



107 Anos de Genética Quantitativa: do Legado de Fisher aos Desafios Futuros

Prof. Kaio Olimpio das Graças Dias (kaio.o.dias@ufv.br)

Departmento de Biologia Geral

Universidade Federal de Viçosa

Outline

- ▶ Gregor Johann Mendel
 - ▶ Mendel's Laws of Inheritance
 - ▶ Mendel: From genes to genome
- ▶ Mendelian-Biometrician Controversy
 - ▶ Regression Toward Mediocrity
- ▶ The Infinitesimal Model
 - ▶ Decomposition of the genotypic value
 - ▶ Influence of a locus on the phenotype
- ▶ Mixed Models
 - ▶ Merging Mixed Models and Quantitative Genetics
- ▶ References



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE VIÇOSA

PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E MELHORAMENTO

Campus Universitário - Viçosa, MG – 36570-900 - Telefone: (31) 3612-2354 -E-mail: gen@ufv.br

Edital EXTRAORDINÁRIO de seleção para admissão ao Doutorado em Genética e Melhoramento (PPGGM/UFV) em 2025/2.

O Programa de Pós-Graduação em Genética e Melhoramento da Universidade Federal de Viçosa (PPGGM/UFV) faz saber as pessoas interessadas em participarem do presente edital **EXTRAORDINÁRIO**, o qual segue as regras definidas pela Pró-Reitoria de Pesquisa e Pós-Graduação da Universidade Federal de Viçosa (PPG/UFV), conforme endereço <http://www.ppg.ufv.br/>.

1. DA INSCRIÇÃO

No ato da inscrição no Processo Seletivo **EXTRAORDINÁRIO** ao doutorado do PPGGM/UFV 2025/2, no endereço eletrônico <https://gpa.ufv.br/gpa-web/editais/1234/>, o/a candidato/candidata deverá fazer a opção por uma das **MODALIDADES DE INSCRIÇÃO**, conforme descrito a seguir:

2. CRONOGRAMA

ATIVIDADE	DATA
INSCRIÇÕES	25/08/2025 – 05/09/2025
DATA LIMITE PARA SOLICITAÇÃO DA ISENÇÃO DO PAGAMENTO DA TAXA DE INSCRIÇÃO*	01/09/2025*
RESULTADO PRELIMINAR	Até 10/09/2025
RECURSO	Até 24 horas após a divulgação do resultado preliminar
RESULTADO FINAL	Até 12/09/2025
MATRÍCULA	Até 18/09/2025 para início do doutorado em setembro de 2025 OU Entre 22/09/2025 a 16/10/2025 para início do doutorado em outubro de 2025

CURSO	VAGAS DE AMPLA CONCORRÊNCIA	VAGAS DE POLÍTICAS AFIRMATIVAS	TOTAL*
DOUTORADO	3	2	5

* O número total de vagas não é garantia da existência de bolsas e pode sofrer alterações com base no número de bolsas destinadas pelas agências de fomento.



Phenotypic Variation

UFV



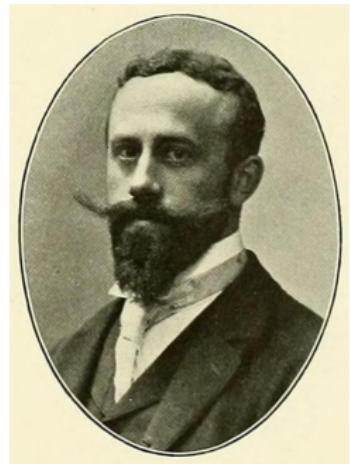
1900: Rediscovery of Mendel's Work



Hugo de Vries



Carl Correns



Erich Tschermak

Gregor Mendel's Pea Plant Experiment

UFV

	Seed Color	Seed Shape	Pod Color	Pod Shape	Flower Color	Flower Position	Plant Height
Dominant Traits	Yellow	Round	Green	Inflated (Full)	Purple	Axial	Tall
Recessive Traits	Green	Wrinkled	Yellow	Constricted (Flat)	White	Terminal	Short

