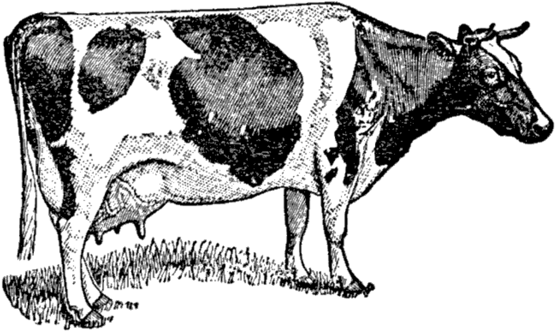
Modelling QTL on BTA06 in dairy cattle using random regression test day model

Tomasz Suchocki
&
Joanna Szyda

Wroclaw University of Environmental
and Life Sciences

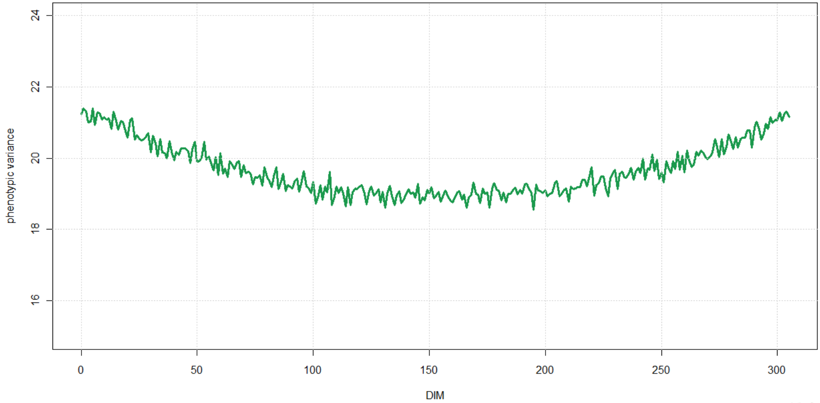
Outline

1. Motivation
2. Data Set
3. Methods
4. Results

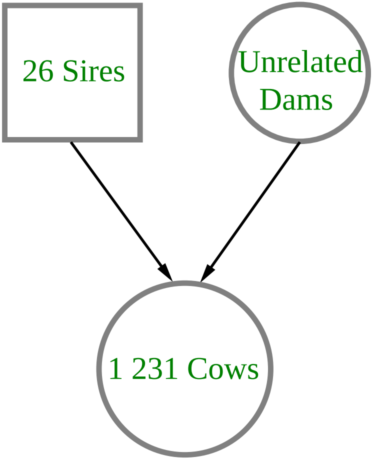


Main Aim

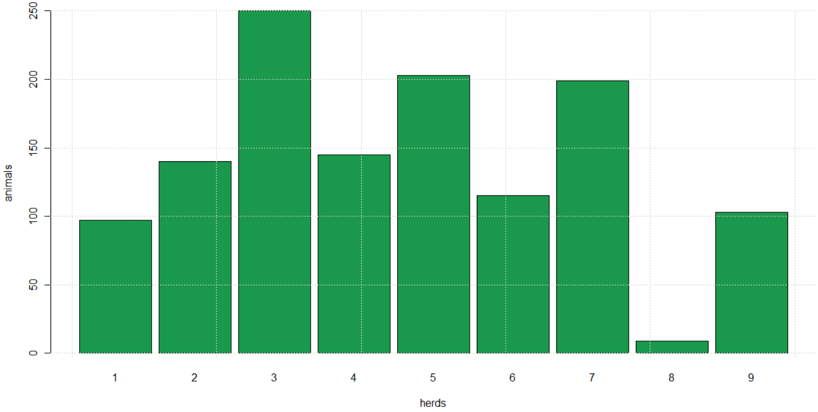
Checking if gene effects are vary or constant in time



