

ESTIMATING EFFECTIVE POPULATION SIZE FROM GENOMIC RATES OF COANCESTRY AND INBREEDING

Raúl López de la Torre¹

Fernández J, Morales-González E, Fernández A, Toro M A, Caballero A and Villanueva B

¹Department of Animal Breeding, INIA-CSIC

raul.lopez@inia.csic.es



Background

- **Effective population size (N_e)**
 - Fundamental parameter in **quantitative genetics** and **evolutionary biology**
 - It plays a crucial role in **animal breeding** and **conservation** programs

*Size of an **idealized population** that would experience the **same rate of inbreeding** as in the population under consideration (random mating, no mutation, migration or selection)*

- Normally, $N_e \ll$ census size (N)



Background

- N_e determines the magnitude of variability loss and the increase in inbreeding in the population
 - **Genetic variability** -- Determines the capacity for adaptation and response to selection
➡ **coancestry coefficient (f)**
 - **Inbreeding depression** -- Reduction in the mean of quantitative traits due to inbreeding
➡ **inbreeding coefficient (F)**

$$N_e = 1/2\Delta f$$
$$N_e = 1/2\Delta F$$

Background

- Coancestry and inbreeding coefficients can be obtained from:

Pedigree ➡ Expected values

Genomic information ➡ Realized values

- Many proposed **genomic coefficients**
- Extensive research comparing different genomic coefficients for various purposes
 - *Identify gain and loss of genetic variability, estimate inbreeding depression, etc*

No comparisons of genomic coancestry and inbreeding for estimating N_e

Objectives

- 1 Evaluate the evolution of N_e using different genomic measures of coancestry and inbreeding - **monitoring** the population

Random Contributions (RC)

parents contribute randomly to the next generation

$$N_e \approx N$$

Equalized Contributions (EC)

all parents have the same number of offspring in the next generation

$$N_e \approx 2N$$

Using **stochastic simulations**



Objectives

- 2 Evaluate N_e using different **genomic coancestry matrices** for **managing** the population

Management through the Optimal Contribution Method (OC)

method that optimizes the contributions of candidates to the next generation by minimizing the global coancestry

Minimize $\mathbf{c}' \boldsymbol{\theta} \mathbf{c}$

- $\mathbf{c}$ is vector of proportions of offspring left by each candidate
- $\boldsymbol{\theta}$ is the coancestry matrix.

➡ Minimizes f

➡ Minimizes Δf

➡ Maximizes N_e

Using **stochastic simulations**



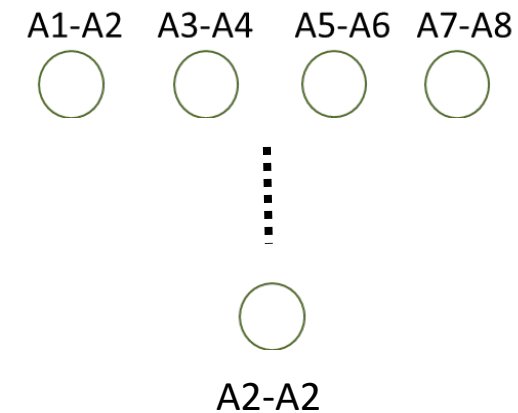
Materials & Methods

- **Base population** ➡ Mutation-drift equilibrium

Size ➡ 100 individuals

Genome {
20 chromosomes (1 Morgan each)
~50,000 **biallelic loci** (SNPs)
Used for monitoring and managing
10,000 **multiallelic loci** ➡ *Unique alleles at $t = 0$*
*Used for obtaining **IBD values** of f and F
(considered as the true values)*

- Monitoring and managing ➡ 10 discrete generations
- 100 replicates



Genomic coancestry matrices

- Using multiallelic loci

$$\theta_{IBD}$$

Locus-by-locus similarity between two individuals

- Using SNPs

- θ_{LH1}
- θ_{LH2}

Deviations from Hardy-Weinberg proportions

Genomic relationship matrix of Li & Horvitz (1953)

