

MODELLING GENETIC EPISTASIS BETWEEN SELECTED CANDIDATE GENES FOR MILK PRODUCTION TRAITS IN JERSEY CATTLE

Joanna Szyda, Anna Gontarek, Tomasz Suchocki

Statistical Genetics Group, Wrocław University of Life Sciences, Poland



Jolanta Komisarek

Agricultural University of Poznań, Poland



- 1. Material**
- 2. Models**
- 3. Results**
- 4. Theoretical consideration**
- 5. Conclusions**

ANIMALS

- 192 Jersey cows

TRAITS

- **production traits:** milk-, fat-, protein- yield, fat-, protein- content
- 305-day lactation yields
- 3 parities

GENES

- leptin receptor (LEPR) BTA03
T945M
 - leptin gene (LEP) BTA04
Y7F, R25C, A80V, C/T sub. -963
 - diacylglycerol acyltransferase 1 (DGAT1) BTA14
K232A
 - butyrophilin (BTN) BTA23
F16Y, P35Q, K468R
-
- no missing genotypes

GENERAL MODEL

$$y = X_1\beta + X_2q + Z_1\alpha + Z_2p + e$$

$\beta =$ $\left[\begin{array}{l} \textit{birth year} \\ \textit{parity} \\ \textit{calving year} * \textit{season} \end{array} \right]$

$q =$ $\left[\begin{array}{l} \textit{LEPR} \\ \textit{LEP} \\ \textit{DGAT1} \\ \textit{BTN1A1} \end{array} \right]$

$$\alpha \sim N(0, A\sigma_\alpha^2) \quad \sigma_\alpha^2 = 30\% \sigma_y^2$$

$$p \sim N(0, I\sigma_p^2) \quad \sigma_p^2 = 20\% \sigma_y^2$$

$$e \sim N(0, I\sigma_e^2) \quad \sigma_e^2 = 50\% \sigma_y^2$$

PARAMETERISATION OF GENE EFFECTS

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\mathbf{q} + \mathbf{Z}_1\boldsymbol{\alpha} + \mathbf{Z}_2\mathbf{p} + \mathbf{e}$$

$$\mathbf{X}_{2a} \in \{2P(2), P(2) - P(1), -2P(1)\}$$

$$\mathbf{X}_{2d} \in \left\{ \frac{1}{P(11)}, -\frac{2}{P(12)}, \frac{1}{P(22)} \right\}$$

$$\mathbf{X}_{2\text{int}} \in \{x_{ai}x_{aj}\}$$

Cockerham, 1954; Kao and Zeng, 2002

PARAMETERISATION OF GENE EFFECTS

$$y = \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\mathbf{q} + \mathbf{Z}_1\boldsymbol{\alpha} + \mathbf{Z}_2\mathbf{p} + \mathbf{e}$$

