Faculty of Agricultural and Nutritional Science

Christian-Albrechts-University Kiel
Institute of Animal Breeding and Husbandry

Genome-wide association studies for production traits in pooled pig F_2 designs

I. Blaj[1], J. Tetens[1], S. Preuß[2], R. Wellmann[2], J. Bennewitz[2], G. Thaller[1]

[1]Institute of Animal Breeding and Husbandry, Christian Albrecht University, Kiel, Germany
[2]Institute of Animal Breeding and Husbandry, University of Hohenheim, Stuttgart, Germany

EAAP Annual Meeting 2016, Belfast, UK
Session 67: Free communications in genetics
1st of September 2016

DFG Deutsche Forschungsgemeinschaft

Outline

- Introduction
- Objectives
- Materials and Methods
- Results and Discussion
  - LD decay
  - Single cross analysis, meta-analysis and joint analysis
- Conclusions and Perspectives

Introduction

- gene mapping experiments in livestock
  - ➢ genetic architecture of quantitative traits
  - ➢ genetic markers to facilitate breeding progress
- several F_2 resource populations have been established and analyzed

resolution
precision
power

- **meioses exploited**
- **number of individuals included**
- **marker density**

- **LD structure**
  - ➢ **the length of the LD blocks can be reduced by pooling several F_2 crosses (*Bennewitz and Wellmann, 2014*)**

Objectives

- three-generation experimental populations
  - **Piétrain** x Large White, **Piétrain** x (Large White x Landrace) – European breeds cross
  - Meishan x **Piétrain,** Wild boar x **Piétrain** – Asian/European breeds cross

- phenotypes: average daily gain (ADG), back fat thickness (BFT), meat to fat ratio (MFR)

- combine data from two experimental F_2 crosses
  - **structural identification of short chromosomal regions that show evidence for trait association**

Materials and Methods

- total of **2,554 animals**
  - 1,894 individuals European breeds cross PxLW/(LWxL)
  - 660 individuals Asian/European breeds cross M/WxP

- $\boldsymbol{P}$ **/** $\boldsymbol{F_1}$ **/** $\boldsymbol{F_2}$ **genotyped** with PorcineSNP60 BeadChip (Illumina)
  - SNP chromosomal positions - current pig genome assembly (Sscrofa build 10.2)

- phenotypes were measured using similar methods and standardized techniques (*Müller et al. 2000, Borchers 2002*)

Materials and Methods

GWAS Workflow

Individual cross

Meta-analysis of the individual crosses

Pooled pre-corrected data

- mixed linear model (MLM) based association analysis (GCTA version 1.26.0, Yang et al, 2011)

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{g} + \varepsilon \text{ with } \mathrm{V} = \frac{\boldsymbol{W}\boldsymbol{W}'}{\boldsymbol{N}}\,\boldsymbol{\sigma_g}^2 + \mathrm{I}\,\boldsymbol{\sigma_\varepsilon}^2 = \mathrm{A}\,\boldsymbol{\sigma_g}^2 + \mathrm{I}\,\boldsymbol{\sigma_\varepsilon}^2$$

- Fixed effects

Cross/FE	European breeds cross	Asian/European breeds cross
ADG	stable, slaughtering period	sex, cross
BFT	sex, stable, slaughtering period, weight at slaughter	sex, slaughtering period, weight at slaughter, age at slaughter, cross
MFR	sex, stable, slaughtering period, birth weight	sex, cross

Materials and Methods

GWAS Workflow

Individual cross

Meta-analysis of the individual crosses

Pooled pre-corrected data

- METAL version 2011, Willer et al 2010
- sample based approach
- analytical strategy

Input	N_i– sample size for study i P_i– p-value for study i Δ_i - direction of effect for study i
Intermediate Statistics	$Z_i = \Phi^{-1}(P_i/2) * \text{sign}(\Delta_i)$ $w_i = \sqrt{N_i}$
Overall Z-score	$Z = \frac{\Sigma_i Z_i w_i}{\sqrt{\Sigma_i {w_i}^2}}$
Overall P-value	$P = 2\Phi\ (\lvert -Z \rvert)$

