A 2-step strategy to infer genome-wide associations for endoparasite traits in local DSN cattle

Katharina May[1,2], Carsten Scheper[1], Kerstin Brügemann[1], Tong Yin[1], Christina Strube[2], Paula Korkuć[3], Gudrun A. Brockmann[3], Sven König[1]

[1]**Institute of Animal Breeding and Genetics, Justus-Liebig-University of Gießen, 35390 Gießen, Germany,**
[2]**Institute for Parasitology, University of Veterinary Medicine Hannover, Buenteweg 17, 30559 Hanover, Germany,**
[3]**Department for Animal Breeding Biology and Molecular Genetics, Faculty of Live Science, Humboldt-Universität of Berlin, 10115 Berlin, Germany**

Endoparasite infections in dairy cows

- Infections with endoparasites lead to reduced milk production and detrimental impacts on fertility
  (e.g., Charlier et al., 2005; Dank et al., 2015)

    → high economic losses

- Increase of anthelmintic resistance

    (e.g., Gasbarre et al., 2014; Sutherland et al., 2011)

- Increasing interest in local breeds being best adapted to various environmental conditions

JLU

Why genomic studies in endangered DSN cattle?

- „*Deutsches Schwarzbuntes Niederungsrind* " /Black Pied cattle
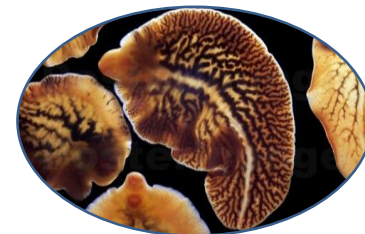
- One of the founding breeds of Holstein Friesian

Better fertility
and productive life
(Biedermann et al., 2005)

More resistant to metabolic
disease (Jaeger et al., 2016)

More robust under harsh
environmental conditions
(Al-Kanaan et al., 2016)

Less susceptibility to endoparasite infections?

Background and preliminary work

- Faecal examination of 1166 dairy cows (2006 observations),17 herds

- 4 different Black and White cattle selection lines

- HF milk, HF pasture, HF New Zealand, dual-purpose cows (DSN)

1. Quantitative-genetic parameters for endoparasite traits:

	Endoparasite trait	$h^2 \pm SE$	
	Gastrointestinal nematodes (GIN)	$0.05 - 0.06 \pm 0.04$	$r_g = 1.00$
	Lungworm larvae (*Dictyocaulus viviparus*)	0.05 ± 0.04	$r_g = -0.10$
	Liver fluke (*Fasciola hepatica*)	0.33 ± 0.06	$r_g = 0.03$

Genomic scales for endoparasite traits

2. Genome-wide association study (GWAS)
for endoparasite traits in DSN

a) Identification of SNP markers being associated with endoparasite resistance

b) Analysis of potential candidate genes
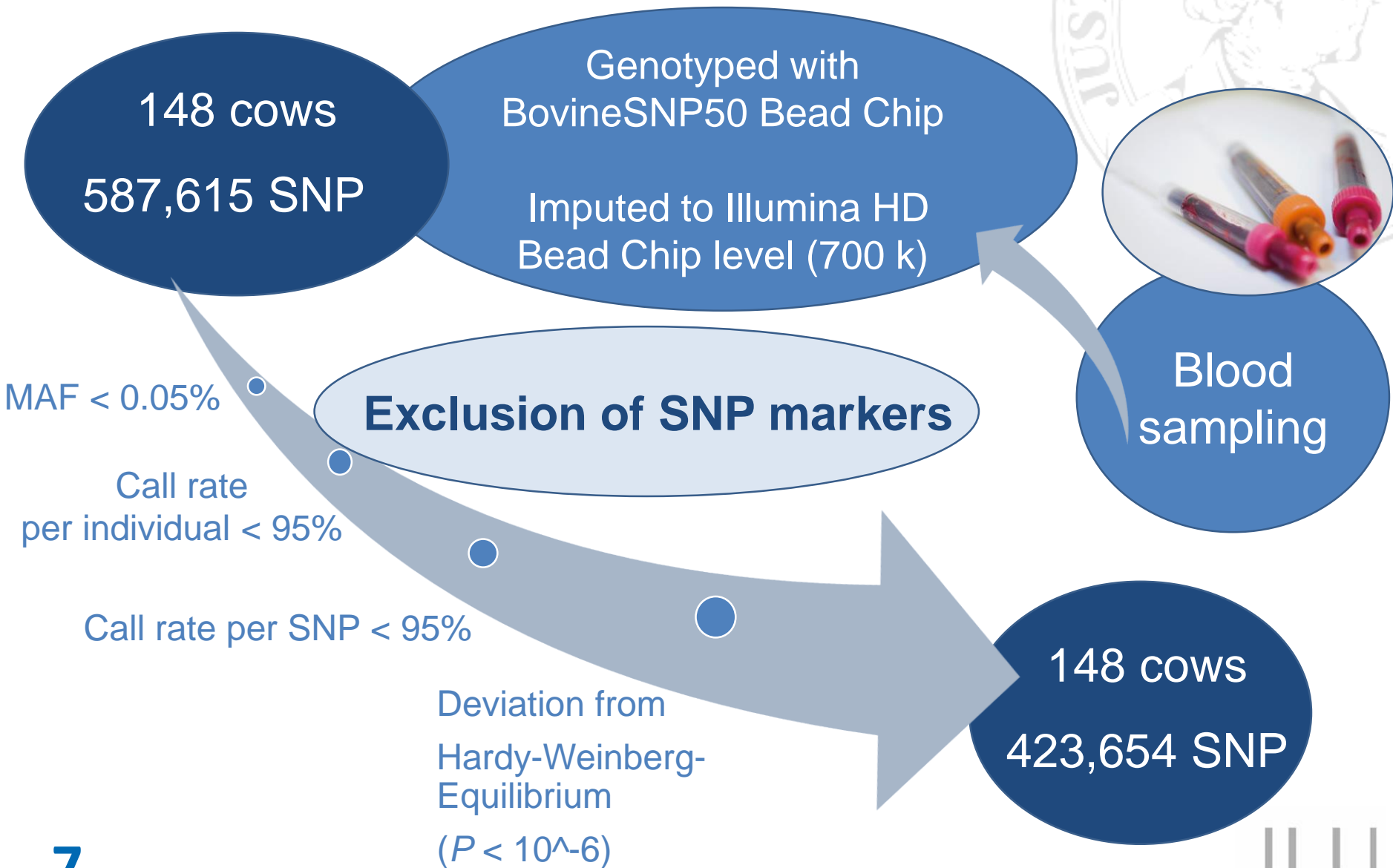
c) Pathway analyses

