SABRE

CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Gene prioritization using text mining and protein-protein interaction in livestock species

Li Jiang
Faculty of Agricultural Sciences
Aarhus university, Denmark

4th SABRE Conference / 60th Annual Meeting EAAP, Barcelona, 24-27 August 2009

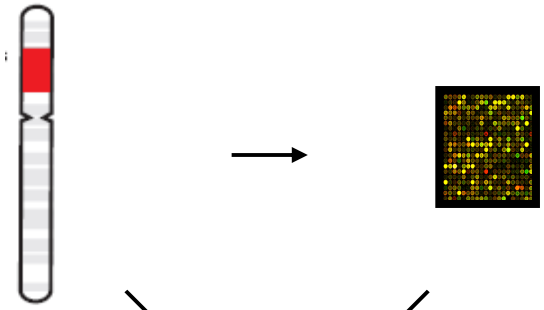
Food Quality and Safety

Find causal genes in complex traits

Linkage study

Expression study

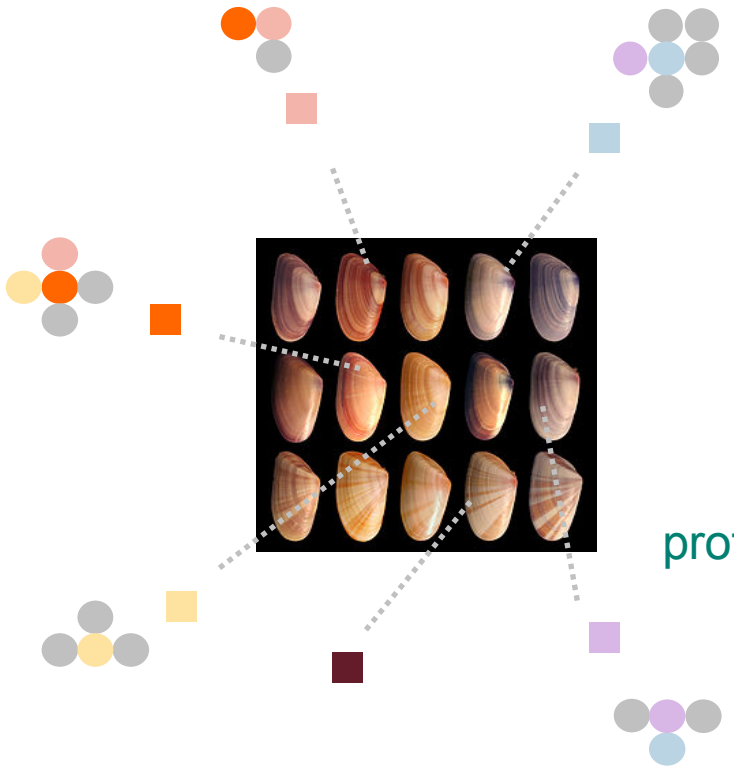
Whole genome



Candidate gene list

Gene prioritization

Phenotype and genotype links in complex traits



phenotype

gene linked to phenotype

protein linked to phenotype

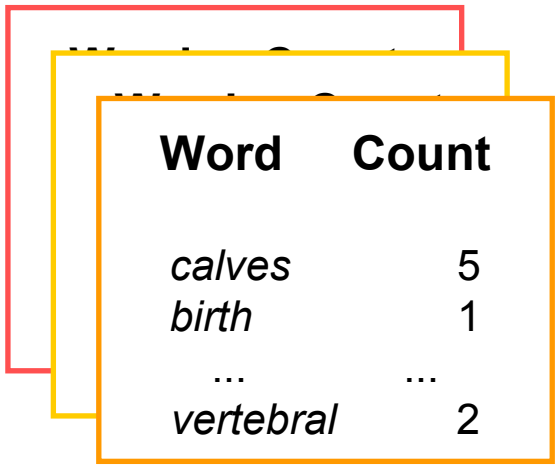
protein not linked to phenotype

Measure phenotype similarity after text mining

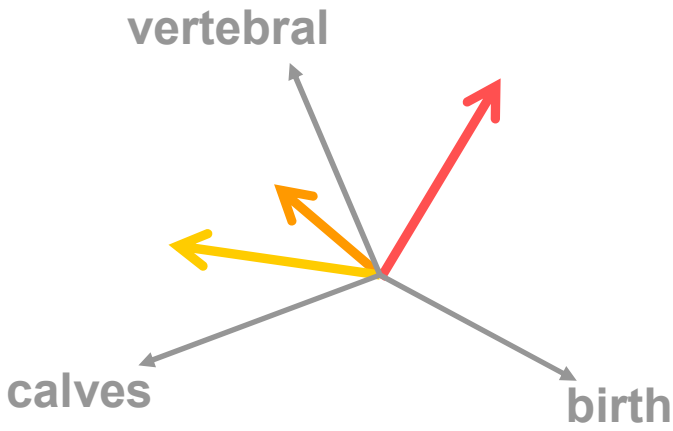
Complex vertebral malformation in Holstein calves

Jorgen S. Agerholm, Christian Bendixen, Ole Andersen, Jens Arnbjerg

Abstract. A recently observed lethal congenital defect of purebred Holstein *calves* is reported. Eighteen genetically related *calves* were necropsied. One calf had been aborted on gestation day 159, and the others were delivered between day 250 and day 285. *Birth* weights were reduced. The defect was characterized by shortening of the cervical and thoracic parts of the *vertebral* column due to multiple hemivertebrae, fused andmisshaped vertebrae, and scoliosis. Symmetrical flexures of the carpal joints and the metacarpophalangeal joint in combination with a slight lateral rotation of the phalanges also were present. Similar low-grade arthrogryposis was present in the posterior limbs. Fifty percent of the *calves* had heart malformation. Other malformations occurred in a few *calves*. Complex *vertebral* malformation (CVM) is proposed as the designation for this defect. A genetic etiology is indicated because cases occurred following breeding between genetically related individuals. Two common ancestors were found; both were former elite sires of US Holstein origin. Because of the widespread international use of semen from sires occurring in the pedigrees of affected *calves*, CVM is expected to occur in several countries.



