Non-additive effects in the genetic evaluation

Zulma G. Vitezica

INPT / INRAE - UMR GenPhySE, Toulouse, France

zulma.vitezica@toulouse-inp.fr

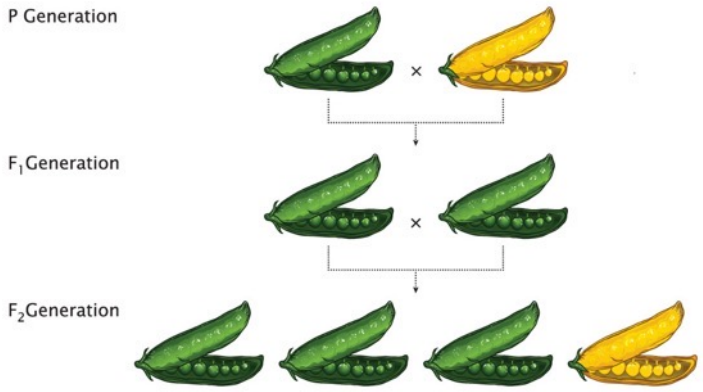
Plan

- Biological vs. statistical effects
  - Why statistical effects matter
- New models accounting for non-additive effects
  - GBLUP, GDBLUP, and its extensions
  - Inbreeding depression
- Is this any useful?
  - Extra accuracy in predictions
  - Variance components
  - Mate allocation
- Conclusions

Biological *vs.* statistical effects

Biological effects

The terms ‘dominance/epistasis’ describe apparent distortions of mendelian segregation ratios that were due to one gene masking the effects of another



Example of dominance

Genotypes at locus 1	Genotypes at locus 2		
	BB	Bb	bb
AA	🐴	🐴	🐴
Aa	🐴	🐴	🐴
aa	🐴	🐴	🐴

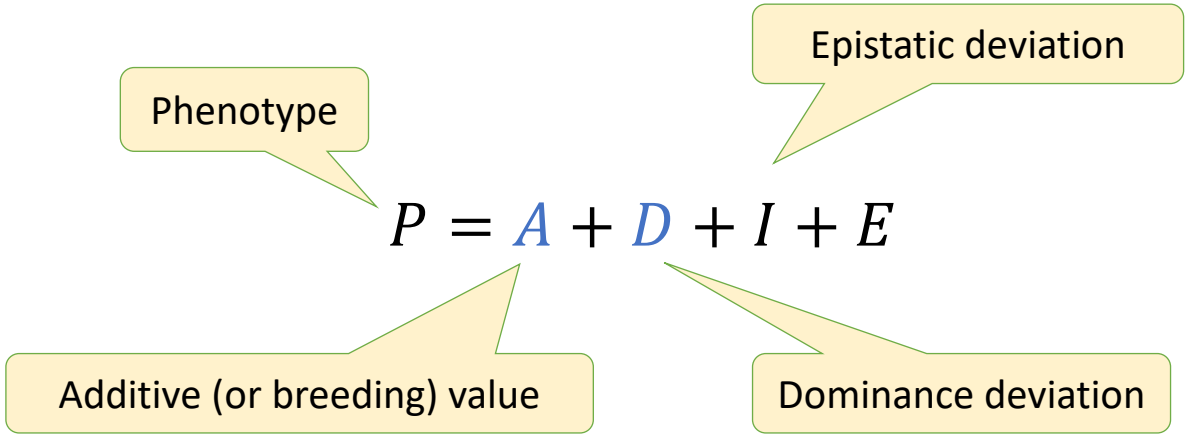
Example of epistasis: dominance-by-dominance two-locus epistasis

Biological effects

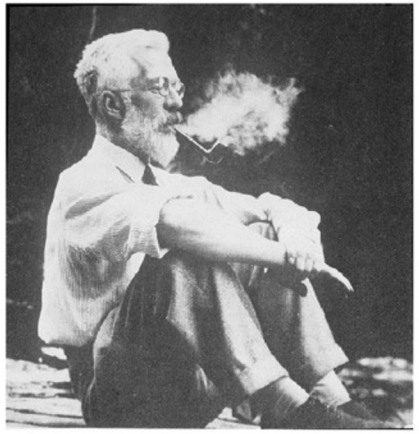
- Unfortunately **we don’t know all gene actions & pathways**
- For many purposes, we need to make educated guesses
- Guesses include:
  - predicting phenotype of progeny (Genetic evaluation)
  - Is this genome region interesting? (GWAS)
  - What happened in this genome region? (selection footprints)
- For these practical purposes, we use statistical models

Statistical effects

- Fisher's described dominance and epistasis as deviations from additivity in a linear statistical model
- Statistical effects (dominance & epistasis) are a population phenomenon
- Genetic model

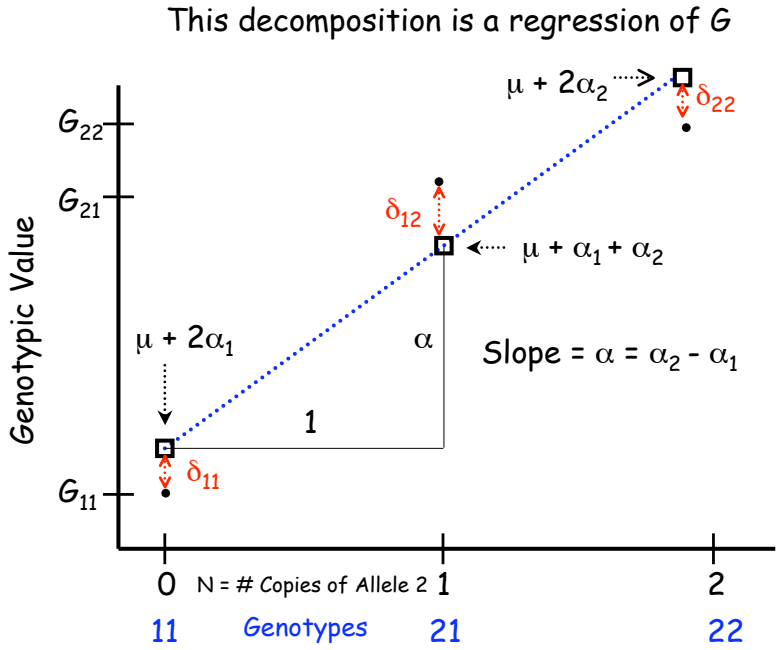


Statistical effects



Fisher (1918) explained that the substitution effect of one allele is the regression of phenotype on genotype

$$\alpha = (\boldsymbol{z}'\boldsymbol{z})^{-1}\boldsymbol{z}'\boldsymbol{y}, \quad \boldsymbol{z} = \begin{cases} \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \end{cases}$$

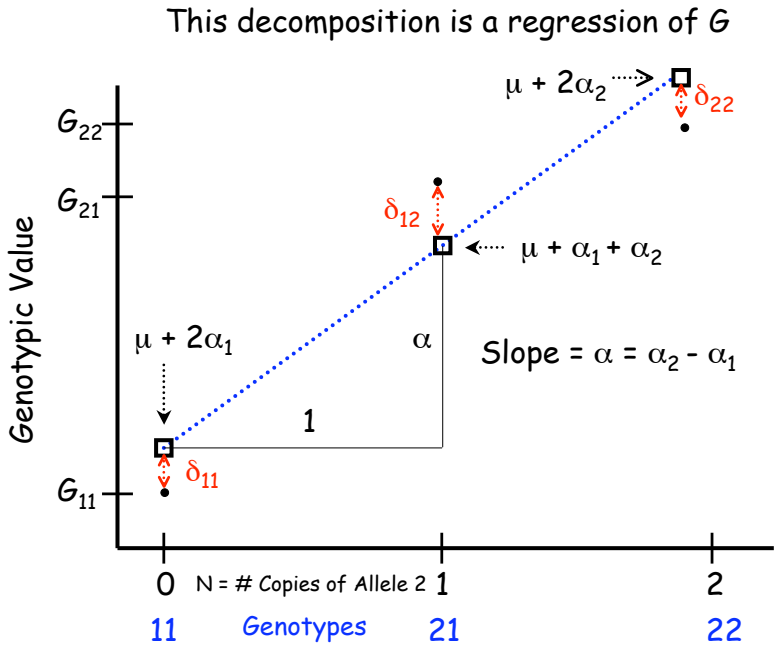


Walsh B., 2013

- Dominance deviations are essentially residuals
- Dominance deviations are the difference for a genotype (in red) between the genotypic value and its prediction from 2 alleles.