Selective genotyping approach

- Collection of blood samples from 148 **dual-purpose DSN cows**

- 3 herds

Selection criteria:

1. Herd prevalence for gastrointestinal nematodes

2. Selection of most resistant and most susceptible cows within herds

   - 50% of cows being infected

   - 50% being not infected

3. Minimization of genetic relationships within and between groups

JLU

Genotyping, Imputing and Quality control

148 cows

587,615 SNP

Genotyped with BovineSNP50 Bead Chip

Imputed to Illumina HD Bead Chip level (700 k)

Blood sampling

Exclusion of SNP markers

MAF < 0.05%

Call rate per individual < 95%

Call rate per SNP < 95%

Deviation from Hardy-Weinberg-Equilibrium ($P < 10^{-6}$)

148 cows

423,654 SNP

Software PLINK, Purcell et al., 2007

JLU

Pre-correction of endoparasite traits

1. Calculation of residuals for endoparasite traits using the whole dataset of 1166 cows

$$y_{ijklm} = \mu + Farm_i + Parity_j + GenLin_k + SP_l + Lstage_m + e_{ijklm}$$

y_{ijklm} = observations for endoparasite traits (n = 2006)

μ = overall mean effect

$Farm_i$ = fixed effect of the ith farm

$Parity_j$ = fixed effect of the jth parity number

$GenLin_k$ = fixed effect of the kth genetic line

SP_l = fixed effect of the lth sampling period

$Lstage_m$ = fixed effect of the mth lactation stage

e_{ijklm} = random residual effect

JLU

Association analysis in GCTA (Yang et al., 2011)

2. Implementation of the GWAS in GCTA using residuals

$$y_i = a_i + SNP_i + e_i$$

y_i = corrected phenotypic values

(residuals of endoparasite traits)

a_i = random polygenetic additive-genetic effect on the basis of genomic relationship matrix

SNP_i = fixed single-locus SNP effect

e_i = random residual effect

9

Manhattan-Plot for gastrointestinal nematodes

-- Bonferroni-corrected genome-wide significance threshold ($p = 4.47 \times 10^{-7}$)

$\cdots$ Suggestive candidate threshold ($p = 1e-04$)
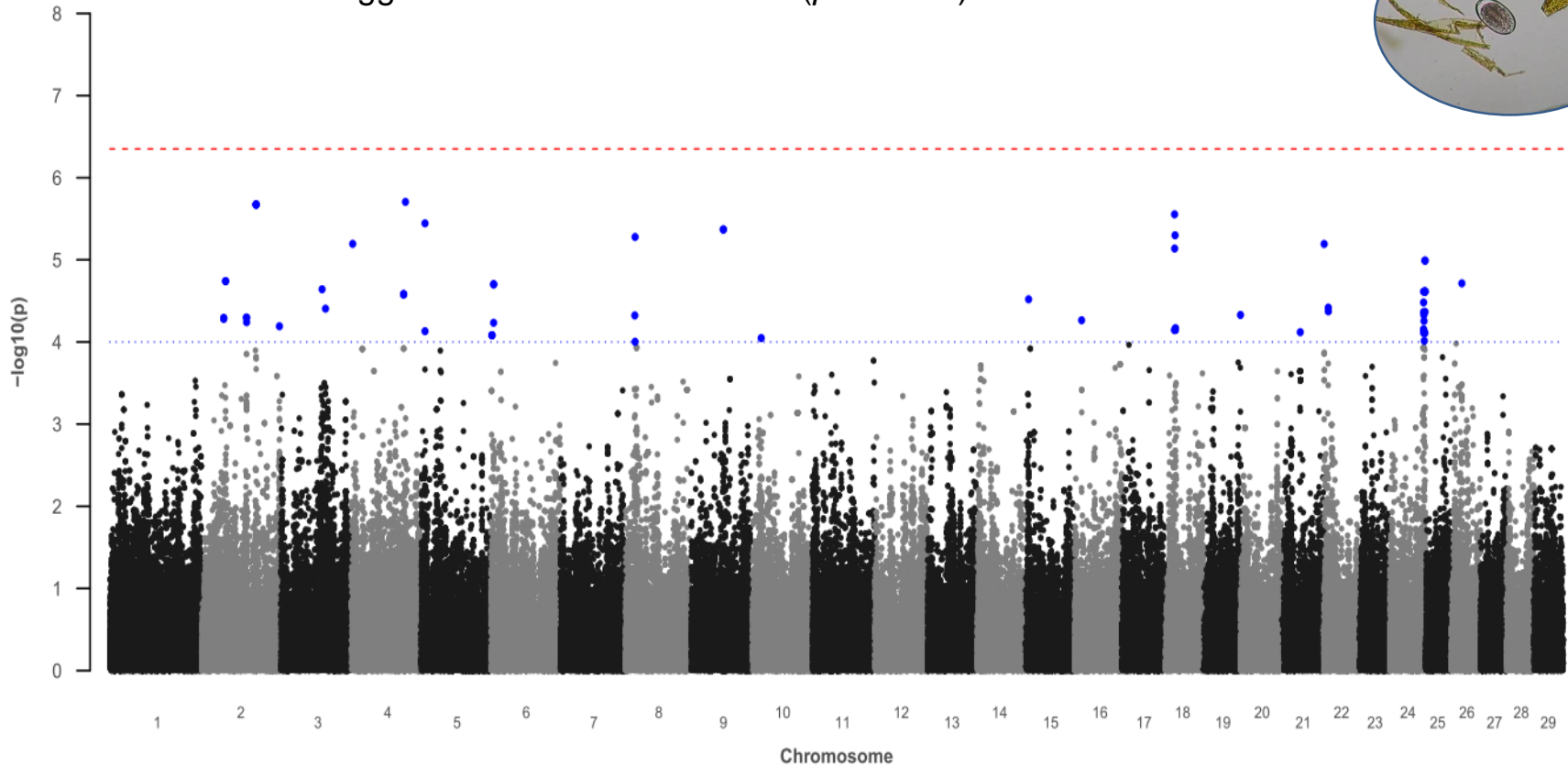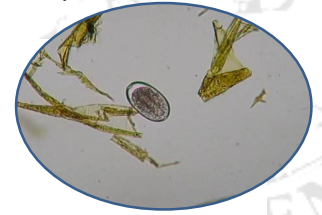
- 68 variants reached suggestive candidate threshold ($p_{Cand} = 1 \times 10^{-4}$)