Word	Count
calves	5
birth	1
...	...
vertebral	2

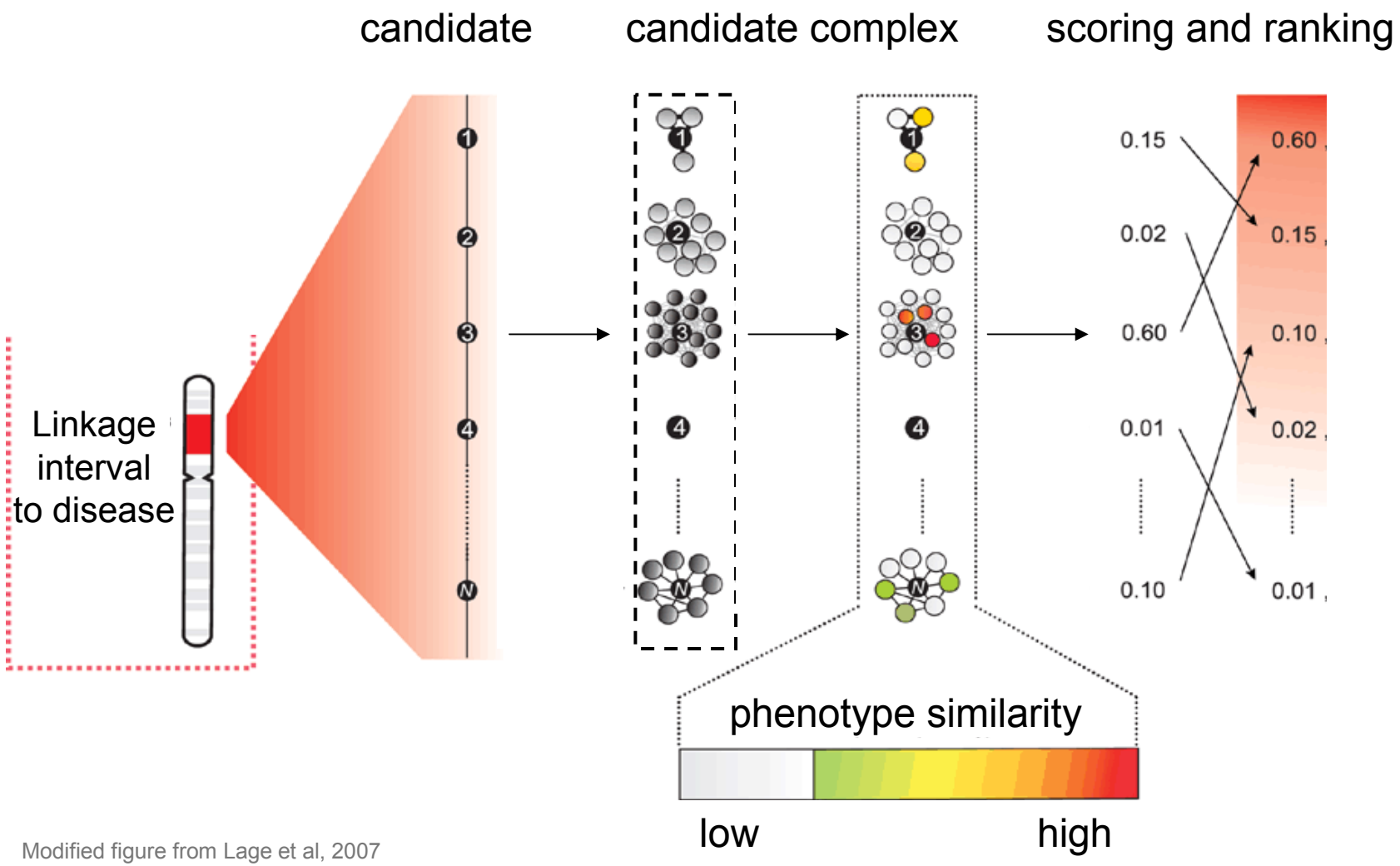


AARHUS UNIVERSITET

SABRE

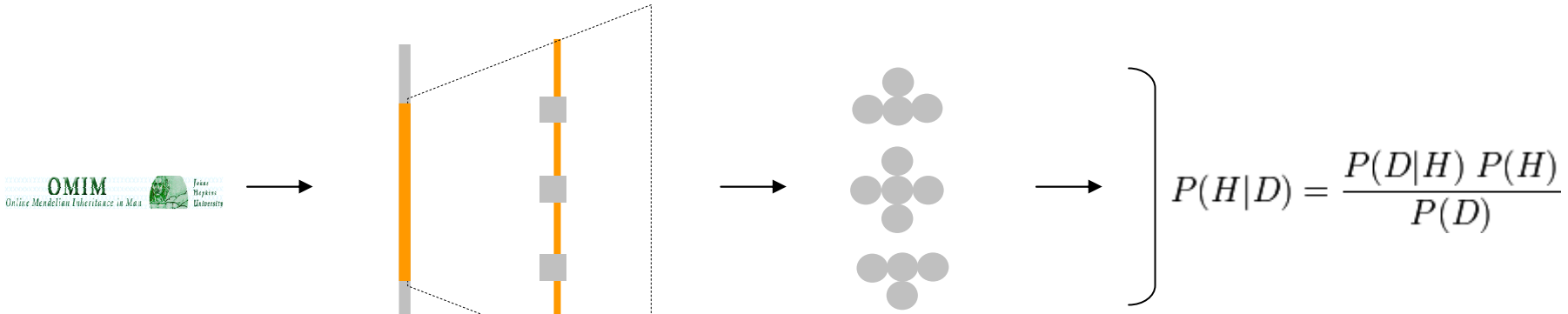
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Gene prioritization using phenome-interactome



