MODELLING THE GENETIC BASIS OF RESPONSE CURVES

FRED VAN EEUWIJK, MARCOS MALOSETTI & MARTIN BOER

WAGENINGEN UR
For quality of life

Overview

- QTL mapping for simple traits by regression and mixed models
- QTL mapping of earlier estimated curve parameters
- Modelling genotypic response curves as functions of QTLs
- (Example(s))

QTL mapping for simple traits by regression and mixed models

A brief review

WAGENINGEN UR

QTL mapping as a partitioning the genotypic effect

$$\underline{P}_{ir} = \mu + G_i + \underline{\varepsilon}_{ir}$$

Phenotype = Mean + Genotype + Error

$$\underline{P}_{ir} = \mu + x_i \alpha + \underline{G}_i^* + \underline{\varepsilon}_{ir}$$

Marker regression

Simple Interval Mapping

Phenotype = Mean + QTL + Residual genotypic effect + Error

x_i = genetic predictor : function of QTL genotype probabilities given flanking marker genotypes and chromosome position

α = QTL effect

Multiple QTL models and their identification

$$\underline{P}_{ir} = \mu + \sum_{q=1}^{n_Q} x_{iq}\alpha_q + \underline{G}_i^* + \underline{\varepsilon}_{ir}$$

Phenotype = Mean + Multiple QTLs + Residual genotypic effect + Error

$$\underline{P}_{ir} = \mu + \sum_{c \in C} x_{ic}\alpha_c + \boxed{x_{iq}\alpha_q} + \underline{G}_i^* + \underline{\varepsilon}_{ir}$$

Phenotype = Mean +

Co-factors +

Putative QTL +

Residual genotypic effect +

Error

QTL mapping of earlier estimated curve parameters for GxE data

Common models for response curves

WAGENINGEN UR

Response models with environmental characterizations on the basis of the phenotype

A reference model: the two-way ANOVA model with interaction

$$\underline{\overline{P}}_{ij} = \mu + G_i + E_j + (GE)_{ij}$$

Regressing individual responses on an environmental index

$$\underline{\overline{P}}_{ij} = \mu + G_i + E_j + \boxed{b_i} E_j + \underline{\delta}_{ij}$$

$$\underline{\overline{P}}_{ij} = \mu + G_i + E_j + \boxed{b_i} E_j + \boxed{c_i} E_j^2 + \underline{\delta}_{ij}$$

Parameters to be mapped

Regressing individual responses on multiple environmental indices

$$\underline{\overline{P}}_{ij} = \mu + G_i + E_j + \sum_{k=1}^{K} \boxed{a_{ki}} b_{kj} + \underline{\delta}_{ij}$$

Response models with explicit environmental characterizations

Factorial regression

$$\underline{\overline{P}}_{ij} = \mu + G_i + E_j + \boxed{b_i} z_j + \underline{\delta}_{ij}$$

$$\underline{\overline{P}}_{ij} = \mu + G_i + E_j + \boxed{b_{1i}} z_j + \boxed{b_{2i}} z_j^2 + \underline{\delta}_{ij}$$

Non-linear regression

$$\underline{\overline{P}}_{ij} = f(\boxed{\boldsymbol{\theta}_i,}\ z_j) + \underline{\delta}_{ij}$$

Parameters to be mapped

Choice of response curve parameters for QTL mapping

- Parameters from response curve formulations derived from physiological knowledge (mechanisms for growth) are preferred
- Parameters from regressions on explicit physical and biological characterizations are preferred
- However, when curve parameters are mapped one by one, without taking into account correlations between the parameters, it is not immediately obvious that QTL mapping of parameters obtained from physiological models will be more powerful than QTL mapping of parameters from GxE models