- 11 candidate genes

JLU

Manhattan-Plot for liver flukes (*F. hepatica*)

-- Bonferroni-corrected genome-wide significance threshold (p = 4.47 x 10^{-7})

··· Suggestive candidate threshold (p = 1e-04)
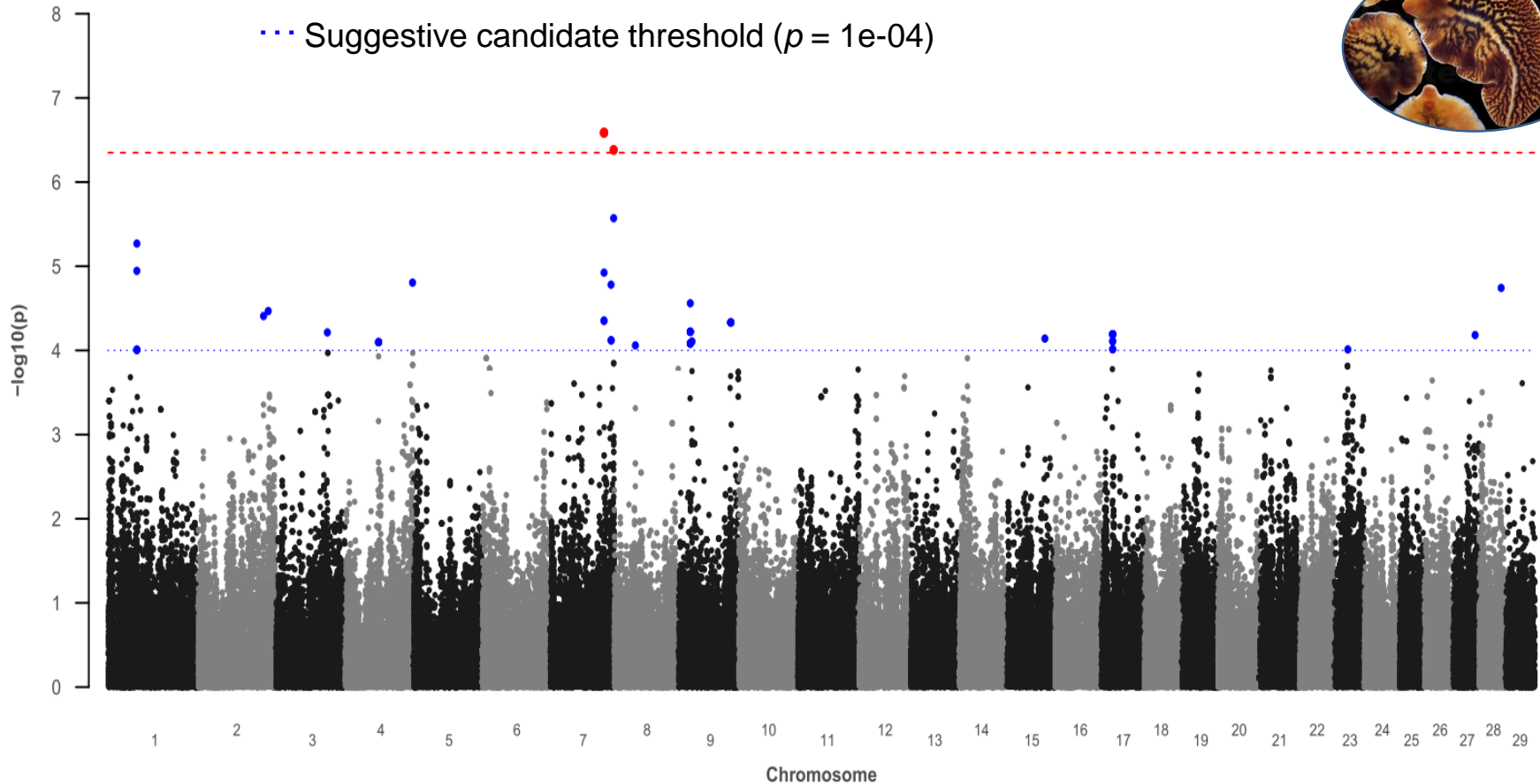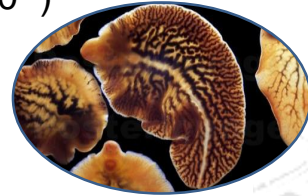
- 3 variants reached p_{Bonf} and 53 variants reached p_{Cand}

- 7 candidate genes

JLU

Gene annotation: Liver fluke (*F. hepatica*)

··· Candidate threshold (p = 1e-04)   – – Bonferroni corrected genome-wide threshold
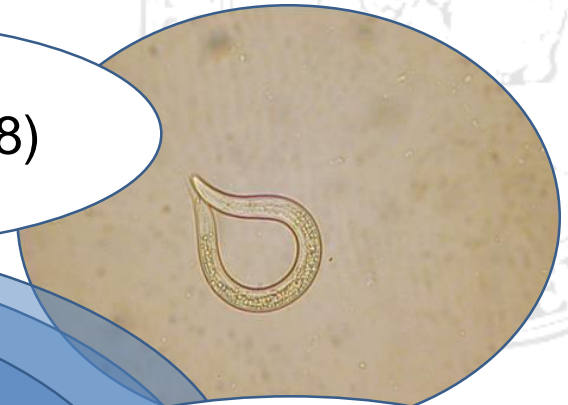(p = 4.47 x 10^{-7})