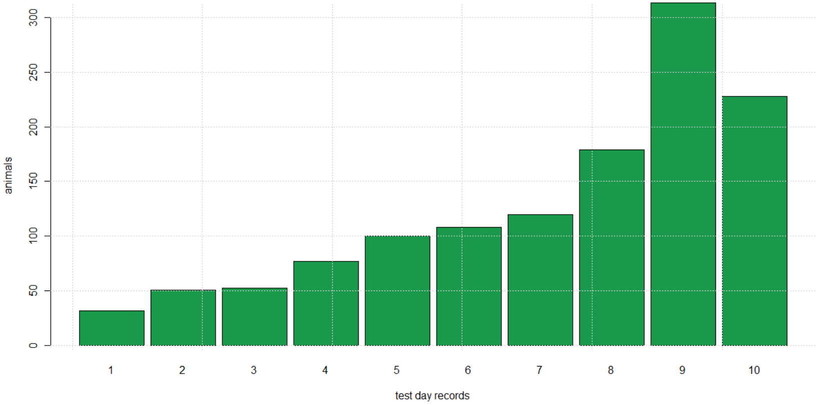
Material – animals



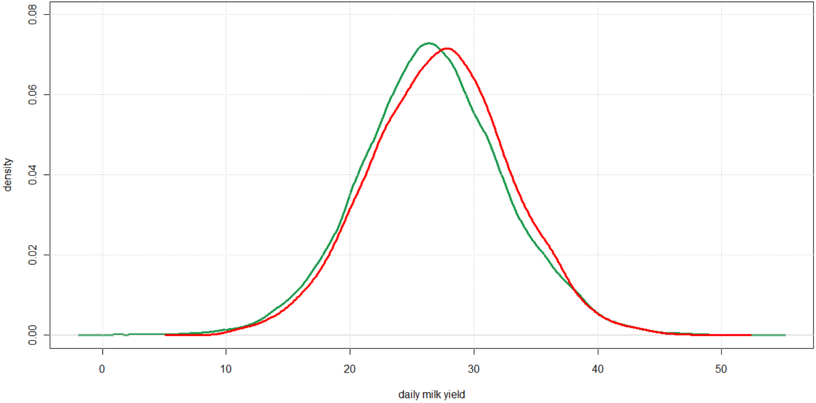
Material – animals



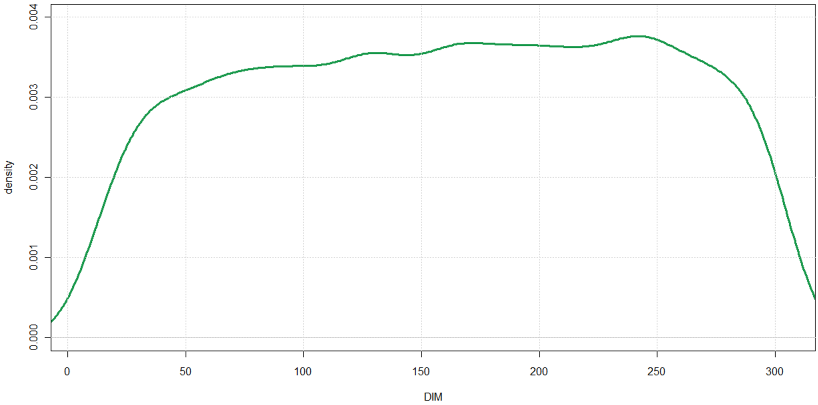
Material – animals

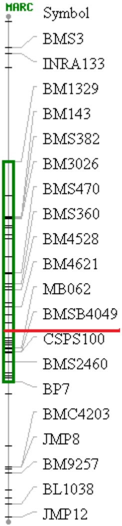


Material – animals



Material – animals





No.	Marker	Map distance (cM)	Alleles
1	BM1329	0.00	6
2	BMS2508	8.54	7
3	BMS1242	17.44	7
4	BM143	18.33	7
5	BMS518	23.57	4
6	BM4322	28.47	7
7	BMS470	32.00	4
8	ILSTS097	37.04	2
9	RM028	43.79	4
10	BM415	46.56	7
11	ILSTS035	51.87	9
12	ILSTS087	54.34	3
13	BMS2460	58.06	4
14	BP7	63.50	5

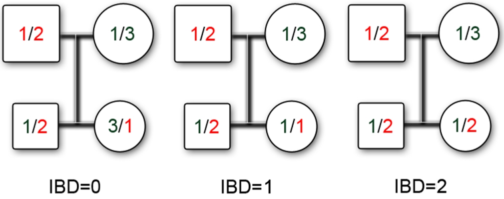
First model – α and ρ constant in time

$$y = X\beta + Z_\alpha\alpha + Z_\rho\rho + \epsilon,$$

where

- y – vector of lactation milk yields
- β – vector of fixed effects
- α – random additive polygenic effect, $\alpha \sim \mathcal{N}\left(0, A\sigma_\alpha^2\right)$
- ρ – random QTL effect, $\rho \sim \mathcal{N}\left(0, IBD\sigma_\rho^2\right)$
- ϵ – residual, $\epsilon \sim \mathcal{N}\left(0, I\sigma_\epsilon^2\right)$

