

EVALUACIONES GENÓMICAS

Luis Varona

Genética Cuantitativa y Mejora Animal

Universidad de Zaragoza



**GOBIERNO
DE ESPAÑA**

**MINISTERIO
DE MEDIO AMBIENTE
Y MEDIO RURAL Y MARINO**

MEJORA GENÉTICA

SELECCIÓN

$$R = \frac{i \times r(a, \hat{a}) \times \sigma_a}{L}$$

Elección de reproductores

CRUZAMIENTO

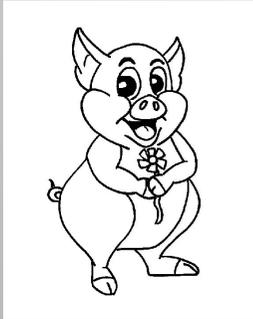
Complementariedad
Heterosis

SELECCIÓN

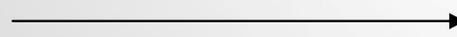
Evaluación de reproductores (a_i)

1. Selección individual
2. Índice de selección
3. Mejor predictor lineal insesgado (BLUP)
4. Selección asistida por genes y marcadores
5. Selección genómica

SELECCIÓN INDIVIDUAL



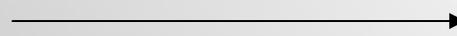