Materials and Methods

GWAS Workflow

Individual cross

Meta-analysis of the individual crosses

Pooled pre-corrected data

- phenotypes pre-corrected in the individual crosses
- MLM: $\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{g} + \varepsilon$

cross effect – 2 classes

Cross/FE	European breeds cross	Asian/European breeds cross
ADG	stable, slaughtering period	sex
BFT	sex, stable, slaughtering period, weight at slaughter	sex, slaughtering period, weight at slaughter, age at slaughter
MFR	sex, stable, slaughtering period, birth weight	sex

Results and Discussion

Tab. 1: Descriptive statistics and heritabilities (h^2_{SNP} and $h^2_{pedigree}$) for average daily gain (ADG), back fat thickness (BFT) and meat to fat ratio (MFR**)**

Cross	Trait	N	mean	sd	min	max	h^2_{SNP}	$h^2_{pedigree}$
European	**ADG[g]**	1769	675.90	92.73	311.0	1039.0	0.35	0.47
Asian/European	**ADG[g]**	595	559.40	124.19	125.0	951.0	0.44	0.24
European	**BFT[mm]**	1766	27.49	3.84	16.00	42.30	0.43	0.43
Asian/European	**BFT[mm]**	595	19.44	6.93	3.70	43.30	0.47	0.56
European	**MFR**	1765	0.38	0.10	0.14	0.85	0.46	0.36
Asian/European	**MFR**	593	0.62	0.21	0.19	1.39	0.51	0.44