- θ_{VR1}

Rare alleles contributed more to f and F measures

Genomic relationship matrix as described in VanRaden (2008) – Method 1

- θ_{VR2}

Rare alleles contributed even more

Genomic relationship matrix as described in VanRaden (2008) – Method 2

- θ_{YA1}

Decrease the variance of f_{VR1}

Genomic relationship matrix as described in Zhang *et al.* (2022)

- θ_{YA2}

Decrease the variance of f_{VR2}

Genomic relationship matrix as described in Yang *et al.* (2010)

Genomic coancestry and inbreeding coefficients

- f obtained from the average f_{ij} (including self-coancestries, f_{ii})
- F obtained from the average F_i , taking F_i from the diagonal elements (self-coancestries, f_{ii})

	1	2	3	4
1	f_{11}	f_{12}	f_{13}	f_{14}
2	f_{21}	f_{22}	f_{23}	f_{24}
3	f_{31}	f_{32}	f_{33}	f_{34}
4	f_{41}	f_{42}	f_{43}	f_{44}

$$F_i = 2 f_{ii} - 1$$

Parameters evaluated

Objective 1

RC

EC



N_e estimated from **different genomic Δf** and **ΔF**
Estimates of N_e are compared to true N_e (IBD)

Objective 2

OC

True N_e estimated when **using different genomic coancestry matrices** in the Optimal Contribution Method

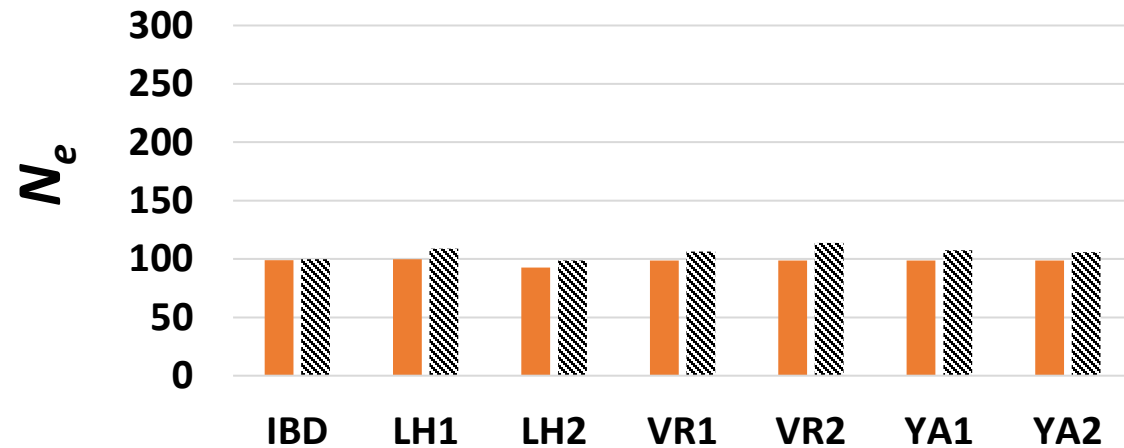
Coefficient of variation of N_e across replicates