Gregor Mendel's Pea Plant Experiment

UFV

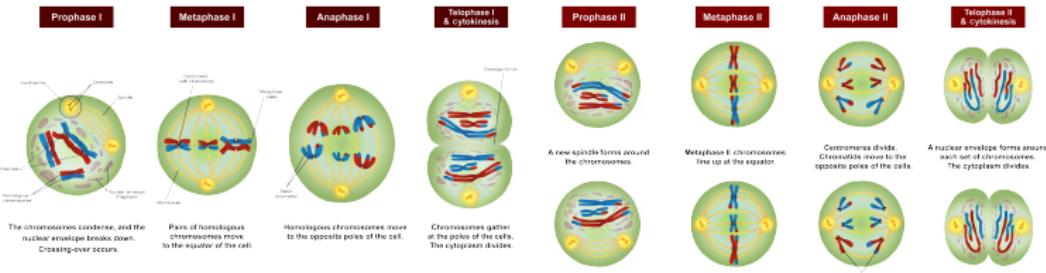
Results

Parental Phenotype	F1	F2	Proportion
1. Seed Shape (Round x Wrinkled)	Round	5474 Round; 1850 Wrinkled	2.96:1
2. Seed Color (Yellow x Green)	Yellow	6022 Yellow; 2001 Green	3.01:1
3. Flower Color (Purple x White)	Purple	705 Purple; 224 White	3.15:1
4. Pod Shape (Inflated x Constricted)	Inflated	882 Inflated; 299 Constricted	2.95:1
5. Pod Colors (Yellow x Green)	Green	428 Green; 152 Yellow	2.82:1
6. Flower Position (Axial x Terminal)	Axial	651 Axial; 207 Terminal	3.14:1
7. Plant Height (Tall x Short)	Tall	787 Tall; 277 Short	2.84:1

Mendel's Laws

What is the chromosomal basis of Mendel's laws?

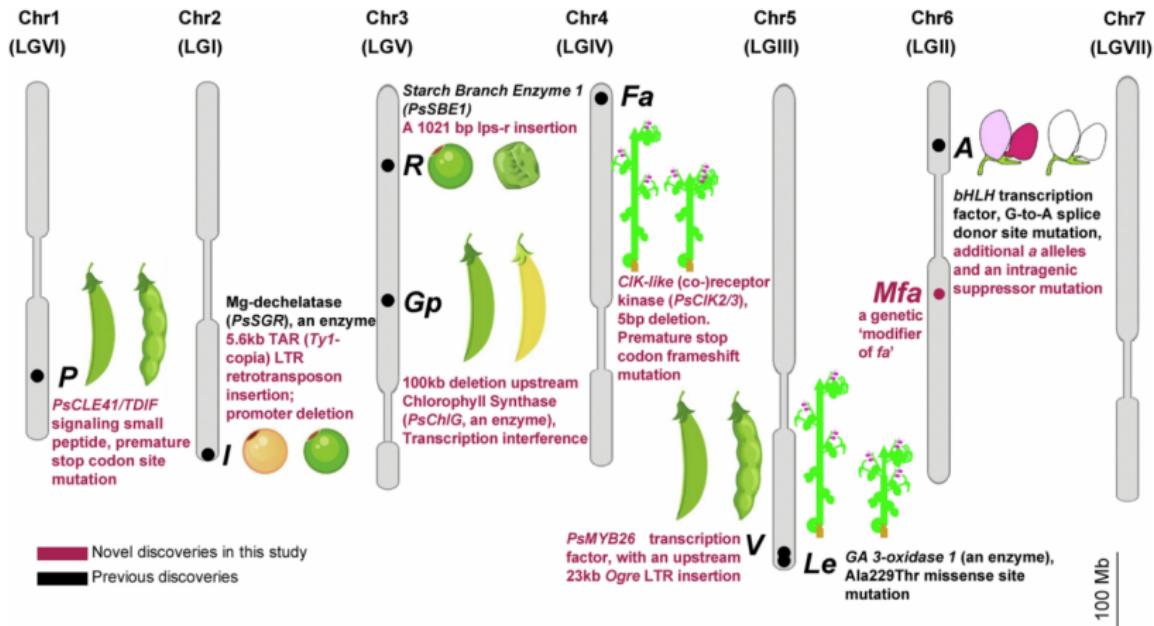
- ▶ Law of Segregation
- ▶ Law of Independent Assortment



