Session 53

corresponding author: M. Erbe
e-mail: malena.erbe@tierzucht.tum.de

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Power and robustness of three whole genome association mapping approaches in selected populations

M. Erbe[1,2], F. Ytournel[1], E. Pimentel[1], A.R. Sharifi[1] and H. Simianer[1]

[1] Group of Animal Breeding and Genetics, Department of Animal Sciences, Georg-August-University, Goettingen

[2] Chair of Animal Breeding, TUM, Freising

Introduction

- selection and consequentially small effective population size in livestock populations
  - spurious associations: associations between two loci which are not closely linked
  - ➔ can affect results of mapping experiments
- whole genome association mapping
  - ➔ no prior knowledge about QTL position
  - no possibility to decide if significant SNP is correct or false positive

➔suitable methods should combine sufficient power with an acceptable number of false positive signals

Simulation overview I

1000 unrelated individuals (500 ♂, 500 ♀)

1000 generations with N = 1000 (random mating)

Bottleneck: 100 individuals (50 ♂, 50 ♀) chosen randomly

10 generations with N = 100 (random mating)

1 generation with N = 1000 (random mating)

1000 individuals (so-called “founder generation”)

Simulation overview II

➔ **500 individuals** (250 ♂, 250 ♀) chosen randomly out of the founder generation (sample differing in each replicate)

15 generations (N = 500) ➔ **different scenarios**

Scenario A	**Scenario B**	**Scenario C**
		
Idealized random mating population	**Limited number of parents**	**Directional selection**
(250 ♂, 250 ♀)	(25 ♂, 200 ♀)	(25 ♂, 200 ♀)

Simulation

- Animals for whole genome association mapping:

  2500 individuals of the last five generations

  ➔ full pedigree, genotypic and phenotypic information

- **10,000 SNPs** on 10 chromosomes of each 1 Morgan length
- **50 biallelic QTL**

  ➔ only additive effects

- Set of SNPs and QTL was the same in all replicates.
- Heritability was assumed to be approximately 0.3.

Approaches

- **Single Marker Regression (SMR)**

  ➔ simple linear regression of phenotypes on genotypes of single marker

  ➔ no correction for family effects

- Genomewide Rapid Association Using Mixed Model and Regression (GRAMMAR) (Aulchenko et al., 2007)
- Quantitative Transmission Disequilibrium Test applied to the Mendelian Sampling Term (MTDT) (Simianer and Pimentel, 2009)

Approaches

- Single Marker Regression (SMR)
- **Genomewide Rapid Association Using Mixed Model and Regression (GRAMMAR)** (Aulchenko et al., 2007)

  ➔ previously estimated residual terms (free of family correlations) as dependent variable in a single marker regression

- Quantitative Transmission Disequilibrium Test applied to the Mendelian Sampling Term (MTDT) (Simianer and Pimentel, 2009)

Approaches

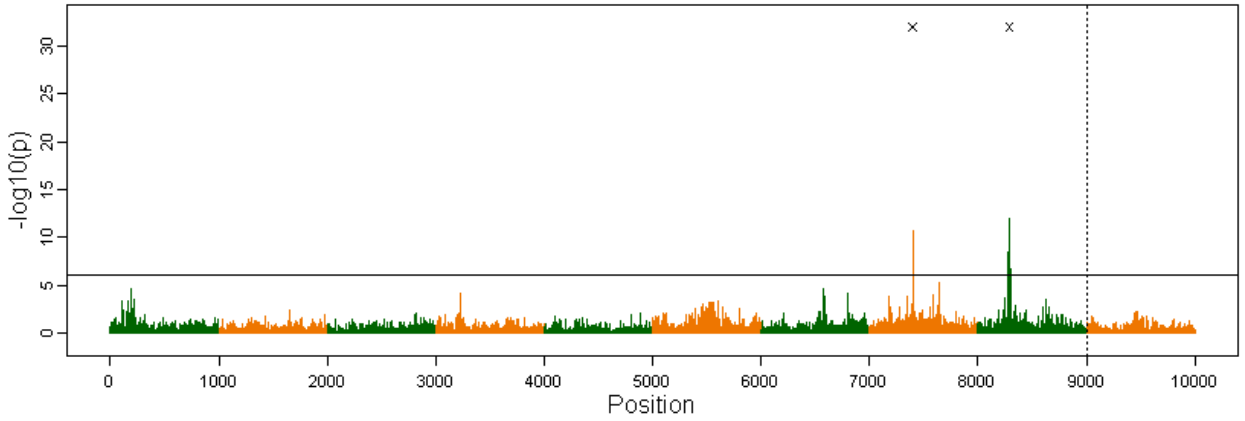
- Single Marker Regression (SMR)
- Genomewide Rapid Association Using Mixed Model and Regression (GRAMMAR) (Aulchenko et al., 2007)