y_1



$$\hat{a}_1 = h^2(y_1 - \mu)$$



y_2



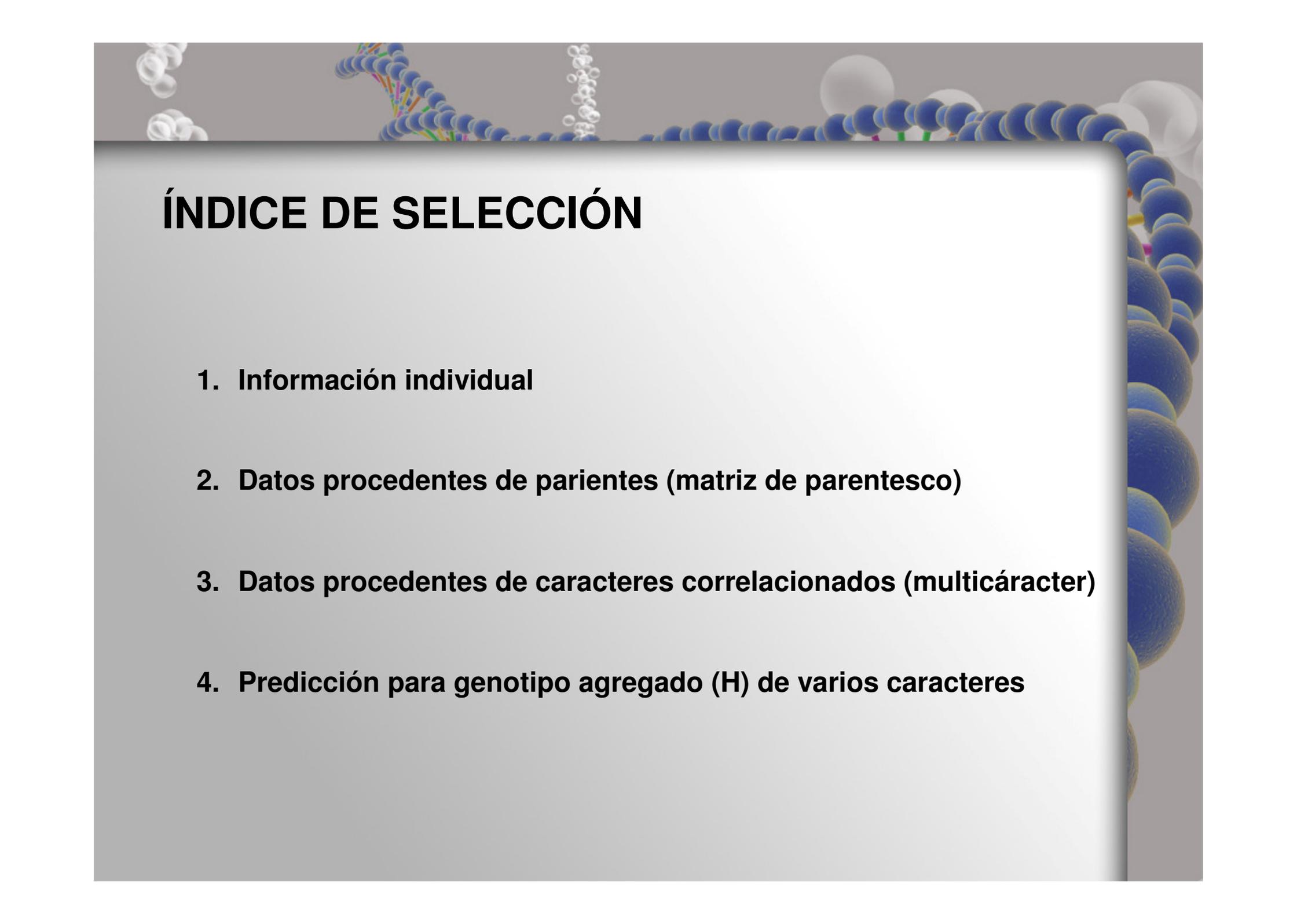
$$\hat{a}_2 = h^2(y_2 - \mu)$$



y_3

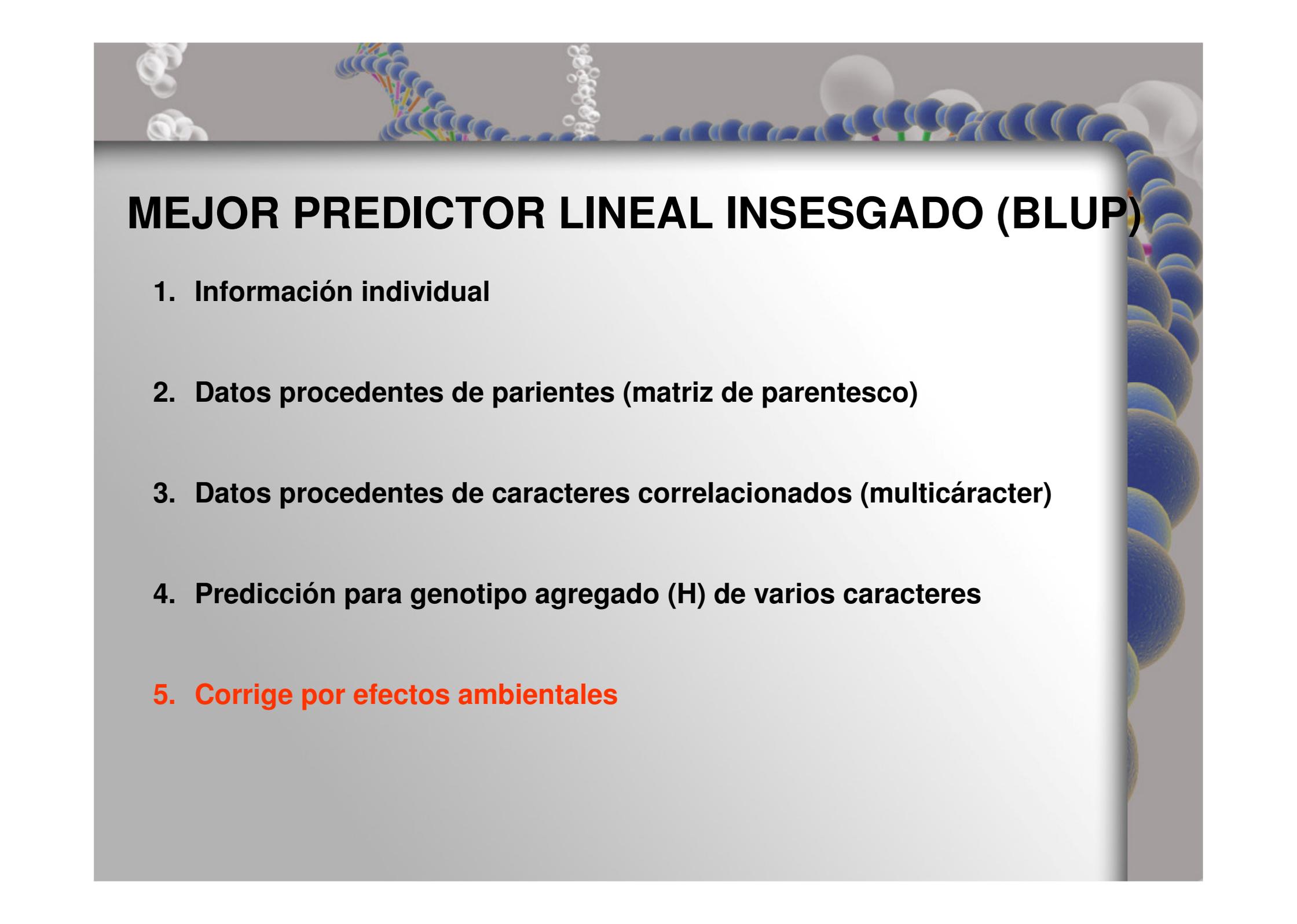


$$\hat{a}_3 = h^2(y_3 - \mu)$$



ÍNDICE DE SELECCIÓN

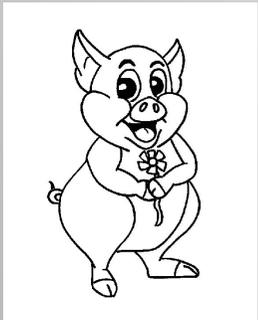
- 1. Información individual**
- 2. Datos procedentes de parientes (matriz de parentesco)**
- 3. Datos procedentes de caracteres correlacionados (multicáriter)**
- 4. Predicción para genotipo agregado (H) de varios caracteres**



MEJOR PREDICTOR LINEAL INSESGADO (BLUP)

1. Información individual
2. Datos procedentes de parientes (matriz de parentesco)
3. Datos procedentes de caracteres correlacionados (multicáriter)