Results – Objective 1 - Monitoring

$t = 10$

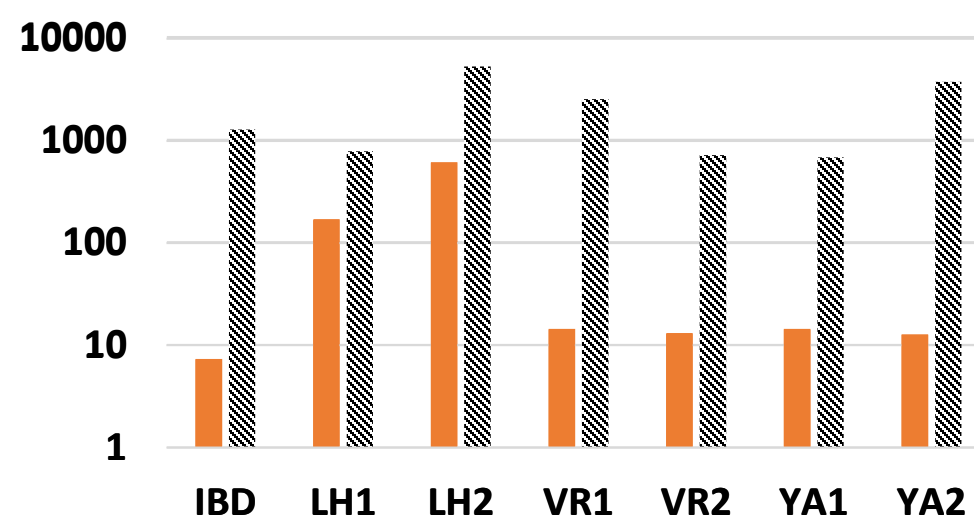
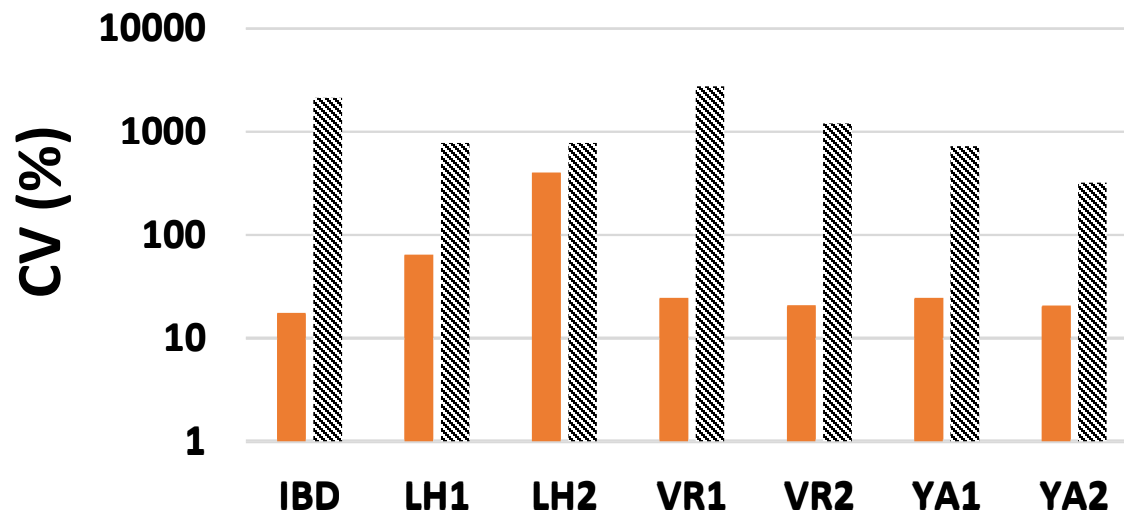
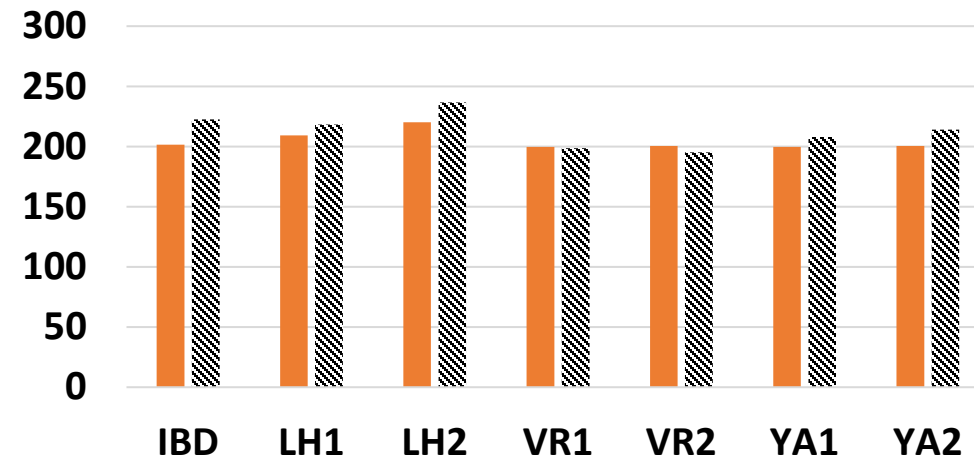
RC