Statistical effects

- Why is α relevant & how does it take care of non-additive gene action?
  - The statistical definition doesn't care how α "works"
  - By definition, α potentially includes biological dominance and epistasis
  - Because individuals pass on gametes (and not complete genotypes) to offspring:
  - α describes how much you gain by selecting an allele (in either natural or artificial selection)



Walsh B., 2013

Example pairwise epistasis

$$\alpha_1 =$$

	Biological
a_A	Additive
$+(1-2p_1)d_A$	Dominant
$+(p_2-q_2)i$	Additive x additive
$+(2p_2q_2)j$	Additive x dominant
$+(1-2p_1)(p_2-q_2)l$	Dominant x additive
$+2p_2q_2(1-2p_1)k$	Dominant x dominant

	BB	Bb	bb
AA	$y_1=m+a_A+a_B+i$	$y_4=m+a_A+d_B+j$	$y_7=m+a_A-a_B-i$
Aa	$y_2=m+d_A+a_B+l$	$y_5=m+d_A+d_B+k$	$y_8=m+d_A-a_B-l$
aa	$y_3=m-a_A+a_B-i$	$y_6=m-a_A+d_B-j$	$y_9=m-a_A-a_B+i$

$$\mu = m+a_A(p_1-q_1)+a_B(p_2-q_2)+2p_1q_1d_A+2p_2q_2d_B+(p_1-q_1)(p_2-q_2)i+2(p_1-q_1)p_2q_2j+2p_1q_1(p_2-q_2)l+4p_1q_1p_2q_2k$$

$$\alpha_1 = a_A+d_A(1-2p_1)+(p_2-q_2)i+2p_2q_2j+(1-2p_1)(p_2-q_2)l+2p_2q_2k(1-2p_1)$$

$$\alpha_2 = a_B+d_B(1-2p_2)+(p_1-q_1)i+2p_1q_1l+(1-2p_2)(p_1-q_1)j+2p_1q_1k(1-2p_2)$$

$$d_1 = (d_A-l)+2p_2(l+k)-2kp_2^2$$

$$d_2 = (d_B-j)+2p_1(j+k)-2kp_1^2$$

$$\alpha_1\alpha_2 = (i+j+k+l)-2p_1(j+k)-2p_2(j+k)-2kp_1p_2$$

$$\alpha_1 d_2 = (j+k)-2kp_1$$

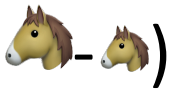
$$d_1\alpha_2 = (1+k)-2kp_2$$

$$d_1d_2 = k$$

Toro, 2017

Statistical & biological effects

Genotypes at locus 1	Genotypes at locus 2		
	BB	Bb	bb
AA			
Aa			
aa			

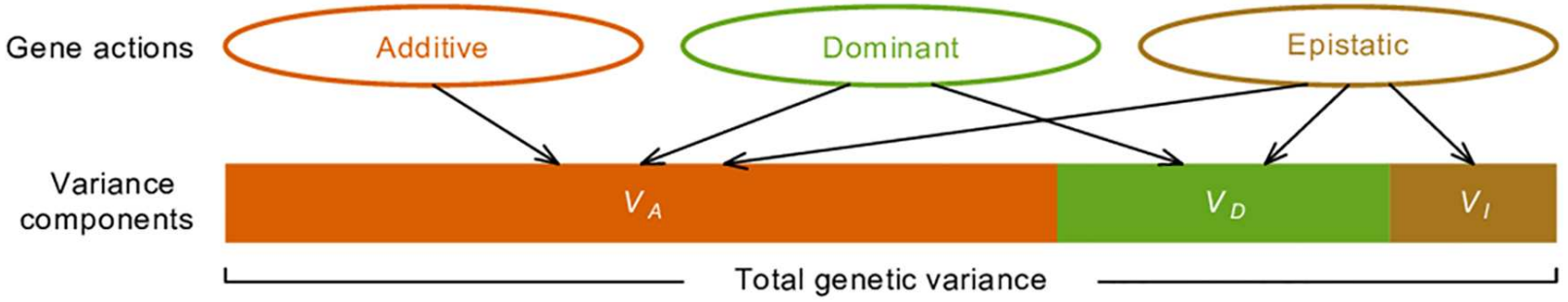
$$\alpha_1 = 2p_2q_2(1-2p_1)(\text{  })$$

Being a “Big” horse is determined by biological **dominance-by-dominance two-locus epistasis**

Genotypes at locus 1	Genotypes at locus 2		
	BB	Bb	bb
AA	μ	μ	μ
Aa	μ	$\mu + \{dd\}_{12}$	μ
aa	μ	μ	μ

Statistical & biological effects

- In the classical $V_A + V_D + V_I$ partition,
  - Additive biological gene actions contribute only to V_A, while
  - Both biological dominant and biological epistatic gene actions contribute to multiple variance components



Huang & Mackay, 2016

  - There is no correspondence between the kind of biological gene action and the variance component

What to do with all these math?

- In absence of knowing true action genes, this gives tools
- α (statistical additive effect) says
  - how much do you improve if you select me
  - Big α = interesting locus
- d_i^* (statistical dominance effect) says
  - For whatever reason, the heterozygote here is interesting
  - Perhaps we can mate these two animals here and maximize it
- $(\alpha\alpha)_{ij}$ (statistical epistatic effect) says
  - Somehow the fates of these two loci are bound together

What to do with all these math?

- α (statistical additive effect) is the ONLY component involved in selection, because only individual alleles are transmitted from parents to descendants
- d_i^* (statistical dominance effect) and $(\alpha\alpha)_{ij}$ (statistical epistatic effect) also contribute to the total genetic value and to the expected phenotype of the crosses/hybrid, but not to selection, because the allele/gene combinations are not transmitted to the descendants

New models accounting for non-additive effects

GBLUP, GDBLUP, and its extensions

Inbreeding depression

“Mixed model” based prediction

- We use quantitative genetic theory to build relationship matrices
- Then we fit them into mixed model

Genomic prediction with non-additive effects

1. We need to construct a linear model based on SNP genotypes
2. Write orthogonal incidence matrices for additive, dominant, additive x additive, additive x dominant… SNP effects
    1. This yields SNP-BLUP or RR-BLUP kind of models but they are cumbersome for epistasis
3. Equivalently, define relationship matrices
    1. High order matrices are products of low order matrices
    2. The whole theory stems from
        1. VanRaden 2008 (A),
        2. Vitezica et al., 2013 **(A+D)**
        3. Vitezica et al., 2017 (**A+D+AxA + any epistatic interactions**)
        4. González-Diéguez et al. (2021) (**A+D+AxA + any epistatic interactions in hybrid crops)**
4. Use a Mixed Model with relationship matrices

This is doable if all individuals are genotyped

- **There is no Single Step GBLUP for dominance or epistasis**

Genomic prediction with non-additive effects

- Recipe:
  1. Define incidence matrices **Z** for α and **W** for d^* , e.g.

