pomocou Pearsonovho korelačného koeficienta súčinu momentov (Hill a Robertson, 1968):

$$r_{LD}^2 = \frac{(p_{AB} - p_A p_B)^2}{p_A(1 - p_A)p_B(1 - p_B)}$$

kde p_{AB} je rozdiel medzi frekvenciou gamét nesúcich alely A a B na dvoch lokusoch a p_A a p_B sú výsledné frekvencie týchto alel

- nadobúda hodnoty od 0 (väzbová rovnováha) po 1 (kompletná väzbová nerovnováha)

HILL, W.G. – ROBERTSON, A. 1968. Linkage disequilibrium in finite populations. In *Theor Appl Genet*, vol. 38, pp. 226-231.

Väzbová nerovnováha a efektívna veľkosť populácií ŽGZ

Efektívna veľkosť populácie (N_e)

- vzťah medzi variabilitou LD a N_e sa najčastejšie používa na odhad historickej, ale aj súčasnej efektívnej veľkosti populácie → N_{eLD}
- N_{eLD} reálnej populácie X s pozorovanou LD pre určitý úsek na genóme (dĺžkový interval v Mb alebo kb) je definovaná ako veľkosť hypotetickej ideálnej populácie, ktorá má v rovnovážnom stave rovnakú úroveň LD v danom úseku, aká je pozorovaná v skutočnej uzavretej populácii



Co-funded by
the European Union

Väzbová nerovnováha a efektívna veľkosť populácií ŽGZ

Efektívna veľkosť populácie (N_e)

- možno ju vypočítať nasledovne (Corbin et al., 2012):

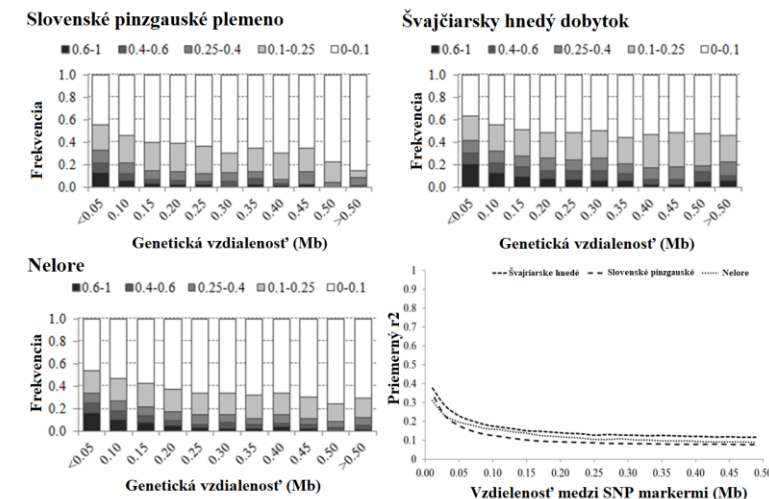
$$N_{eLD} = \frac{1}{kc} \left[\frac{1}{r_{LD}^2 - \frac{1}{n_g}} - \alpha \right]$$

kde k vyjadruje model dedivosti ($k = 4$ v prípade autozomálnych lokusov), c je fyzická vzdialenosť medzi SNP markermi v Morganoch, n_g je veľkosť vzorky gamét (dvojnásobok počtu testovaných jedincov) a α je faktor zohľadňujúci pôsobenie mutácií

- historická N_{eLD} : funkcia času a fyzickej vzdialenosti medzi dvoma markermi za predpokladu konštantného lineárneho rastu N_{eLD} s časom vyjadreným minulými generáciami
- môže nadobúdať hodnoty od 0 po n
- najčastejšie využívané programy: LDNE, SneP, GONE

CORBIN, L.J. et al. 2012. Estimation of historical effective population size using linkage disequilibrium with marker data. In *J Anim Breed Genet*, vol. 129, pp. 257-270.

KASARDA, Radovan - MORAVČÍKOVÁ, Nina - MESZÁROŠ, G. - TRAKOVICKÁ, Anna - KADLEČÍK, Ondrej. Genetic divergence of cattle populations based on genomic information. In *Scientia agriculturae bohemica*. ISSN 1211-3174, 2016, vol. 47, no. 3, s. 113-117.



