FBN

LEIBNIZ INSTITUTE
FOR FARM ANIMAL BIOLOGY

Member of Leibniz Association

69th Annual Meeting of the European Federation of Animal Science,
Dubrovnik, Croatia, 27th to 31st August 2018

Posterior Probabilities of Genome-Wide Associations in Backcross Families

Manuela Reichelt, Manfred Mayer, Friedrich Teuscher, Norbert Reinsch

Leibniz Institute for Farm Animal Biology

WPPA

WPPA – Window Posterior Probability of Association

$\overline{m}$ - Vector of means of estimated posterior marker effects
$\hat{m}$ - Vector of estimated marker effects in one MCMC iteration

$\overline{sq} = {\Sigma_i \overline{m}_i^2}/{p}$ - Expected genetic variance reflected by markers (ignoring Linkage Disequilibrium)

Ratio: $q = \frac{observation\ in\ iteration\ t}{expected\ variance}$

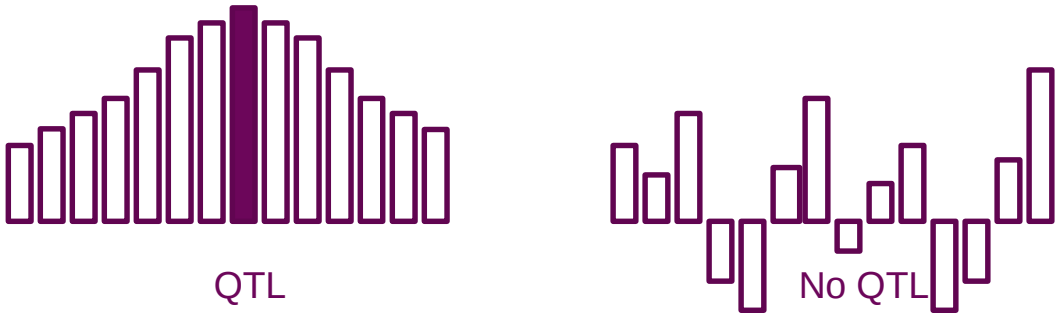
q is sampled
Posterior probabilities for q in a certain interval (window) is determined
$q > 1$ - Association

WPPA

Why windows?

Effect of a QTL is reflected by several adjacent markers jointly

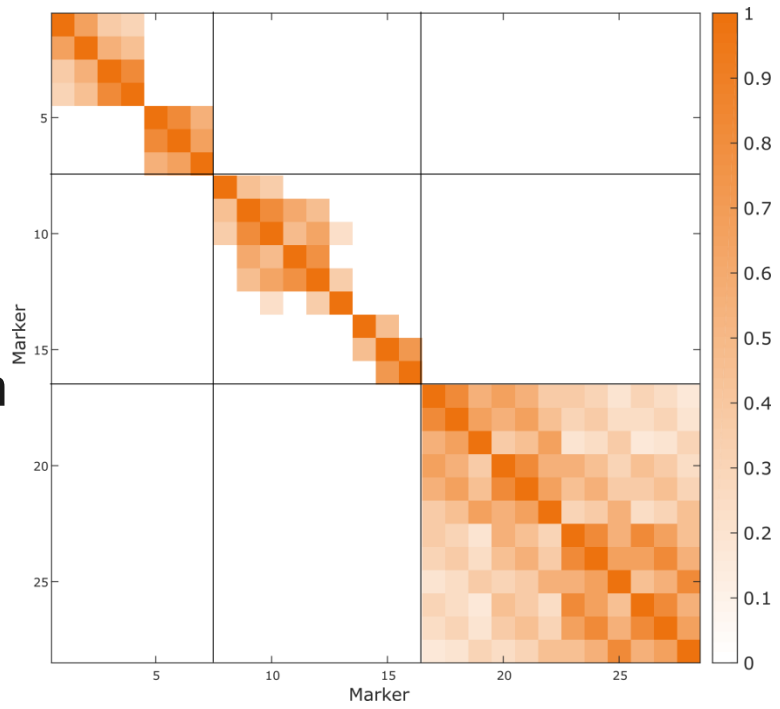
Marker effects



Windows reflect the effect of chromosome segments
No Linkage Disequilibrium – no covariance between markers is considered

