

Two Reasons to Model Phenotype Mean and Variance in QTL Mapping

Robert Corty

UNC Chapel Hill

BCB Colloquium

January 23, 2017

Overview

Background

Mean QTL Mapping

Variance QTL Mapping

Unsolicited Advice

Acknowledgements

Software: CRAN package `vqtl`

Slides:

github.com/rcorty/BCB_colloquium

Collaborators:

Lisa Tarantino

Ethan Lange

Leslie Lange

Laura Raffield

Valdar lab:

Will Valdar

Greg Keele

Dan Oreper

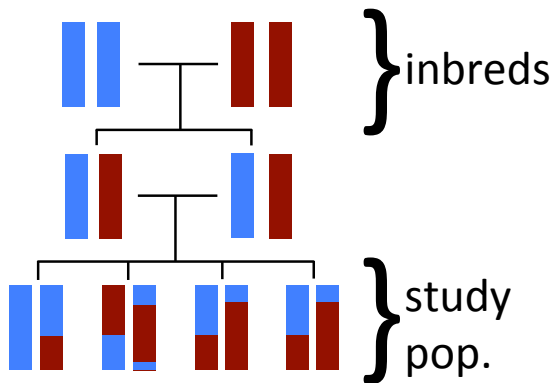
Paul Maurizio

Wes Crouse

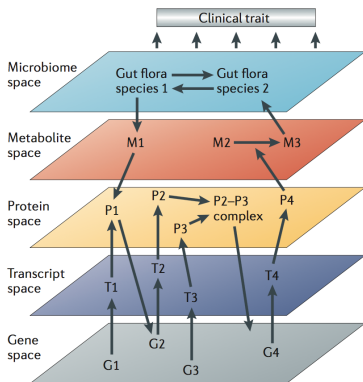
Yanwei Cai

Background

F2 Intercross Mapping Population



Current Approach to QTL Mapping

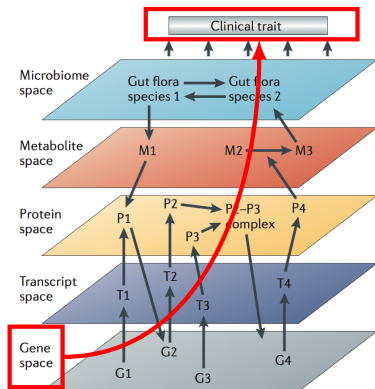


Truth:

$$y_i = f(g_i, t_i, e_i, \dots)$$

Civelek, 2014

Current Approach to QTL Mapping



Truth:

$$y_i = f(g_i, t_i, e_i, \dots)$$

Statistical Model:

$$y_i = m_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2)$$

with

$$m_i = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{q}_i^T \boldsymbol{\alpha}$$

Civelek, 2014

DGLM Model

Constant variance

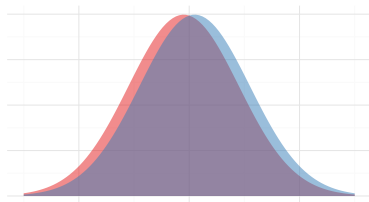
QTL Mapping Model:

$$y_i = m_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2)$$

with

$$m_i = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{q}_i^T \boldsymbol{\alpha}$$



DGLM Model

Constant variance

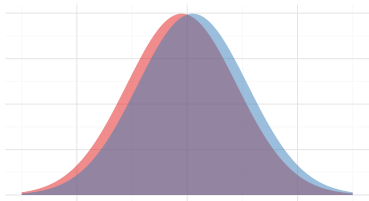
QTL Mapping Model:

$$y_i = m_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2)$$

with

$$m_i = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{q}_i^T \boldsymbol{\alpha}$$



Heterogeneous variance

QTL Mapping Model:

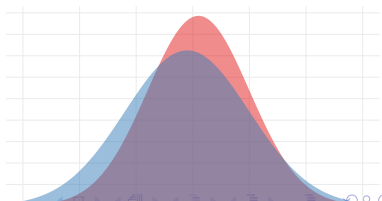
$$y_i = m_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \exp(v_i)^2)$$

with

$$m_i = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{q}_i^T \boldsymbol{\alpha}$$

$$v_i = \mathbf{z}_i^T \boldsymbol{\gamma} + \mathbf{q}_i^T \boldsymbol{\theta}$$



