Genome-Wide Association Study for Leg Disorders in Austrian Braunvieh and Fleckvieh

Barbara Kosinska-Selbi, Tomasz Suchocki, Magdalena Fraszczak, Christa Egger-Danner, Hermann Schwarzenbacher, Joanna Szyda

EAAP, 2019

UNIVERSITAS RERUM NATURALIUM WRATISLAVIENSIS

THETA
Statistical Genetics Group
Institute of Animal Genetics

ZuchtData
EDV-DIENSTLEISTUNGEN GMBH

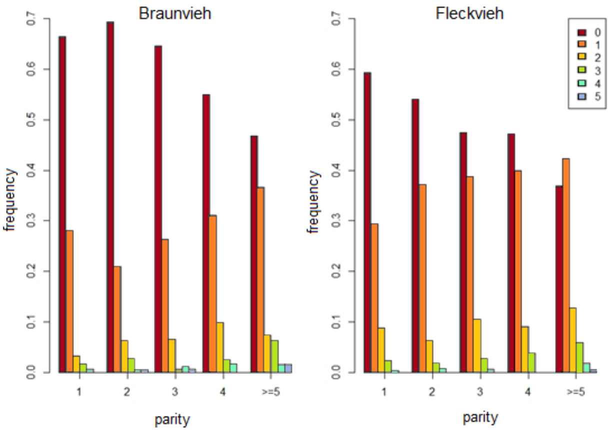
Objectives

Mining the **genetic heterogeneity** between Braunvieh and Fleckvieh

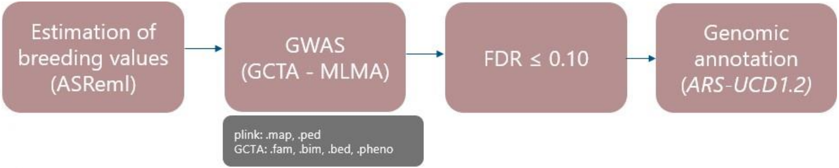
Material

1. Cows
    - 1 999 Fleckvieh (FLV), 985 Braunvieh (BSW)
2. SNPs
    - Geneseek Genomic Profiler HD BeadChip
    - 76 932 SNPs
    - 74 762 SNPs (MAF 0.01; call rate 99 %)
3. Phenotype
    - total number of leg disorders until DIM 300^{th}
    - EBV

Material


Braunvieh
Fleckvieh
frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
1
2
3
4
>=5
parity
0
1
2
3
4
5

Methods/genome-wide association study (GWAS)



Estimation of breeding values (ASReml)

GWAS (GCTA - MLMA)

FDR ≤ 0.10

Genomic annotation (*ARS-UCD1.2*)

plink: .map, .ped
GCTA: .fam, .bim, .bed, .pheno

Methods/EBV prediction

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z}_u\boldsymbol{u} + \boldsymbol{Z}_v\boldsymbol{v} + \boldsymbol{Z}_p\boldsymbol{p} + \boldsymbol{\varepsilon}$$

$\boldsymbol{y}$	total number of leg disorders scored till DIM 300
$\boldsymbol{\beta}$	fixed effects : a general mean, breed (Braunvieh or Fleckvieh), parity (1,2,3,4 or >4), calving year-season, hoof status recording code (four levels)
$\boldsymbol{u}$	$\sim N(0, \boldsymbol{A}\sigma_u^2)$ EBV
$\boldsymbol{v}$	$\sim N(0, \boldsymbol{I}\sigma_v^2)$ random veterinarian effect
$\boldsymbol{p}$	$\sim N(0, \boldsymbol{I}\sigma_p^2)$ random permanent environmental effect
$\boldsymbol{\varepsilon}$	$\sim N(0, \boldsymbol{I}\sigma_\epsilon^2)$ residual

Methods/GWAS model

$$\boldsymbol{u} = \mu + \boldsymbol{X}\boldsymbol{b} + \boldsymbol{Z}\boldsymbol{g} + \boldsymbol{e}$$