Mendel: From genes to genome

Nature Genetics 2025

From: [Genomic and genetic insights into Mendel's pea genes](#)



Mendelian-Biometrician Controversy

Biometrists

- ▶ Karl Pearson and Raphael Weldon
- ▶ Rejected Mendelism as a theory of inheritance

Mendelism

- ▶ William Bateson
- ▶ Laws of segregation and independent assortment

Continuous Variation for Height

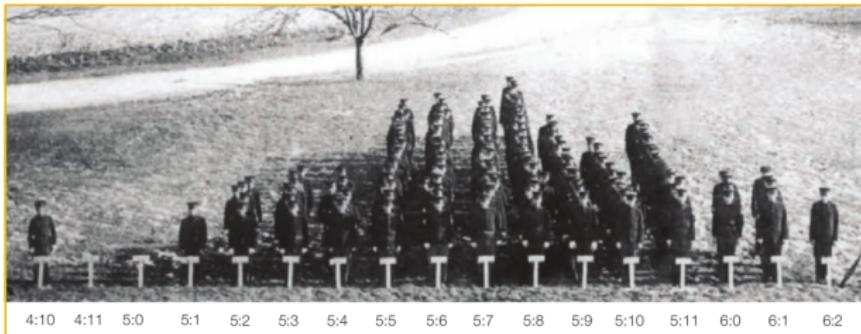
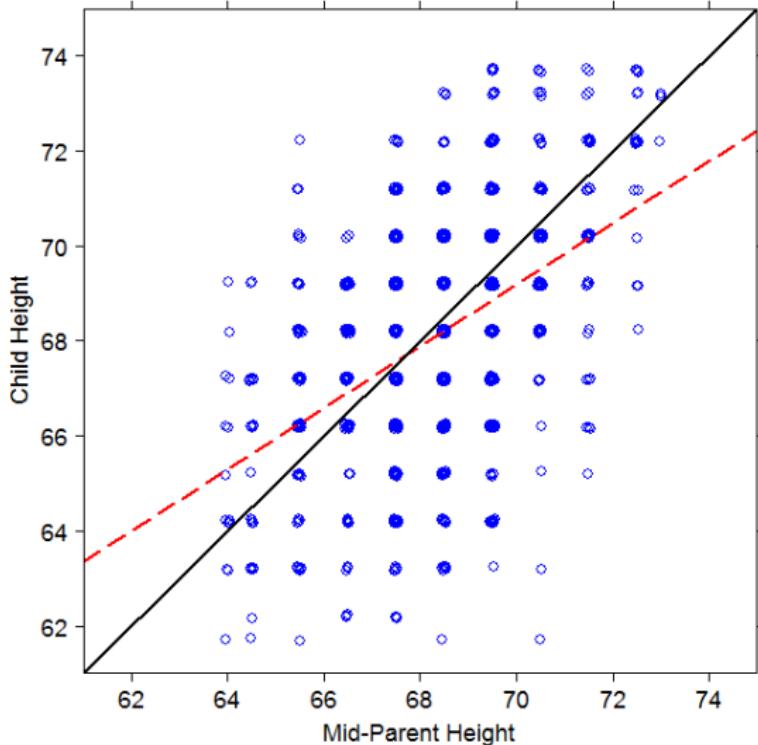


FIGURE 1-6 Students at the Connecticut Agriculture College in 1914 show a range of heights. Ronald Fisher proposed that continuously variable traits like human height are controlled by multiple Mendelian genes. [A. F. Blakeslee, "Corn and Men," *Journal of Heredity* 5, 11, 1914, 511–518.]

Source: Griffiths et al., 2015

Regression Toward Mediocrity

Francis Galton



XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. By R. A. Fisher, B.A. Communicated by Professor J. ARTHUR THOMSON. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

CONTENTS.