4. Predicción para genotipo agregado (H) de varios caracteres
5. Corrige por efectos ambientales

BLUP UNIVARIANTE



$$y_1 = b + a_1 + e_1$$

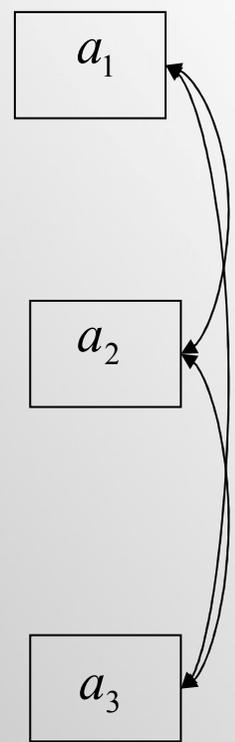


$$y_2 = b + a_2 + e_2$$



$$y_3 = b + a_3 + e_3$$

Parentesco



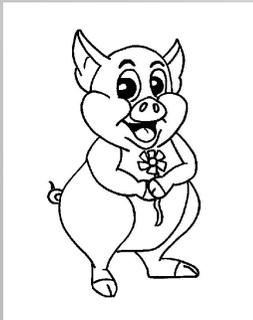
BLUP UNIVARIANTE

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

$$\alpha = \frac{\sigma_e^2}{\sigma_a^2}$$

BLUP MULTICARACTER



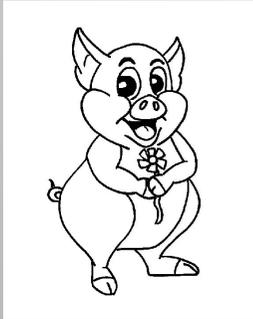
Parentesco

$$\begin{array}{l}
 y_{11} \\
 y_{12}
 \end{array}
 =
 \begin{array}{l}
 b_1 \\
 b_2
 \end{array}
 +
 \begin{array}{l}
 a_{11} \\
 a_{12}
 \end{array}
 r_g
 +
 \begin{array}{l}
 e_{11} \\
 e_{12}
 \end{array}
 r_e$$

