

Detecting Association of SNPs and Prediction of Future Yields for Growth Trait Using Mixed Model with Orthogonal Polynomials

Tomasz Suchocki & Joanna Szyda

Wrocław University of Environmental and Life Sciences
Institute of Animal Genetics

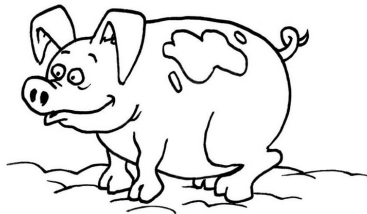


Goals

- association
- prediction growth trait at time point 600 (phenotyped and non-phenotyped animals)

Methods

- longitudinal mixed model
- Legendre polynomials for fixed and random effects



Single SNP model

$$y_i(t) = \mu(t) + x_i\alpha(t) + \xi_i(t) + \epsilon_i,$$

- $y_i(t)$ trait value for i th individual at time t
- $\alpha(t)$ additive effect of SNP at time t
- $x_i \in \{-1, 0, 1\}$ SNP genotype indicator for i th individual
- $\xi_i(t)$ permanent environmental effect for i th individual at time t
- ϵ_i error for i th individual

Methods: Basic Model

- SNP variance at time t

$$\sigma_g^2(t) = \sigma_x^2 \alpha^2(t) = \frac{1}{2} \alpha^2(t)$$

- phenotypic variance at time t

$$\sigma_p^2(t) = \sigma_g^2(t) + \sigma_\xi^2(t) + \sigma_\epsilon^2 = \frac{1}{2} \alpha^2(t) + \sigma_\xi^2(t) + \sigma_\epsilon^2$$

- heritability at time t

$$h^2(t) = \frac{\sigma_g^2(t)}{\sigma_p^2(t)} = \frac{\frac{1}{2} \alpha^2(t)}{\frac{1}{2} \alpha^2(t) + \sigma_\xi^2(t) + \sigma_\epsilon^2}$$

Orthogonal polynomials

$$\Psi(\tau) = [\Psi_0(\tau), \Psi_1(\tau), \dots, \Psi_r(\tau)]$$

- Ψ_r Legendre polynomial of order $r = 3$
- τ standardized time point $\left(\tau = 2 \cdot \frac{t-t_0}{t_{max}-t_0} - 1\right)$

Methods: Extended notation

- single SNP model with Legendre polynomials

$$y_i(\tau) = \Psi(\tau)\mu + x_i\Psi(\tau)\alpha + \Psi(\tau)\xi_i + \epsilon_i$$

ξ_i is i.i.d. $\mathcal{N}(0, \Sigma)$, where Σ is a $(r + 1) \times (r + 1)$ covariance matrix

- heritability

$$h^2(\tau) = \frac{\Psi(\tau) \left[\frac{1}{2} \alpha \alpha^T \right] \Psi^T(\tau)}{\Psi(\tau) \left[\frac{1}{2} \alpha \alpha^T + \Sigma \right] \Psi^T(\tau) + \sigma_\epsilon^2}$$

- estimation $[\mu, \alpha, \Sigma, \sigma_\epsilon^2]$ using EM algorithm
- 615 iterations (on average)