- **Quantitative Transmission Disequilibrium Test applied to the Mendelian Sampling Term (MTDT)** (Simianer and Pimentel, 2009)

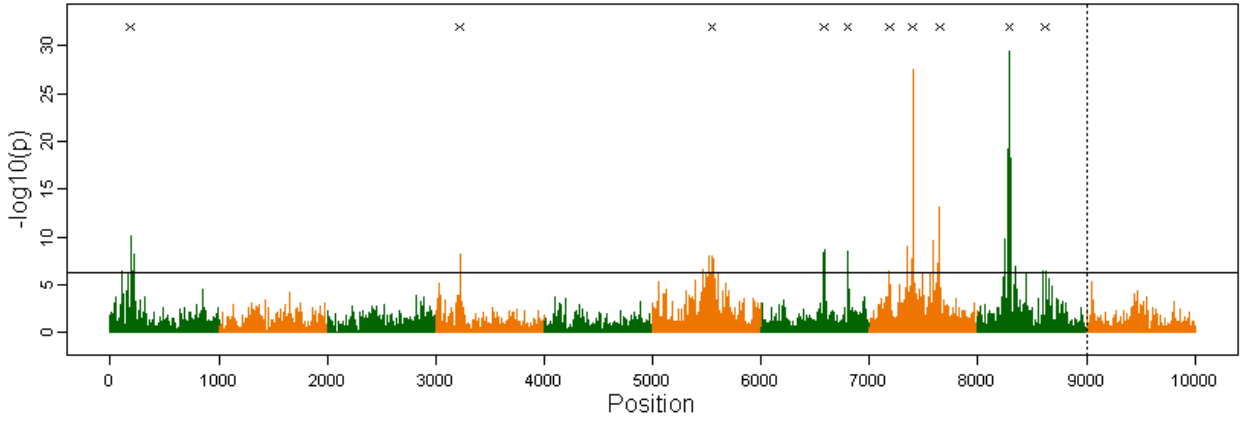
  ➔ estimated Mendelian sampling terms derived from the estimated breeding values