WPPA-R

- In certain types of populations linkage disequilibrium is large and wiedespread across chromosomes
- Such as in a backcross between inbred lines
- Covariance matrix $\boldsymbol{R}$ can be defined (Bonk et al., 2016)
- Reflects genetic distances between markers
- Linkage phase of markers in parents



WPPA-R

Matrix of thresholds
Example: 4 markers, equal distances of 0.01 M

$$\boldsymbol{R} = \begin{bmatrix} 1 & 0.98 & 0.96 & 0.94 \\ 0.98 & 1 & 0.98 & 0.96 \\ 0.96 & 0.98 & 1 & 0.98 \\ 0.94 & 0.96 & 0.98 & 1 \end{bmatrix}$$

$$\boldsymbol{T} = \boldsymbol{R}^0 \cdot \overline{cp} + \boldsymbol{I} \cdot \overline{sq} = \begin{bmatrix} 0 & 0.98 & 0.96 & 0.94 \\ 0.98 & 0 & 0.98 & 0.96 \\ 0.96 & 0.98 & 0 & 0.98 \\ 0.94 & 0.96 & 0.98 & 0 \end{bmatrix} \cdot \overline{cp} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \overline{sq}$$

$\overline{sq} = {\sum_i \bar{m}_i^2}/{p}$ (average genetic variance per marker)

$\overline{cp} = {\sum_i \sum_j \bar{m}_i \bar{m}_j}/{(p^2 - p)}$ (total covariance is equally distributed across all marker combinations)

WPPA-R

Matrix of threshold values for windows
Example: 3 markers in a window, 4 windows

$$\boldsymbol{K} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\boldsymbol{T}_w = \boldsymbol{K}' \cdot \boldsymbol{T} \cdot \boldsymbol{K}$$

Matrix of expected variances and covariances of all defined windows and their combinations

WPPA-R

Observation in MCMC iteration t, estimated marker effects $\hat{m}_i$
Example: 3 markers in a window, 4 windows

$$\boldsymbol{K}_t = \begin{bmatrix} \hat{m}_1 & \hat{m}_1 & 0 & 0 \\ \hat{m}_2 & \hat{m}_2 & \hat{m}_2 & 0 \\ 0 & \hat{m}_3 & \hat{m}_3 & \hat{m}_3 \\ 0 & 0 & \hat{m}_4 & \hat{m}_4 \end{bmatrix}$$

$$\boldsymbol{T}_{wt} = \boldsymbol{K}_t' \cdot \boldsymbol{R} \cdot \boldsymbol{K}_t$$