$$Z_{ij} = \begin{cases} 2-2p \\ 1-2p \\ 0-2p \end{cases} \quad \text{and} \quad W_{ij} = \begin{cases} -2q^2 \\ 2pq \\ -2p^2 \end{cases} \quad \text{for genotypes} \begin{cases} AA \\ Aa \\ aa \end{cases}$$

  2. Relationship matrices are:
     - $\boldsymbol{G_A} = \dfrac{\boldsymbol{ZZ'}}{2\sum p_i q_i}$ for individual additive effects (GEBVs)
     - $\boldsymbol{D} = \boldsymbol{G_D} = \dfrac{\boldsymbol{WW'}}{4\sum (p_i q_i)^2}$ for dominance deviations

Use in Mixed Model: GD-BLUP

Genomic prediction with non-additive effects

- Recipe:
  2. Relationship matrices are:
     - $\boldsymbol{G_A} = \boldsymbol{ZZ'}/2\sum p_i q_i$ for individual additive effects (GEBVs)
     - $\boldsymbol{D} = \boldsymbol{G_D} = \boldsymbol{WW'}/4\sum (p_i q_i)^2$ for dominance deviations

     - $\boldsymbol{G_{AA}} = \boldsymbol{G_A} \odot \boldsymbol{G_A}/\text{mean}(\text{diag}(\boldsymbol{G_A} \odot \boldsymbol{G_A}))$ for additive x additive
     - $\boldsymbol{G_{AD}} = \boldsymbol{G_A} \odot \boldsymbol{G_D}/\text{mean}(\text{diag}(\boldsymbol{G_A} \odot \boldsymbol{G_D}))$ for additive x dominant
     - …
     - … e.g. $\boldsymbol{G_{AAD}} = \boldsymbol{G_A} \odot \boldsymbol{G_A} \odot \boldsymbol{G_D}/\text{mean}(\text{diag}(\boldsymbol{G_A} \odot \boldsymbol{G_A} \odot \boldsymbol{G_D}))$

Genomic prediction with non-additive effects

- Recipe:
  2. Relationship matrices are:
     - $\boldsymbol{G_{AD}} = \boldsymbol{G_A} \odot \boldsymbol{G_D} / \text{mean}(\text{diag}(\boldsymbol{G_A} \odot \boldsymbol{G_D}))$ for additive x dominant

Genomic **additive** relationship matrix

Genomic **dominant** relationship matrix

$$\boldsymbol{G}_{AD} = \frac{\boldsymbol{G}_A \odot \boldsymbol{G}_D}{tr(\boldsymbol{G}_A \odot \boldsymbol{G}_D)/n}$$

Use in Mixed Model: GDI-BLUP

A standardization based on the trace of the relationship matrices is needed.

Genomic prediction with non-additive effects

- Recipe:
  - Then use these matrices in (G)(D)(I)BLUP / REML

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{g}_A + \boldsymbol{g}_D + \boldsymbol{g}_{AA} + \boldsymbol{g}_{AD} + \boldsymbol{g}_{DD} + \cdots (+\boldsymbol{pe}) \ldots + \boldsymbol{e}$$

$$Var(\boldsymbol{g}_A) = \boldsymbol{G}_A\sigma_A^2; Var(\boldsymbol{g}_D) = \boldsymbol{D}\sigma_D^2; Var(\boldsymbol{g}_{AA}) = \boldsymbol{G}_{AA}\sigma_{AA}^2$$

- $\boldsymbol{pe}$ is the permanent environmental effect
  - captures remaining genetic effects (e.g. AxAxAxA…) in repeated records (such as analysis of milk yield)
- The matrices of higher orders $\boldsymbol{G}_{AA}, \boldsymbol{G}_{AAA}, \boldsymbol{G}_{AAAA}$ are increasingly less informative and at some point they're not worth fitting.

Genomic prediction with non-additive effects – crosses in hybrid crops

- In hybrid crops like maize, the cultivated plant is usually an F1 hybrid which is the cross of two homozygote lines, each from a different population (“heterotic group”)
- Parental homozygote lines are homozygous at all loci
- This generates a particular partition of additive, dominance and epistasis *across* and *within* heterotic groups

Genomic prediction of hybrid crops allows disentangling dominance and epistasis

David González-Diéguez [1,*], Andrés Legarra[1], Alain Charcosset[2], Laurence Moreau[2], Christina Lehermeier [3], Simon Teyssèdre[3], and Zulma G. Vitezica [1]

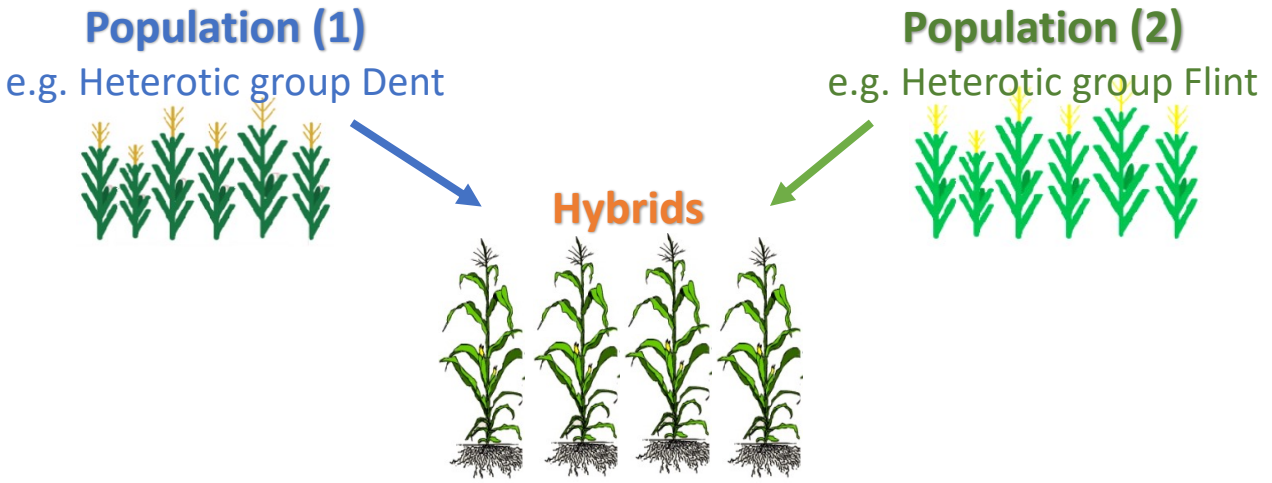
GENETICS, 2021, **218(1)**, iyab026

DOI: 10.1093/genetics/iyab026

Genomic prediction in hybrid crops

- Hybrid crops from pure lines
  - E.g. maize: population 1 is “Dent” and population 2 is “Flint”
  - The effects (GCA and SCA) are defined “according to parental origin”
  - We define **Z**-matrices within each heterotic groups
  - **W**-matrix is defined in the hybrid

David González-Diéguez



Genomic prediction in hybrid crops

- Recipe:
  1. For each locus,

  define incidence matrices $\mathbf{Z}_1$ for α_1 (pop 1), $\mathbf{Z}_2$ for α_2 (pop 2) and $\mathbf{W}$ for d^* (in hybrids)

$$Z_{1_{ij}} = \begin{cases} (1-p_1) \\ (\quad -p_1) \end{cases} \text{ for genotypes } \begin{cases} B_1B_1 \\ b_1b_1 \end{cases}, \qquad Z_{2_{ij}} = \begin{cases} (1-p_2) \\ (\quad -p_2) \end{cases} \text{ for genotypes } \begin{cases} B_2B_2 \\ b_2b_2 \end{cases}$$

and

$$W_{ij} = \begin{cases} -2q_1q_2 \\ 2q_1p_2 \\ 2p_1q_2 \\ -2p_1p_2 \end{cases} \text{ for genotypes } \begin{cases} B_1B_2 \\ B_1b_2 \\ b_1B_2 \\ b_1b_2 \end{cases}$$