Activated leukocyte cell adhesion molecule (**ALCAM**) gene
Immune response mechanisms, e.g. cell adhesion interactions of T cells
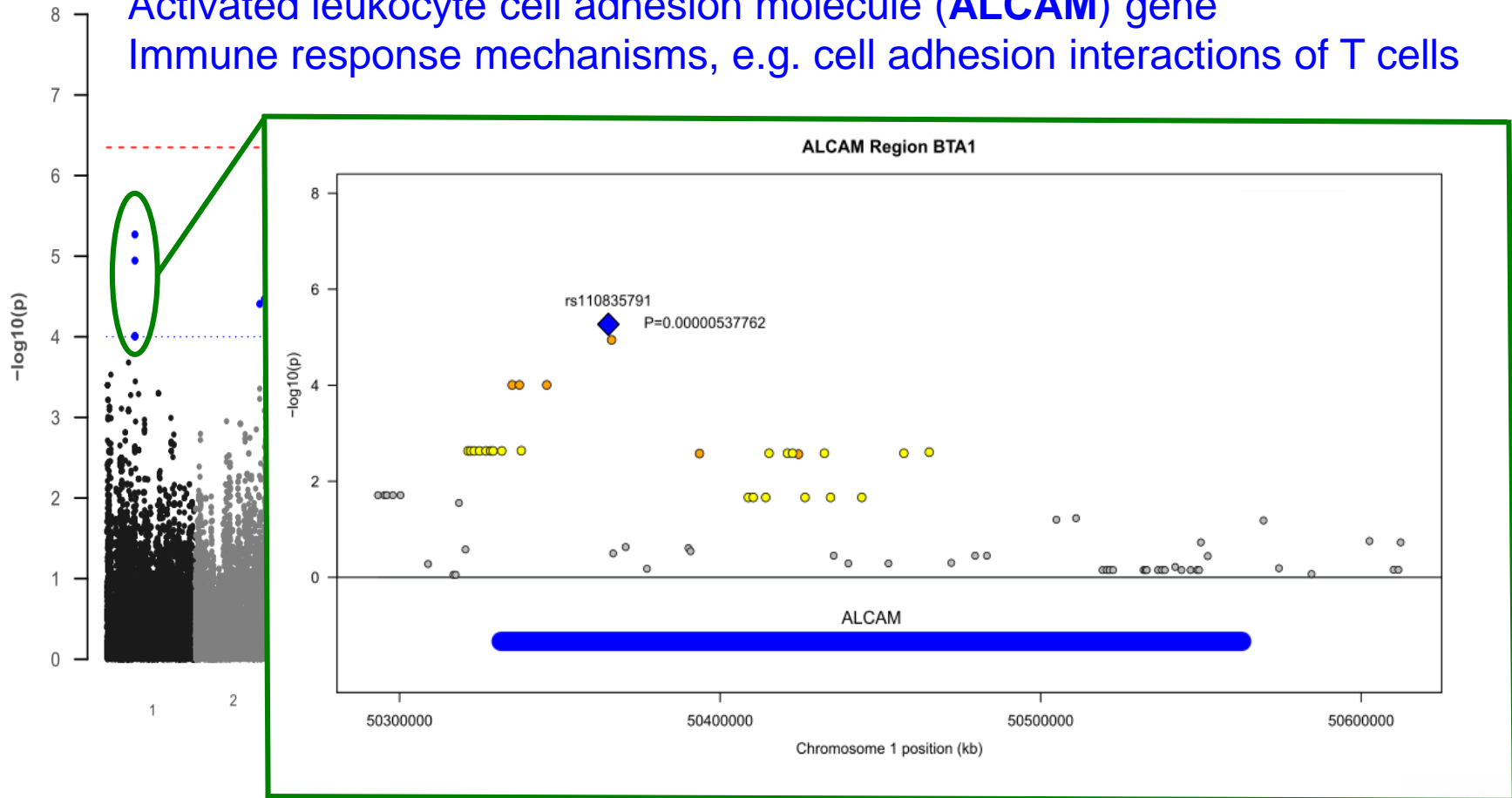
Cell adhesion interactions of T cells

(BTA 28)

Cell adhesion molecules

PI3K/Akt signaling pathway

gen signaling pathway

immunological pathways

KCNJ3 (BTA2)
EGFR (BTA 22)