Technical aspects – IBD



- Calculated using Citius ($http : //jay.au.poznan.pl/ \sim mcszyd/citius/citius.pdf$)
- Coefficient of relatedness = expected fraction of alleles shared that are IBD

$$\lambda = \frac{1}{2} P(IBD = 1) + P(IBD = 2)$$

Second model – α varies in time

$$y_{ij} = \beta + \sum_{m=0}^{k_\alpha} \alpha_{im}\phi_m(\tau_{ij}) + \rho_i + \xi_i + \epsilon_{ij},$$

where

- y_{ij} – jth test day record for individual i
- $\phi_m(\tau_{ij})$ – mth Legendre polynomial at time point τ_{ij}
- α_{im} – mth random regression coefficient for additive polygenic effect, $\alpha \sim \mathcal{N}(0, A \otimes G_\alpha)$
- ξ – random permanent environmental effect, $\xi \sim \mathcal{N}\left(0, I\sigma_\xi^2\right)$

Third model – α, ρ and ξ vary in time

$$y_{ij} = \beta + \sum_{m=0}^{k_\alpha} \alpha_{im}\phi_m(\tau_{ij}) + \sum_{m=0}^{k_\rho} \rho_{im}\phi_m(\tau_{ij}) + \sum_{m=0}^{k_\xi} \xi_{im}\phi_m(\tau_{ij}) + \epsilon_{ij},$$

where

- ρ_{im} – mth random regression coefficient for QTL effect, $\rho \sim \mathcal{N}\left(0, IBD \otimes G_\rho\right)$
- ξ_{im} – mth random regression coefficient for permanent environmental effect, $\xi \sim \mathcal{N}\left(0, I \otimes P_\xi\right)$

Likelihood

- Model 1

$$
\begin{aligned}
-2\log L &= \mathrm{const} + n\log\sigma_\epsilon^2 + \log|A| + n\log\sigma_\alpha^2 + \log|IBD| + \\
&+ n\log\sigma_\rho^2 + n\log\sigma_\xi^2 + log|C| + y^T S y
\end{aligned}
$$

Likelihood

- Model 1

$$
\begin{aligned}
-2\log L &= \text{const} + n\log\sigma_\epsilon^2 + \log|A| + n\log\sigma_\alpha^2 + \log|IBD| + \\
&+ n\log\sigma_\rho^2 + n\log\sigma_\xi^2 + log|C| + y^T S y
\end{aligned}
$$

- Model 2

$$
\begin{aligned}
-2\log L &= \text{const} + n\log\sigma_\epsilon^2 + 4\log|A| + n\log|G_\alpha| + \log|IBD| + \\
&+ n\log\sigma_\rho^2 + n\log\sigma_\xi^2 + log|C| + y^T S y
\end{aligned}
$$

Likelihood

- Model 1

$$
\begin{aligned}
-2\log L &= \text{const} + n\log\sigma_\epsilon^2 + \log|A| + n\log\sigma_\alpha^2 + \log|IBD| + \\
&+ n\log\sigma_\rho^2 + n\log\sigma_\xi^2 + log|C| + y^T S y
\end{aligned}
$$

- Model 2

$$
\begin{aligned}
-2\log L &= \text{const} + n\log\sigma_\epsilon^2 + 4\log|A| + n\log|G_\alpha| + \log|IBD| + \\
&+ n\log\sigma_\rho^2 + n\log\sigma_\xi^2 + log|C| + y^T S y
\end{aligned}
$$

- Model 3

$$
\begin{aligned}
-2\log L &= \text{const} + n\log\sigma_\epsilon^2 + 4\log|A| + n\log|G_\alpha| + 4\log|IBD| + \\
&+ n\log|G_\rho| + n\log|G_\xi| + log|C| + y^T S y
\end{aligned}
$$

- $S = V^{-1} - V^{-1}X(X^T V^{-1} X)^{-1} X^T V^{-1}$

Technical aspects – EM

- Model 1 – Estimating 3 parameters

$$\sigma_\alpha^2,\ \sigma_\rho^2 \text{ and } \sigma_\epsilon^2$$

- Model 2 – Estimating 13 parameters

$$G_\alpha \text{ (Matrix } 4 \times 4\text{)},\ \sigma_\rho^2,\ \sigma_\xi^2 \text{ and } \sigma_\epsilon^2$$