New models accounting for non-additive effects

GBLUP, GDBLUP, and its extensions

Inbreeding depression

Inbreeding / heterosis

- Inbreeding depression is the decline in biological fitness (viability, fertility, …) as a consequence of inbreeding

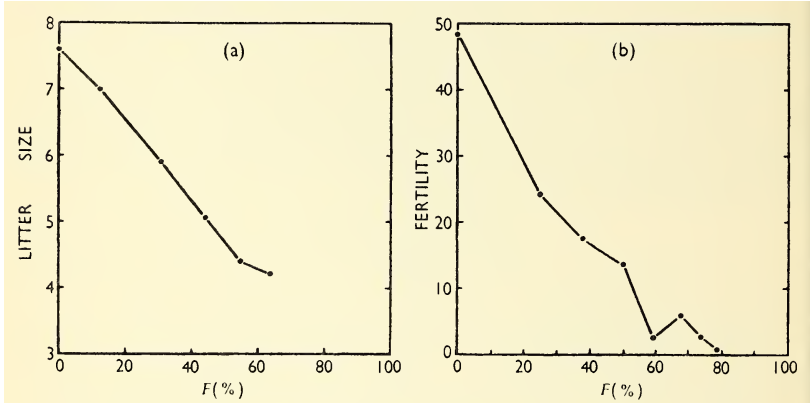


FIG. 14.1. Examples of inbreeding depression affecting fertility. (*a*) Litter-size in mice (original data). Mean number born alive in 1st litters, plotted against the coefficient of inbreeding of the litters.

(Falconer, 1981)

- This phenomenon may be explained by directional dominance.
- Directional dominance, e.g. the heterozygote is usually “better”

(Lynch & Walsh, 1998)

Inbreeding/ heterosis

- If heterosis or inbreeding depression, $E(\boldsymbol{d}) = \mathbf{1}\mu_D$ with $\mu_D > 0$
- Statistically this translates into a regression on a measure *F* of homozygosity ($\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{F}b + \boldsymbol{g}_A + \boldsymbol{g}_D + \cdots + \boldsymbol{e}$)
  - Across individual markers: “genomic inbreeding” (Silio et al 2013; Xiang et al 2016)
  - In blocks: ROHs (long ROHs are better because inbreeding has not been purged)
- Ignoring inbreeding/heterosis may inflate estimates of dominance variance
- Including inbreeding/heterosis allows finer estimates of EBV

Results?

OK, so we have this nice theory, what now?

- Is this any useful?
  - Extra accuracy in predictions
  - Variance components
  - Mate allocation

Example in pigs

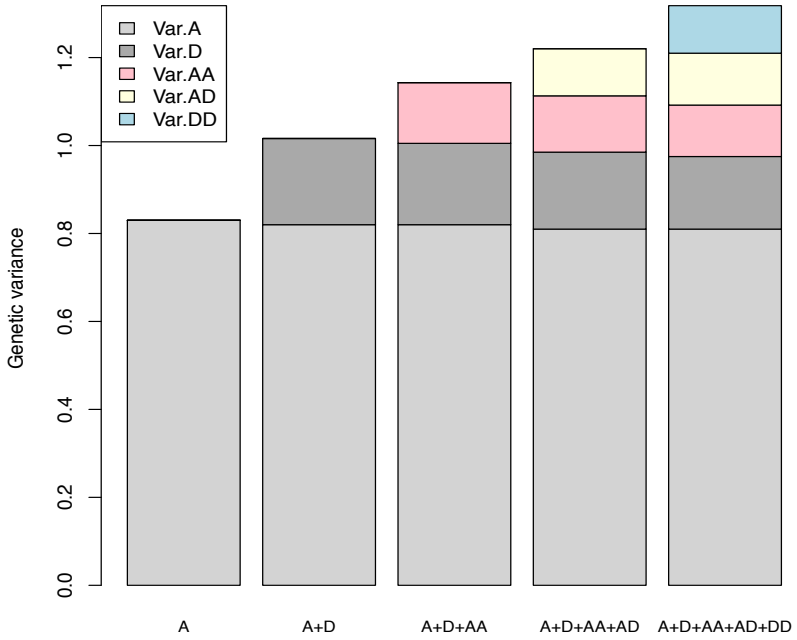
$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{F}b + \boldsymbol{g}_A + \boldsymbol{g}_D + \boldsymbol{g}_{AA} + \boldsymbol{g}_{DA} + \boldsymbol{g}_{DD} + \boldsymbol{pe} + \boldsymbol{e}$$

- Small variances for non-additive effects
- The model is empirically orthogonal: variance component estimates do not change by adding an extra term
- Inclusion of dominance/epistasis did not increase the accuracy of prediction of breeding values

PIC®

Litter size
12.7 ± 3.1

Genus plc (Hendersonville, TN, USA)
3,619 genotyped sows 13,369 records
38,779 SNPs

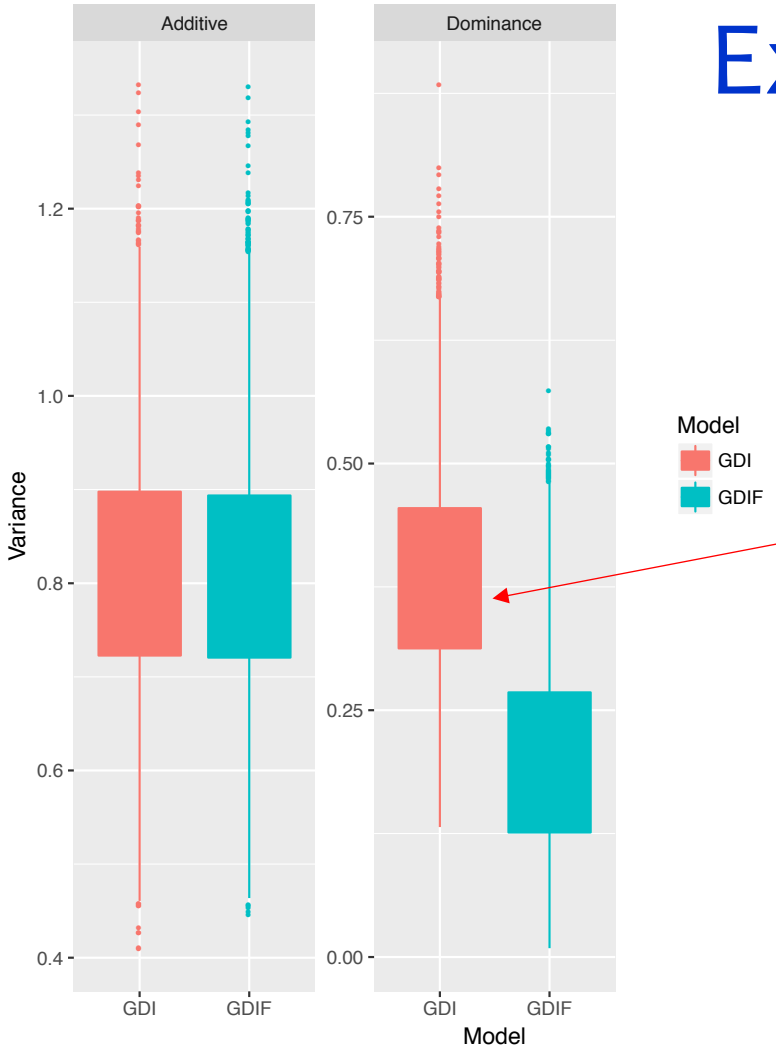


From Genus

Vitezica et al., 2018.