Results and Discussion

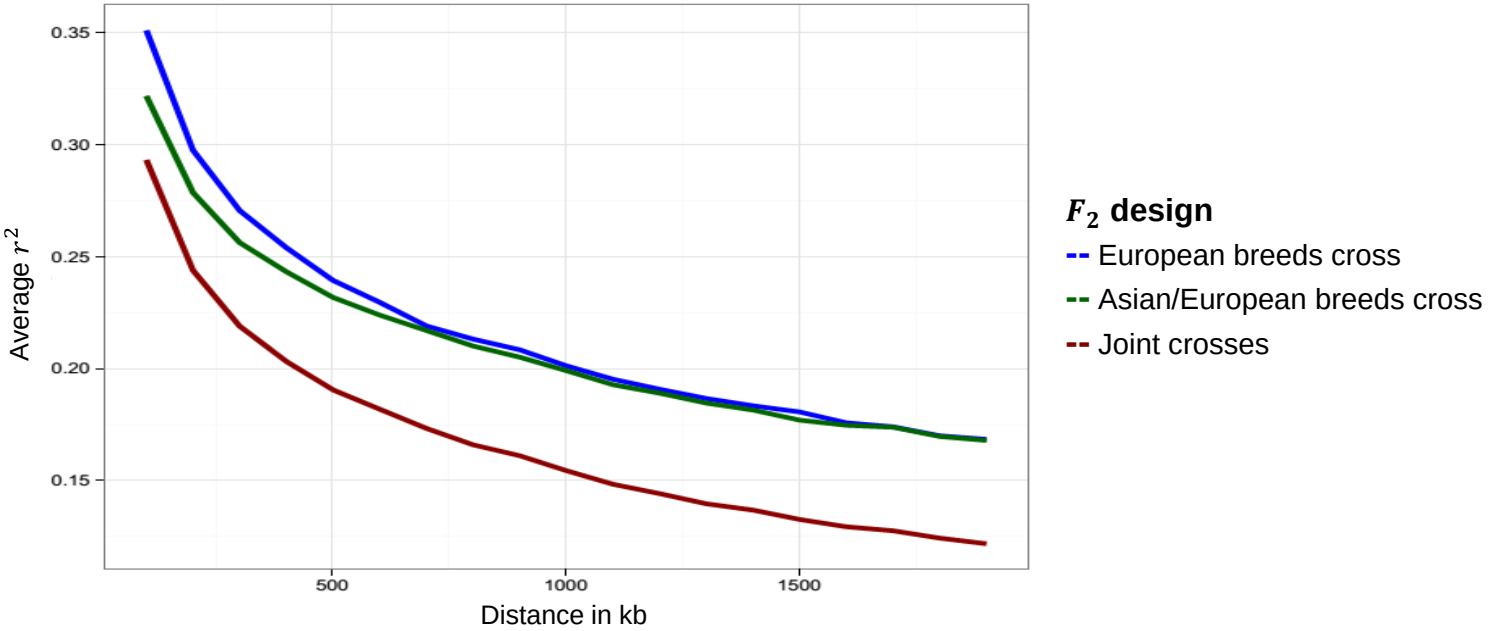

Fig. 1: LD decay over physical distance

ADG – average daily gain

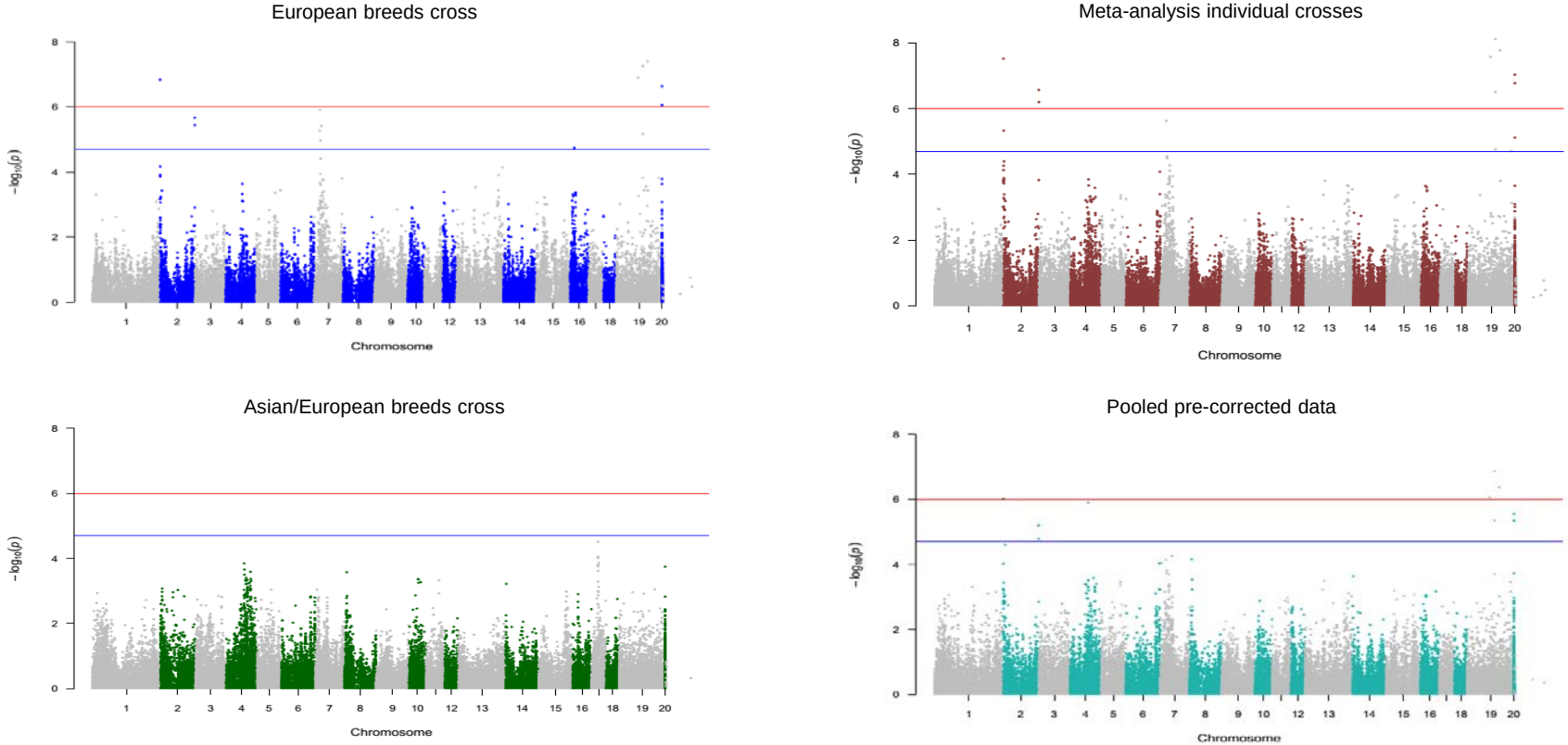

Fig. 2: Manhattan plot of genome-wide association studies for average daily gain. The red line indicates the Bonferroni–corrected significance threshold ($P=1.1x10^{-6}$) and the blue line indicates the threshold ($P=2.2x10^{-5}$) for suggestive SNPs.