  ➔ estimated Mendelian sampling terms used as phenotypic information for a QTDT

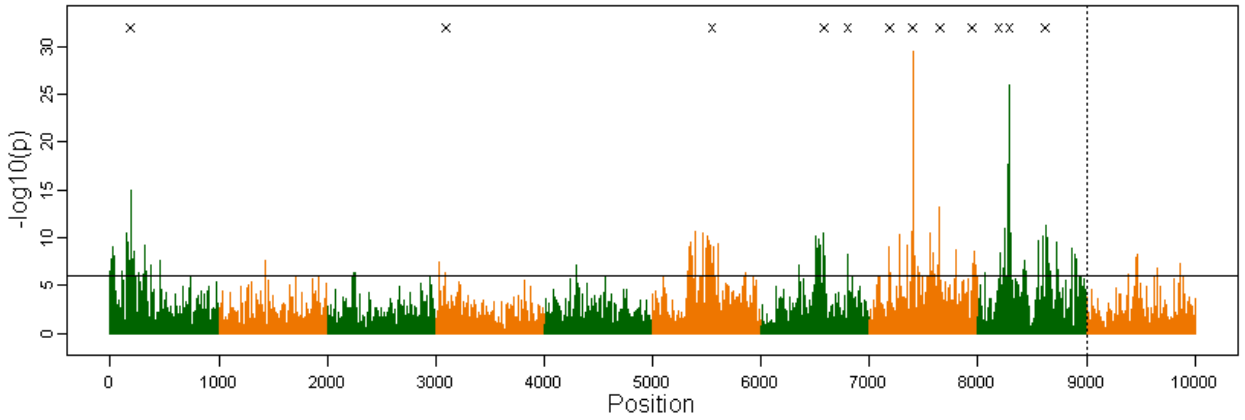