Modelling genotypic responses as functions of QTLs

Integrating response models and QTL models

WAGENINGEN UR

Models with environment specific QTLs

$$\underline{P}_{ij} = \mu + E_j + x_i \alpha_j + \underline{\delta}_{ij}$$

i for genotype

j for environment

Phenotype = Mean +

Environment +

Environment specific QTLs +

Residual

$$\underline{P}_{ij} = \mu + E_j + \boxed{\sum_{q=1}^{n_Q} x_{iq} \alpha_{jq}} + \underline{\delta}_{ij}$$

$$\underline{P}_{ij} = \mu + E_j + \sum_{q^*=1}^{n_{Q^*}} x_{iq^*} \alpha_{q^*} + \boxed{\sum_{q=1}^{n_Q} x_{iq} \alpha_{jq}} + \underline{\delta}_{ij}$$

QTL effect regressed on environmental variable

$$\underline{P}_{ij} = \mu + E_j + x_i \alpha_j + \underline{\delta}_{ij}$$

$$\underline{\alpha}_j = a_0 + a_1 z_j + \underline{\alpha}^*_j$$

$$\underline{P}_{ij} = \mu + E_j + x_i a_0 + x_i a_1 z_j + x_i \underline{\alpha}^*_j + \underline{\delta}_{ij}$$

Phenotype = Mean +
Environment +
QTL main effect +
Predictable QTL x E +
Unpredictable QTL x E +
Residual (= genotypic effect + plot error)

QTL effect regressed on environmental variable for non-linear response curves

$$\underline{\overline{P}}_{ij} = f(\boldsymbol{\theta}_i, z_j) + \underline{\delta}_{ij}$$

$$\underline{\theta}_i^k = \mu^k + \sum_{q=1}^{n_{Q^k}} x_{iq}\alpha_q^k + \underline{\varepsilon}_i^k$$

For each non-linear parameter a separate QTL model can be developed

QTL analyses for CIMMYT drought stress program in maize: mapping response curve parameters & mixed model QTL analysis

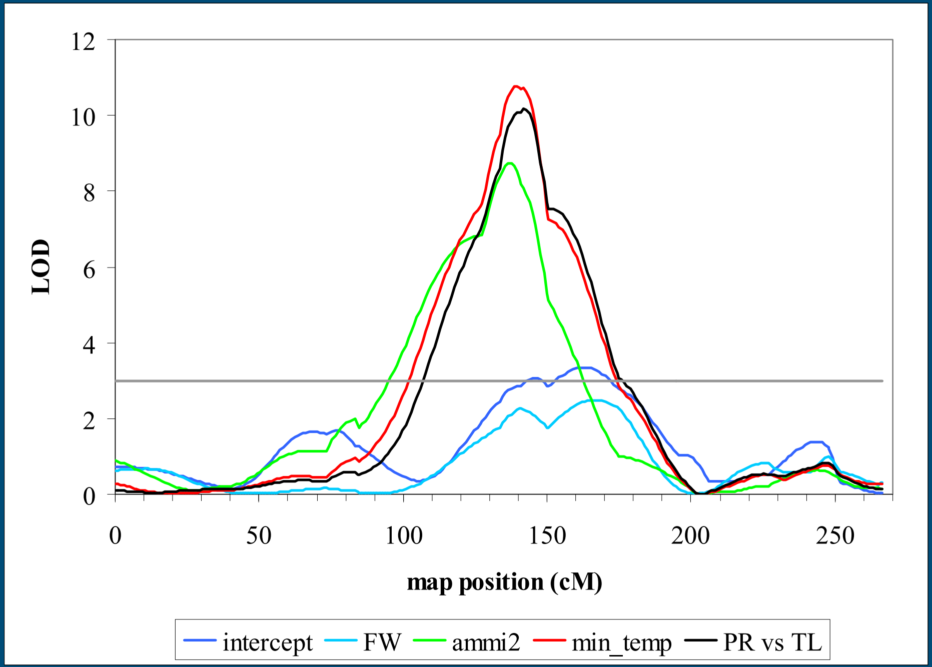
M. Malosetti, J. Crossa, M. Vargas,
J.M. Ribaut, M.P. Boer & F.A. van Eeuwijk

WAGENINGEN UR

CIMMYT maize (C1 population)