BFT – back fat thickness

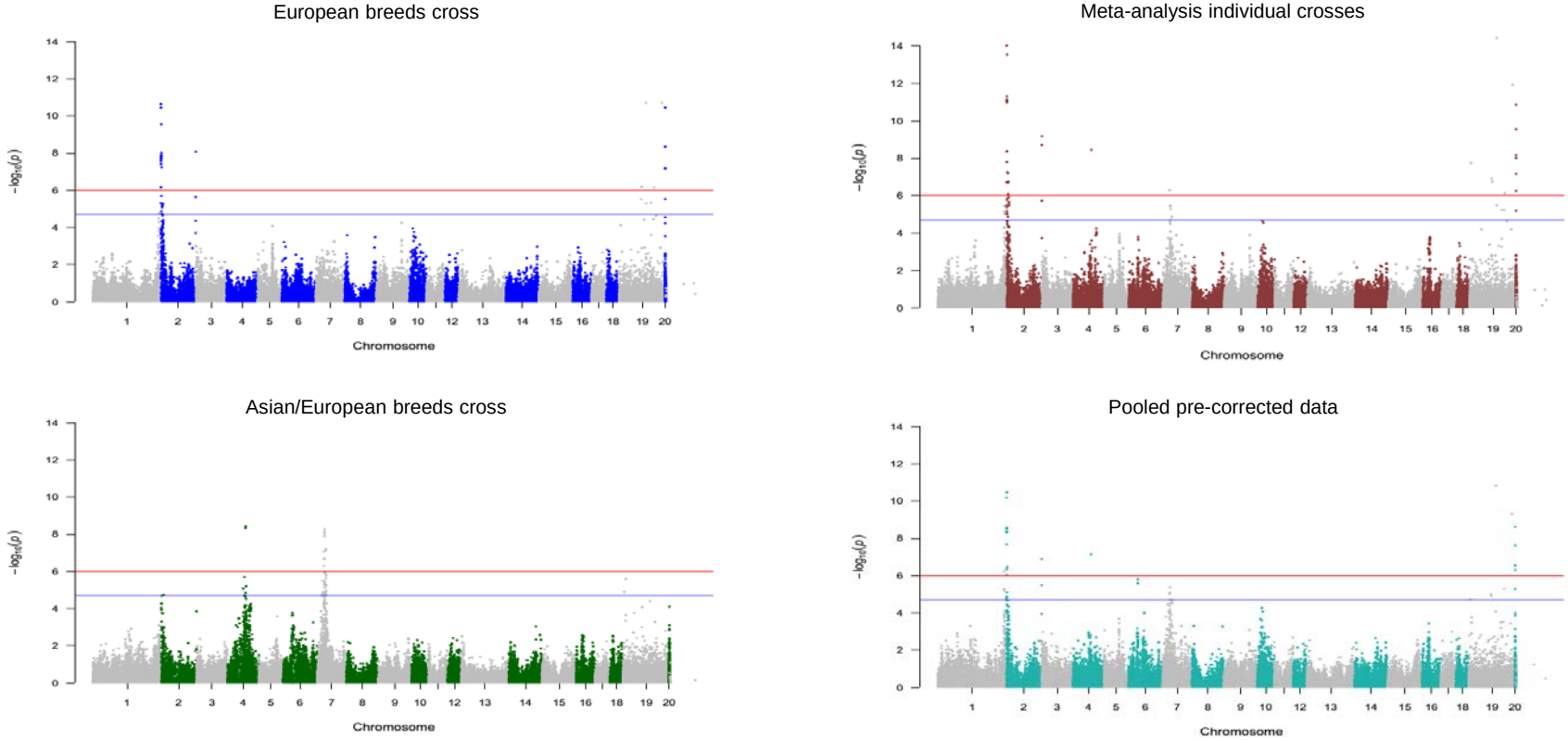

Fig. 3: Manhattan plot of genome-wide association studies for back fat thickness. The red line indicates the Bonferroni–corrected significance threshold ($P=1.1x10^{-6}$) and the blue line indicates the threshold ($P=2.2x10^{-5}$) for suggestive SNPs.