GRAMMAR



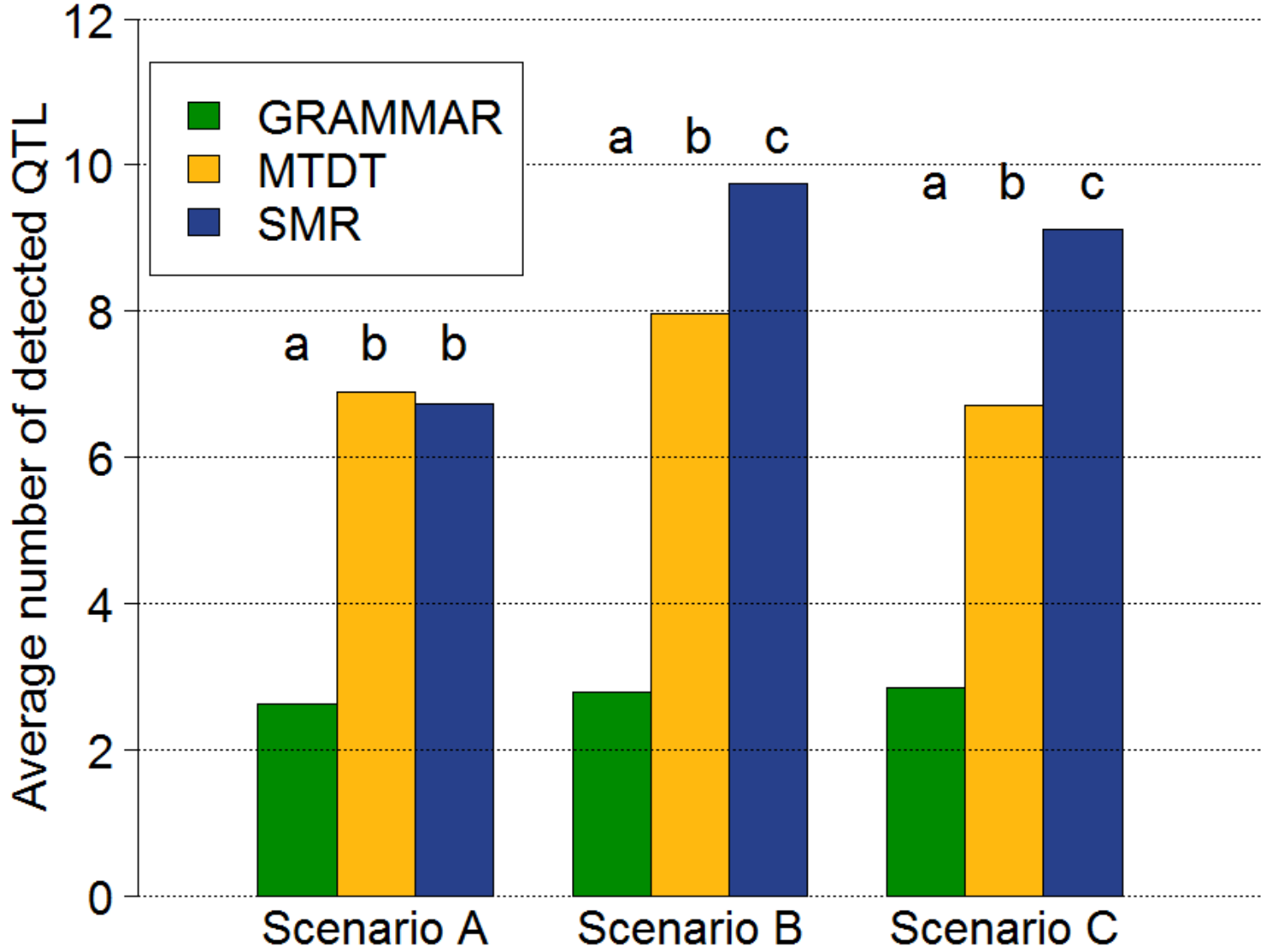
MTDT



SMR

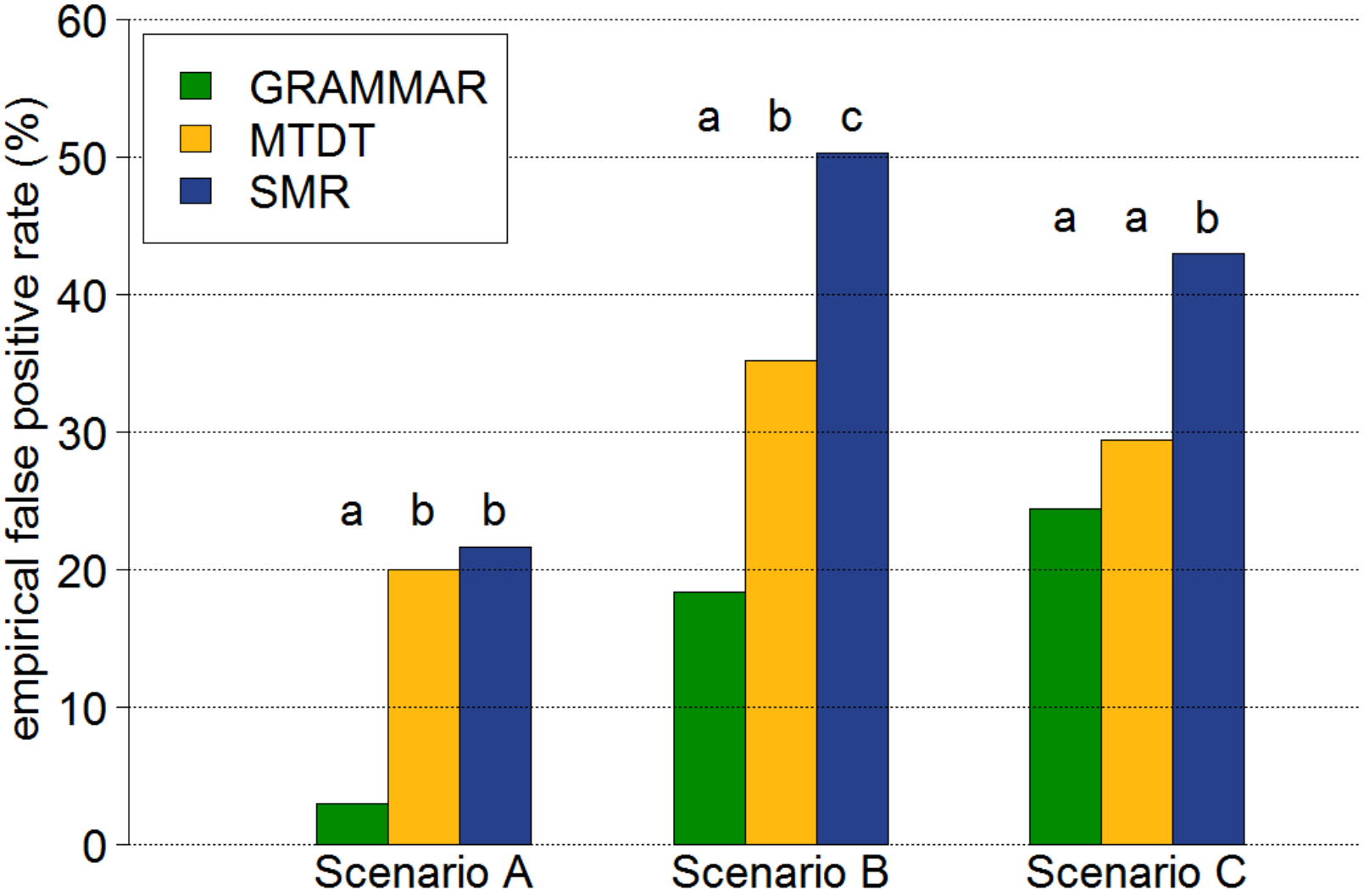