$\boldsymbol{u}$ EBV

$\boldsymbol{\mu}$ general mean

$\boldsymbol{b}$ additive effect of all single SNP

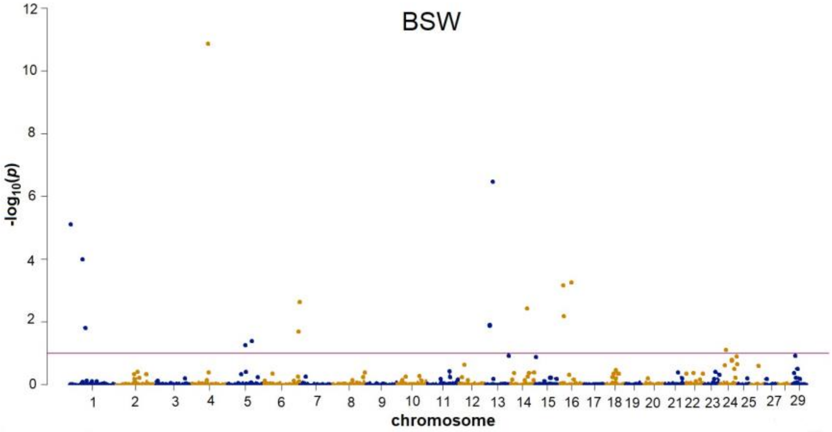
$\boldsymbol{X}$ design matrix {0, 1, 2}

$\boldsymbol{g} \sim N\left(0, \boldsymbol{G}\sigma_g^2\right)$ cumulated effect of all remaining SNPs,

$\boldsymbol{G}$ corresponds to the genomic covariance matrix between cows calculated based on SNPs.

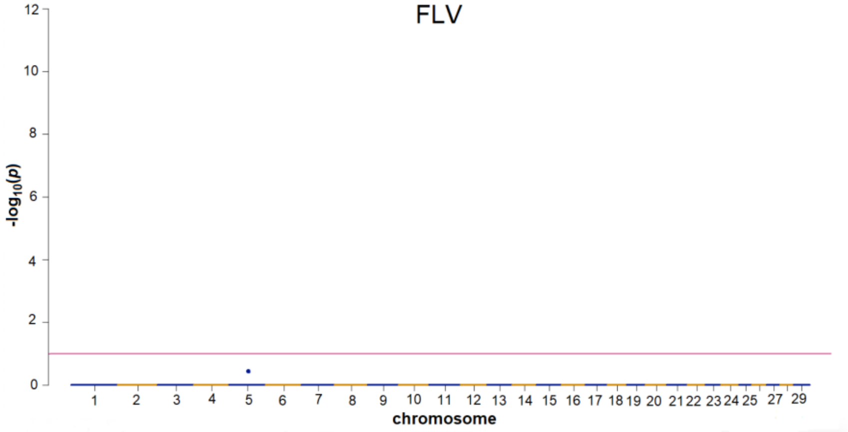
$\boldsymbol{e} \sim N(0, \boldsymbol{I}\sigma_e^2)$ residual