$$\begin{array}{l}
 y_{21} \\
 y_{22}
 \end{array}
 =
 \begin{array}{l}
 b_1 \\
 b_2
 \end{array}
 +
 \begin{array}{l}
 a_{21} \\
 a_{22}
 \end{array}
 r_g
 +
 \begin{array}{l}
 e_{21} \\
 e_{22}
 \end{array}
 r_e$$

$$\begin{array}{l}
 y_{31} \\
 y_{32}
 \end{array}
 =
 \begin{array}{l}
 b_1 \\
 b_2
 \end{array}
 +
 \begin{array}{l}
 a_{31} \\
 a_{32}
 \end{array}
 r_g
 +
 \begin{array}{l}
 e_{21} \\
 e_{22}
 \end{array}
 r_e$$

BLUP MULTICARACTER



$$\hat{H}_1 = w_1 \hat{a}_{11} + w_2 \hat{a}_{12}$$



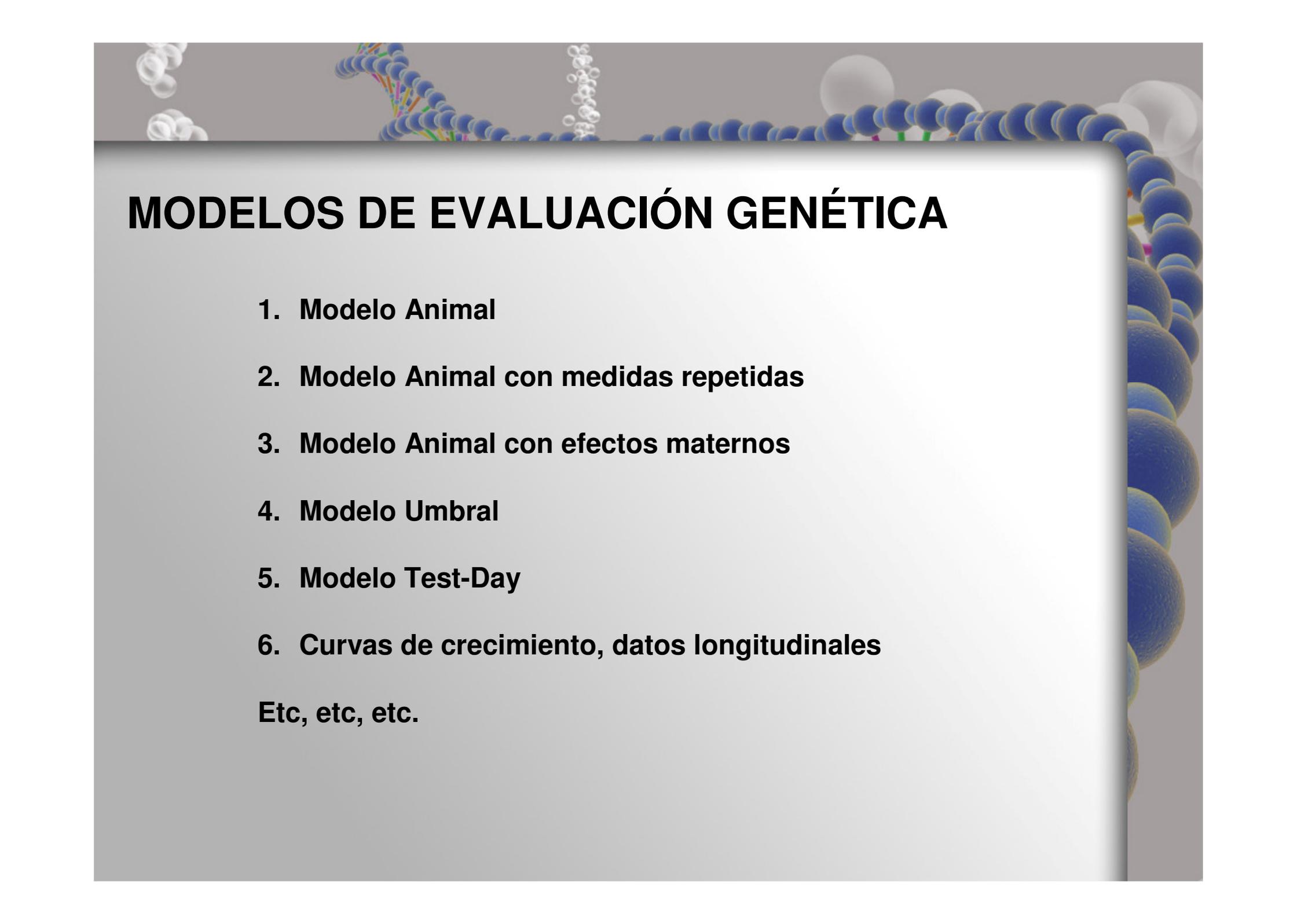
$$\hat{H}_2 = w_1 \hat{a}_{21} + w_2 \hat{a}_{22}$$