Detected QTL - Power



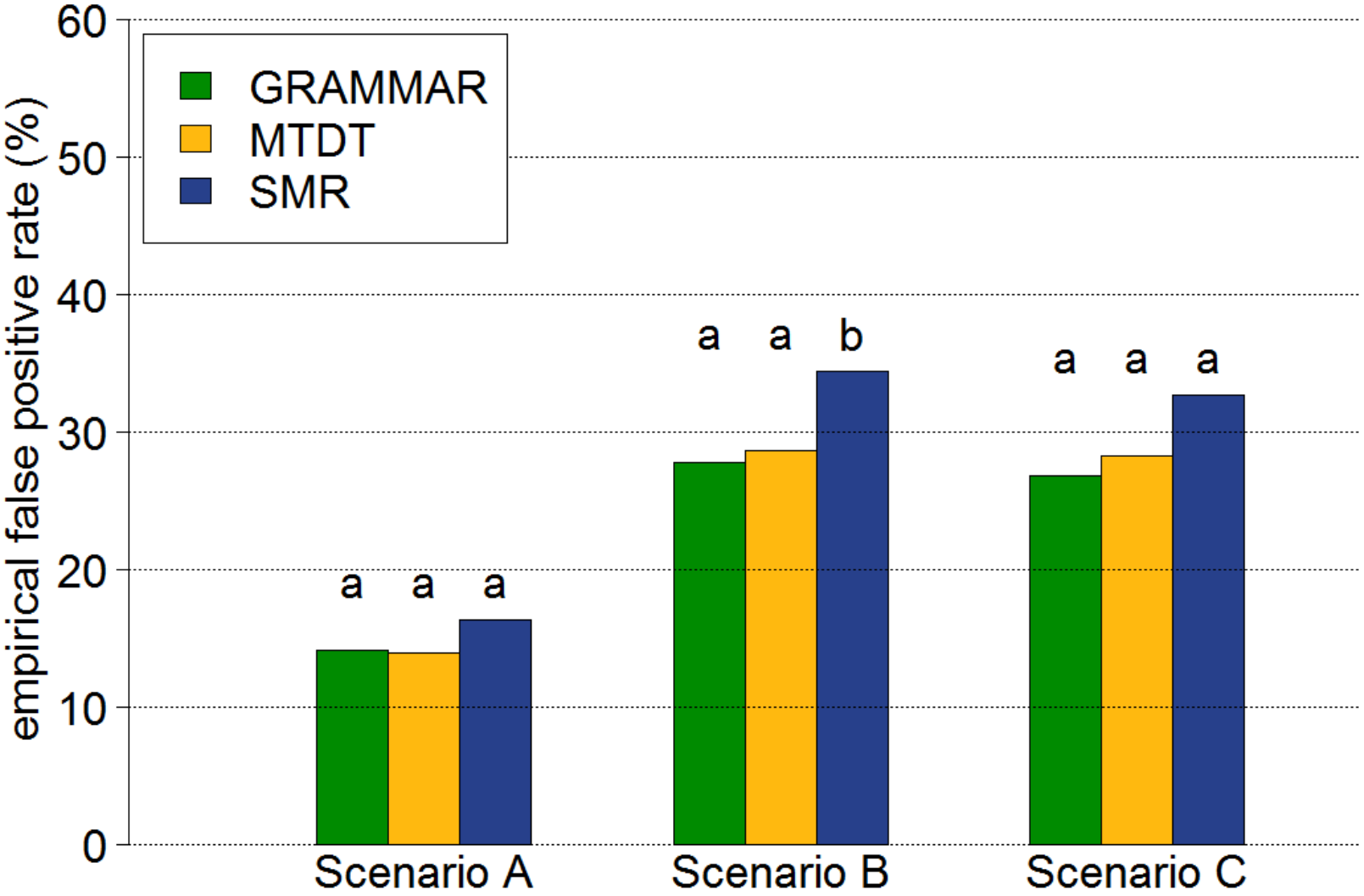
a, b, c: significantly different on a 1% error level