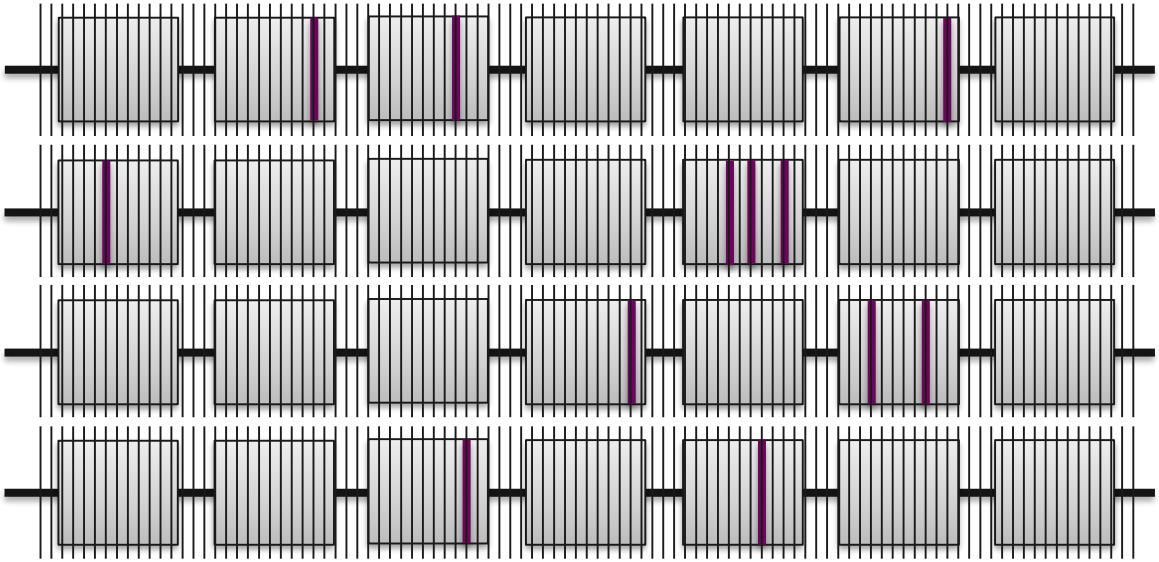
Matrix of observed variances and covariances of all defined windows and their combinations