- Response
  - Yield
- Environments
  - 8 trials = 8 managed stress environments, intermediate and severe drought stress (IS, SS), low and high nitrogen (LN, HN), no stress
    - 1992, 1994, 1996
    - 2 locations (TI, PR)
    - Winter and summer seasons
- Genotypes
  - 211 F2 derived F3 lines
- Covariables
  - Genotypes
    - 132 marker loci
  - Environments
    - Min. and max. temperature, radiation, rain and number of sun hours for vegetative, flowering and grain filling stages

LOD profiles for various response parameters



Parameters mapped:

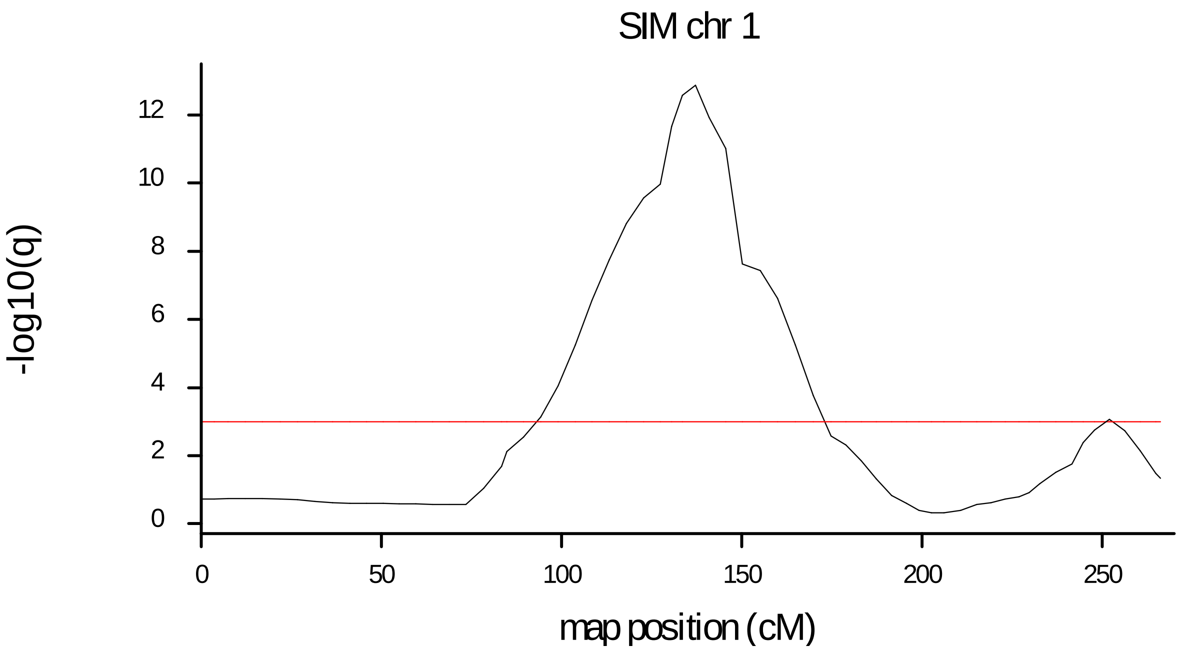
•Slope and intercept regression on environmental index