Empirical false positive rate –
Bonferroni corrected significance threshold



a, b, c: significantly different on a 1% error level

Empirical false positive rate –
standardized conditions (number of detected QTL=5)



a, b: significantly different on a 1% error level

Conclusions

- **SMR:** results should be considered with great caution
- **GRAMMAR** and **MTDT** avoid false positive signals
  ➔ suitable for genome scans in livestock populations
- **GRAMMAR:** efforts have to be made to find an appropriate significance threshold
- **Suggestion:**
  - use of a fast approach (GRAMMAR or MTDT) for a first screen of the genome
  - use of more refined methods (both for modelling and for determining the significance threshold) in candidate regions

Thank you for your attention!

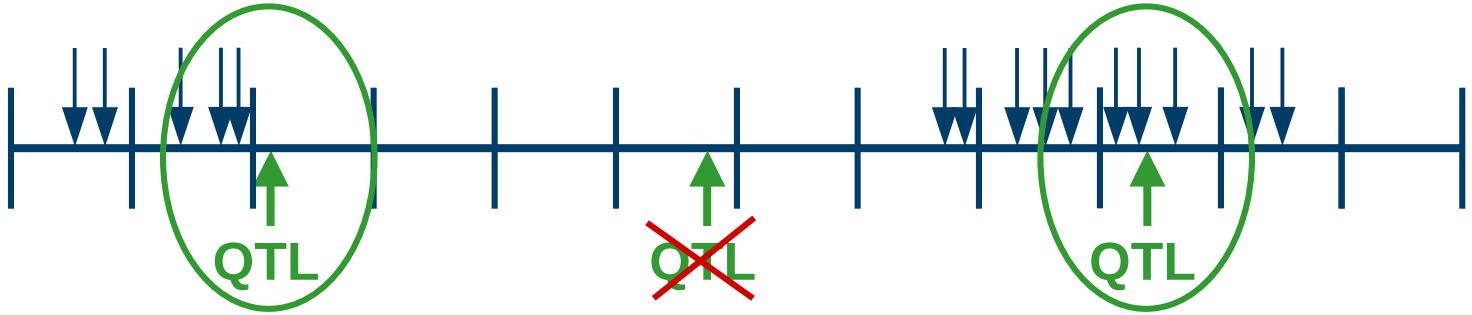
References

Aulchenko, Y. S., D.-J. de Koning, and C. Haley, 2007. Genomewide Rapid Association Using Mixed Model and Regression: A Fast and Simple Method For Genomewide Pedigree-Based Quantitative Trait Loci Association Analysis. Genetics 177: 577-585.

Simianer, H., and E. C. G. Pimentel, 2009. Robust QTL fine mapping by applying a quantitative transmission disequilibrium test to the Mendelian sampling term. J. Anim. Breed. Genet. (accepted)

Definitions

- 100 intervals per chromosome each of one cM length
- **Detected QTL:**



- **Detected False Associations**