Results/GWAS



EAAP, 2019

Results/GWAS



Results/GWAS

SNP	Position [bp]**	Additive effect	FDR*	Annotation**
ARS-BFGL-NGS-92033	1:3,303,269	0.018	$7.5 \cdot 10^{-6}$	intergenic, between MIS18A and HUNK
BTA-89698-no-rs	1:43,542,488	0.012	0.0001	intergenic, between DCBLD2 and COL8A1
ARS-BFGL-NGS-6521	1:50,767,507	0.007	0.0150	intergenic, between CBLB and CCDC54
BovineHD0400014448	4:52,028,036	0.029	$1.4 \cdot 10^{-11}$	intergenic, between **CAV2** and TES
BovineHD0400014458	4:52,079,221	0.029	$1.4 \cdot 10^{-11}$	
Hapmap48066-BTA-73690	5:61,220,624	0.010	0.0530	intergenic, closest to NEDD1
ARS-BFGL-NGS-85328	5:81,769,685	0.008	0.0430	intergenic, between CCDC91 and **PTHLH**
ARS-BFGL-NGS-103113	6:115,208,599	0.016	0.0210	intergenic, between ADRA2C and LRPAP1
ARS-BFGL-NGS-100768	6:114,116,280	0.010	0.0020	intron of SORCS2
ARS-BFGL-NGS-25175	13:13,590,662	0.008	0.0120	intergenic, closest to CELF2
ARS-BFGL-NGS-101509	13:23,590,146	0.019	$3.3 \cdot 10^{-7}$	intergenic, between SPAG6 and PIP4K2A
ARS-BFGL-NGS-63852	14:55,768,446	0.010	0.0040	intergenic, between TMEM74 and EMC2
Hapmap55901-rs29024589	16:12,125,227	0.010	0.0007	intergenic between B3GALT2 and GLRX2
ARS-BFGL-NGS-109246	16:12,280,122	0.008	0.0060	intergenic, between UCHL5 and RGS2
Hapmap51828-BTA-38538	16:36,037,389	0.007	0.0006	intergenic, between RGS7 and XCL1
BovineHD2400006669	24:24,273,191	0.002	0.0820	intergenic, between CCDC178 and KLHL14

*False Discovery Rate

**ARS UCD1.2.assembly

Methods/principal components analysis

genotypes

PCA
Matrix size
50x50 SNPs

Separation of
BSW and FLV

Methods/Local principal components analysis between breeds/ Machalanobis distance

$$\boldsymbol{D_M}=\sqrt{\boldsymbol{d'V^{-1}d}}$$

$$\boldsymbol{d} = [\ \bar{\varepsilon}_{1BSW} - \bar{\varepsilon}_{1FLV}, \bar{\varepsilon}_{2BSW} - \bar{\varepsilon}_{2FLV}, \dots, \bar{\varepsilon}_{10BSW} - \bar{\varepsilon}_{10FLV}\ ]$$

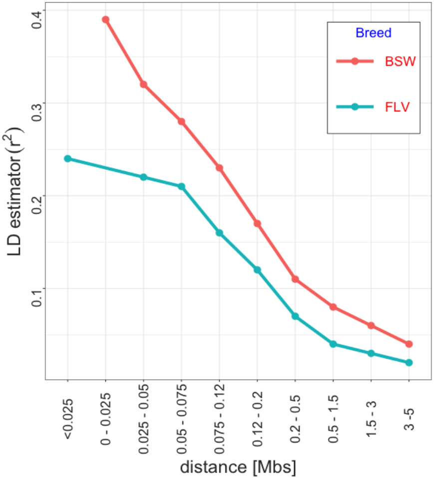
$\bar{\varepsilon}_1, \bar{\varepsilon}_2, \dots, \bar{\varepsilon}_{10}$ - eigenvectors

$\boldsymbol{V}$ - pooled covariance matrix of all eigenvectors

$$\boldsymbol{T} = \frac{n_B n_F}{n_B + n_F} \cdot \frac{n_B + n_F - 11}{10(n_B + n_F - 2)} \cdot \boldsymbol{d'V^{-1}d} \sim F_{10, n_B + n_F - 11}$$