Credits

J. R. Statist. Soc. B (1989)
51, No. 1, pp. 47–60

Generalized Linear Models with Varying Dispersion

By GORDON K. SMYTH†

University of California, Santa Barbara, USA

Copyright © 2011 by the Genetics Society of America
DOI: 10.1534/genetics.111.127068

Detecting Major Genetic Loci Controlling Phenotypic Variability in Experimental Crosses

Lars Rönnegård^{*,*1} and William Valdar[†]

^{}Statistics Unit, Dalarna University, SE-781 70 Borlänge, Sweden and [†]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7265*

Mean QTL Mapping

Nuisance Variance Heterogeneity

Could some levels of nuisance covariates yield more precise observations than others?

- ▶ Technician
- ▶ Day
- ▶ Apparatus
- ▶ Sex of model organism

Up-weight those observations.

Down-weight the (otherwise) high leverage points.

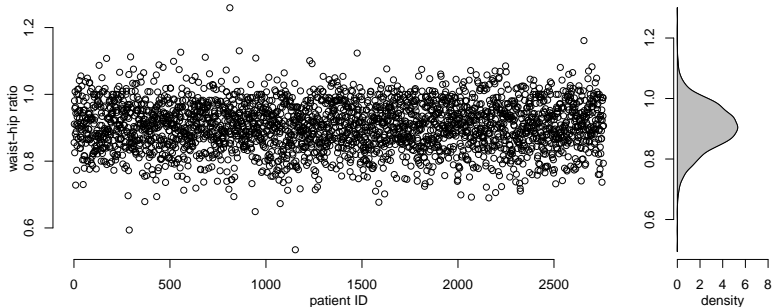
Significance vs. Zero-ness

$$p > 0.05 \not\Rightarrow \beta = 0$$

If we wouldn't trust a result that requires excluding some covariate, we should probably model it.