$N_e \approx 100$

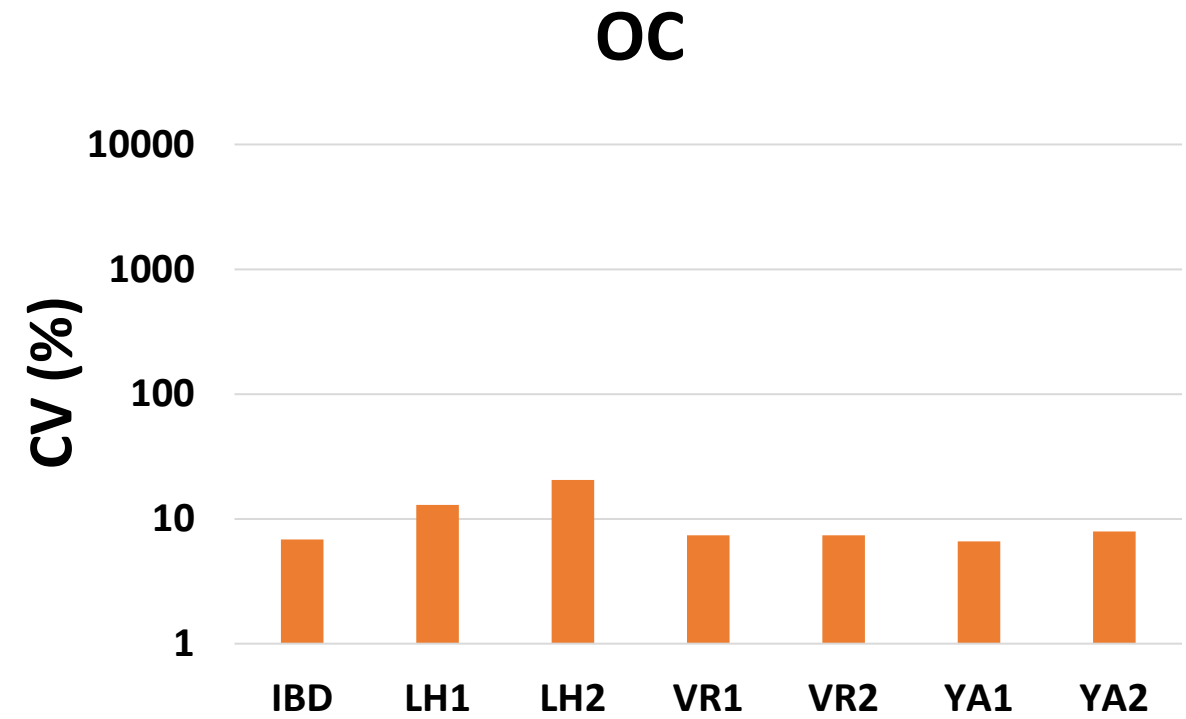
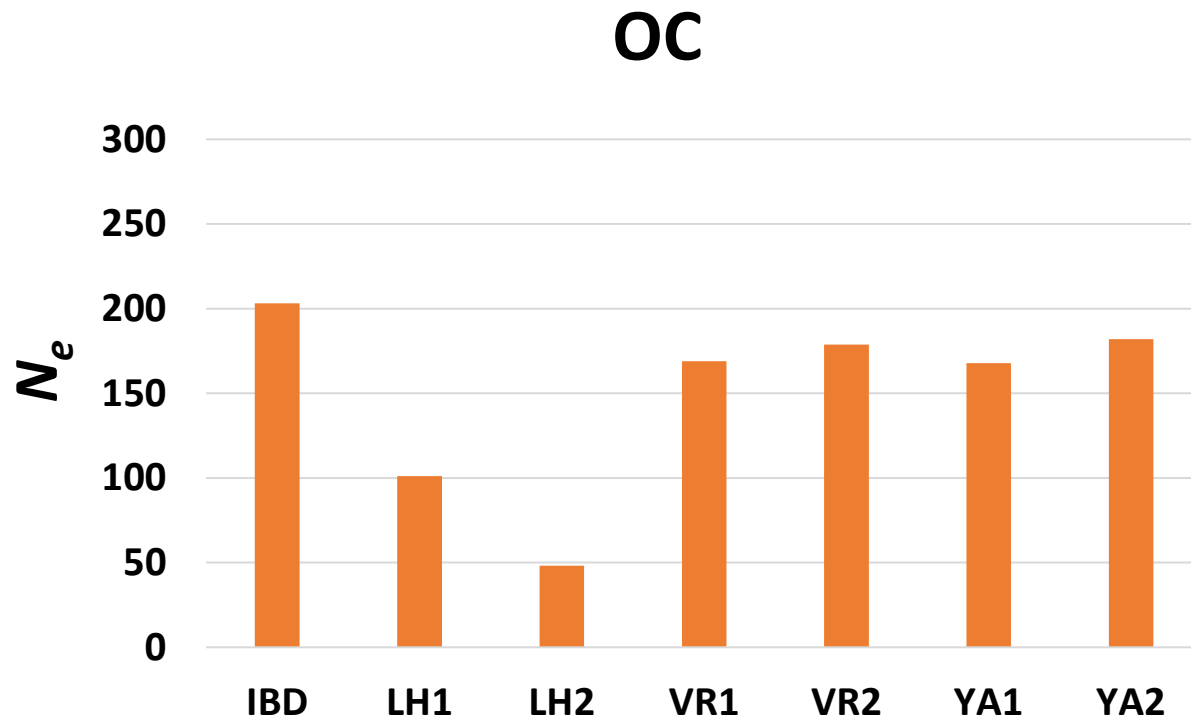