MFR – meat to fat ratio

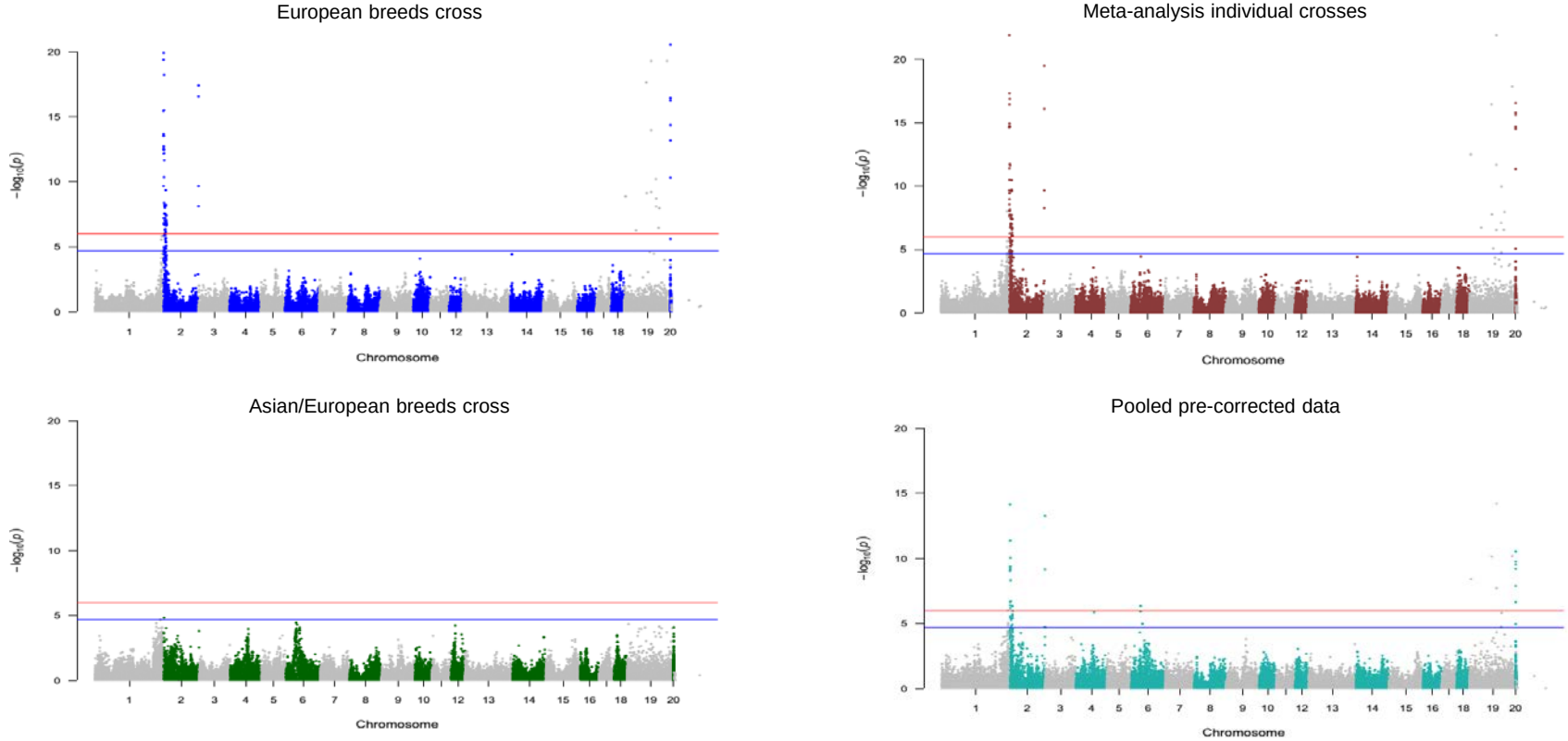

Fig. 4: Manhattan plot of genome-wide association studies for meat to fat ratio. The red line indicates the Bonferroni–corrected significance threshold ($P=1.1x10^{-6}$) and the blue line indicates the threshold ($P=2.2x10^{-5}$) for suggestive SNPs.

Conclusions and Perspectives

Conclusions

- the meta-analysis was generally more powerful in detecting more precise locations and higher significance levels in the combined crosses vs. single cross
- association levels in pooled pre-corrected data were lower than in the meta-analysis
- common underlying variants that show a different frequency between the two crosses
- chromosomes showing significant evidence for trait association in the meta-analysis
  - **ADG** - *SSC2;* **BFT** - *SSC2, SSC4, SSC7;* **MFR** - *SSC1, SSC2*

Perspectives

- heterogeneous residual variance to be modelled in the joint analysis
- sequencing of the P → imputation → Whole-genome sequence based association studies

Thank you for your attention!

Questions?

*****The authors would like to thank the German Research Foundation (*DFG*) for funding.