CDH2 (BTA 24)
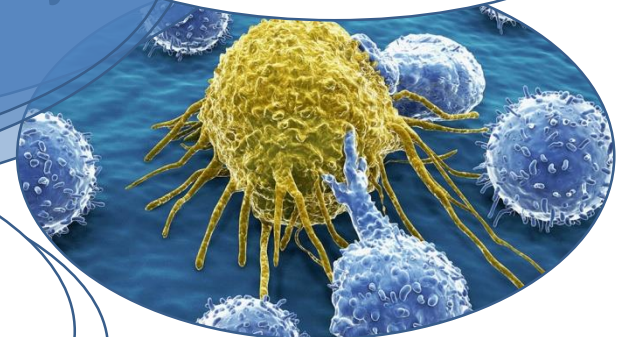
Important functions in cellular immune response, involved in host defense mechanisms during infections with *Neospora caninum*
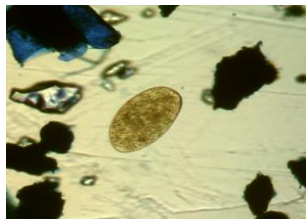
JLU
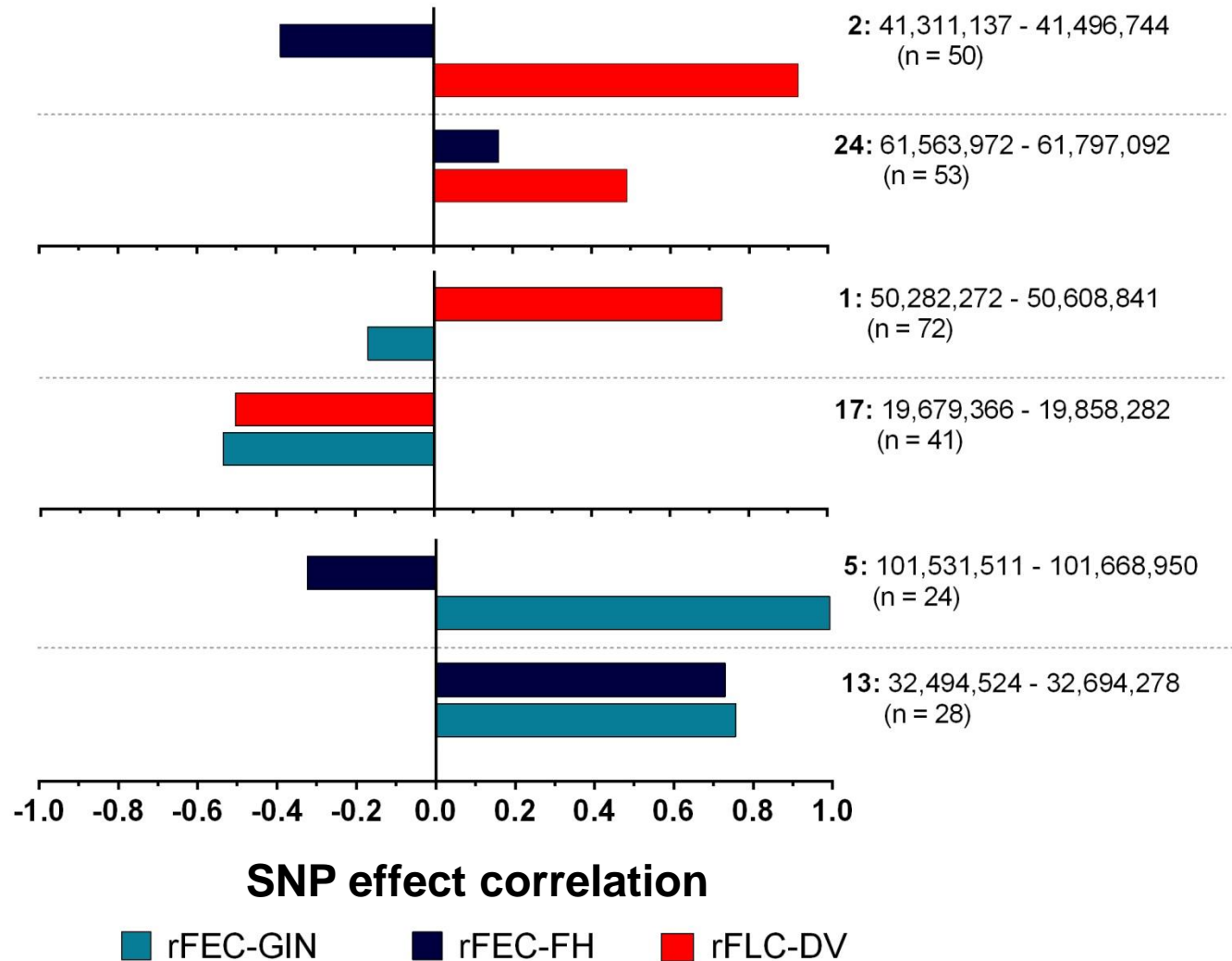
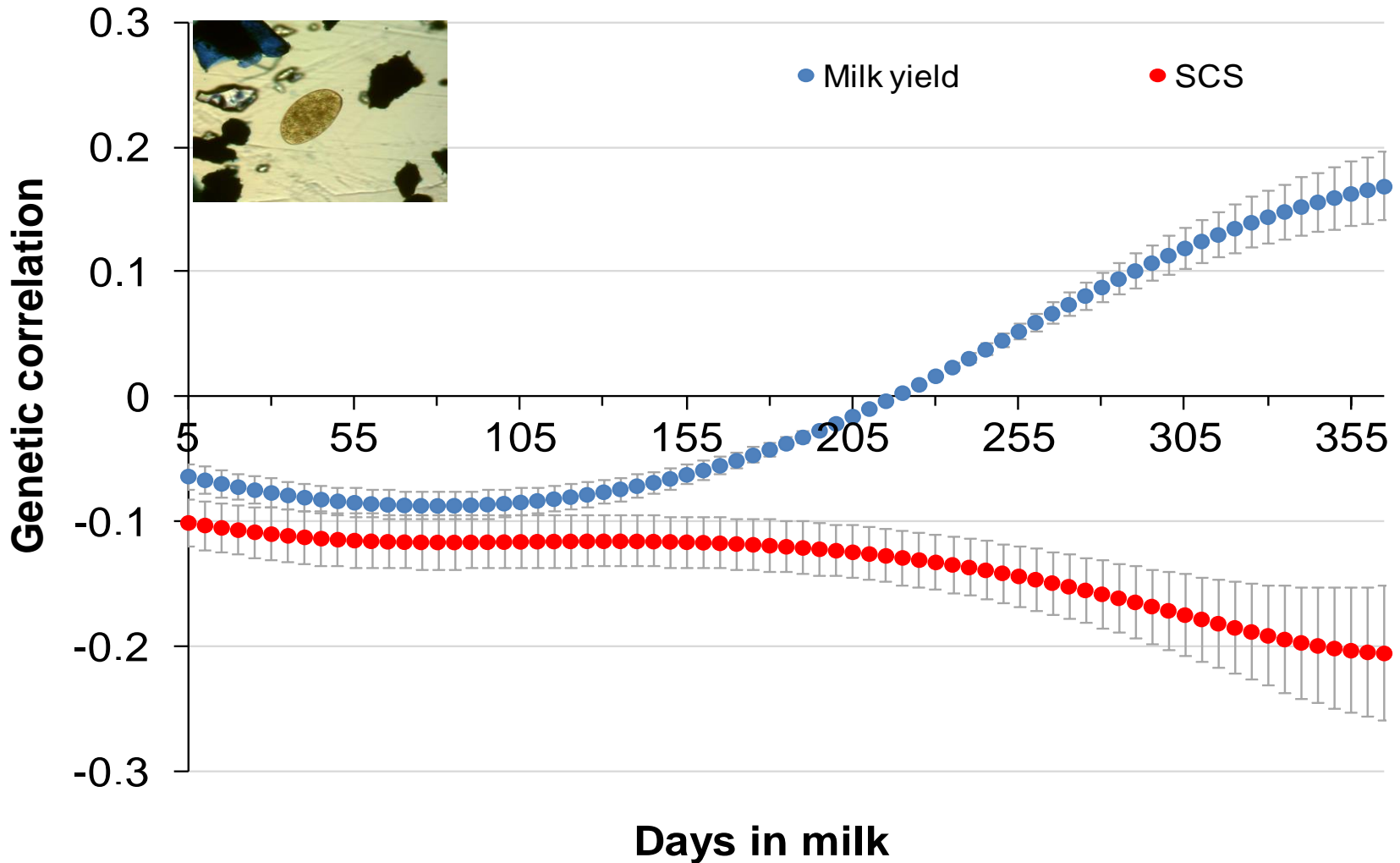# SNP effect correlations within endoparasite traits