Simulation

„Mouse“ genome

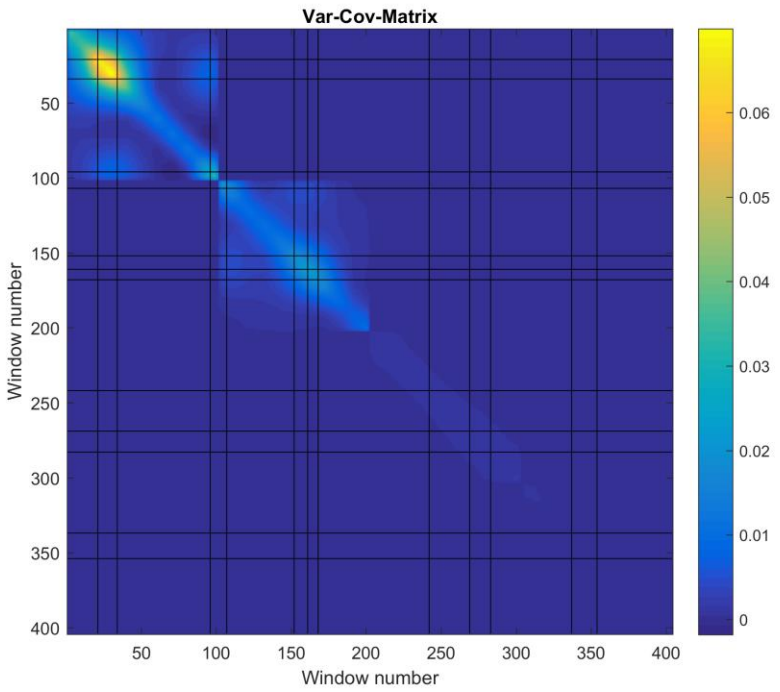
20 chromosomes

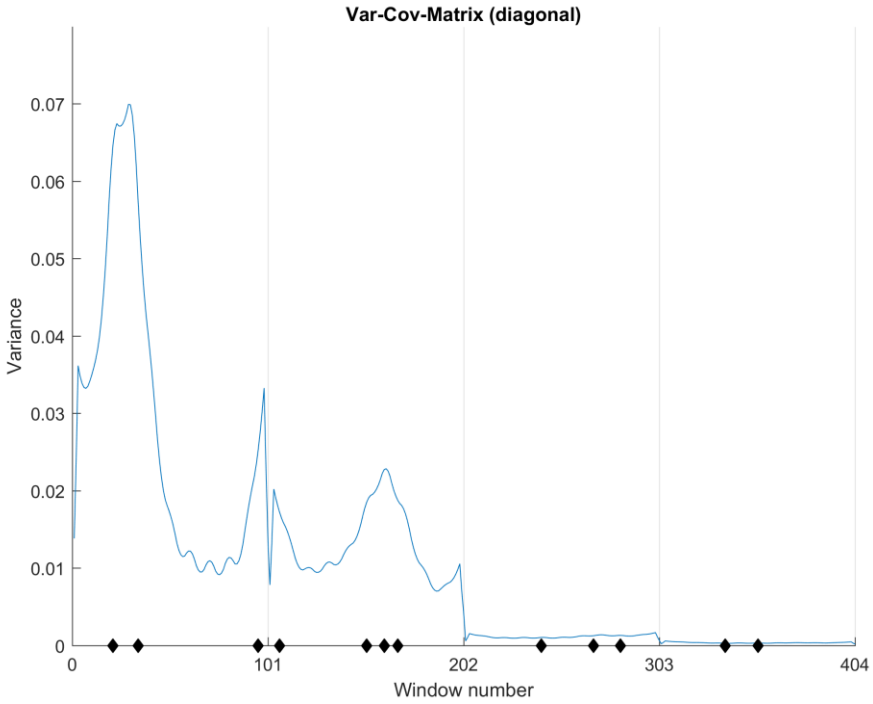
Length of 1 M
101 equally spaced markers
12 QTLs
Heritabilities $h^2 \in \{0.17,\ 0.29,\ 0.70\}$
200 experiments



Simulation $h^2 = 0.17$

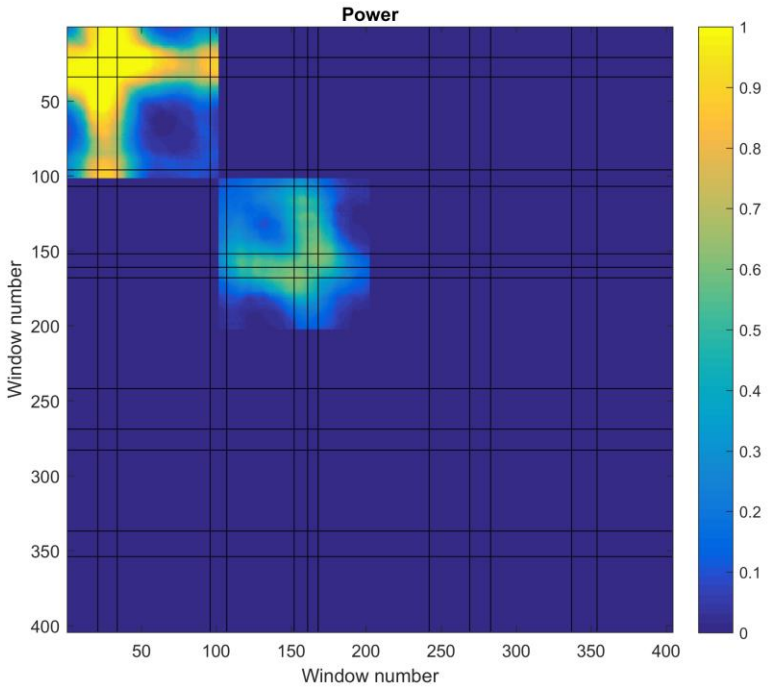
Mean (over all repetitions) estimated variances and covariances