PAGE		PAGE	
1. The superposition of factors distributed independently	401	15. Homozygosity and multiple allelomorphism	416
2. Phase frequency in each array	402	16. Coupling	418
3. Parental regression	403	17. Theories of marital correlation; ancestral correlations	419
4. Dominance deviations	403	18. Ancestral correlations (second and third theories)	421
5. Correlation for parent; genetic correlations	404	19. Numerical values of association	421
6. Fraternal correlation	405	20. Fraternal correlation	422
7. Correlations for other relatives	406	21. Numerical values for environment and dominance ratio; analysis of variance	423
8. Epistacy	408	22. Other relatives	424
9. assortative mating	410	23. Numerical values (third theory)	425
10. Frequency of phases	410	24. Comparison of results	427
11. Association of factors	411	25. Interpretation of dominance ratio (diagrams)	428
12. Conditions of equilibrium	413	26. Summary	432
13. Nature of association	413		
14. Multiple allelomorphism	415		

Several attempts have already been made to interpret the well-established results of biometry in accordance with the Mendelian scheme of inheritance. It is here attempted to ascertain the biometrical properties of a population of a more general type than has hitherto been examined, inheritance in which follows this scheme. It is hoped that in this way it will be possible to make a more exact analysis of the causes of human variability. The great body of available statistics show us that the deviations of a human measurement from its mean follow very closely the Normal Law of Errors, and, therefore, that the variability may be uniformly measured by the standard deviation corresponding to the square root of the mean square error. When there are two independent causes of variability capable of producing in an otherwise uniform population distributions with standard deviations σ_1 and σ_2 , it is found that the distribution, when both causes act together, has a standard deviation $\sqrt{\sigma_1^2 + \sigma_2^2}$. It is therefore desirable in analysing the causes of variability to deal with the square of the standard deviation as the measure of variability. We shall term this quantity the Variance of the normal population to which it refers, and we may now ascribe to the constituent causes fractions or percentages of the total variance which they together produce. It is desirable on the one hand that the elementary ideas at the basis of the calculus of correlations should be clearly understood, and easily expressed in ordinary language, and on the other that loose phrases about the "percentage of causation,"

The Infinitesimal Model

UFV

FISHER Model - 1918



Theoretical Population Biology 118 (2017) 50–73

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb



The infinitesimal model: Definition, derivation, and implications



N.H. Barton^a, A.M. Etheridge^b, A. Véber^{c,*}

^a Institute of Science and Technology, Am Campus 1, A-3400 Klosterneuburg, Austria

^b Department of Statistics, University of Oxford, 24–29 St Giles, Oxford OX1 3LB, UK

^c Centre de Mathématiques Appliquées, École Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France

ARTICLE INFO

Article history:

Received 18 May 2017

Available online 11 July 2017

Keywords:

Infinitesimal model

Selection

Epistasis

Quantitative genetics

ABSTRACT

Our focus here is on the *infinitesimal model*. In this model, one or several quantitative traits are described as the sum of a genetic and a non-genetic component, the first being distributed within families as a normal random variable centred at the average of the parental genetic components, and with a variance independent of the parental traits. Thus, the variance that segregates within families is not perturbed by selection, and can be predicted from the variance components. This does not necessarily imply that the trait distribution across the whole population should be Gaussian, and indeed selection or population structure may have a substantial effect on the overall trait distribution. One of our main aims is to identify some general conditions on the allelic effects for the infinitesimal model to be accurate. We first review the long history of the infinitesimal model in quantitative genetics. Then we formulate the model at the

Allele and genotype frequencies

- ▶ One locus, two alleles
- ▶ Genotypes: B_1B_1 , B_1B_2 , B_2B_2
- ▶ Genotype frequencies: P_{11} , P_{12} , P_{22}
- ▶ Allele frequencies: p_1 , p_2

$$p_1 = P_{11} + \frac{1}{2}P_{12}$$

$$p_2 = P_{22} + \frac{1}{2}P_{12}$$

Hardy-Weinberg Equilibrium

$$p_1^2, 2p_1p_2, p_2^2$$

Influence of a locus on the phenotype

- $z = G + E$



k is the dominance

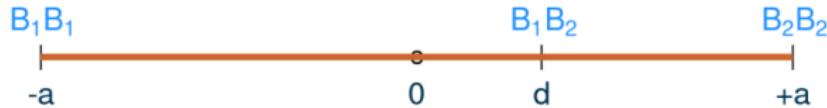
- $k = 0$ **No dominance**
- $0 < k < 1$ = **Partial dominance**
- $k = 1$ or $k = -1$ **Full dominance**
- $k > 1$ **Overdominance**