ROI für GIN

ROI for liver flukes

ROI for lungworms

2: 41,311,137 - 41,496,744 (n = 50)

24: 61,563,972 - 61,797,092 (n = 53)

1: 50,282,272 - 50,608,841 (n = 72)

17: 19,679,366 - 19,858,282 (n = 41)

5: 101,531,511 - 101,668,950 (n = 24)

13: 32,494,524 - 32,694,278 (n = 28)

SNP effect correlation

rFEC-GIN   rFEC-FH   rFLC-DV

… and between endoparasite and test-day traits

Conclusions

- The 2-step approach using pre-corrected phenotype data based on a larger dataset is a valid strategy for a small number of phenotypes from small-sized cattle populations

- 53 potential candidate genes were related to endoparasite resistance

- 7 pathways associated with immune response mechanisms to endoparasite infections or involved in host-pathogen interactions for candidate genes located on BTA 1, 2, 6, 13, 22, 24 and 28

- Joint genetic basis for the two nematodes GIN and *D. viviparus* (high correlations on genetic and genomic scales, same genes (*NAV3*) and pathways)

- Negative correlations between GIN or *F. hepatica* and SCS within most genomic regions

JLU

Bundesministerium
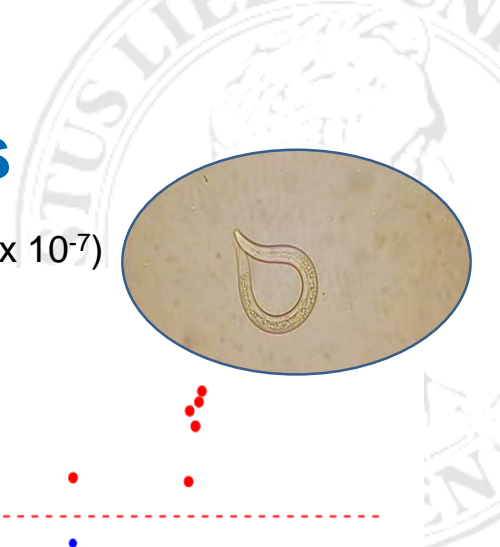für Ernährung
und Landwirtschaft

Thanks for your attention!
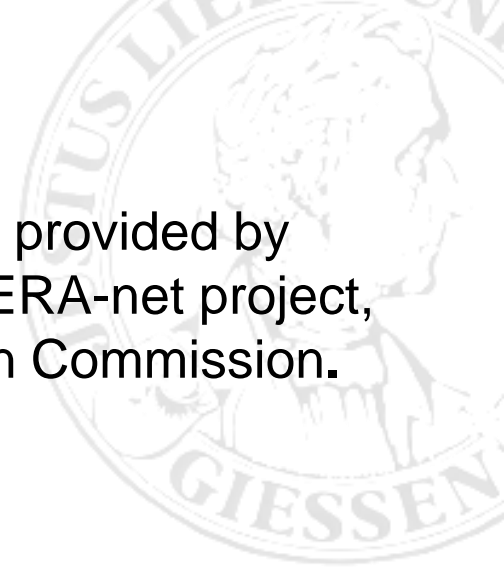
Manhattan-Plot for bovine lungworms

- - Bonferroni-corrected genome-wide significance threshold ($p = 4.47 \times 10^{-7}$)