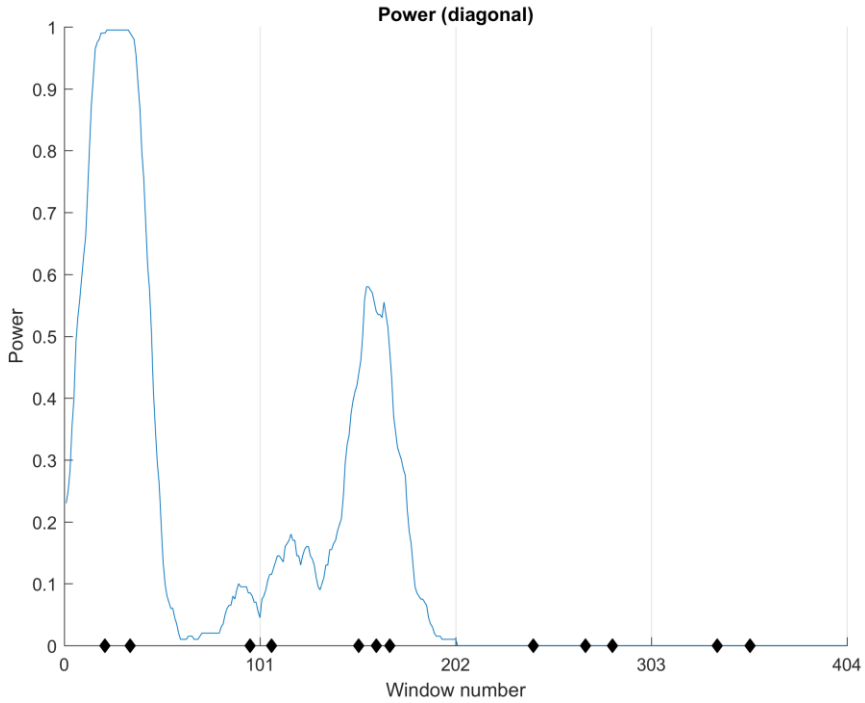




Simulation $h^2 = 0.17$

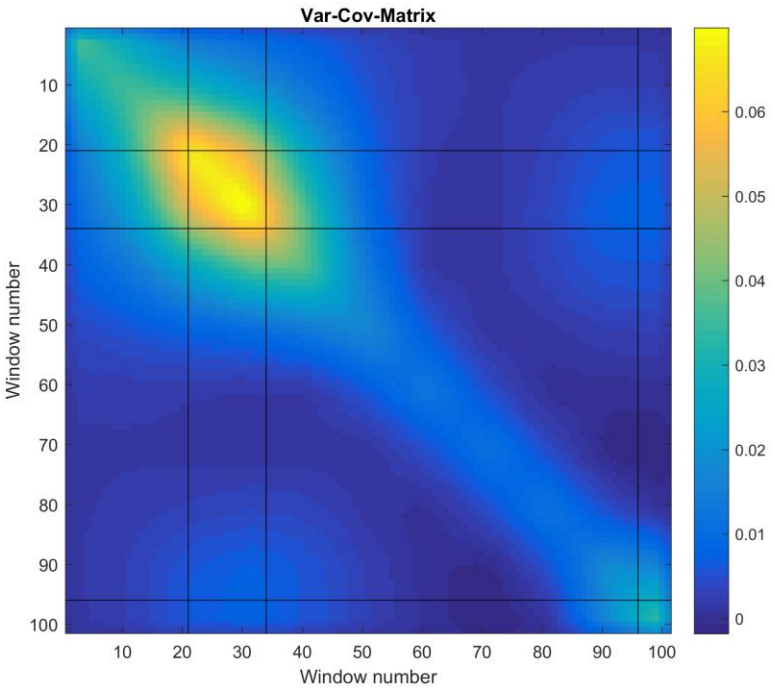
Mean (over all repetitions) posterior probabilities to exceed the threshold („Power“)