- no genes



- additive

$$\mathbf{q}^T = \left[\sum_{i=1}^4 a_i \right]$$

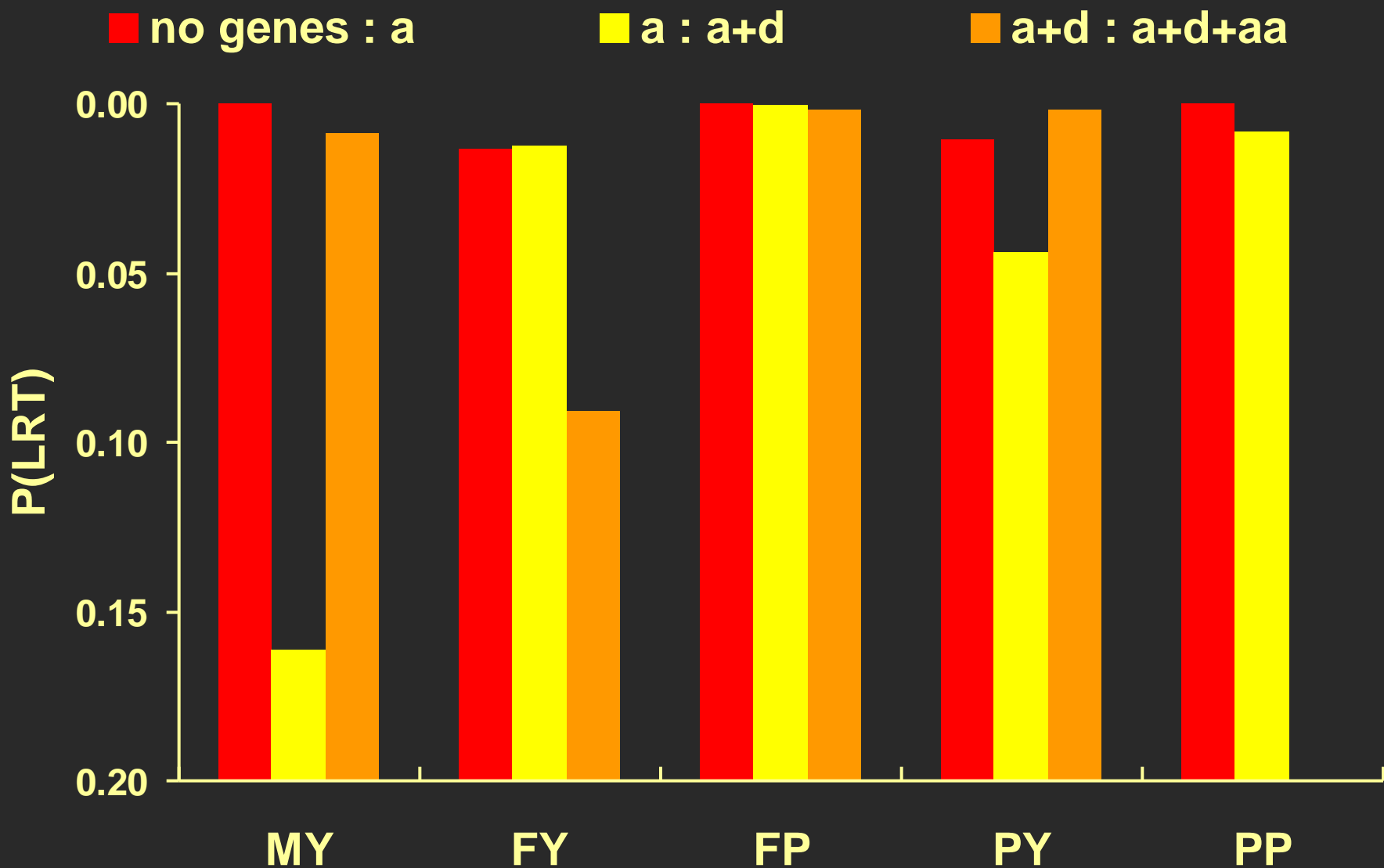
- additive and dominance

$$\mathbf{q}^T = \left[\sum_{i=1}^4 a_i + \sum_{i=1}^4 d_i \right]$$

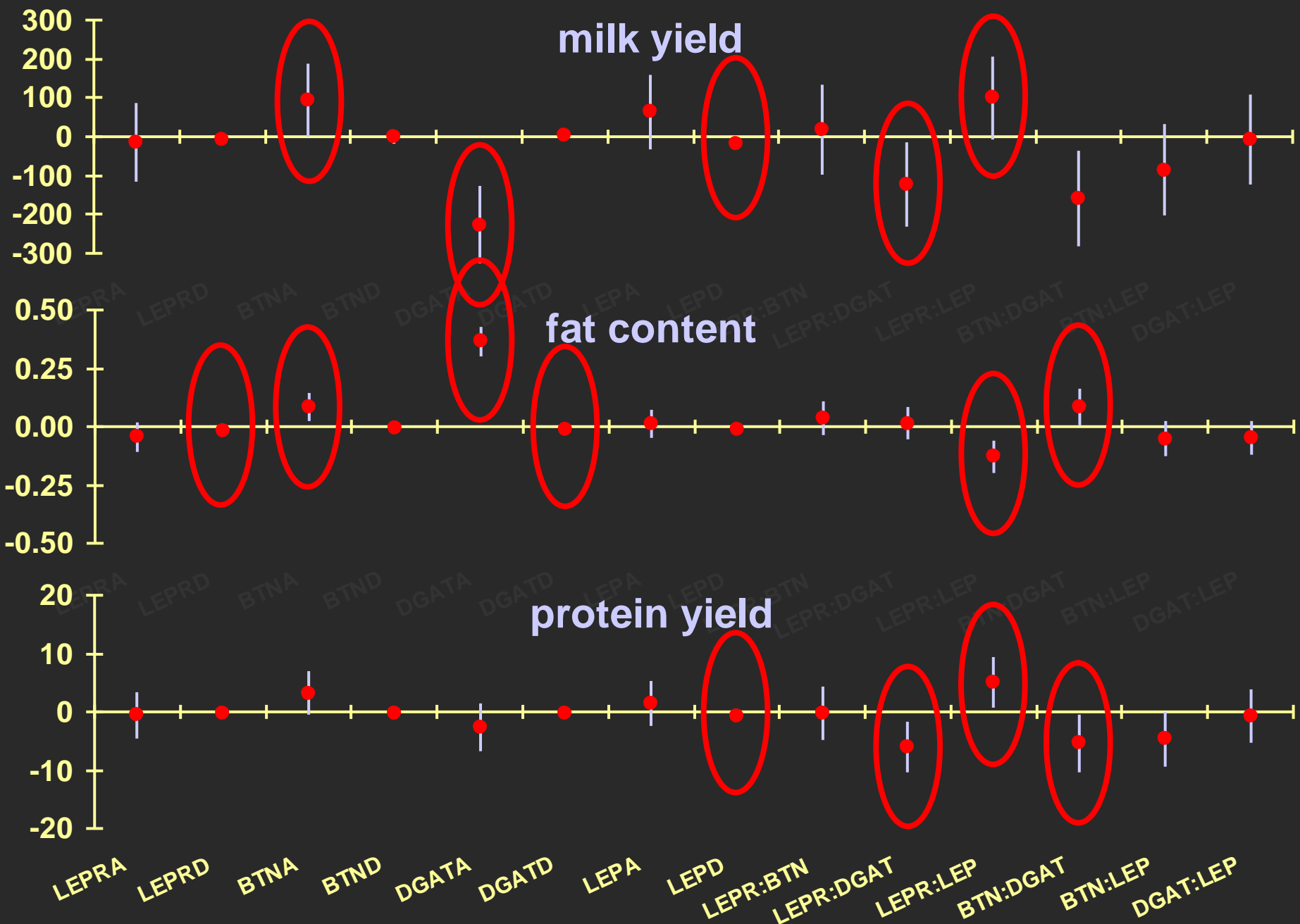
- additive, dominance and additive*additive

$$\mathbf{q}^T = \left[\sum_{i=1}^4 a_i + \sum_{i=1}^4 d_i + \sum_{i=1}^4 \sum_{j=i}^4 a_i a_j \right]$$

MODEL COMPARISON



GENE EFFECTS ESTIMATES



MODEL PARAMETERISATION

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta} + \mathbf{X}_2\mathbf{q} + \mathbf{Z}_1\boldsymbol{\alpha} + \mathbf{Z}_2\mathbf{p} + \mathbf{e}$$

Kao & Zeng F2 model

$$a_i = \frac{1}{2}(G_{i11} - G_{i22}) \quad d_i = 2G_{i12}$$

$$\mathbf{X}_a \in \{1, 0, -1\} \quad \mathbf{X}_d \in \left\{-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right\} \quad \mathbf{X}_{\text{int}} \in \{x_{ai}x_{aj}\}$$