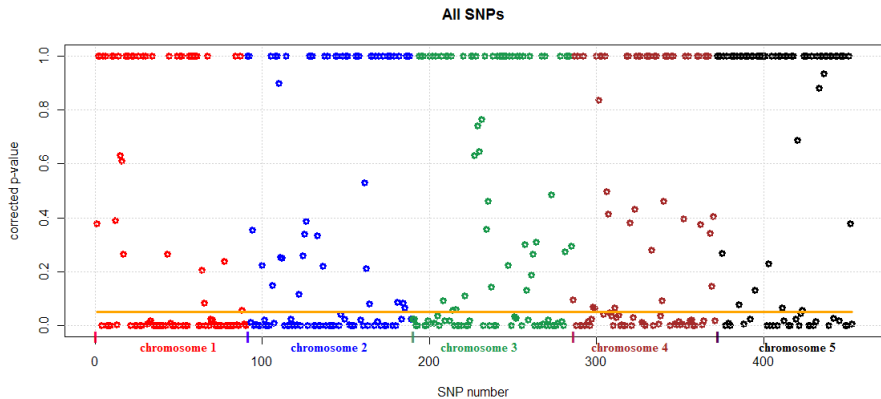
Hypothesis

$$H_0 : \Psi\alpha = 0 \quad \text{vs} \quad H_1 : \Psi\alpha \neq 0$$

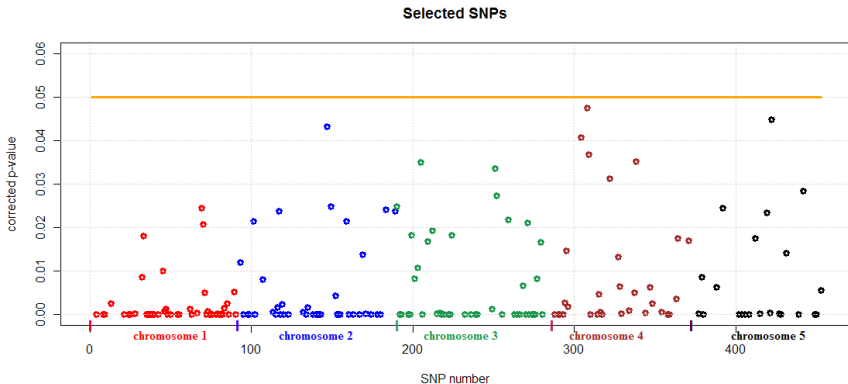
Test statistic

$$\lambda_0 = -2 \ln \frac{L(\hat{\theta}_0)}{L(\hat{\theta})} \sim \chi_4^2.$$

Results: Detecting association



Results: Detecting association

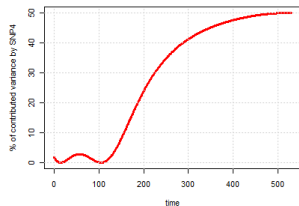
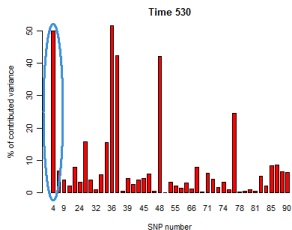
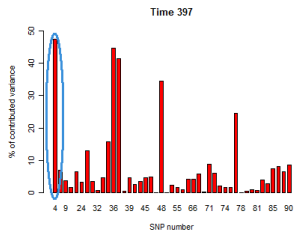
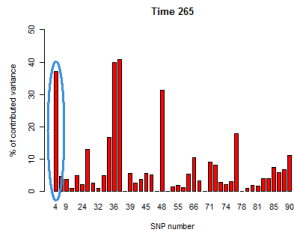
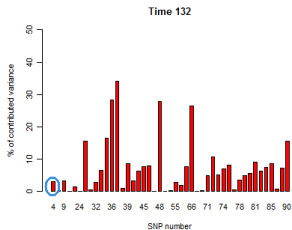
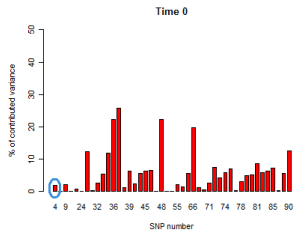