Example in pigs

From Genus



Posterior distributions of additive and dominance genetic variances for model including (GDIF) or not (GDI) genomic inbreeding

Without including inbreeding depression in the model, dominance variance was overestimated

This has long been known for pedigree analysis (e.g. DeBoer and Hoeschele, 1993).

Vitezica et al., 2018.

Example in beef cattle

Carolina Garcia-Baccino

American Angus Association
19,375 genotyped males
39,245 SNPs

Journal of Animal Science, 2020, 1–7
doi:10.1093/jas/skz384
Advance Access publication December 23, 2019
Received: 29 August 2019 and Accepted: 19 December 2019
Animal Genetics and Genomics

ANIMAL GENETICS AND GENOMICS

Estimating dominance genetic variances for growth traits in American Angus males using genomic models

Carolina A. Garcia-Baccino,[*,1] Daniela A. L. Lourenco,[†] Stephen Miller,[‡] Rodolfo J. C. Cantet,[*,§] and Zulma G. Vitezica[#]

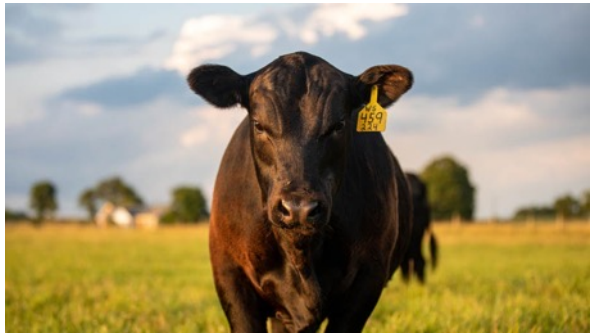
[*]Departamento de Producción Animal, Facultad de Agronomía, Universidad de Buenos Aires, C1417DSQ Buenos Aires, Argentina, [†]Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602, [‡]Angus Genetics Inc., St. Joseph, MO 64506, [§]INPA, UBA-CONICET, C1427CWO Buenos Aires, Argentina, [#]INP ENSAT, UMR 1388 GenPhySE, 31326 Castanet-Tolosan, France

Small variances for non additive effects
Inclusion of dominance in the model did not increase the accuracy of prediction of breeding values

Table 2. Estimates of additive, dominance deviation, and residual variance components (σ^2_A, σ^2_D, σ^2_e) and heritability for growth traits using MG and MGD models

Trait[1]	Model[2]	σ^2_A	σ^2_D	h^2_A	h^2_D	(σ^2_D/σ^2_A)	σ^2_e
BW	MG	6.27 (0.33)	—	0.25	—	—	18.82 (0.24)
	MGD	6.28 (0.33)	0.18 (0.15)	0.25	0.01	0.03	18.65 (0.28)
WW	MG	222.75 (14.61)	—	0.16	—	—	1186.28 (14.26)
	MGD	223.55 (14.82)	10.02 (4.98)	0.16	0.01	0.04	1176.88 (14.86)
PWG	MG	270.76 (20.42)	—	0.16	—	—	1388.81 (19.87)
	MGD	270.30 (21.94)	21.68 (10.95)	0.16	0.01	0.08	1369.01 (26.00)

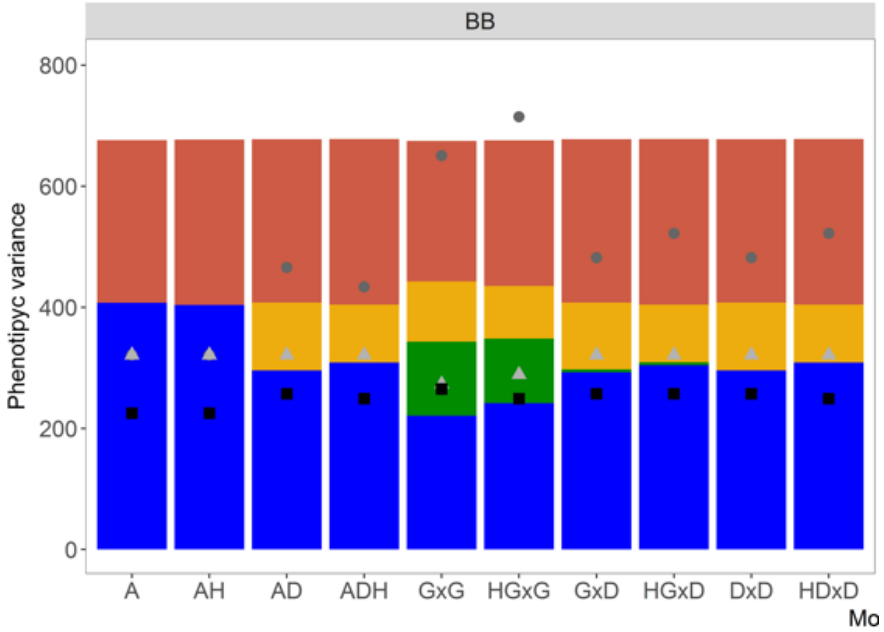
[1]BW, birth weight; WW, weaning weight; PWG, postweaning gain.
[2]MG, model including only additive effects; MGD, model including both additive and dominant effects.
The results are given as estimate (in parenthesis SE); $h^2_A = \sigma^2_A/\sigma^2_P$ and $h^2_D = \sigma^2_D/\sigma^2_P$, where σ^2_P is the phenotypic variance.



From AAA

Example in beef cattle

2,111 **Australian Brahman** (BB) cows and bulls
Genotyped with 770,000 SNPs
Body yearling weight



From ABBA

Small variances for non additive effects

Without including inbreeding depression in the model, dominance variance was overestimated

Raidan et al., 2018

Results

- Inclusion of dominance/epistasis
  - does not increase the accuracy of prediction of breeding values (Ertl et al., 2014; Xiang et al, 2016; Esfandyari et al., 2016; Moghaddar and van der Werf, 2017, González-Diéguez et al., 2019, Garcia-Baccino et al., 2020 – Pégard et al., 2020, González-Diéguez et al., 2021)
  - with the exception of Aliloo et al. (2016) (for fat yield in Holstein)
- Inclusion of inbreeding depression/heterosis effect
  - does increase predictive ability (Xiang et al., 2016) in pigs
  - and in maize (Roth et al., 2022)
- Fitting non-orthogonal models or non fitting inbreeding
  - Biases in variance component estimation (Vitezica et al. 2013; 2018)

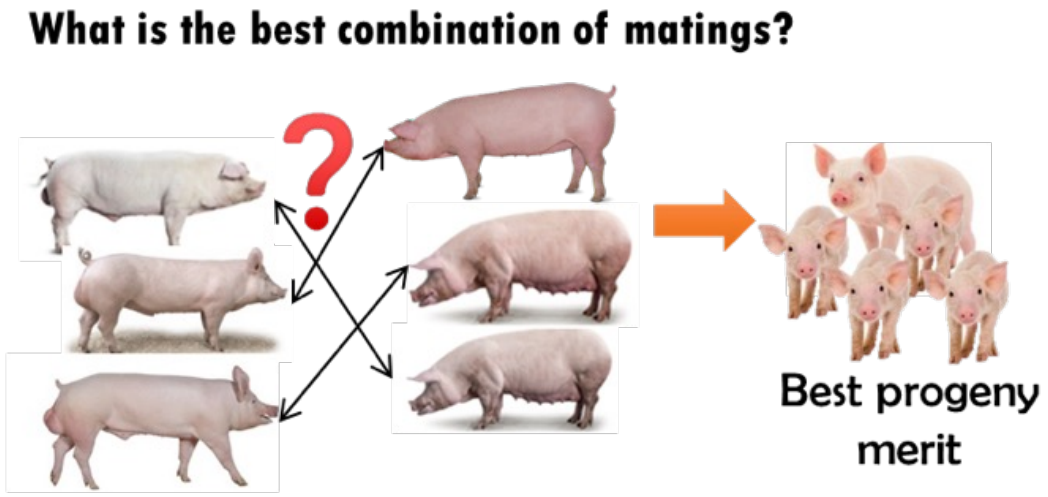
Results?