•Scores of principal component analysis on interaction in GxE table

•Slopes for minimum temperature during flowering

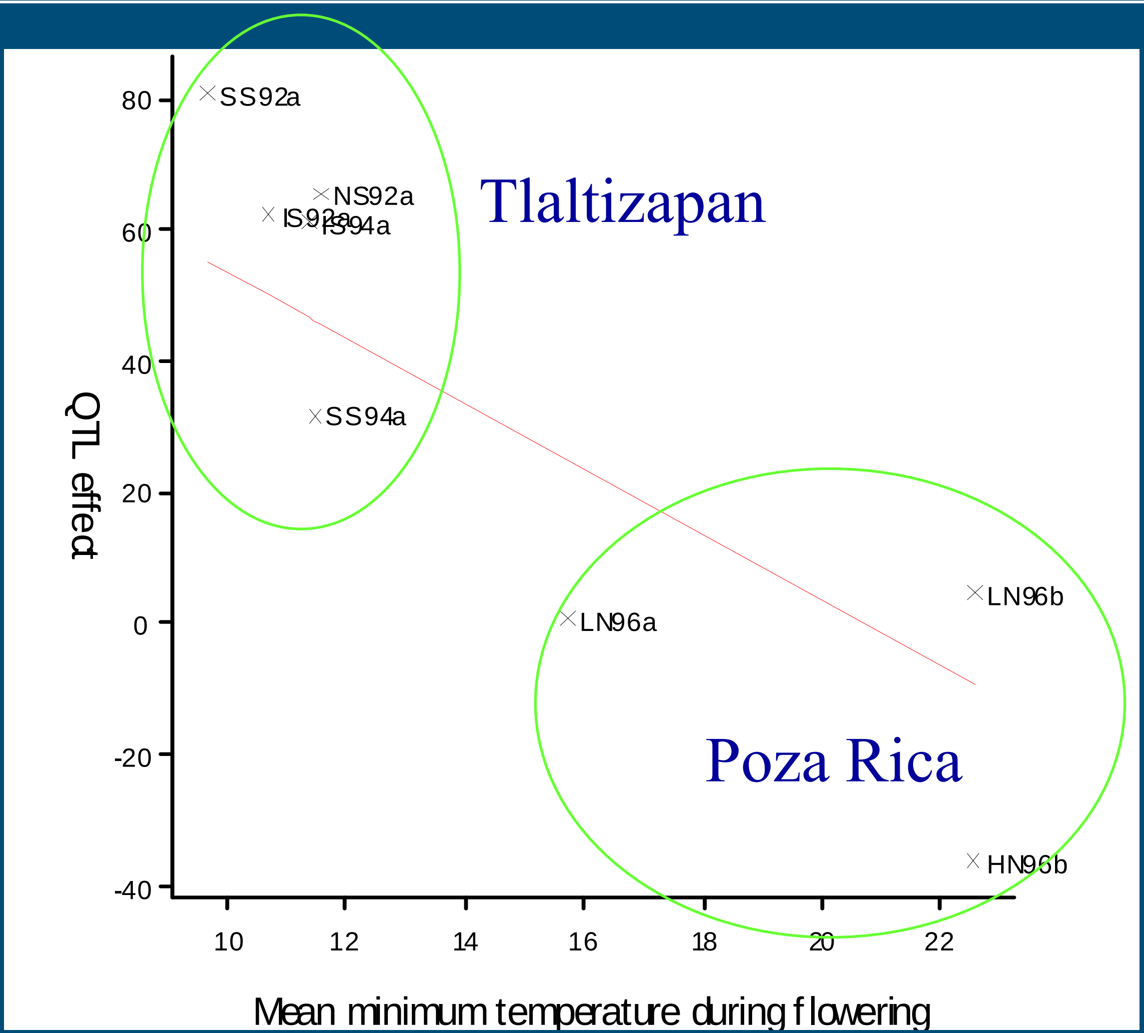
•Contrast Tlaltizapan - Poza Rica

Results mixed model scan with environment specific QTLs



Env	QTL effect (ton ha^{-1})		SE
NS92a	0.644	*	0.207
IS92a	0.623	*	0.154
SS92a	0.831	*	0.128
IS94a	0.605	*	0.153
SS94a	0.311	ns	0.160
LN96a	0.010	ns	0.071
HN96b	-0.371	*	0.157
LN96b	0.053	ns	0.067
* p<0.05; ns: not significant			

Regression of QTLxE on min. temperature during flowering



Interpretation QTLxE

- Having the allele of parent 1 or 2 for the yield QTL on chromosome 1, 140 cM, will imply a yield increase or reduction, respectively, that increases with each degree Celsius that the minimum temperature during flowering increases

Conclusions

- QTL mapping of response curves can be done
  - by mapping of estimated parameters
  - within integrated response and QTL model
- Both approaches lead to similar conclusions when error in estimation and correlation between parameter estimates is small
- Integrated approach is generally more reliable and transparent

QTL Methodology for Response Curves on the Basis of Nonlinear Mixed Models, With an Illustration to Senescence in Potato