Example 1

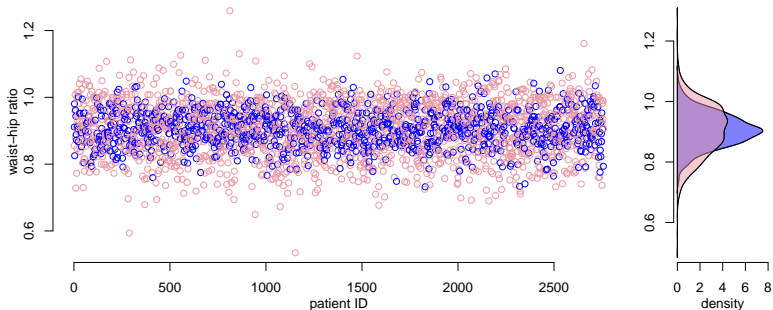
Waist-to-hip ratio in ARIC study



Implications for study design

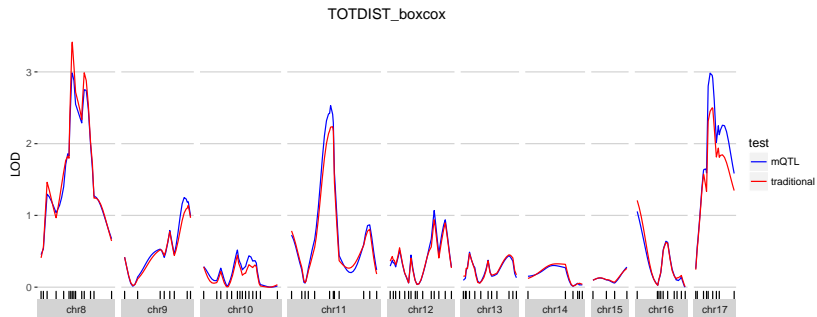
Example 1

Waist-to-hip ratio in ARIC study



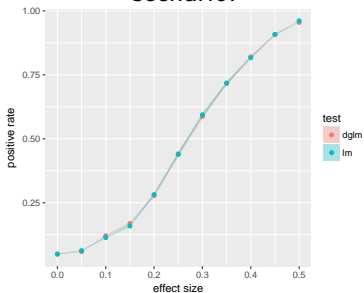
Implications for study design

Example 2

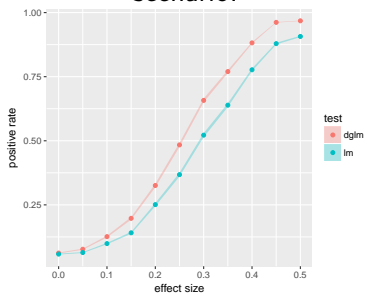


Implications for Power

Homogenous variance scenario:

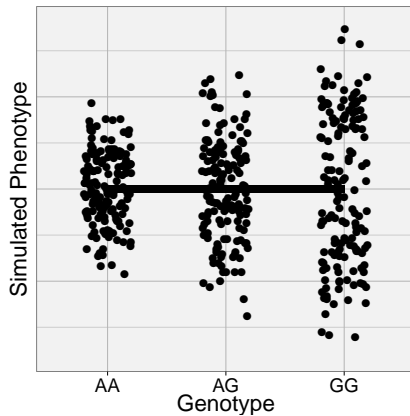


Heterogeneous variance scenario:

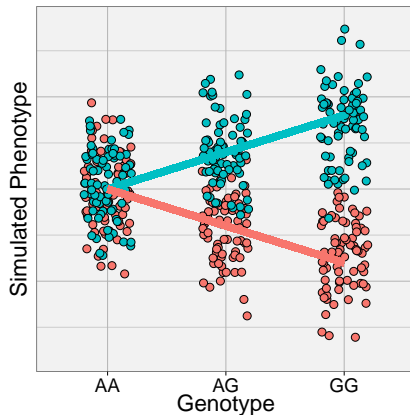


Variance QTL Mapping

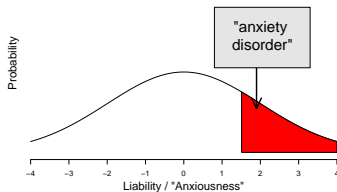
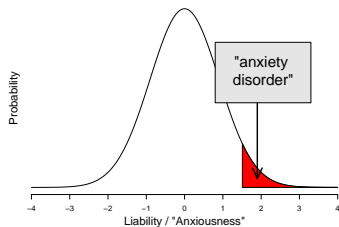
GxE



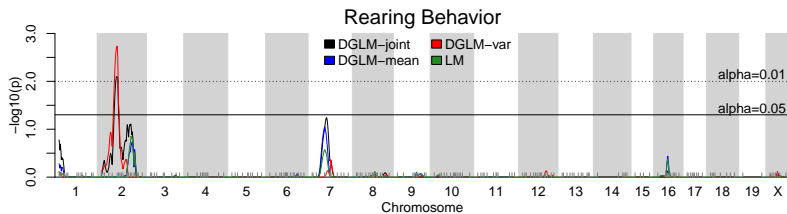
GxE



Liability-Threshold Model

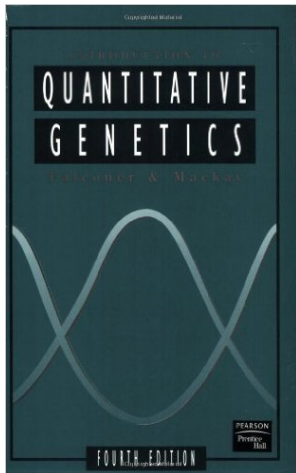


Example 1



Unsolicited Advice

Unsolicited Advice



- ▶ Work for someone you like.
- ▶ Do an easy project first.
- ▶ Read Falconer.

Gratten 2016, Nature Genetics

Risk of psychiatric illness from advanced paternal age is not predominantly from *de novo* mutations

Jacob Gratten¹, Naami R Wenz¹, Wouter J Peyrot², John J McGrath^{3,5}, Peter M Visscher^{1,4,7} & Michael E Goddard^{1,7}