Porovnanie podielu markerov v rámci piatich kategórií r_{LD}^2 a historická efektívna veľkosť populácie (Kasarda et al., 2016)



Co-funded by
the European Union

Genetická štruktúra populácií

- rôznorodosť populácií na vnútro aj medzipopulačnej úrovni
- genetická vzdialenosť - stupeň génovej rozdielnosti (genomickej rozdielnosti napr. vo frekvenciách alel) medzi druhmi alebo populáciami, ktorú je možné kvantifikovať na základe rôznych štatistických prístupov
- Najčastejšie využívané parametre:
 - Neiove genetické vzdialenosť a Wrightov F_{ST} index
 - analýza hlavných komponentov
 - Bayesianská analýza stupňa genetického premiešania/toku génov medzi populáciami



Co-funded by
the European Union

Genetická štruktúra populácií

Neiove genetické vzdialenosti

- predpokladá, že ak sú dve populácie vykazujúce nízke genetické vzdialenosti podobné, tak s vysokou mierou spoľahlivosti majú spoločných predkov
- možno ju vypočítať (Nei, 1972):

$$D = -\ln I ; I = \frac{J_{xy}}{\sqrt{J_x J_y}}$$

kde I je normalizovaná zhoda génov (alebo genetická zhoda) medzi populáciou X a Y

- nadobúda hodnoty od 0 (populácie/jedinci sú geneticky identické) po 1 (populácie /jedinci sú geneticky rozdielne)
- na výpočet možno použiť program R (balíky StAMPP, Poppr a i.)

NEI, M. 1972. Genetic distance between populations. In Am Nat, vol. 106, pp. 283-285.



Co-funded by
the European Union

Genetická štruktúra populácií

Wrightov F_{ST} index

- indikátor intenzity fragmentácie populácie vyjadrenej poklesom heterozygotnosti v subpopuláciách v dôsledku pôsobenia genetického driftu
- možno ho vypočítať pomocou vzorca (Weir a Cockerham, 1984):

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

kde H_T vyjadruje očakávanú heterozygotnosť metapopulácie a H_S je priemerná heterozygotnosť v subpopuláciách

- nadobúda hodnoty od 0 (populácie sú geneticky identické) po 1 (populácie sú geneticky rozdielne) $\rightarrow F_{ST} > 0,25$ populácie možno považovať za geneticky diferencované
- na výpočet možno použiť program R (balíky StAMPP, Poppr a i.)

WEIR, B.S. – COCKERHAM, C.C. 1984. Estimating F-statistics for the analysis of population structure. In *Evolution*, vol. 38, pp. 1358-1370.



Co-funded by
the European Union

Genetická štruktúra populácií

Analýza hlavných komponentov

- populárna viacrozmerná štatistická metóda, ktorá našla využitie v rôznych vedných odvetviach, vrátane populačnej genetiky
- znázornenie nadrozmerných dát, napr. genomických informácií o jedincoch alebo populáciách, v menšom počte dimenzií

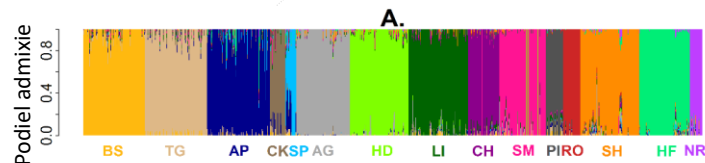
Bayesianská analýza genetického premiešania – admixie

- admixia alebo premiešanie populácií je jav, ku ktorému môže dôjsť následkom introgresie a hybridizácie jedincov, populácií alebo druhov
- Bayesianská štatistika - odvetvie modernej štatistiky, ktoré pracuje s podmienenou pravdepodobnosťou a umožňuje spresniť pravdepodobnosť východiskovej hypotézy v slede ako sa objavujú ďalšie relevantné skutočnosti

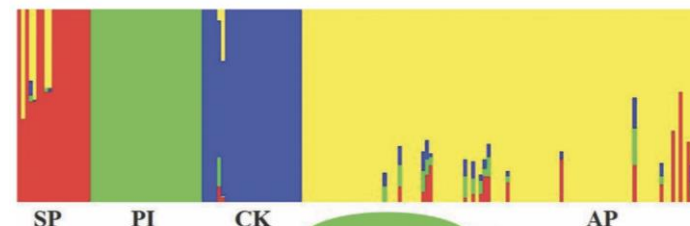
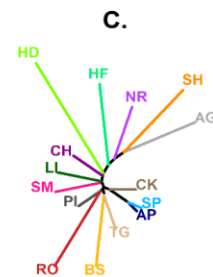
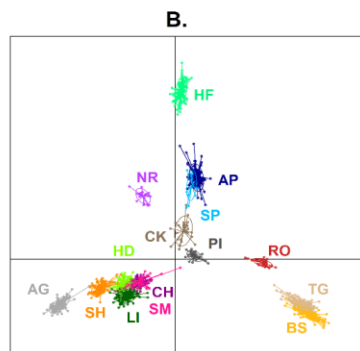