EC

$N_e \approx 200$



Coancestry
Inbreeding



Conclusions

- N_e is **more estable** when is calculated from **rate of coancestry** → **lowest CV**

Objective 1 - RC and EC

- θ_{YA2} and θ_{VR2} → Closest to IBD Lowest variation
- θ_{LH1} and θ_{LH2} → Furthest to IBD Highest variation

Objective 2 - OC

- θ_{YA2} and θ_{VR2} → Highest N_e Lowest variation
- θ_{LH1} and θ_{LH2} → Lowest N_e Highest variation

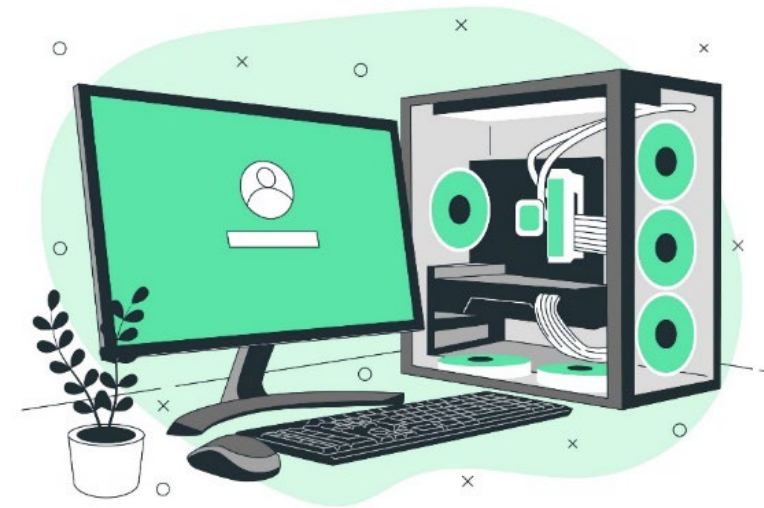
This study highlights the importance of choosing an appropriate measure of genomic coancestry in management programs

Acknowledgments

Thank you very much for your attention!