OK, so we have this nice theory, what now?

- Is this any useful?
  - Extra accuracy in predictions
  - Variance components
  - Mate allocation

Mate allocation: theory

- What happens if I mate *i* and *j* so that the product has an extraordinarily good phenotype (=dominance deviation)?

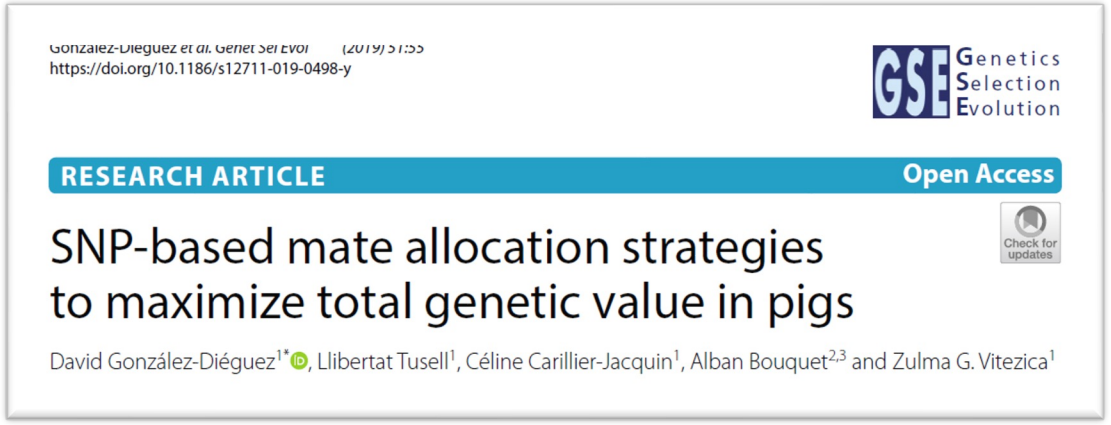


Example in pigs (within breed)

David González-Diéguez

France Genetic Porc
Age at 100 kg (AGE), Backfat depth (BD), Average piglet weight at birth (APWL)
39,353 SNPs

Trait	Boars	Sows	Genotyped animals	Number of records	Mean (SD)
AGE (days)	789	2179	2968	2968	149.03 (9.36)
BD (mm)	1007	2675	3682	3682	11.20 (1.68)
APWL (g)	1446	1226	2672	3297	1321.73 (213)


González-Diéguez et al. Genet Sel Evol (2019) 51:55
https://doi.org/10.1186/s12711-019-0498-y

Genetics Selection Evolution

RESEARCH ARTICLE Open Access

SNP-based mate allocation strategies to maximize total genetic value in pigs

David González-Diéguez[1*], Llibertat Tusell[1], Céline Carillier-Jacquin[1], Alban Bouquet[2,3] and Zulma G. Vitezica[1]

Landrace français

Example in pigs (within breed)

Estimation of variance components: σ_A^2 σ_D^2

- Model GD : additive + dominance + genomic inbreeding

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Fb} + \boldsymbol{Zu} + \boldsymbol{Zv} + \boldsymbol{e}$$

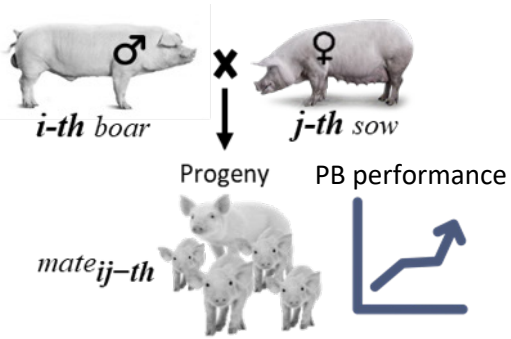
$\boldsymbol{F}$ is a vector of genomic inbreeding coefficients

b is the inbreeding depression parameter

$\boldsymbol{u} \sim N(\mathbf{0}, \boldsymbol{G}\sigma_A^2)$, $\boldsymbol{G}$ built as in VanRaden (2008)

$\boldsymbol{v} \sim N(\mathbf{0}, \boldsymbol{D}\sigma_D^2)$, $\boldsymbol{D}$ built as in Vitezica *et al.* (2013)

Within breed



- Model G : only additive + genomic inbreeding

remlf90 software (Misztal *et al.* 2012)

Estimation of additive and dominant SNP effects: $\hat{a}$ and $\hat{d}$

- BLUP-SNP model including dominance and genomic inbreeding

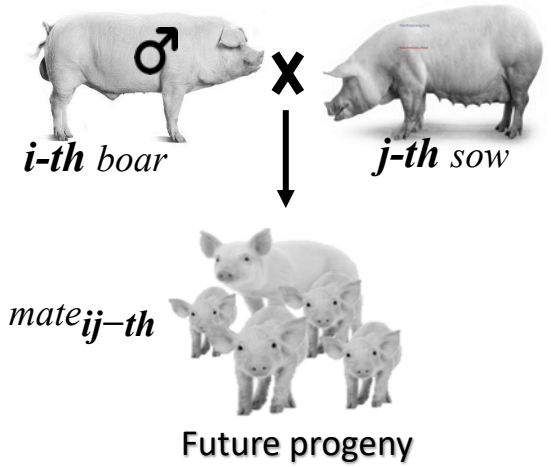
GS3 software (Legarra *et al.* 2011)

Example in pigs (within breed)

Prediction of expected progeny values (Toro and Varona 2010):

- Prediction of the total genetic values (g_{ij}) of the mating

$$\hat{g}_{ij} = \sum_k \left[P_{ijk}(CC)\hat{a}_k + P_{ijk}(CT)\hat{d}_k + P_{ijk}(TT)(-\hat{a}_k) \right]$$



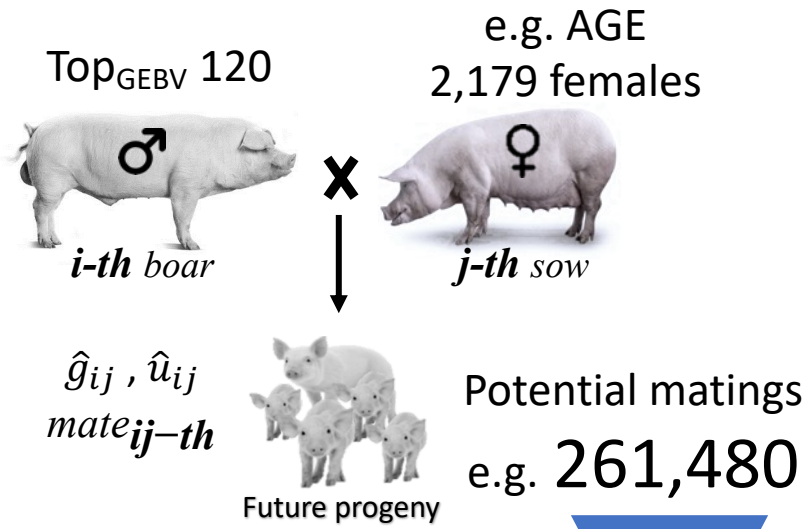
- Prediction of the breeding values (u_{ij}) of the progeny

$$\hat{u}_{ij} = \sum_k \left[P_{ijk}(CC)(2-2p_k)\hat{\alpha}_k + P_{ijk}(CT)(1-2p_k)\hat{\alpha}_k + P_{ijk}(TT)(-2p_k)\hat{\alpha}_k \right]$$

$$\hat{\alpha}_k = \hat{a}_k + \hat{d}_k(q_k - p_k)$$

Example in pigs (within breed)