n_x number of cows representing each breed

Comparison of LD patterns



Methods/Differences in pairwise LD structure between breeds

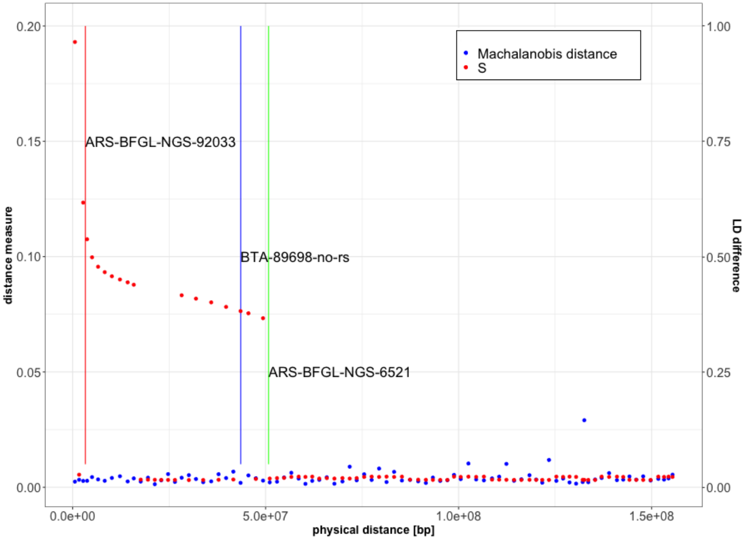
$$\boldsymbol{S} = \left[\sum_{i=1}^{n}[(\boldsymbol{v}_{i21} + \boldsymbol{v}_{i22}) - (\boldsymbol{v}_{i12} + \boldsymbol{v}_{i21})]^2\right] + \left[\sum_{i=1}^{n}[(\boldsymbol{v}_{i11} + \boldsymbol{v}_{i12}) - (\boldsymbol{v}_{i21} + \boldsymbol{v}_{i22})]^2\right]$$

$\boldsymbol{S}$ – general meausure of variability

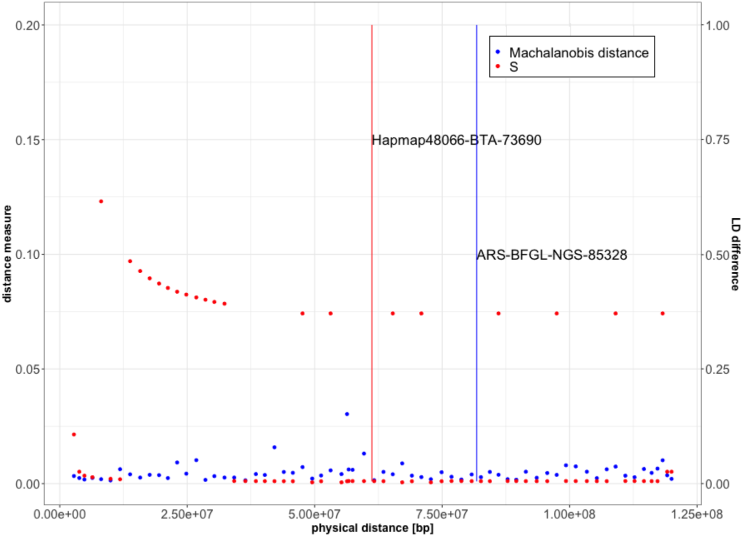
$\boldsymbol{v}_{ij}$ – product of linkage disequilibrium covariance matrix of breed$_{\mathrm{i}}$
and the vector of eigenvectors corresponding to breed$_{\mathrm{j}}$

(Garcia, 2012)