··· Suggestive candidate threshold ($p = $ 1e-04)
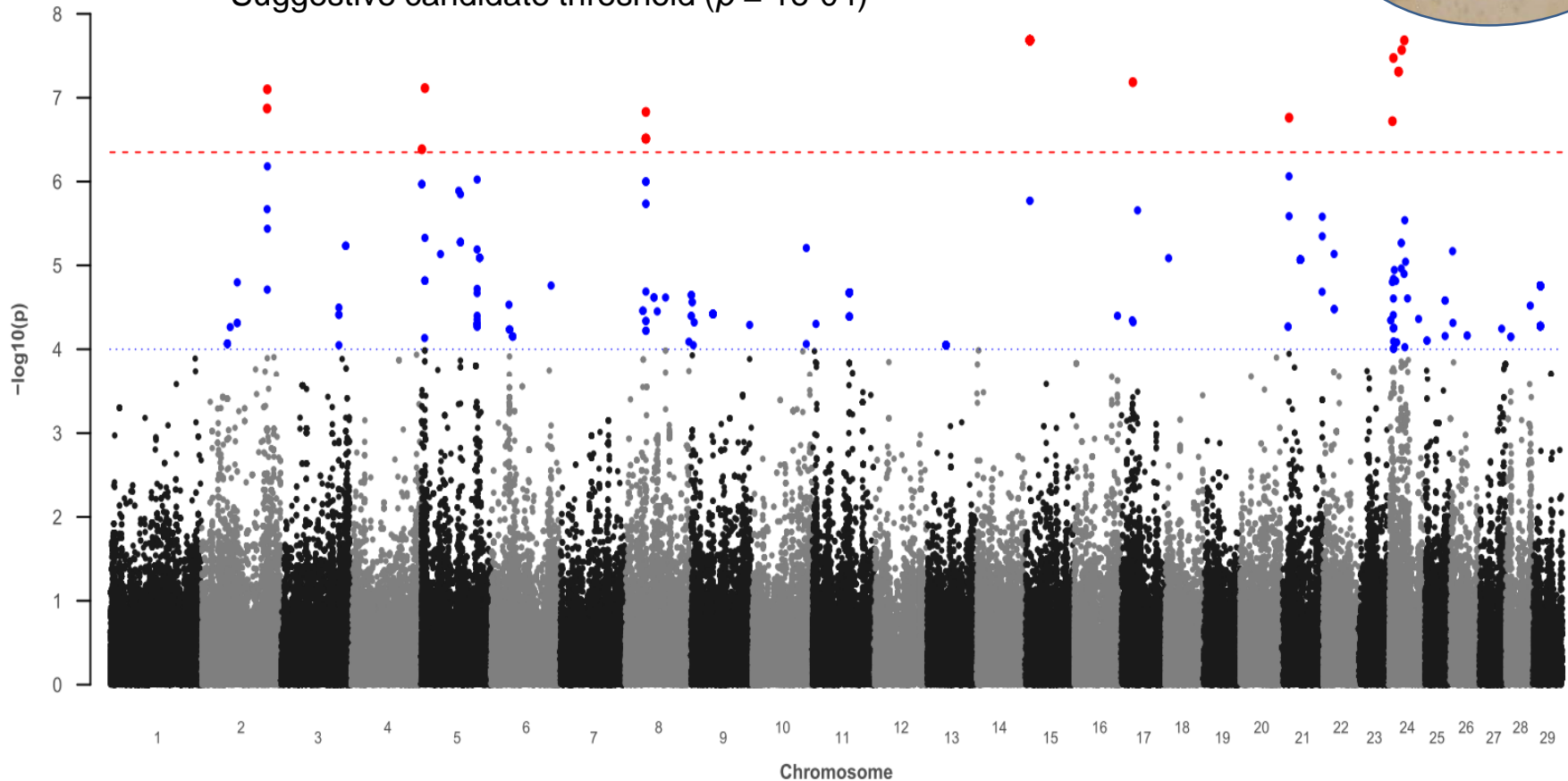
- 41 variants reached p_{Bonf} and 311 variants reached p_{Cand}
- 35 candidate genes

18

Genotyping and Imputing

- Genotyped with BovineSNP50 Bead Chip (50k SNP chip)

- Imputing to Illumina HD Bead Chip level (i.e. 700,000 SNPs) using a multi-breed reference panel of 2188 animals

- The reference panel comprised:

➢ 48 DSN animals genotyped with the Illumina HD Bead Chip array (Illumina Inc., San Diego, CA, USA)

➢ 2140 sequenced animals from the 1000 bull genome project database (Daetwyler et al., 2014) downscaled to Illumina HD Bead Chip density including 30 sequenced DSN animals