Influence of a locus on the phenotype

- Lynch & Walsh



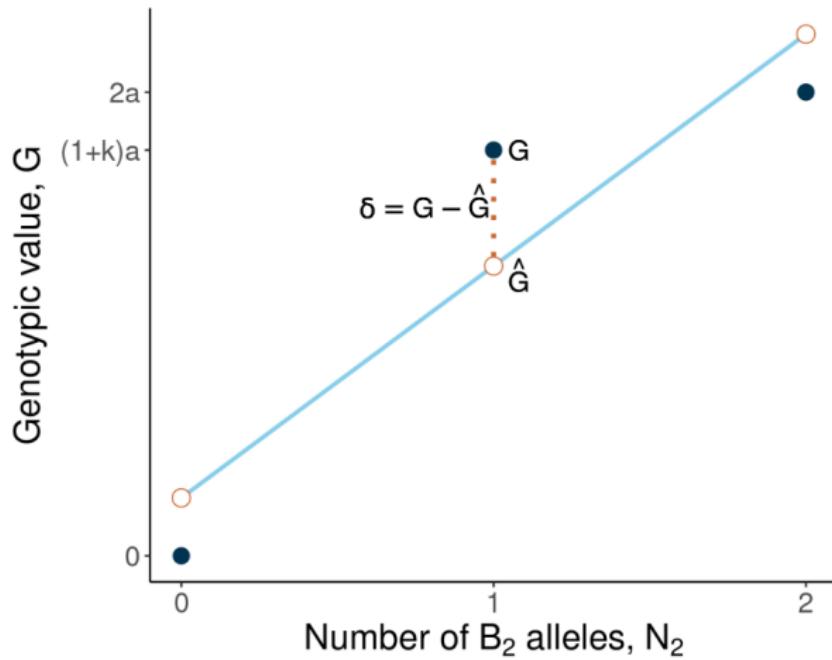
- Falconer & Mackay



Values of d

- $d = 0$ No dominance
- $d = a$ Full dominance
- $d \geq a$ Overdominance

- ▶ Number of copies of a particular allele



Model

- Genetic value of B_iB_j is given by:

$$G_{ij} = \hat{G}_{ij} + \delta_{ij} = \mu_G + N_1\alpha_1 + N_2\alpha_2 + \delta_{ij}$$

Constraint: $N_1 = 2 - N_2$

$$G_{ij} = \mu_G + (2 - N_2)\alpha_1 + N_2\alpha_2 + \delta_{ij}$$

$$G_{ij} = (\mu_G + 2\alpha_1) + N_2(\alpha_2 - \alpha_1) + \delta_{ij}$$

$$G_{ij} = \iota + N_2\alpha + \delta_{ij}$$

Partitioning the genetic variance

UFV

- Given that $G = \hat{G} + \delta$, the total genetic variance is

$$\sigma_G^2 = \sigma^2(\hat{G} + \delta) = \sigma^2(\hat{G}) + 2\sigma(\hat{G}, \delta) + \sigma^2(\delta)$$

\hat{G} and δ are uncorrelated, then:

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2$$

- σ_A^2 additive genetic variance
- σ_D^2 dominance genetic variance

$$\sigma_A^2 = E(\hat{G}^2) - [E(\hat{G})]^2 = 2p_1 p_2 \alpha^2$$

$$\sigma_D^2 = E(\delta^2) - [E(\delta)]^2 = (2p_1 p_2 a k)^2$$

Partitioning the genetic variance

UFV

- Given that $G = \hat{G} + \delta$, the total genetic variance is

$$\sigma_G^2 = \sigma^2(\hat{G} + \delta) = \sigma^2(\hat{G}) + 2\sigma(\hat{G}, \delta) + \sigma^2(\delta)$$

\hat{G} and δ are uncorrelated, then:

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2$$

- σ_A^2 additive genetic variance
- σ_D^2 dominance genetic variance

$$\sigma_A^2 = E(\hat{G}^2) - [E(\hat{G})]^2 = 2p_1 p_2 \alpha^2$$

$$\sigma_D^2 = E(\delta^2) - [E(\delta)]^2 = (2p_1 p_2 \alpha k)^2$$

Cockerham's model

F_2 - metric model



AN EXTENSION OF THE CONCEPT OF PARTITIONING HEREDITARY VARIANCE FOR ANALYSIS OF COVARIANCES AMONG RELATIVES WHEN EPISTASIS IS PRESENT^{1*}

C. CLARK COCKERHAM

North Carolina State College, Raleigh

Received March 22, 1954

IT is convenient, for purposes of description and analysis, to consider the phenotypic expression of a characteristic as a sum of an hereditary or genotypic value and of an environmental value. If the actual joint results deviate from this linear description (i.e., if interaction effects exist) the breeder or geneticist must exercise caution in extrapolating from his results because in this case the hereditary and environmental values are defined specifically in terms of each other (NELDER 1950). For some characteristics a transformation of scale may help in coming closer to additivity (WRIGHT 1950).