Theoretical and Applied Genetics, 2006

M. Malosetti, R.G.F. Visser, C. Celis-Gamboa & F.A. van Eeuwijk

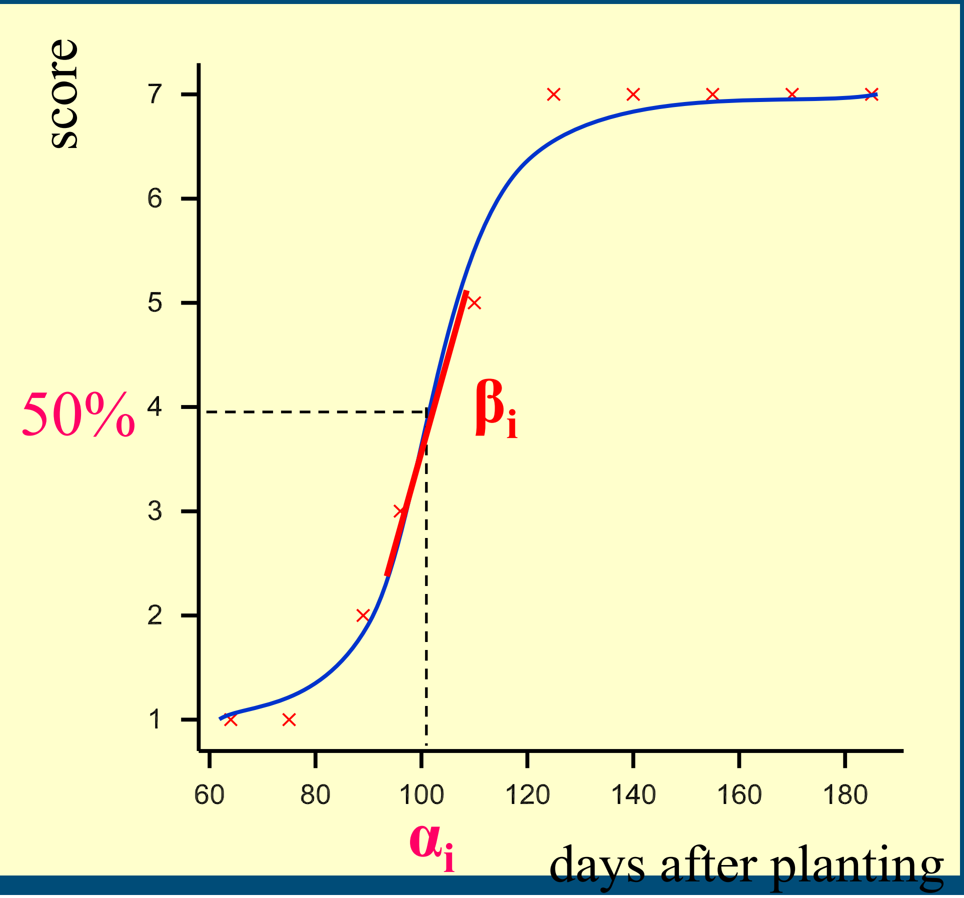
WAGENINGEN UNIVERSITY
PLANT SCIENCES

Data used to illustrate the approach

- *Solanum phureja* x *Solanum tuberosum* backcross (2x)
- 205 genotypes
- Genotypic information:
  - 153 AFLP markers in 12 linkage groups
- Phenotypic information:
  - Senescence evaluation at 11 time points
    Senescence = 1 → plant completely green
    Senescence = 7 → plant completely brown/yellow