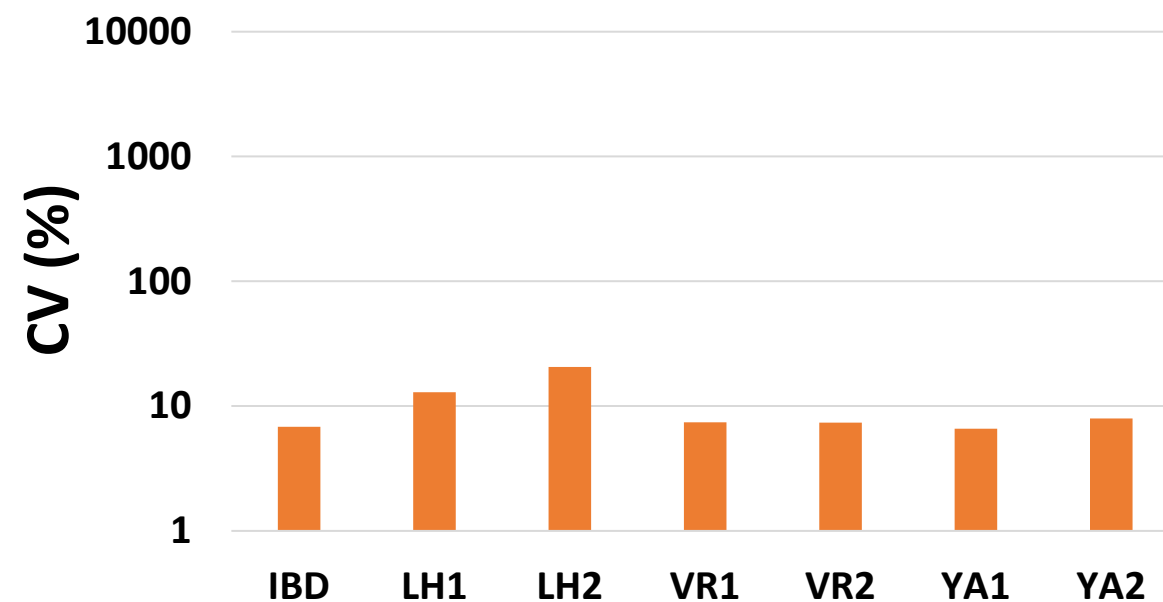
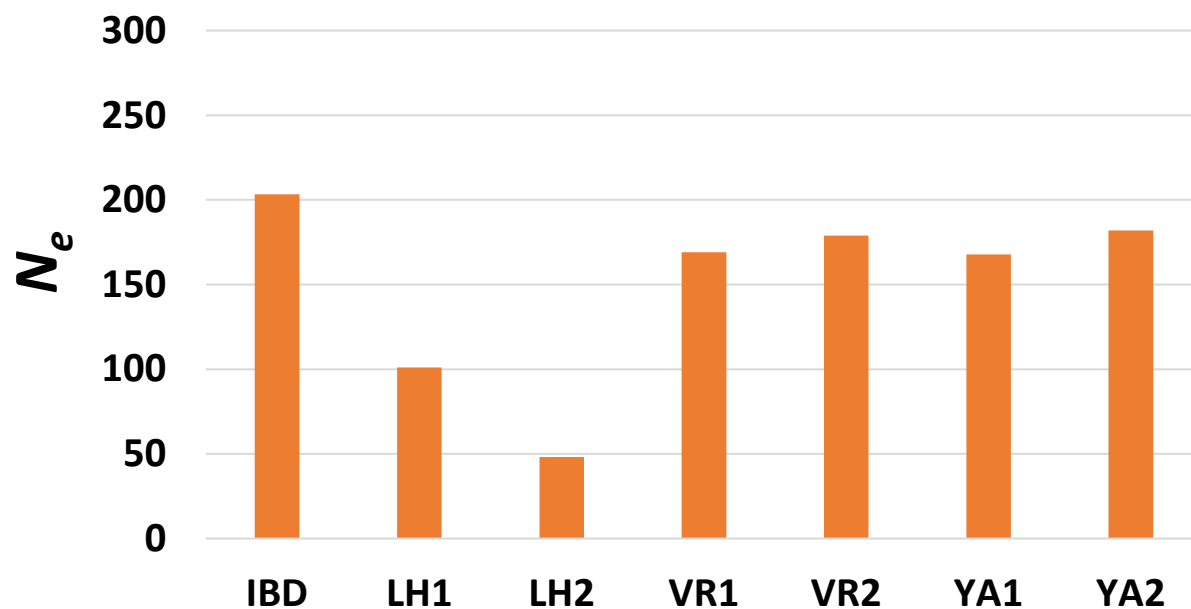
Project PID2020-114426GB-C2

Predoctoral scholarship: PRE2021-096997



Results – Objective 2 - Managing

Coancestry at $t = 10$



Number of individuals contributing

Gen	IBD	LH1	LH2	VR1	VR2	YA1	YA2
1	100	62	32	100	100	100	100
10	99	74	46	92	94	92	95