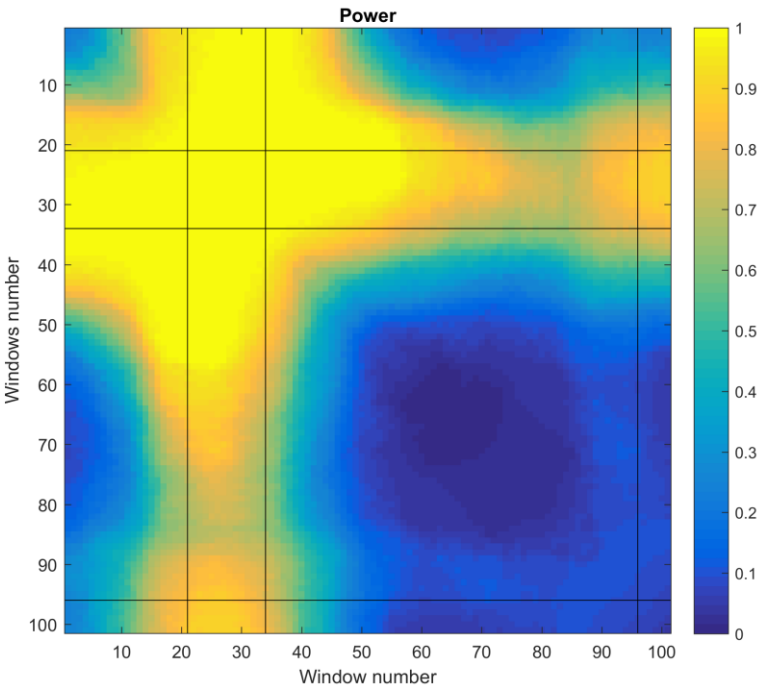


Simulation $h^2 = 0.17$

Estimated variances and covariances



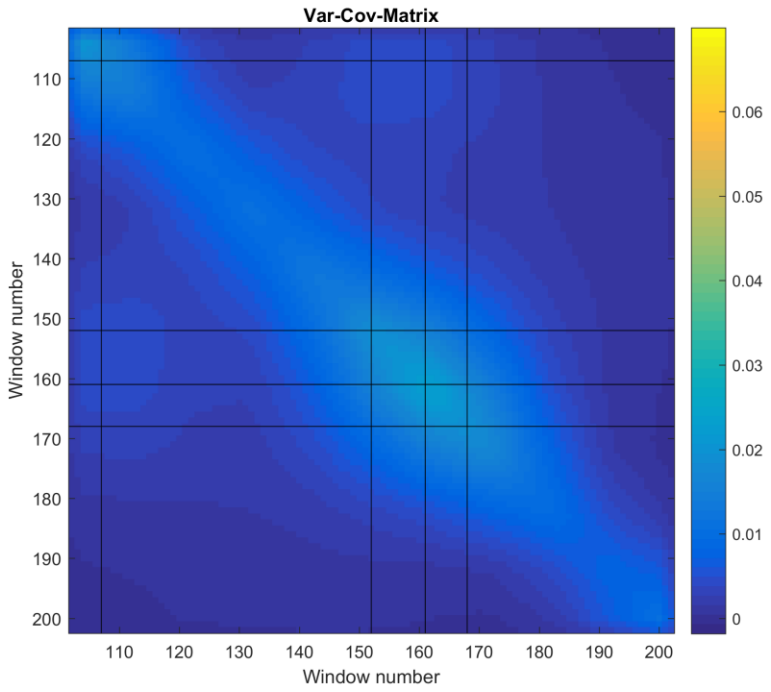
Posterior probabilities



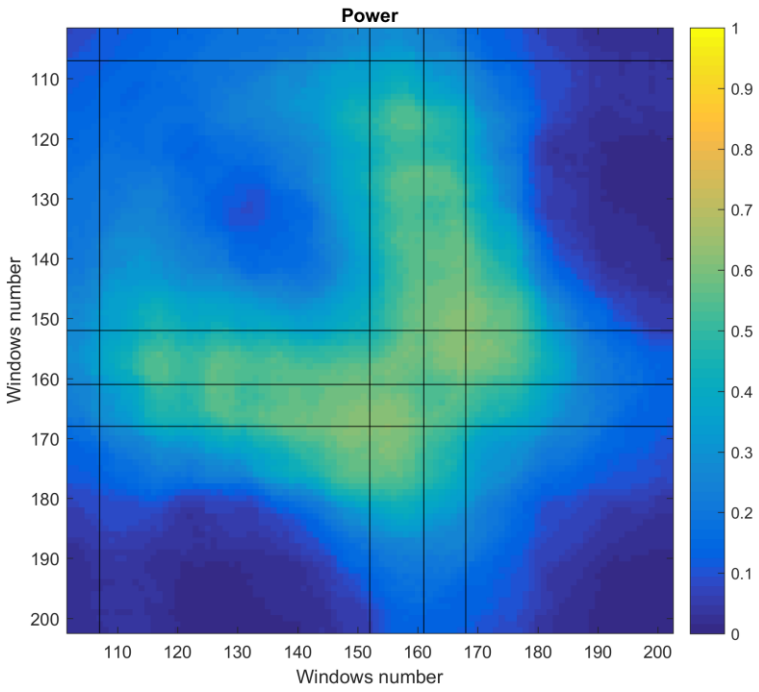
First chromosome: 3 marker effects: 2.0, 2.0, 1.0