Allocation of matings



Two mate allocation strategies:

(1) 600 matings selected on $\hat{u}_{ij} \rightarrow f_{optim}(\hat{u}_{ij})$

(2) 600 matings selected on $\hat{g}_{ij} \rightarrow f_{optim}(\hat{g}_{ij})$

Evaluation of expected genetic gains:

Additive genetic gain (Δu):

- $\Delta u = mean(\hat{u}_{600}) - mean(\hat{u}_{all\_matings})$

Total genetic superiority (Δg):

- $\Delta g = mean(\hat{g}_{600}) - mean(\hat{g}_{all\_matings})$

Optimization by linear programming

R package *lpsolve* (Berkelaar *et al.*, 2004)

Two constraints:

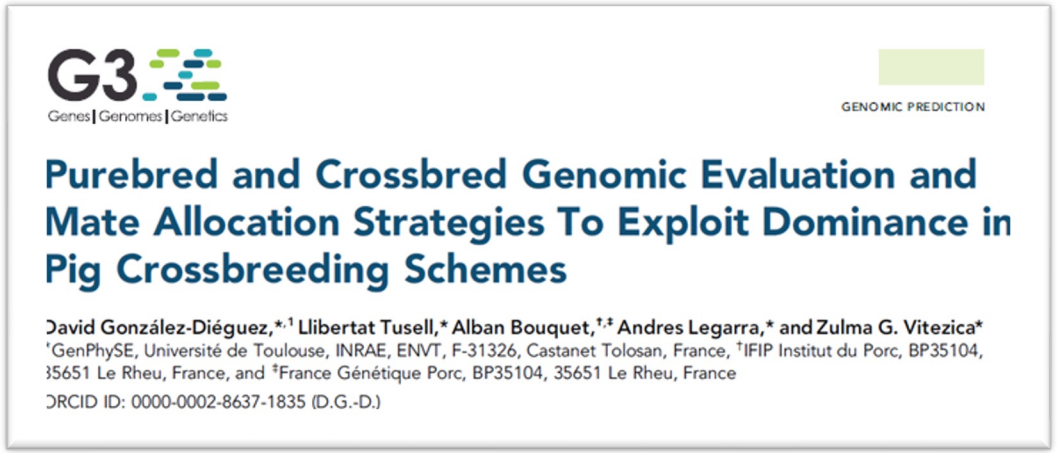
(1) each boar could be mated to up to 15 sows

(2) each sow could not be mated to more than one boar

Example in pigs (across breeds)

Is it possible to boost CB performance by implementing mate allocation in a two-way pig crossbreeding scheme in the long term?

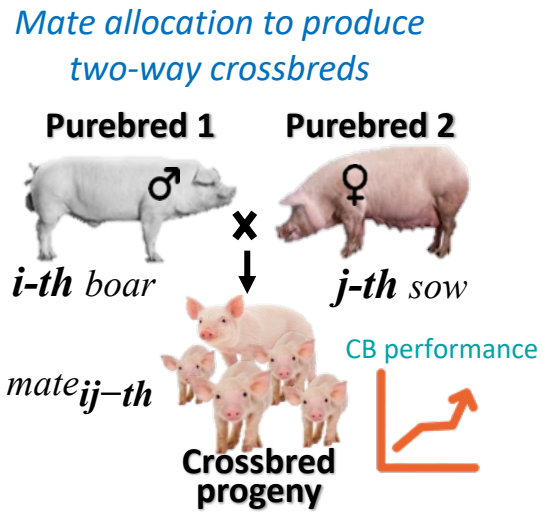
Simulation study (QMSim + Fortran program)
Maternal trait: litter size
Genome: 18 Chr 120 cM each


G3 Genes | Genomes | Genetics

GENOMIC PREDICTION

Purebred and Crossbred Genomic Evaluation and Mate Allocation Strategies To Exploit Dominance in Pig Crossbreeding Schemes

David González-Diéguez,*,[1] Llibertat Tusell,* Alban Bouquet,[†,‡] Andres Legarra,* and Zulma G. Vitezica*
*GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326, Castanet Tolosan, France, [†]IFIP Institut du Porc, BP35104, 35651 Le Rheu, France, and [‡]France Génétique Porc, BP35104, 35651 Le Rheu, France
ORCID ID: 0000-0002-8637-1835 (D.G.-D.)



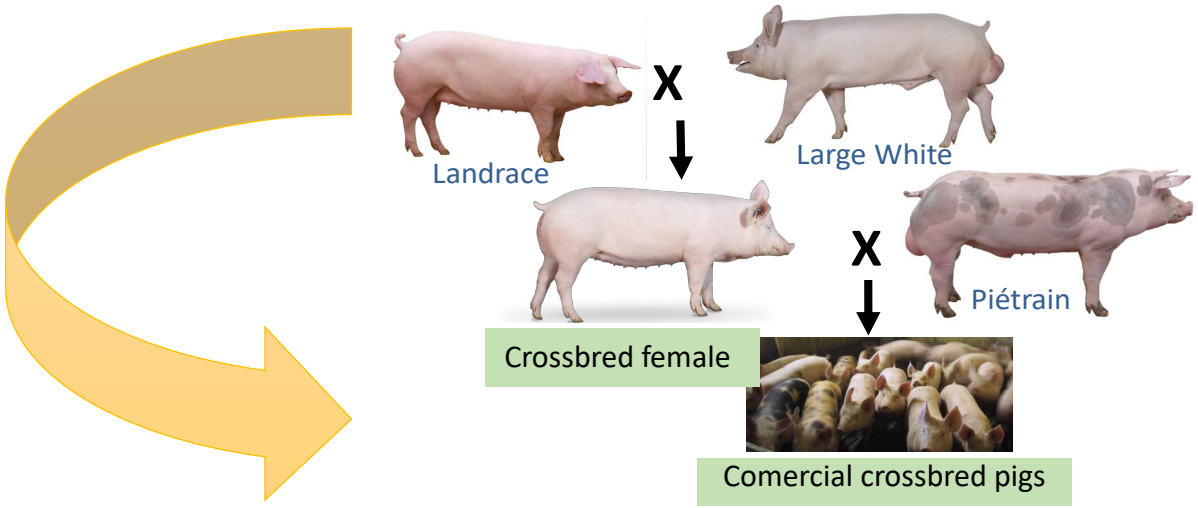
David González-Diéguez

Sargolzaei and Schenkel, 2009

Genetic improvement in pigs

- It uses selection and crossbreeding
- The breeding goal is to improve crossbred (CB) performance, while selection takes place in purebred (PB) animals based on PB performances
- Selection depends on the correlation between PB and CB performance (r_{PC})

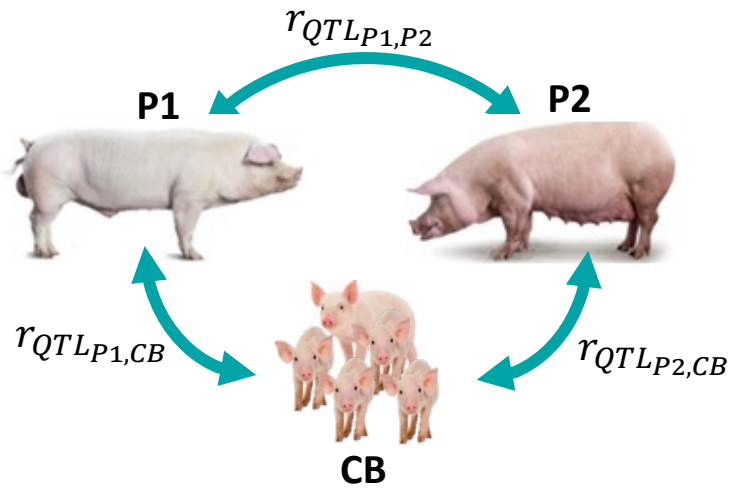
Selection may be suboptimal (GxE)
$r_{PC} < 1$ (~0.7)