1.
$$r^2 = \frac{D^2}{p_1 q_1 p_2 q_2}$$

2. threshold for dropping SNPs is 0.8

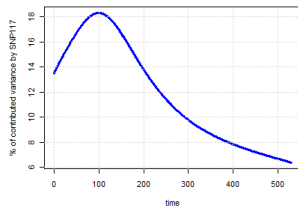
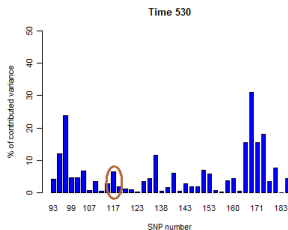
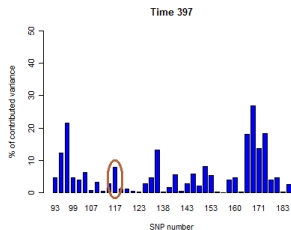
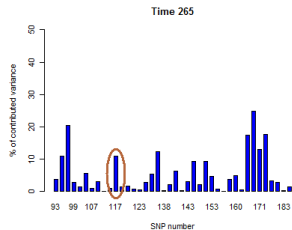
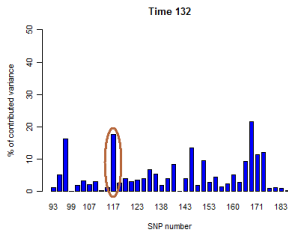
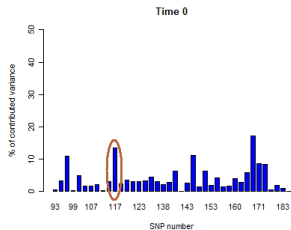
Results: Detecting association

Chromosome 1



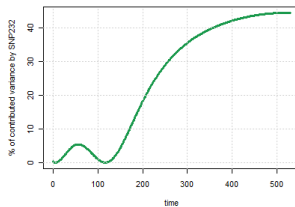
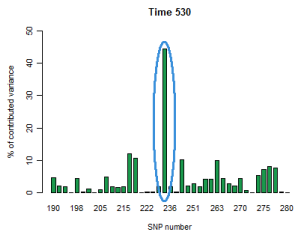
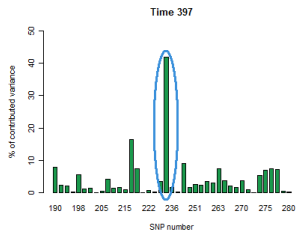
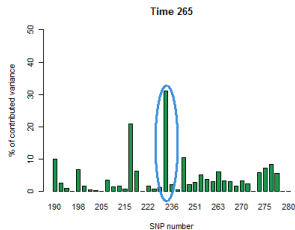
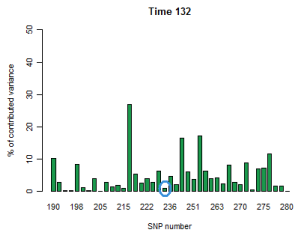
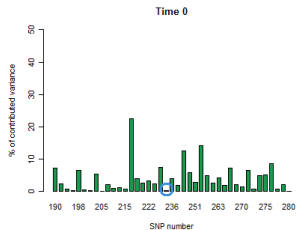
Results: Detecting association

Chromosome 2



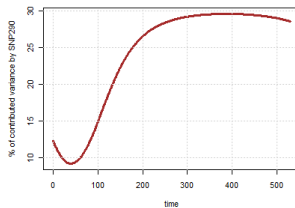
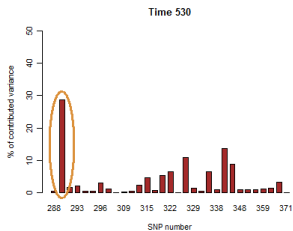
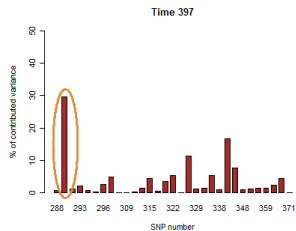
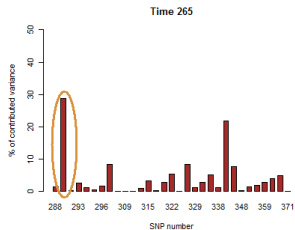
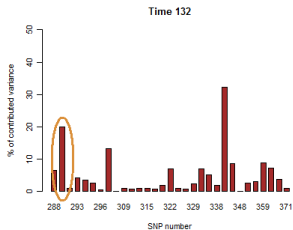
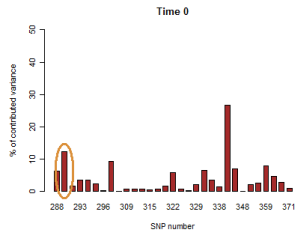
Results: Detecting association