Modified figure from Lage et al, 2007

Challenges and solutions for livestock species



phenotypic description

phenotype-genotype links

scoring model

Free text of reference-dependent

GeneRIF, PubMed abstract

z-score
random-set
enrichment scoring

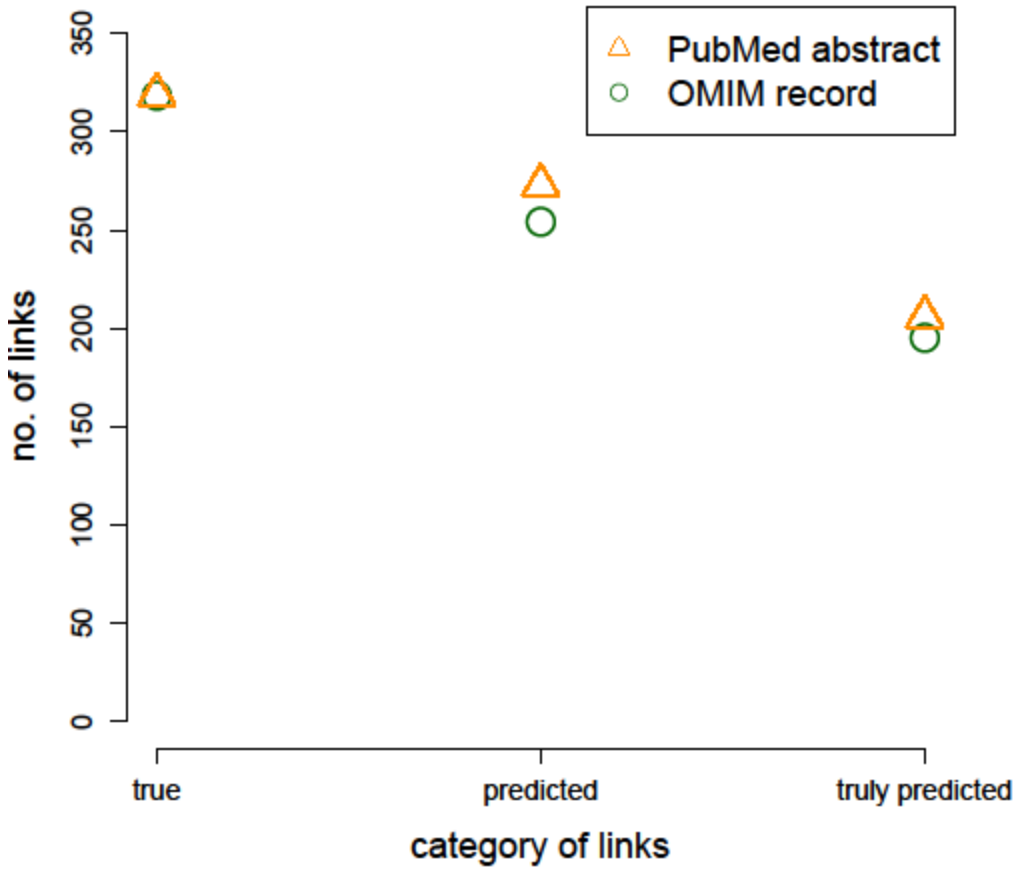
PubMed
A service of the National Library of Medicine and the National Institutes of Health
www.pubmed.gov