Co-funded by
the European Union

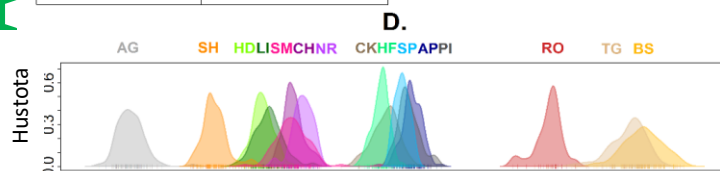
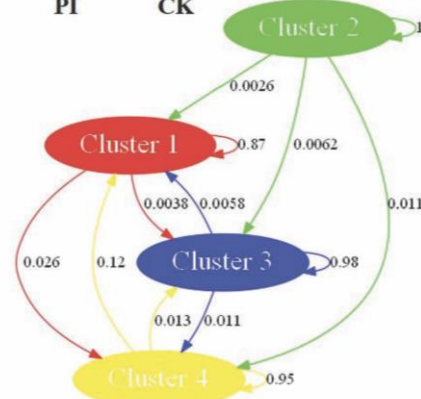
Genetická štruktúra populácií



Grafická vizualizácia populačnej štruktúry a genetických vzťahov medzi 15 európskymi plemenami hovädzieho dobytku (angus-AG, rakúsky pinzgaušský-AP, brown swiss-BS, cika-CK, hereford-HD, holštajnský-HF, charolais-CH, limousine-LI, norwegian red-NR, piedmontese-PI, romagnola-RO, shorthorn-SH, slovenské pinzgaušské-SP, simentálske-SM, tyrol grey-TG) (Kukučková et al., 2017)



Analýza admixie a toku génov medzi 4 plemenami (SP: slovenské pinzgaušské Cluster 1, PI: piedmontese Cluster 2, CK: cika Cluster 3, AP: rakúske pinzgaušské Cluster 4) (Kukučková et al., 2018)



KUKUČKOVÁ, Veronika - MORAVČÍKOVÁ, Nina - FERENČAKOVIČ, Maja - SIMČIČ, Mojca - MÉSZÁROS, Gábor - SÖLKNER, Johann - TRAKOVICKÁ, Anna - KADLEČÍK, Ondrej - CURIK, Ino - KASARDA, Radovan. Genomic characterization of Pinzgau cattle: genetic conservation and breeding perspectives. In *Conservation Genetics*. ISSN 1566-0621, 2017, vol. 18, no. 4, s. 893-910.
KUKUČKOVÁ, Veronika - MORAVČÍKOVÁ, Nina - CURIK, Ino - SIMČIČ, Mojca - MÉSZÁROS, Gábor - KASARDA, Radovan. Genetic diversity of local cattle. In *Acta Biochimica Polonica*. ISSN 0001-527X, 2018, vol. 65, iss. 3, s. 421-424.

Vplyv selekcie na štruktúru genómu

- kvantifikácia aj bez prístupu k fenotypovým informáciám
- identifikácia tzv. selekčných signálov
- závisí od intenzity selekcie, dĺžky jej pôsobenia na genóm, rekombinačnej rýchlosti...
- dva základné prístupy:
 - medzipopulačné/medziplenné rozdiely
 - vnútrapopulačné rozdiely

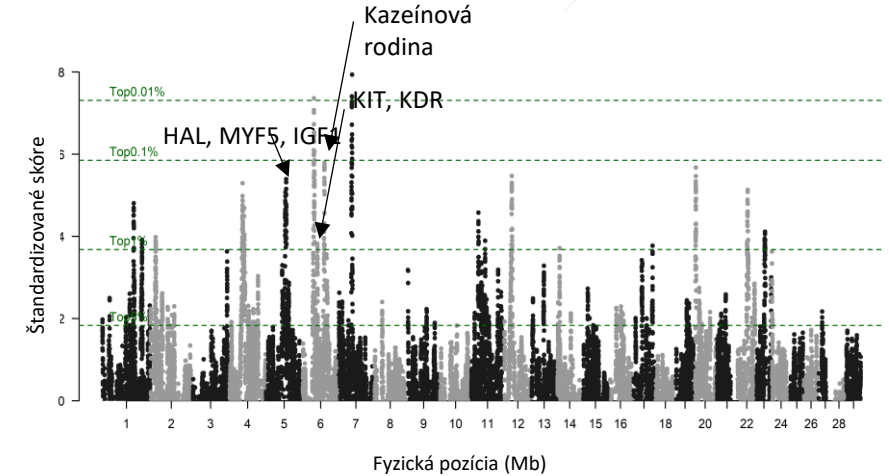


Co-funded by
the European Union

Vplyv selekcie na štruktúru genómu

Medzipopulačné/medziplenné rozdiely

- celogenómový skrining Wrightovho F_{ST} indexu
 - analýza variability vo väzbovej nerovnováhe
 - analýza hlavných komponentov
 - integrované skóre haplotypov
-
- na výpočet možno použiť programy: PLINK v1.9, varLD, pcadapt, rehh...