Chromosome 3



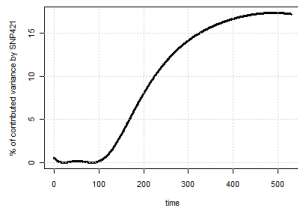
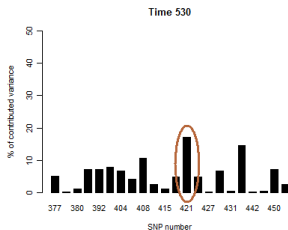
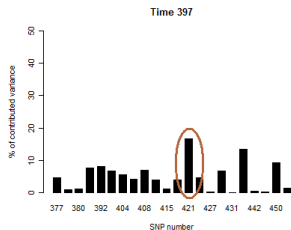
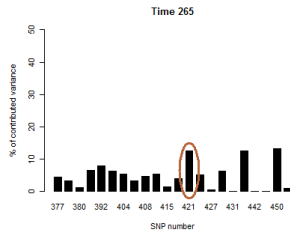
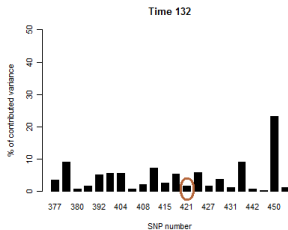
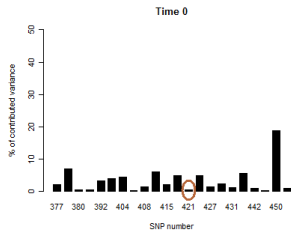
Results: Detecting association

Chromosome 4



Results: Detecting association

Chromosome 5



Methods: Prediction (non-phenotyped animals)

- estimation of parameters μ and α

$$y(\tau) = \Psi(\tau)\mu + [X_1 \otimes \Psi(\tau)]\alpha + \epsilon$$

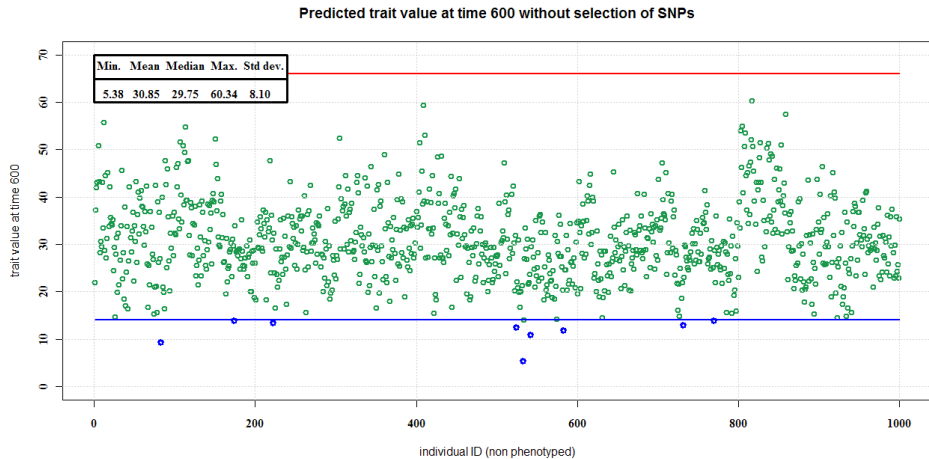
- X_1 incidence matrix of SNP genotypes for phenotyped animals
- τ based on 5 time point $\{0, 132, 265, 397, 530\}$
- $t_{max} = 600$