- Model 3 – Estimating 31 parameters

$$G_\alpha,\ G_\rho,\ P_\xi \text{ and } \sigma_\epsilon^2$$

Technical aspects – EM

- Model 1 – Estimating 3 parameters

$$\sigma_\alpha^2,\ \sigma_\rho^2 \text{ and } \sigma_\epsilon^2$$

- Model 2 – Estimating 13 parameters

$$G_\alpha \text{ (Matrix } 4 \times 4\text{)},\ \sigma_\rho^2,\ \sigma_\xi^2 \text{ and } \sigma_\epsilon^2$$

- Model 3 – Estimating 31 parameters

$$G_\alpha,\ G_\rho,\ P_\xi \text{ and } \sigma_\epsilon^2$$

Gaussian elimination

EM – estimators

- $$\sigma_\alpha^{2[t+1]} = \frac{\widehat{\alpha}'^{[t]} A^{-1} \widehat{\alpha}^{[t]} + tr(A^{-1} C_\alpha^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$$

- $$\sigma_\rho^{2[t+1]} = \frac{\widehat{\rho}'^{[t]} IBD^{-1} \widehat{\rho}^{[t]} + tr(IBD^{-1} C_\rho^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$$

- $$\sigma_\xi^{2[t+1]} = \frac{\widehat{\xi}'^{[t]} \widehat{\xi}^{[t]} + tr(C_\xi^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$$

EM – estimators

- $\sigma_\alpha^{2[t+1]} = \dfrac{\widehat{\alpha}'^{[t]} A^{-1} \widehat{\alpha}^{[t]} + tr(A^{-1} C_\alpha^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$
- $G_{\alpha(kl)}^{[t+1]} = \dfrac{\widehat{\alpha}_k'^{[t]} A^{-1} \widehat{\alpha}_l^{[t]} + tr(A^{-1} C_{\alpha_k \alpha_l}^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$
- $\sigma_\rho^{2[t+1]} = \dfrac{\widehat{\rho}'^{[t]} IBD^{-1} \widehat{\rho}^{[t]} + tr(IBD^{-1} C_\rho^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$
- $G_{\rho(kl)}^{[t+1]} = \dfrac{\widehat{\rho}_k'^{[t]} IBD^{-1} \widehat{\rho}_l^{[t]} + tr(IBD^{-1} C_{\rho_k \rho_l}^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$
- $\sigma_\xi^{2[t+1]} = \dfrac{\widehat{\xi}'^{[t]} \widehat{\xi}^{[t]} + tr(C_\xi^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$
- $P_{\xi(kl)}^{[t+1]} = \dfrac{\widehat{\xi}_k'^{[t]} \widehat{\xi}_l^{[t]} + tr(C_{\xi_k \xi_l}^{-1[t]}) \sigma_\epsilon^{2[t]}}{q}$

EM – estimators

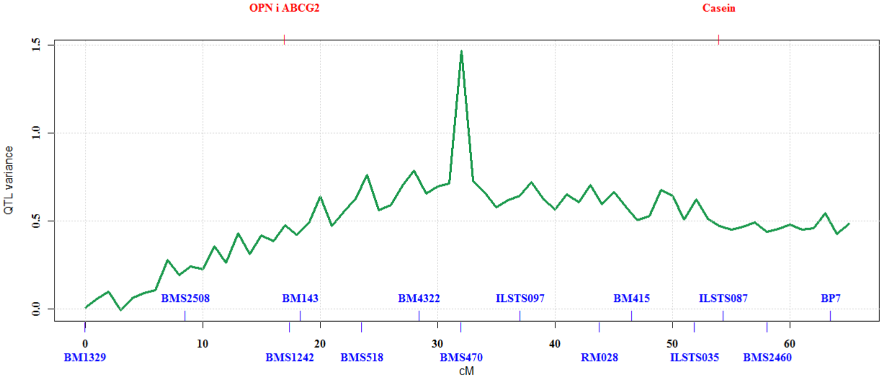
- Model 1

$$\sigma_\epsilon^{2[t+1]} = \frac{\widehat{\epsilon}'^{[t]}\widehat{\epsilon}^{[t]} + tr([X, Z_\alpha, Z_\rho]C^{-1[t]}[X, Z_\alpha, Z_\rho]')\sigma_\epsilon^{2[t]}}{n}$$