Kao & Zeng F_∞ model

$$a_i = \frac{1}{2}(G_{i11} - G_{i22}) \quad d_i = G_{i12} + \frac{1}{2}d_j$$

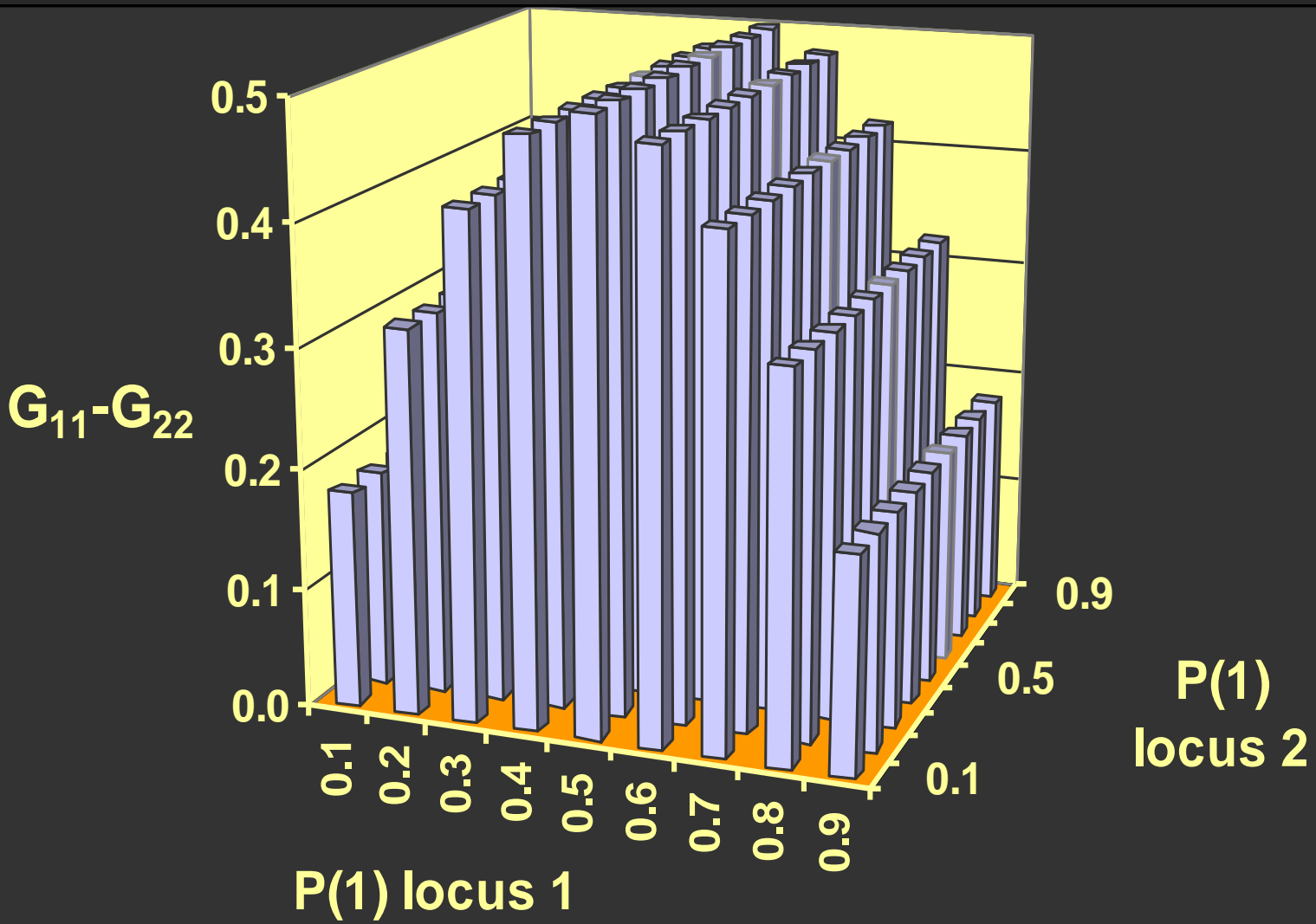
$$\mathbf{X}_a \in \{1, 0, -1\} \quad \mathbf{X}_d \in \{0, 1, 0\} \quad \mathbf{X}_{\text{int}} \in \{x_{ai}x_{aj}\}$$

