

Using QTLs as a genomic feature in Duroc pigs

Pernille Sarup







Genome partitioning of complex traits

- a way integrating multiple layers of experimental and bioinformatic data







PigQTLdb

Browse Search View Maps

CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

Pig QTL data summary

As of *Aug 07,2014*, there have been **10,497** pig QTLs curated into the Pig QTLdb from **416** publications representing **647** different pig traits (*These numbers may have included preliminary releases of newly curated data, therefore may appear to differ from the official release numbers on the main page. They will become consistent on the next official release).*

Top 15 QTL Traits

1. Number of QTL by Traits

- 2. Number of QTL by Trait Types
- 3. Number of QTL by Trait Classes
- 4. Number of QTL by Chromosomes
- 5. Number of QTL by Publishing Years
- 6. Number of QTL by Publishing Journals
- 7. Number of Papers by Publishing Years
- 8. Number of QTL that are private
- 9. Types of structural genome information aligned against the QTL maps

Number of QTL Traits Drip loss 1,017 Loin muscle area 256 216 Average backfat thickness Backfat at last rib 202 186 Average daily gain 175 Age at puberty Shear force 175 Backfat at tenth rib 155 Intramuscular fat content 155 146 Backfat at rump 137 Total number born Carcass length 135 Body weight (birth) 110 Total number born alive 106 LDL cholesterol 104







QTL sets

QTL regions varies in size midpoint ± 250 kb was used

Vertebrate Trait Ontology > 10,000 individual QTL's from 416 studies -> 140 QTL sets







Does this information help?

Evaluation criteria:

Genomic heritability Predictive ability in a breeding relevant setup







DATA

Data from the Danish national pig test centre at Bøgildgård

34,425 Duroc boars with phenotype: Feed Efficiency







Modeling of the phenotype

Effect of environment removed from "Raw" phenotype

 $\mathbf{y} = Sw + Section + \mathbf{pen} + \mathbf{litter} + \mathbf{e}$

Residuals from this model used in the following models







Data for genomic models

Genotypes for 3085 boars SNP60 BeadChip (33,029 SNPs)

Split into Training and Validation set:

Training set: 1814 Duroc boars born 1998-2010







"Simple" modeling of the training dataset using REML:

$$\mathbf{y} = \mathbf{\mu} + \mathbf{g} + \mathbf{e}$$

$$\mathbf{g} \sim \mathrm{N}(0, \mathbf{G}\sigma_g^2)$$
$$\mathbf{e} \sim \mathrm{N}(0, \mathbf{I}\sigma_e^2)$$







"Simple" modeling of the data using REML:

G was constructed using all markers

$\mathbf{G} = \mathbf{W}\mathbf{W}'/\mathbf{m}$

W: centred and scaled genotype matrix m: number of markers







Back solving:

Estimating the individual SNP effects

$$\hat{\mathbf{s}} = \mathbf{W}'(\mathbf{W}\mathbf{W}')^{-1}\hat{\mathbf{g}}$$









- i: individual markers
- $\hat{\mathbf{g}}_i$: genetic value associated to the marker







QTL set summery statistic

$T_{CVSi} = \hat{\mathbf{g}}' \hat{\mathbf{g}}_i = \hat{\mathbf{g}}' \mathbf{W}_i \hat{\mathbf{s}}_i$







QTL set_k









Permutations



Programme (FP7 2007/2013) under grant agreement No. 311794.



Genomic feature model

$$\mathbf{y} = \mathbf{\mu} + \mathbf{g}_1 + \mathbf{g}_2 + \mathbf{e}$$

 $\mathbf{g_1} \sim \mathrm{N}(0, \mathbf{G_1}\sigma_{g_1}^2)$ $\mathbf{g_2} \sim \mathrm{N}(0, \mathbf{G_2}\sigma_{g_2}^2)$ $\mathbf{e} \sim \mathrm{N}(0, \mathbf{I}\sigma_e^2)$







Definition of the genomic feature

- g₁: random genetic effect of markers associated with significant QTLs
- g₂ : random genetic effect of remaining markers

A range of 11 different cut off P-values [0.01-0.9] determined whether a QTL set was included in g_1





Data for gmf models

Genotypes for 3085 boars SNP60 BeadChip (33,029 SNPs)

Split into Training and Validation set:

Training set: 1814 Duroc boars born 1998-2010 Validation set: 1271 Duroc boars born 2012-2014







Results





CENTER FOR QUANTITATIVE GENETICS AND GENOMICS









Predictive Ability

in a breeding relevant setup:

Training set: 1814 Duroc boars born 1998-2010 Validation set: 1271 Duroc boars born 2012-2014

At least 1 generation between boars in training and validation set







Predictive Ability

Correlation between y and $\widehat{\mathbf{g}} = \widehat{\mathbf{g}_1} + \widehat{\mathbf{g}_2}$ for 100 random subsamples of $\frac{1}{5}$ of the validation set

NOT cross validation!















COOPERATION



Conclusion

Including QTL set information makes little difference to genomic heritability

Still it improved PA by 2 percentage points compared to normal gBLUP (17% increase).







Thank you for your attention

Peter Sørensen, Palle D Rhode, Stefan M. Edwards, Just Jensen, Tage Osterse, Mark Henryon, Ole Fredslund Christensen





For providing the data

Funded by the European Union Seventh Framework Programme (FP7 2007/2013) under grant agreement No. 311794.







Trait	# markers	Ρ
leukocyte quantity	314	0.0083
glycolytic potential	129	0.0104
ejaculation trait	25	0.0124
blood leptin amount	26	0.0126
cannon bone circumference	76	0.0143
eating behavior trait	52	0.0171
hindlimb muscle mass	94	0.0375
exploratory behavior trait	20	0.0431
limb conformation trait	11	0.0474
white adipose amount	31	0.0478