- Model 2 & 3

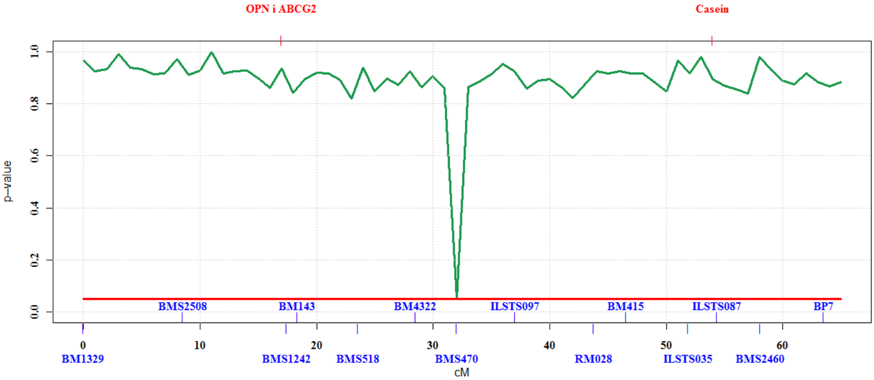
$$\sigma_\epsilon^{2[t+1]} = \frac{\widehat{\epsilon}'^{[t]}\widehat{\epsilon}^{[t]} + tr([X, Z_\alpha, Z_\rho, Z_\xi]C^{-1[t]}[X, Z_\alpha, Z_\rho, Z_\xi]')\sigma_\epsilon^{2[t]}}{n}$$

Results – Model 1

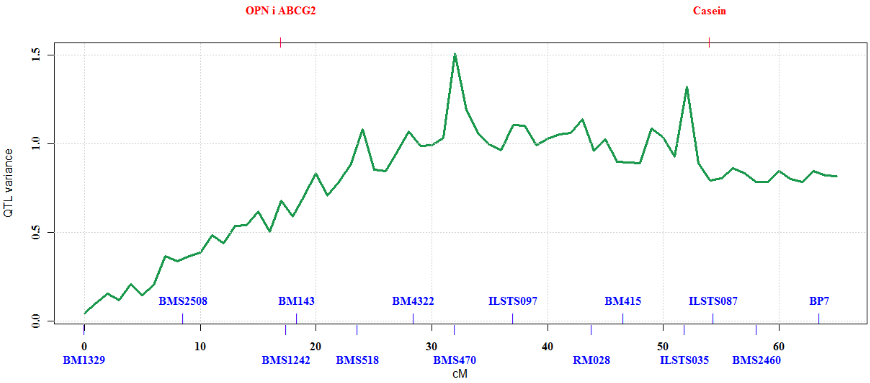


Results – Model 1

LRT $\quad H_0 : \widehat{\sigma}_q^2 = 0$ *vs.* $H_1 : \widehat{\sigma}_q^2 \neq 0$

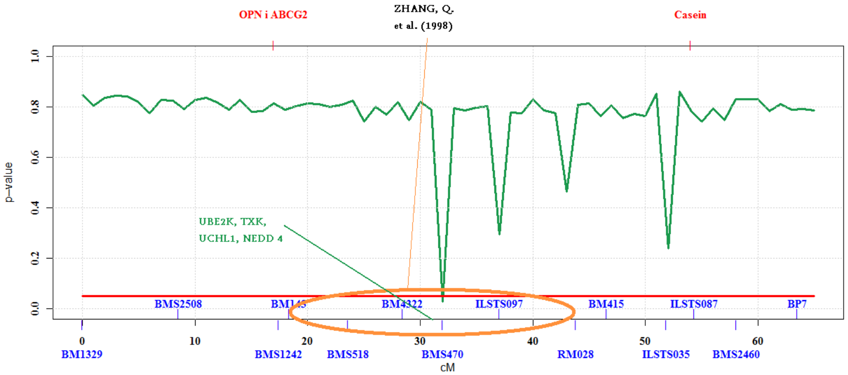


Results – Model 2



Results – Model 2

$$\text{LRT} \quad H_0 : \widehat{\sigma}_q^2 = 0 \ \textit{vs.} \ H_1 : \widehat{\sigma}_q^2 \neq 0$$



Conclusions

- QTL position for milk yield was mapped close to marker BMS470 (Model 1 & Model 2).

Conclusions

- QTL position for milk yield was mapped close to marker BMS470 (Model 1 & Model 2).

- In the region near marker BMS470, QTL affecting milk yield have been reported in several previous studies (Szyda et al. 2003, Ihara et al. 2004 and Chen et al. 2006).

Conclusions

- QTL position for milk yield was mapped close to marker BMS470 (Model 1 & Model 2).

- In the region near marker BMS470, QTL affecting milk yield have been reported in several previous studies (Szyda et al. 2003, Ihara et al. 2004 and Chen et al. 2006).

- Possible position for second QTL was mapped close to marker ILSTS 087 (Model 2).

Thank You!