OPEN ACCESS Freely available online

PLOS GENETICS

Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits

William G. Hill^{1*}, Michael E. Goddard^{2,3}, Peter M. Visscher⁴

1 Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom, **2** Faculty of Land and Food Resources, University of Melbourne, Victoria, Australia, **3** Department of Primary Industries, Victoria, Australia, **4** Queensland Institute of Medical Research, Brisbane, Australia

Abstract

The relative proportion of additive and non-additive variation for complex traits is important in evolutionary biology, medicine, and agriculture. We address a long-standing controversy and paradox about the contribution of non-additive genetic variation, namely that knowledge about biological pathways and gene networks imply that epistasis is important. Yet empirical data across a range of traits and species imply that most genetic variance is additive. We evaluate the evidence from empirical studies of genetic variance components and find that additive variance typically accounts for over half, and often close to 100%, of the total genetic variance. We present new theoretical results, based upon the distribution of allele frequencies under neutral and other population genetic models, that show why this is the case even if there are non-additive effects at the level of gene action. We conclude that interactions at the level of genes are not likely to generate much interaction at the level of variance.

Citation: Hill WG, Goddard ME, Visscher PM (2008) Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. PLoS Genet 4(2): e1000008. doi:10.1371/journal.pgen.1000008

Editor: Trudy F. C. Mackay, North Carolina State University, United States of America

Received: October 3, 2007; **Accepted:** January 23, 2008; **Published:** February 29, 2008

Copyright: © 2008 Hill et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was partially supported by the Australian National Health and Medical Research Council (grants 389892 and 442915) and the Australian Research Council (grant DP0770096). None of the funders had any role in the analyses and interpretation of the data (which were already collected) or in the preparation, review or approval of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: w.g.hill@ed.ac.uk

Mapping Quantitative Traits

Prof. Zhao-Bang Zeng

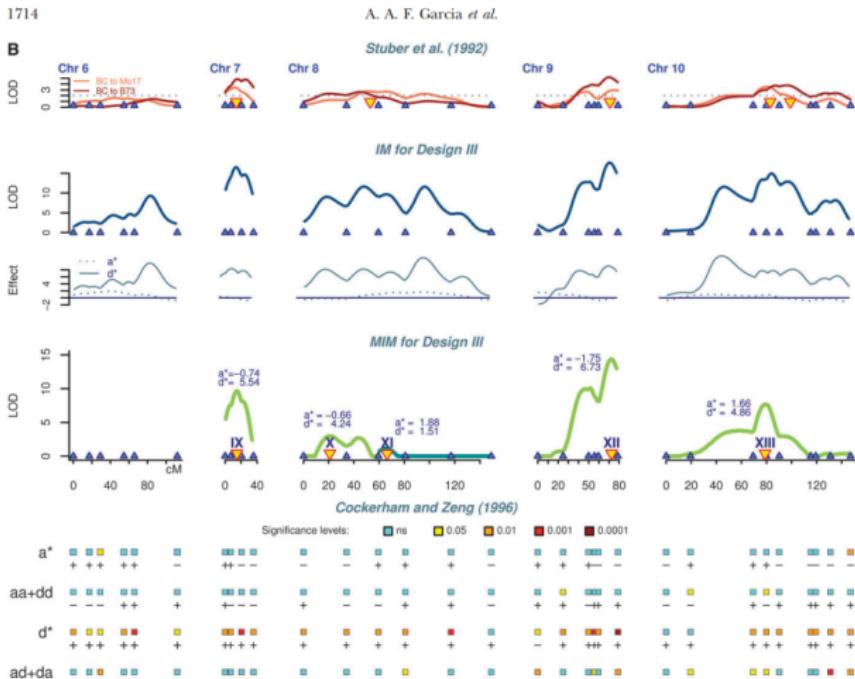


FIGURE 1.—Continued.