JLU
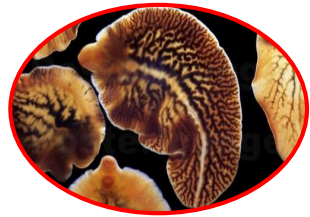
IgG1, IL4, IL10

Interleukin

Cytokines
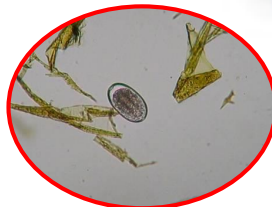
Antigen

Marked antigen

Activated
T lymphocyte

B lymphocyte

Antibody or immune globulines

B cell
development

Plasma cell: antibody
production

IgE

IgA, IgG, IgM

JLU
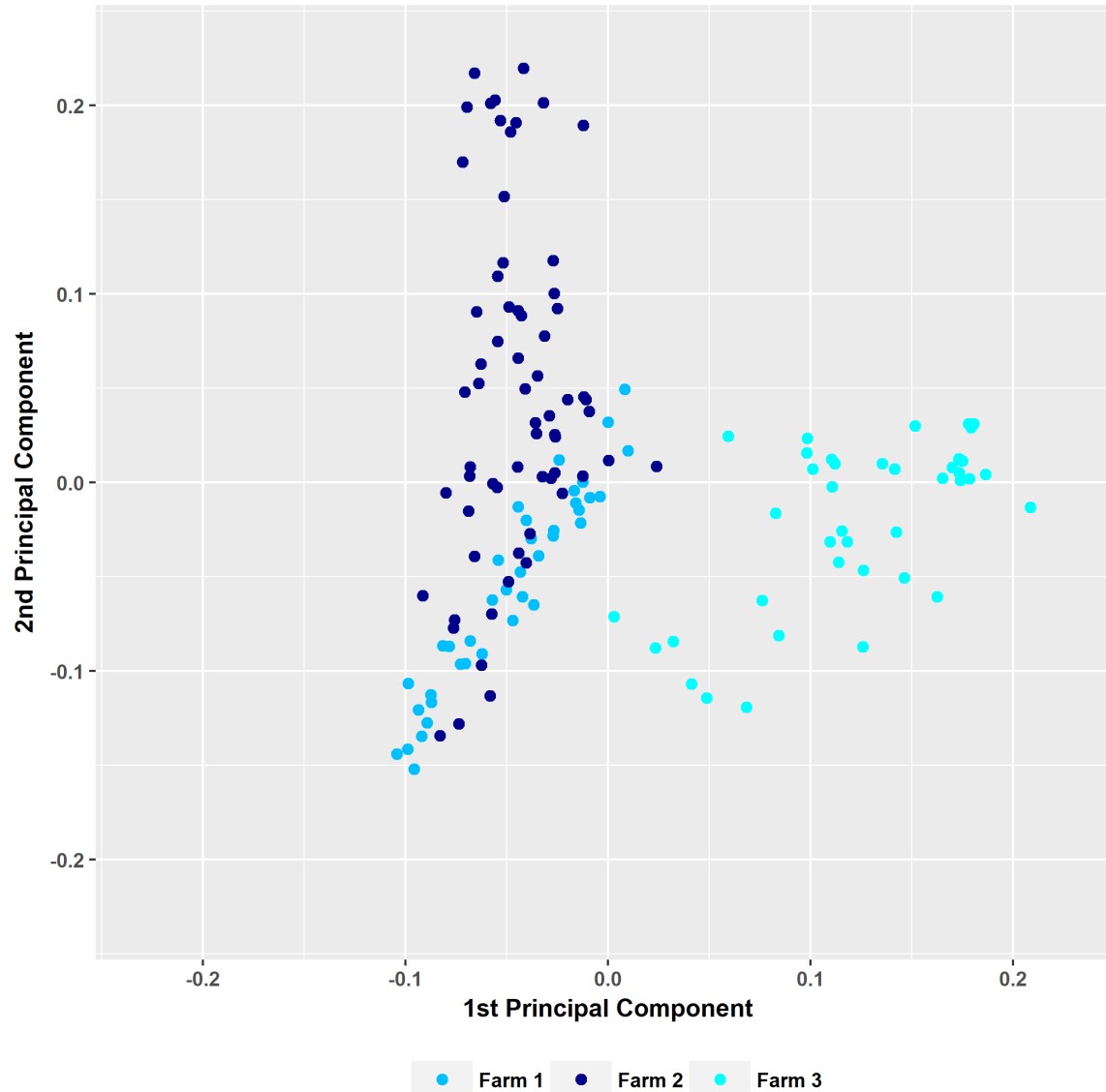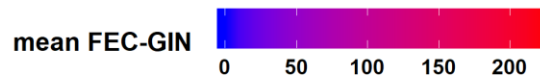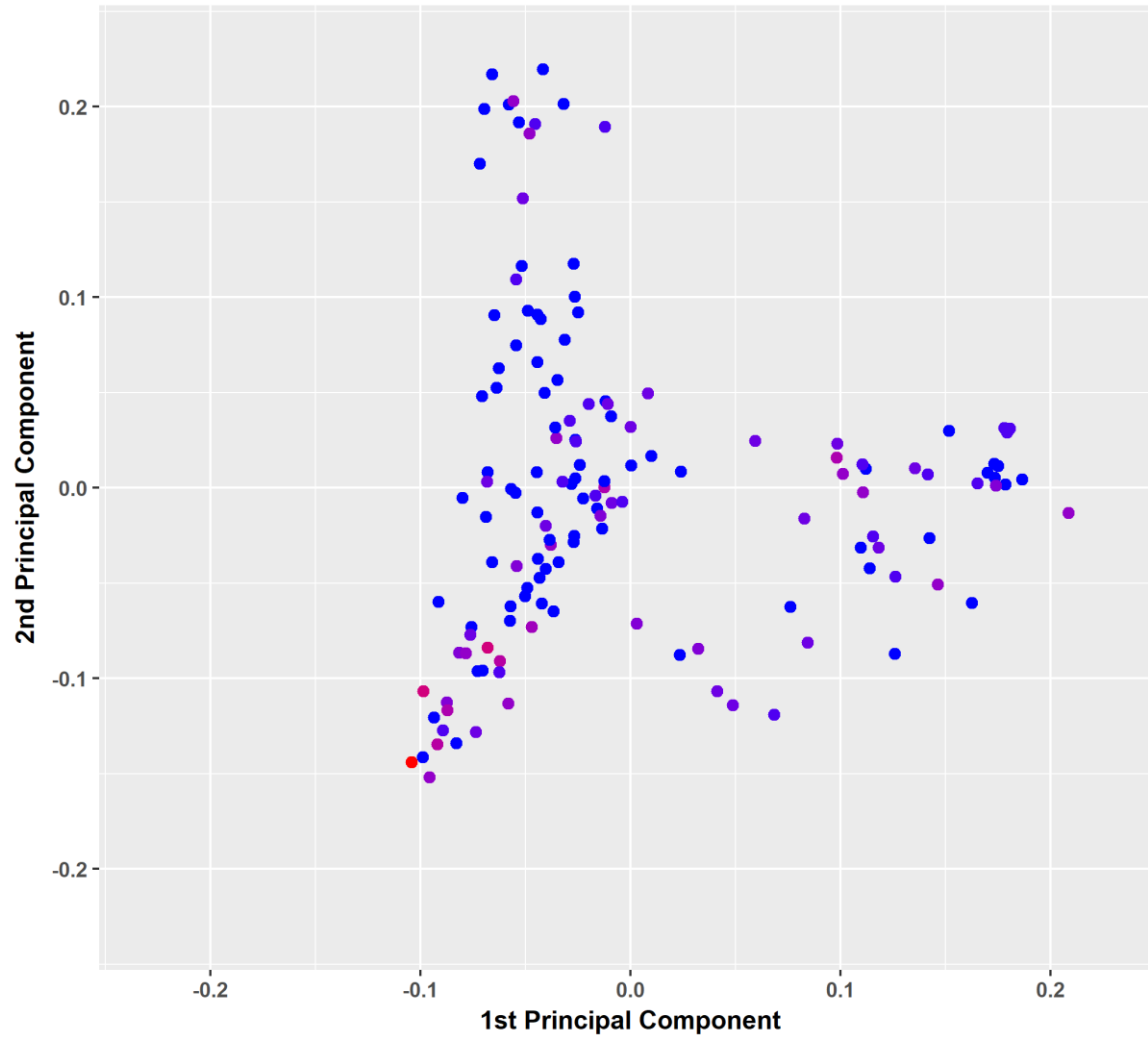
PCA for the three herds

PCA plotted by GIN

PCA plotted by liver flukes (*F. hepatica*)