Cockerham model

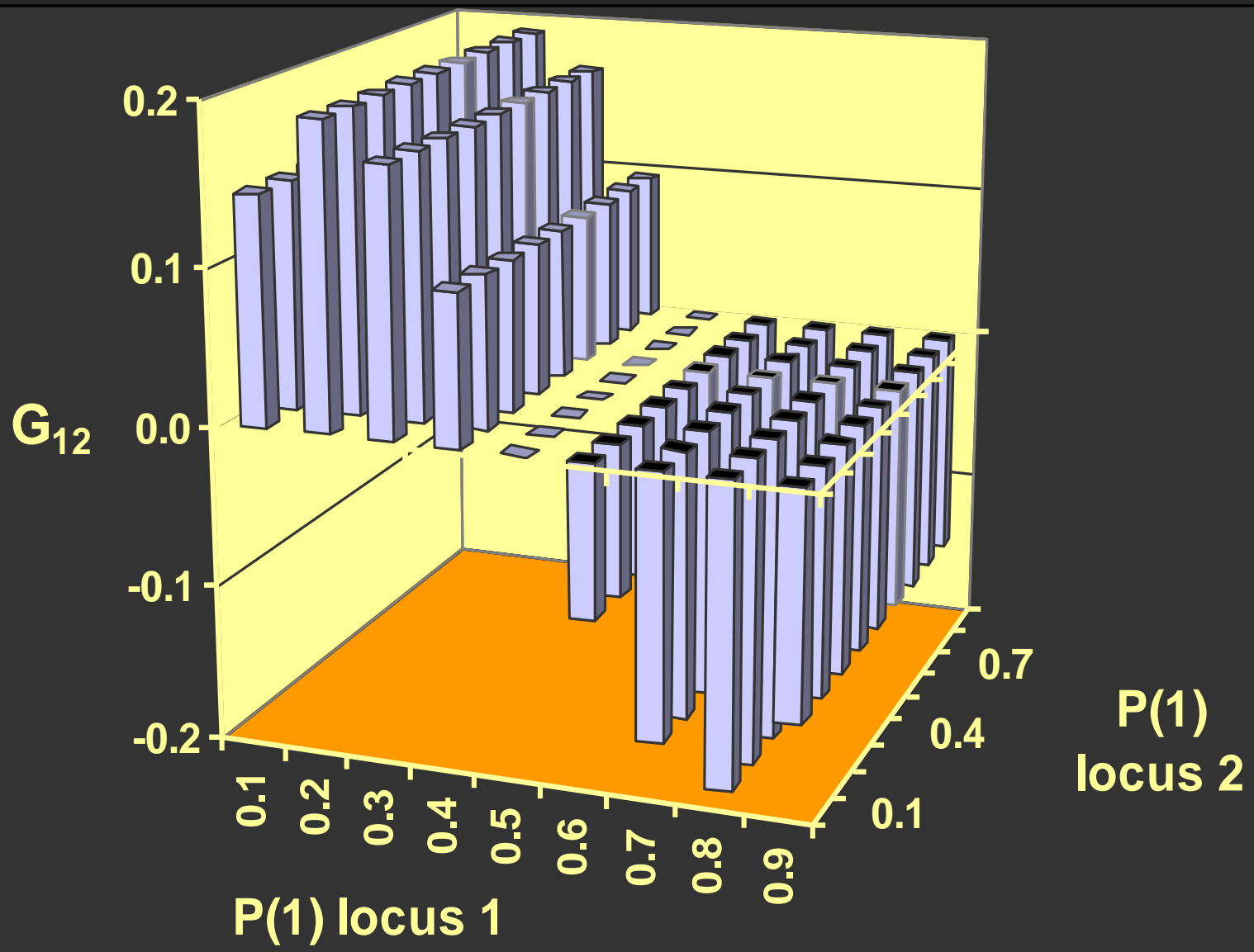
$$a_i = \frac{(G_{i11} - G_{i22})}{P(12)} \quad d_i = [P(1)P(22) - P(2)P(11)]a_i - \frac{G_{i12}}{2}$$

$$\mathbf{X}_a \in \{2P(2), P(2) - P(1), -2P(1)\} \quad \mathbf{X}_d \in \left\{\frac{1}{P(11)}, -\frac{2}{P(12)}, \frac{1}{P(22)}\right\} \quad \mathbf{X}_{\text{int}} \in \{x_{ai}x_{aj}\}$$

MODEL PARAMETERISATION



MODEL PARAMETERISATION



CONCLUSIONS

1. Model selection shows significant epistasis for:

- milk yield → DGAT1-LEPR and DGAT1-BTN
- fat content → LEP-LEPR and DGAT1-BTN
- protein yield → LEP-LEPR, DBAT1-BTN, DGAT1-LEPR

2. Using Cockerham's model

- estimates of additive effects - biased upwards
- correction → $0.5 * P(12)$
- estimates of dominance effects - difficult to interpret
- estimates of main effects of 1 locus - independent of the other loci

Thank you !