WIKIPEDIA
The Free Encyclopedia

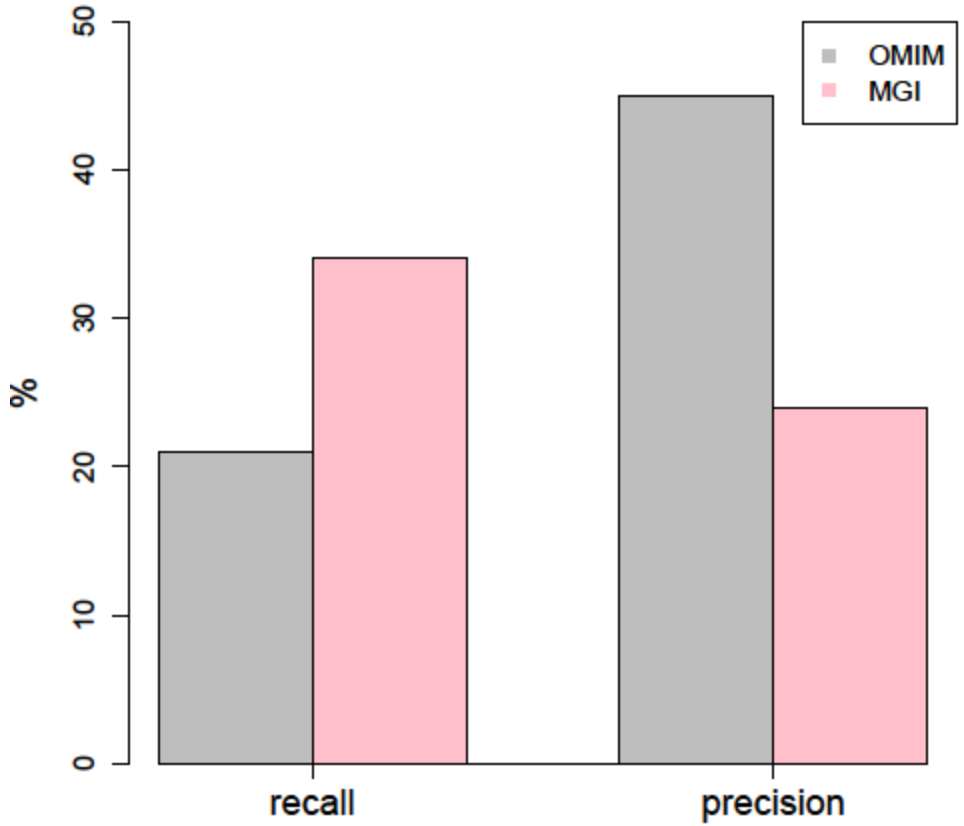
AARHUS UNIVERSITET

SABRE
CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Replace phenotypic description