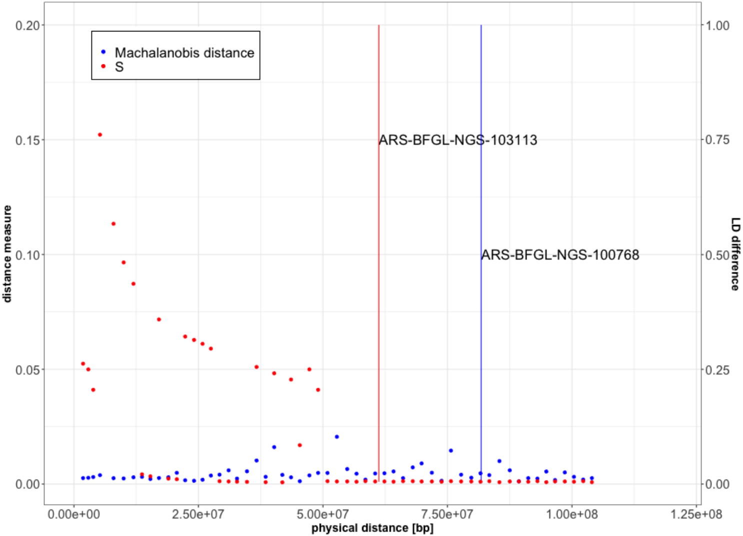
Results/PCA/BTA1



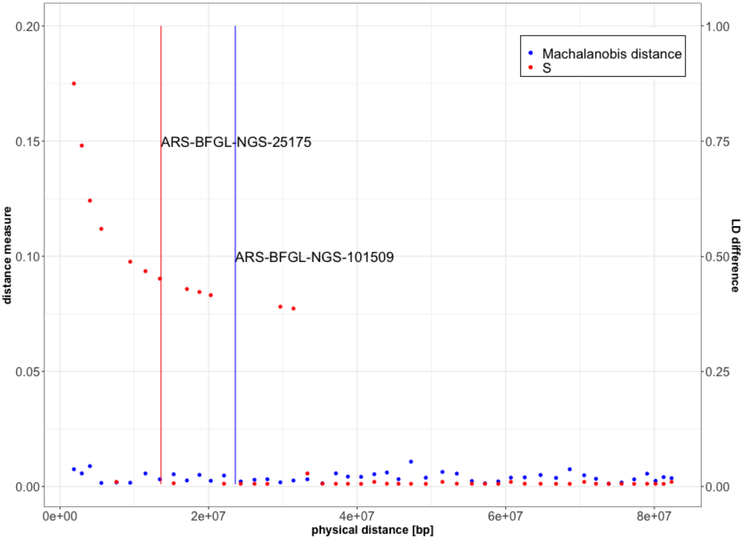
Results/PCA/BTA5



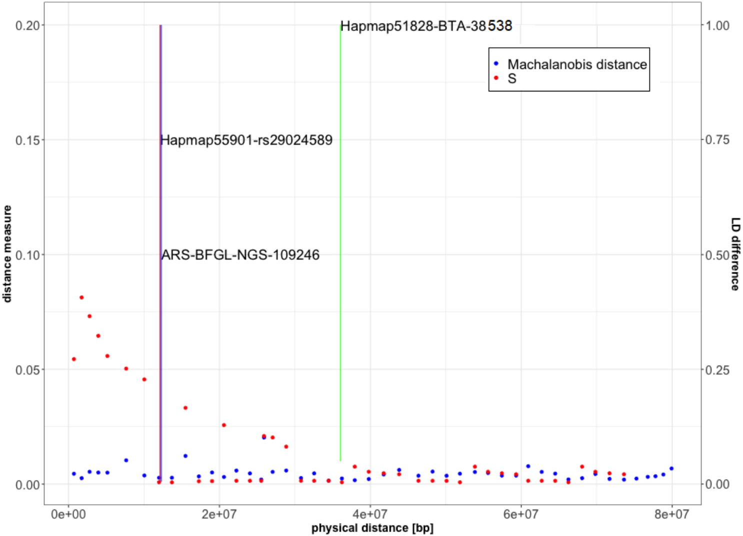
Results/PCA/BTA6



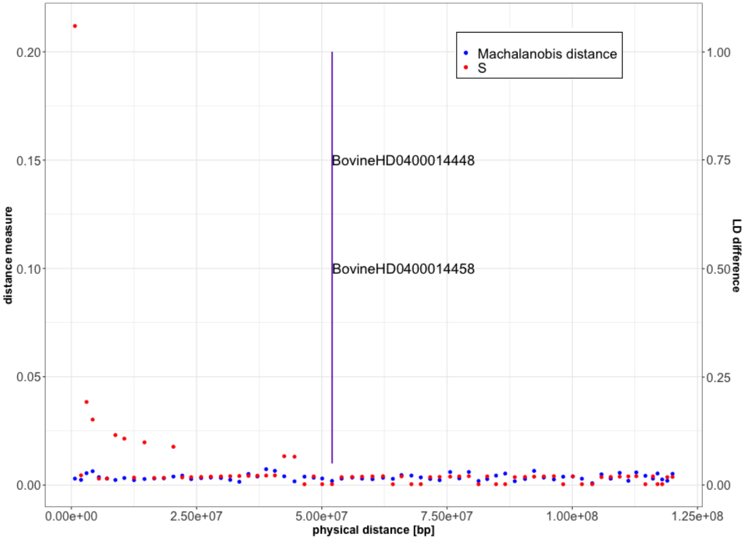
Results/PCA/BTA13



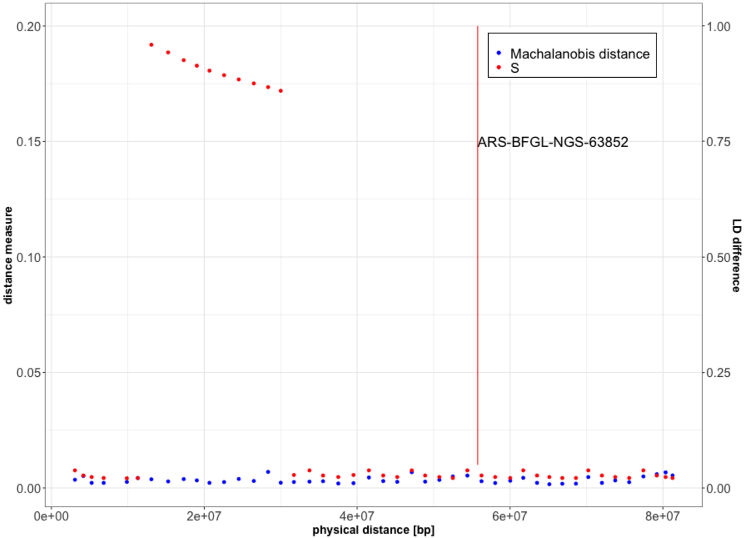
Results/PCA/BTA16



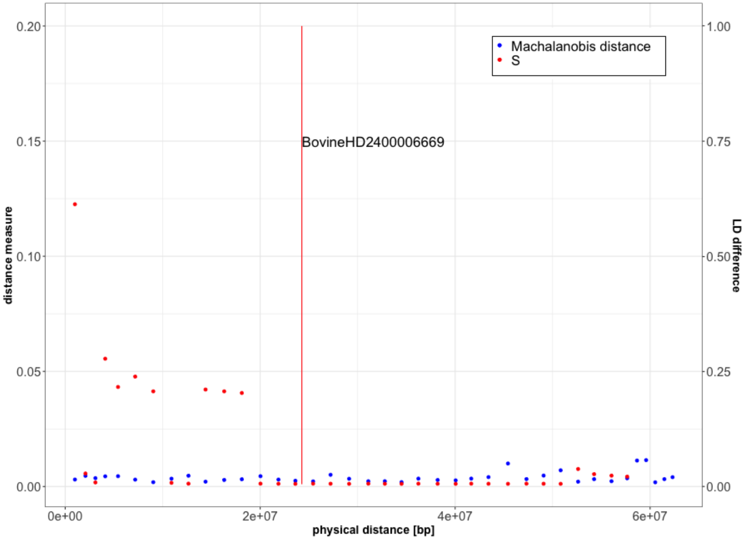
Results/PCA/BTA4



Results/PCA/BTA14



Results/PCA/BTA24



Conclusions

- Genetic heterogeneity in the number of leg disorders between Fleckvieh and Braunvieh
- Differences in LD structure between Fleckvieh and Braunvieh in some genomic regions
- Differences in LD structure partially explain genetic heterogeneity

Acknowledgements

- NCN grant No. 2015/19/B/NZ9/03725 as well as by the Efficient Cow and the Gene2Farm projects
- Biostatistics group, Department of Genetics, Wroclaw University of Environmental and Life Sciences
  - Magdalena Fraszczak
  - Joanna Szyda
  - Tomasz Suchocki
- ZuchtData, Vienna, Austria
  - Christa Egger-Danner
  - Hermann Schwarzenbacher