Functional QTL modeling of phenotypic response curves

$$\underline{P}_{ij} = f(z_j; \theta) + \underline{\varepsilon}_{ij}$$

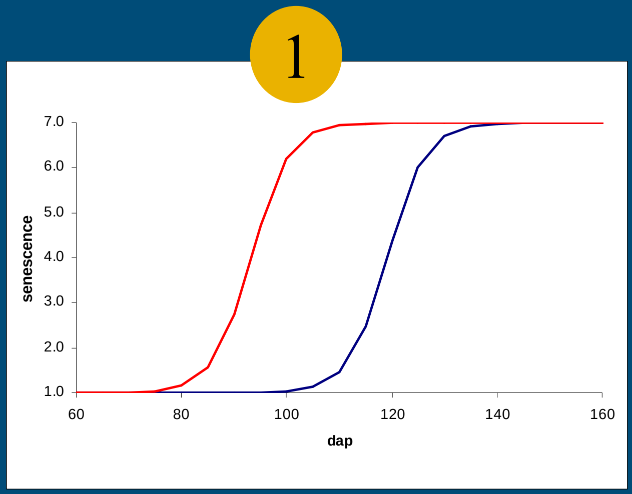


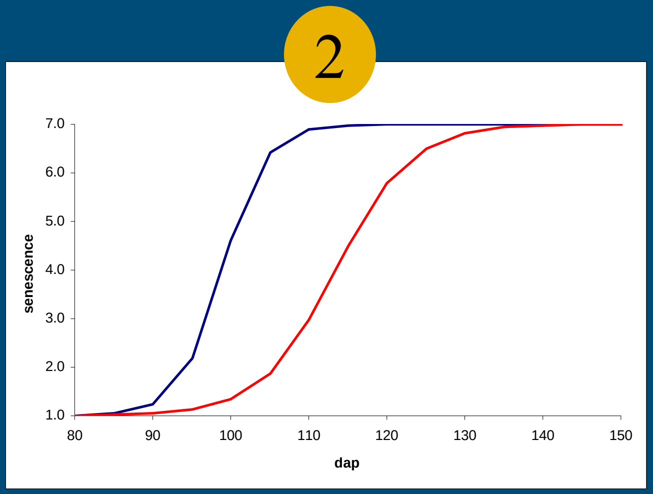
Curve parameters:

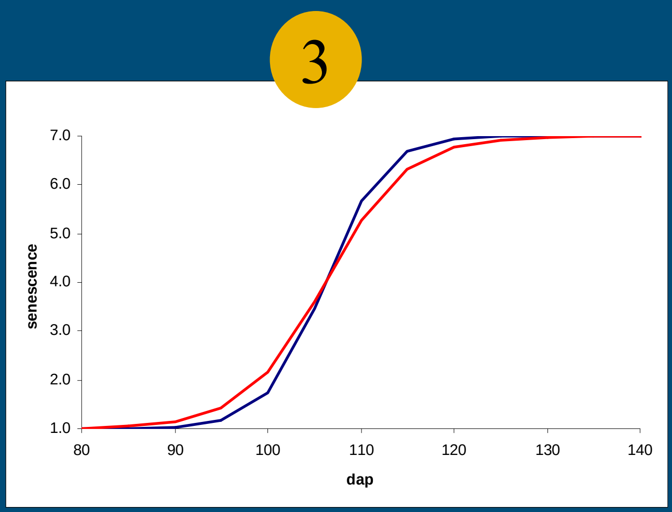
α_i : location inflection point

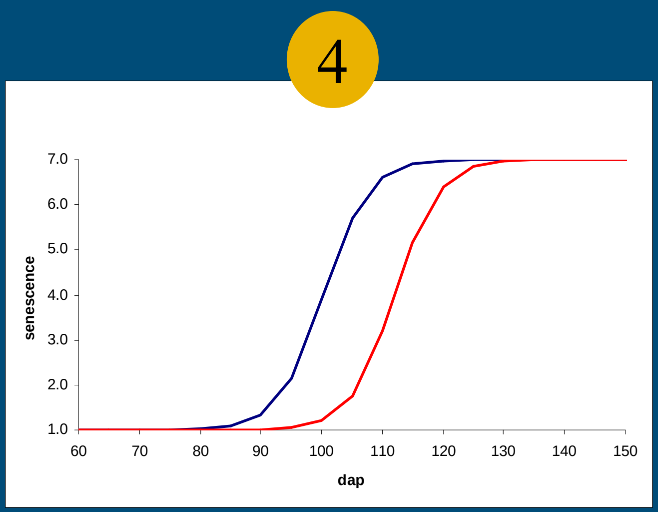
β_i : slope

Effects of QTLs on senescence

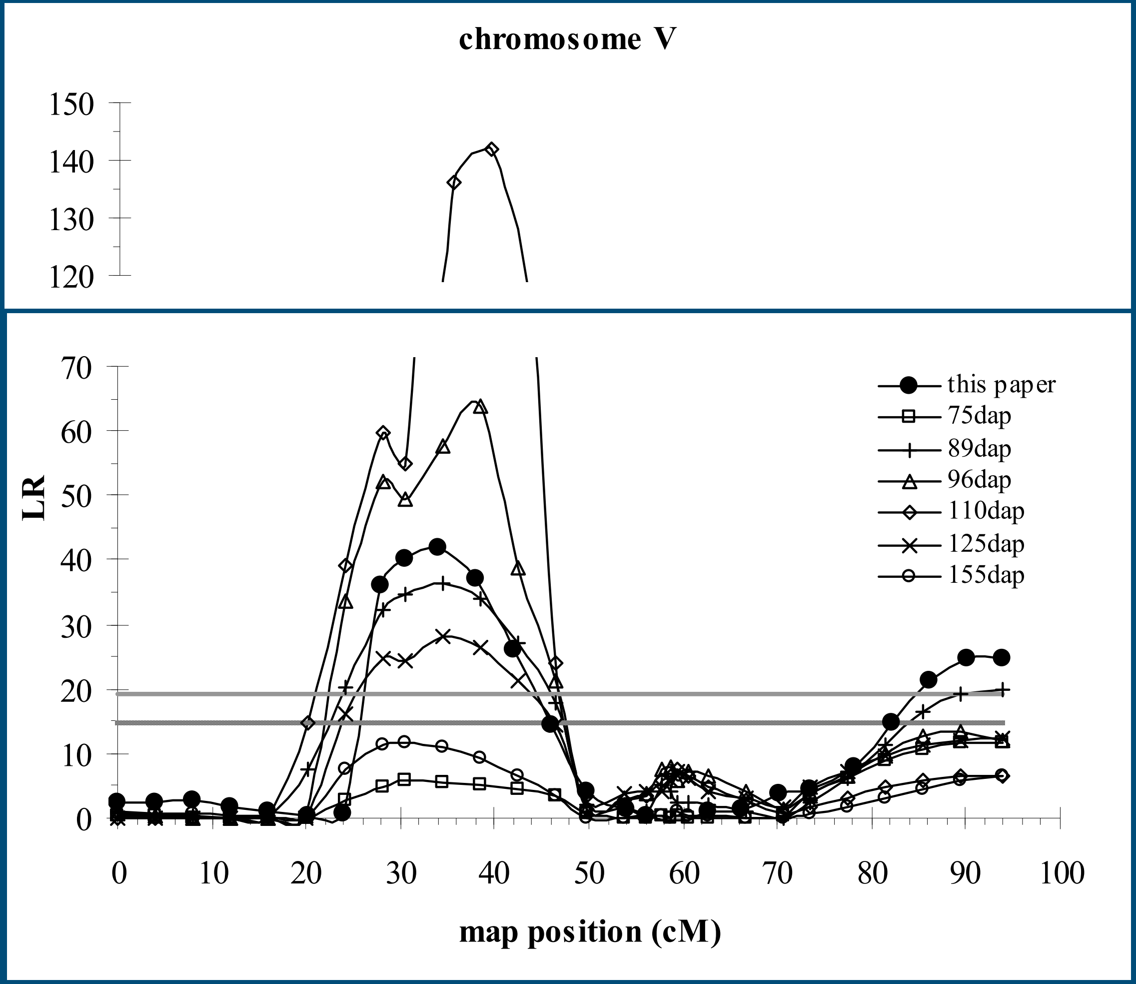








Comparing power of curve approach to analyses per time point



WAGENINGEN UR

Pleiotropy or linkage

$$\underline{y}_{ij} = f\left(t;\quad \alpha_0 + \sum_{l \in L} X_{il}\alpha_l + \underline{a}_i,\quad \beta_0 + \sum_{l \in L^*} X_{il}\beta_l + \underline{b}_i\right) + \underline{e}_{ij}$$