Prediction of Complex Traits

UFV

Published January, 1994

CROP BREEDING, GENETICS & CYTOLOGY

Prediction of Maize Single-Cross Performance Using RFLPs and Information from Related Hybrids

Rex Bernardo*

Copyright © 2001 by the Genetics Society of America

Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps

T. H. E. Meuwissen,^{*} B. J. Hayes[†] and M. E. Goddard^{‡,§}

^{*}*Research Institute of Animal Science and Health, 8200 AB Lelystad, The Netherlands, [†]Victorian Institute of Animal Science, Attwood 3049, Victoria, Australia and [‡]Institute of Land and Food Resources, University of Melbourne, Parkville 3052, Victoria, Australia*

Manuscript received August 17, 2000
Accepted for publication January 17, 2001

Mixed Models

- ① Hierarchical Linear Models
- ② Multilevel Models
- ③ Mixed-effects model

Definition

Is a statistical model containing both fixed effects and random effects

Random

- Large number of population levels
- random behavior

Fixed

- Small or large number of population levels
- Systematic or other non-random behavior

► Genotype, Block, Environments..., ?

Block effects: Fixed or Random?

Consequences of the choice

► Traditional randomized complete block design

"The result that does depend on the choice is the standard error (se) of a treatment mean. When blocks are random, the block variance component contributes to the se of a mean. This corresponds to the view that inferences in a random block model are to new blocks in the study population, while inferences in a fixed block model are restricted to the blocks included in the study. Differences or linear contrasts among treatments are estimated within each block, so the variability among blocks cancels out. The variability among blocks does affect inferences about an individual treatment mean because there is no cancellation of the block effects."

Variety effects: Fixed or Random?



[Smith, Cullis and Thompson, 2005]

The analysis of crop cultivar breeding and evaluation trials : an overview of current mixed model approaches

Journal of Agricultural Science (2005), 143, 449–462

“The present authors believe the choice depends on the aim of the analysis and consideration of the properties of the two types of estimation procedures, namely empirical best linear unbiased prediction (E-BLUP) for random effects and empirical best linear unbiased estimation (E-BLUE) for fixed effects”

Variety effects: Fixed or Random?



[Smith, Cullis and Thompson, 2005]

The analysis of crop cultivar breeding and evaluation trials : an overview of current mixed model approaches

Journal of Agricultural Science (2005), 143, 449–462

“Some statisticians advocate the use of random effects in this setting because they regard that the varieties themselves are a random sample from a population. After some unspecified number of stages of selection, this ceases to be a reasonable assumption so that at this point variety effects are regarded as fixed. The present authors do not adhere to this line of reasoning”

Variety effects: Fixed or Random?



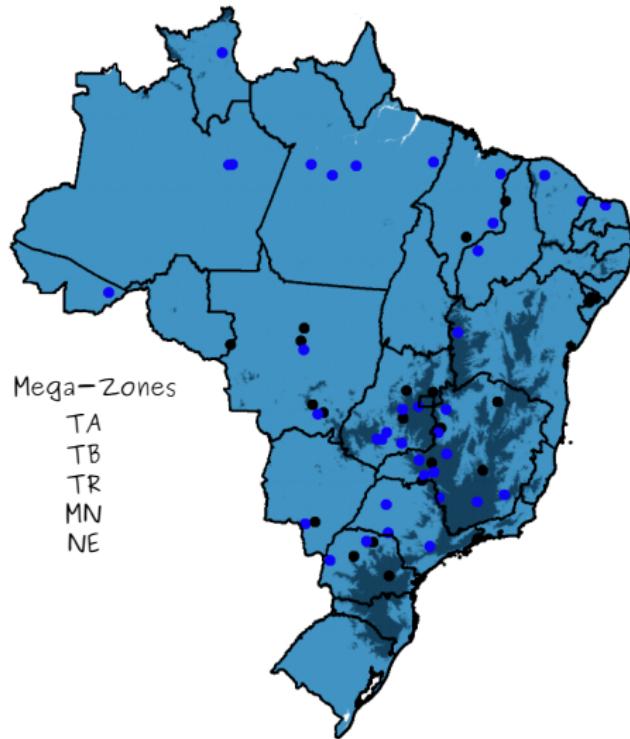
[Smith, Cullis and Thompson, 2005]

The analysis of crop cultivar breeding and evaluation trials : an overview of current mixed model approaches

Journal of Agricultural Science (2005), 143, 449–462

“Of course, with balanced data and orthogonal analyses, the rankings of varieties would be the same in both the fixed and random variety settings. Even so, the present authors still prefer the use of random variety effects since the resultant predictions of genetic gain are more realistic than those based on fixed variety effects.”