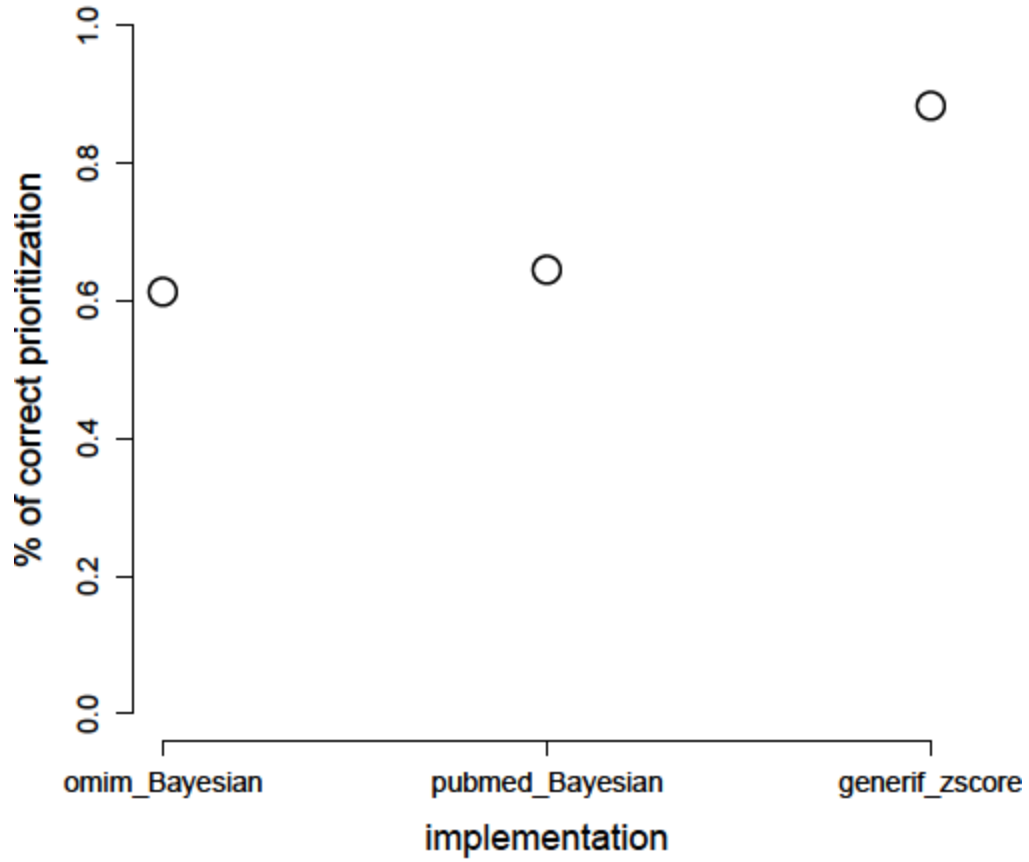


Test Bayesian posterior model in other species



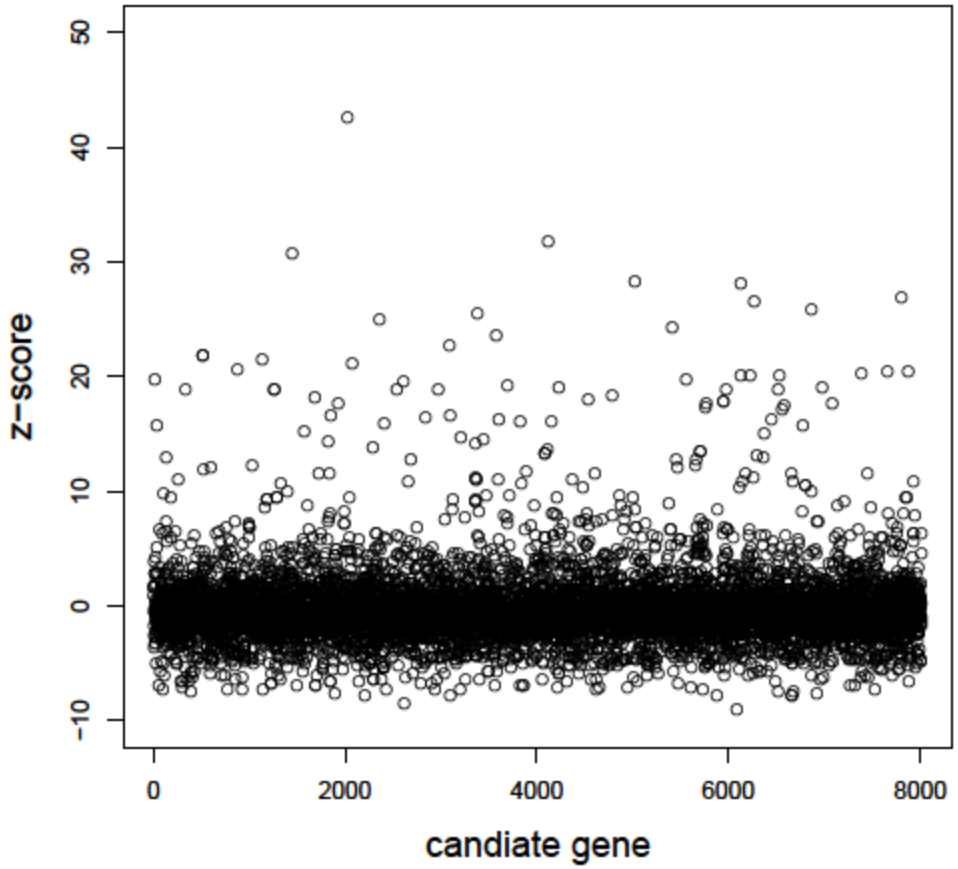
OMIM data is from Lage et al, 2007

Performance of different implementations



Application in bovine mastitis

- WikiPedia
- GeneRIF
- z-score



Conclusions

- Integrating phenotypic text and protein interactions is powerful in candidate genes prioritization regardless of species.
- Data mining reference-dependent text such as PubMed abstracts, GeneRIF and WikiPedia articles provides useful phenotypic information for both traits and genes.
- Less-parametric model using z-score is optimal for flexible utilization of heterogeneous resources.

Acknowledgements

- Peter Sørensen (AU)
- Chris Workman (DTU)
- Bernt Guldbrandtsen (AU)
- These results are part of the SABRE research project that has been co-financed by the European Commission, within the 6th Framework Programme, contract No. FOOD-CT-2006-016250.