- prediction future trait value for i th individual

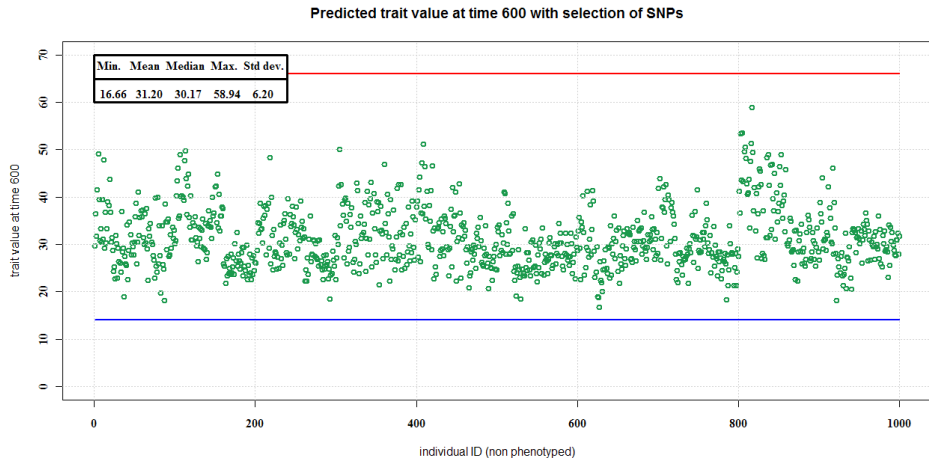
$$\hat{y}_i(1) = \Psi(1)\hat{\mu} + [X_{2i} \otimes \Psi(1)]\hat{\alpha}$$

- X_{2i} incidence vector of SNP genotypes for non-phenotyped i th animal
- $(\tau = 1) \Leftrightarrow (t = 600)$

Results: Prediction (non-phenotyped animals)



Results: Prediction (non-phenotyped animals)



Methods: Prediction (phenotyped animals)

- estimation of parameters μ and α for i th individual

$$y_i(\tau) = \Psi(\tau)\mu + [X_i \otimes \Psi(\tau)]\alpha + \epsilon$$

- X_i : incidence vector of SNP genotypes for i th individual
 - τ based on 5 time point $\{0, 132, 265, 397, 530\}$
 - $t_{max} = 600$
- prediction future trait value for i th individual

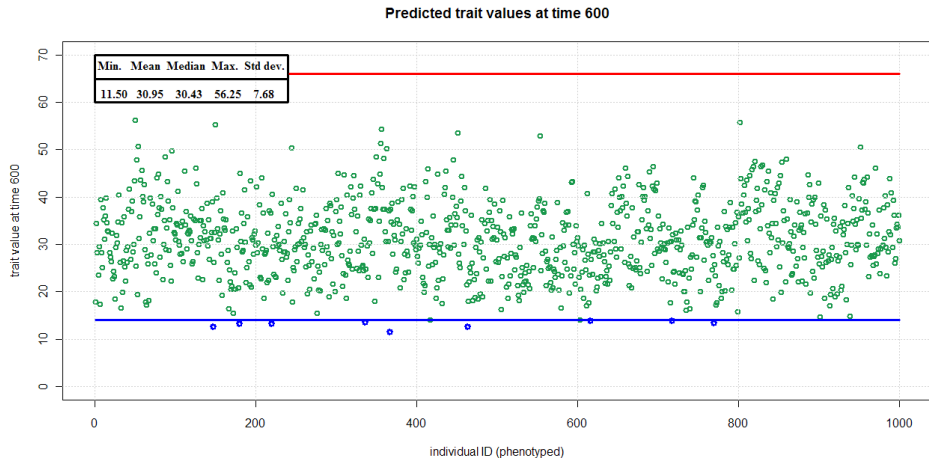
$$\hat{y}_i(1) = \Psi(1)\hat{\mu} + [X_i \otimes \Psi(1)]\hat{\alpha}$$

- $(\tau = 1) \Leftrightarrow (t = 600)$

Results: Prediction (phenotyped animals)



Results: Prediction (phenotyped animals)



Conclusions

- many significant SNPs, which influence the trait
- SNP effect varies across time
- prediction with selection gives results with less variation and no individual exceeds the trait limits (for non-phenotyped animals)
- selection in prediction for phenotyped animals does not change the results

Bibliography

- Runqing Yang, Quan Tian and Shizhong Xu, 2006 Mapping Quantitative Trait Loci for Longitudinal Traits in Line Crosses. *Genetics* **173**: 2339–2356.

Thank You!