Environments effects: Fixed or Random?



Source: Theoretical and Applied Genetics (2020) 133:443–455

Mixed Model

Theory of Linear Mixed Models

► Linear Mixed Model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}$$

Where:

- \mathbf{y} is a vector of the response variable
- \mathbf{X} matrix of the predictor variables
- $\boldsymbol{\beta}$ vector of the fixed-effects regression coefficients
- \mathbf{Z} design matrix for the random effects
- \mathbf{u} vector of random effects
- $\boldsymbol{\epsilon}$ vector of the residuals

Mixed Model

Theory of Linear Mixed Models

- Mixed model equations (MME)

$$\begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} & (\mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1}) \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix}$$

Mixed Model

Theory of Linear Mixed Models

- Mixed model equations (MME)

$$\begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} & (\mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1}) \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix}$$

- Then, we can also show that:

$$\begin{bmatrix} \mathbf{C}_{XX} & \mathbf{C}_{XZ} \\ \mathbf{C}_{ZX} & \mathbf{C}_{ZZ} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_{Xy} \\ \mathbf{c}_{Zy} \end{bmatrix}$$

- Where:

- $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\mathbf{u}}')$
- $\mathbf{C}\hat{\boldsymbol{\theta}} = \mathbf{c}$

Parameter estimation

Mixed model equations

Fixed effects

- Best Linear Unbiased Estimation (eBLUE)

$$\mathbf{c}_{Xy} - \mathbf{C}_{XZ}\mathbf{C}_{ZZ}^{-1}\mathbf{c}_{Zy} =$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{x}' \mathbf{H}^{-1} \mathbf{x})^{-1} \mathbf{x}' \mathbf{H}^{-1} \mathbf{y}$$

Random effects

- Best Linear Unbiased Prediction (eBLUP)

$$\hat{\boldsymbol{u}} = \mathbf{C}_{ZZ}^{-1}(\mathbf{c}_{Zy} - \mathbf{C}_{ZX}\hat{\boldsymbol{\beta}})$$

$$\hat{\boldsymbol{u}} = \mathbf{GZ}' \mathbf{H}^{-1} (\mathbf{y} - \mathbf{x}\hat{\boldsymbol{\beta}})$$

Our Team



GUILHERME BRAVIM

Postgraduate student



JOÃO AMARIO

PhD student



THALES NOGUEIRA

PhD student



GABRIEL BLASQUES

PhD student



JOÃO NETTO

PhD student



THIAGO BROMMONSCHENKEL

PhD student



VALDEIMIRO AMÉRICO

Master's student



BENJAMIM FABRI

Undergraduate student



MIGUEL OLIVEIRA

Undergraduate student



RAFAEL MUZZI

Undergraduate student



References



[Bernardo, 2020]

Reinventing quantitative genetics for plant breeding: something old, something new, something borrowed, something BLUE
Heredity volume 125, pages375–385



[Dias et al. 2022]

Leveraging probability concepts for cultivar recommendation in multi-environment trials Theoretical and Applied Genetics 135 (4), 1385-1399



[Araujo et al. (2024)]

GIS-FA: an approach to integrating thematic maps, factor-analytic, and envirotyping for cultivar targeting

Theoretical and Applied Genetics 137 (4), 1-23



[Dias et al. 2018]

Improving accuracies of genomic predictions for drought tolerance in maize by joint modeling of additive and dominance effects in multi-environment trials

Heredity volume 121, pages24–37



[Chaves et al. 2024]

ProbBreed: a novel tool for calculating the risk of cultivar recommendation in multienvironment trials

G3: Genes, Genomes, Genetics 14 (3), jkae013