Example in pigs (across breeds)

Simulation of heterosis and QTL effects

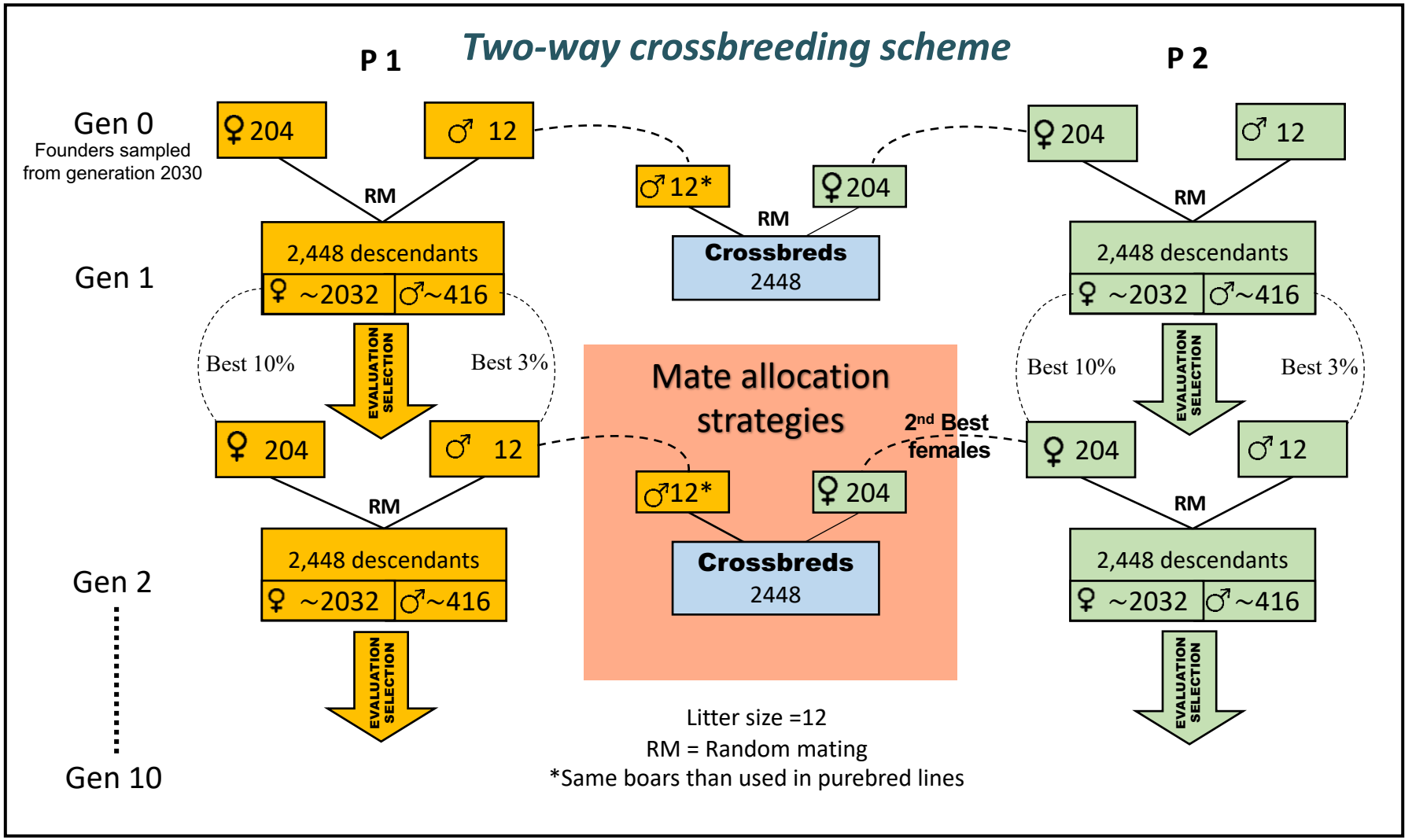
- Maternal trait: “e.g. Litter size” controlled by additive and dominant QTL action (2,500 QTLs)
- **Inbreeding depression** was assumed to be -1 piglet per 10% increase in genomic inbreeding in P1, P2 and CB
- Additive and dominance QTL effects were sampled from a MVN distribution with **correlation between the three populations to account for GxE and GxG**. Landrace and Yorkshire genetic variances were taken from Xiang *et al.* (2016)



Correlation between QTLs (r_{QTL}):

$r_{QTL_{P1,CB}} = r_{QTL_{P1,P2}} = r_{QTL_{P2,CB}} = 0.5$

Example in pigs (across breeds)



Mate allocation: results

- Mate allocation has a small added benefit within-breed and no benefit across-breed
- Selecting PB animals for CB performance using PB and CB data is a good strategy to exploit heterosis and improve crossbred performance, especially if the r_{PC} is low

Some conclusions

- We have a comprehensive theory
- We know how to properly define/estimate non-additive statistical effects
- Inbreeding/heterosis should be fit in the genetic evaluation model
- Fitting dominance and epistatic effects is interesting to correctly appraise genetic variances

Some conclusions

- Dominance and epistasis is not difficult with markers provided all animals ☺ (plants ☺) are genotyped
- In our experience, computational complexity is not an issue (models fit into computers), but convergence and accuracy are an issue (many parameters, little information)

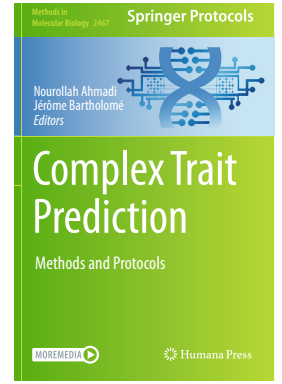
frontiers in Genetics

REVIEW
published: 06 March 2018
doi: 10.3389/fgene.2018.00078

Non-additive Effects in Genomic Selection

Luis Varona[1,2*], *Andres Legarra*[3], *Miguel A. Toro*[4] *and Zulma G. Vitezica*[5]

[1] *Departamento de Anatomía, Embriología y Genética Animal, Universidad de Zaragoza, Zaragoza, Spain,* [2] *Instituto Agroalimentario de Aragón (IA2), Zaragoza, Spain,* [3] *Génétique Physiologie et Systèmes d'Elevage (GenPhySE), Institut National de la Recherche Agronomique de Toulouse, Castanet-Tolosan, France,* [4] *Departamento Producción Agraria, ETS Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, Madrid, Spain,* [5] *Génétique Physiologie et Systèmes d'Elevage (GenPhySE), Université de Toulouse, Castanet-Tolosan, France*



Chapter 8

Genomic Prediction Methods Accounting for Nonadditive Genetic Effects

Luis Varona, Andred Legarra, Miguel A. Toro, and Zulma G. Vitezica

Acknowledgments

Luis Varona
Miguel Angel Toro

Manolo Baselga
Juan Pablo Sanchez
Miguel Perez-Enciso

Daniela Laurenco
Ignacy Misztal

Ole Christensen

Toni Reverter
Fernanda Raidan

Andrés Legarra
Llibertat Tusell
Jean Michel Elsen
Alan Charcosset
Laurence Moreau

Johann Ertl
Eduardo Fernandez
Natalia Forneris
Carolina Garcia-Baccino
Tao Xiang
David Gonzalez-Dieguez

Christina Lehermeier
Simon Teyssedre

Thank you for your attention!