Simulation $h^2 = 0.17$

Estimated variances and covariances



Posterior probabilities



Second chromosome: 4 marker effects: 1.0, 1.0, 0.5, 1.0

Simulation $h^2 = 0.17$

scenario			Max. Power Chr. 1 (3 marker effects: 2.0, 2.0, 1.0)			Max. Power Chr. 2 (4 marker effects: 1.0, 1.0, 0.5, 1.0)		
Nr.	h^2	cM	IN	IB_{1_w}	DB_{1_w}	IN	IB_{1_w}	DB_{1_w}
1	0.17	5	0.07	0.98	1.00	0.02	0.39	0.67
2	0.17	1	0.01	0.66	1.00	0.01	0.16	0.58
4	0.29	5	0.23	1.00	1.00	0.03	0.73	0.94
5	0.29	1	0.02	0.92	1.00	0.01	0.38	0.87
7	0.70	5	0.98	1.00	1.00	0.37	1.0	1.0
8	0.70	1	0.15	1.00	1.00	0.03	0.95	1.0

Simulation $h^2 = 0.17$

scenario			Max. Power Chr. 1 (3 marker effects: 2.0, 2.0, 1.0)			Max. Power Chr. 2 (4 marker effects: 1.0, 1.0, 0.5, 1.0)		
Nr.	h^2	cM	IN	IB_{1_w}	DB_{1_w}	IN	IB_{1_w}	DB_{1_w}
1	0.17	5	0.07	0.98	1.00	0.02	0.39	0.67
2	0.17	1	0.01	0.66	1.00	0.01	0.16	0.58
4	0.29	5	0.23	1.00	1.00	0.03	0.73	0.94
5	0.29	1	0.02	0.92	1.00	0.01	0.38	0.87
7	0.70	5	0.98	1.00	1.00	0.37	1.0	1.0
8	0.70	1	0.15	1.00	1.00	0.03	0.95	1.0

Conclusions and Outlook

- Different methods can be analysed and compared with the same calculations

- Three methods are known:
  1. Inferences drawn of posterior probabilities of marker effects
  2. WPPA
  3. WPPA-R

- Better visualisation of the representation of the genetic architecture with WPPA and WPPA-R

- Covariances help with the analysis of this architectures

Thank you for your attention!

Thank you for your attention

Literature:

- **Bonk S.**, Reichelt M., Teuscher F., Segelke D., Reinsch N. (2016): *Mendelian sampling covariability of marker effects and genetic values.* Genet Sel Evol 48: 36

- **Fernando R.L.**, Toosi A., Garrick D.J., Dekkers J.C.M. (2014) Application of Whole-Genome Prediction Methods for Genome-Wide Association Studies: a Bayesian Approach. Proceedings of the 10th World Congress of Genetics Applied to Livestock Production, 17-22 August, Vancouver, BC, Canada, 201

LEIBNIZ INSTITUTE
FOR FARM ANIMAL BIOLOGY

Leibniz-Institut für Nutztierbiologie FBN

Wilhelm-Stahl-Allee 2
18196 Dummerstorf
Germany

Contact

Manuela Reichelt

Phone: +49 38208 68 914
Fax: +49 38208 68 902
E-Mail: reichelt@fbn-dummerstorf.de
Internet: www.fbn-dummerstorf.de