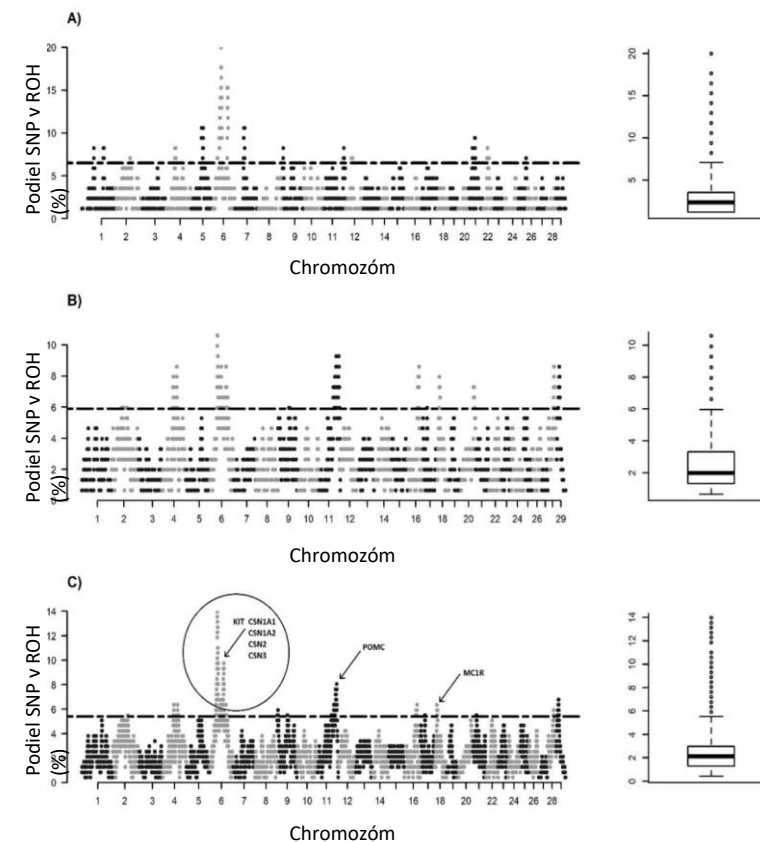


Rozdiely vo väzbovej nerovnováhe v genome slovenského pinzgauského a slovenského strakatého plemena (Moravčíková et al., 2019)

Vplyv selekcie na štruktúru genómu

Vnútropopulačné rozdiely

- distribúcia homozygotných úsekov v genóme
- úroveň väzbovej nerovnováhy
- integrované skóre haplotypov
- na výpočet možno použiť programy: PLINK v1.9, detectRUNS, rehh...



Frekvencia výskytu SNP markerov v ROH úsekoch (%) pre slovenské strakaté (A) a slovenské pinzgauské plemeno (B) a výskyt SNP markerov v ROH v rámci genómu oboch plemien (C) (Kasarda et al., 2018)

KASARDA, Radovan - MORAVČÍKOVÁ, Nina - KADLEČÍK, Ondrej - TRAKOVICKÁ, Anna - CANDRÁK, Juraj. The impact of artificial selection on runs of homozygosity in Slovak Spotted and Pinzgau cattle. In *Slovak journal of animal science*. ISSN 1337-9984, 2018, vol. 51, no. 3, s. 91-103.



Siedlce University
of Natural Sciences
and Humanities



Czech University
of Life Sciences Prague

Partneri:

Ďakujem za pozornosť!



This presentation has been supported by the Erasmus+ KA2 Cooperation Partnerships grant no. 2021-1-SK01-KA220-HED-000032068 "Innovation of the structure and content of study programs in the field of animal genetic and food resources management with the use of digitalisation - Inovácia obsahu a štruktúry študijných programov v oblasti manažmentu živočíšnych genetických a potravinových zdrojov s využitím digitalizácie". The European Commission support for the production of this presentation does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



Nina Moravčíková



nina.moravcikova@uniag.sk



Co-funded by
the European Union