$$\hat{H}_3 = w_1 \hat{a}_{31} + w_2 \hat{a}_{32}$$

PONDERACIÓN

ECONÓMICA



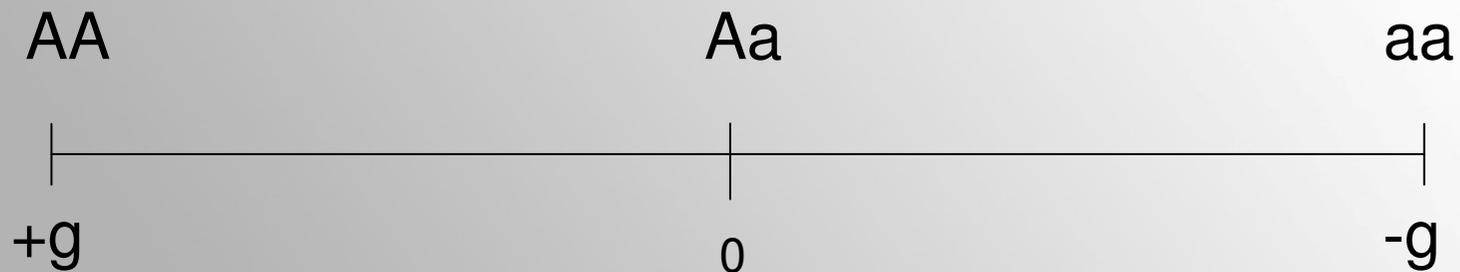
MODELOS DE EVALUACIÓN GENÉTICA

- 1. Modelo Animal**
 - 2. Modelo Animal con medidas repetidas**
 - 3. Modelo Animal con efectos maternos**
 - 4. Modelo Umbral**
 - 5. Modelo Test-Day**
 - 6. Curvas de crecimiento, datos longitudinales**
- Etc, etc, etc.**

POLIMORFISMO: VALORES GENOTÍPICOS

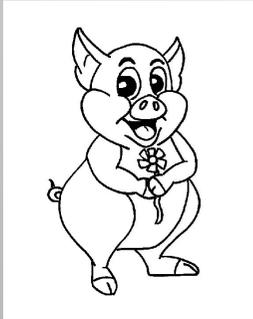
Polimorfismo Causal

Polimorfismo en Desequilibrio de Ligamiento con la mutación causal



SELECCIÓN ASISTIDA POR GENES Y MARCADORES

AA



$$y_1 = b + a_1 + g + e_1$$

Aa



$$y_2 = b + a_2 + 0 + e_2$$

aa



$$y_3 = b + a_3 - g + e_3$$

SELECCIÓN ASISTIDA POR GENES Y MARCADORES

AA



bb

$$y_1 = b + a_1 + g_a - g_b + e_1$$

Aa



BB

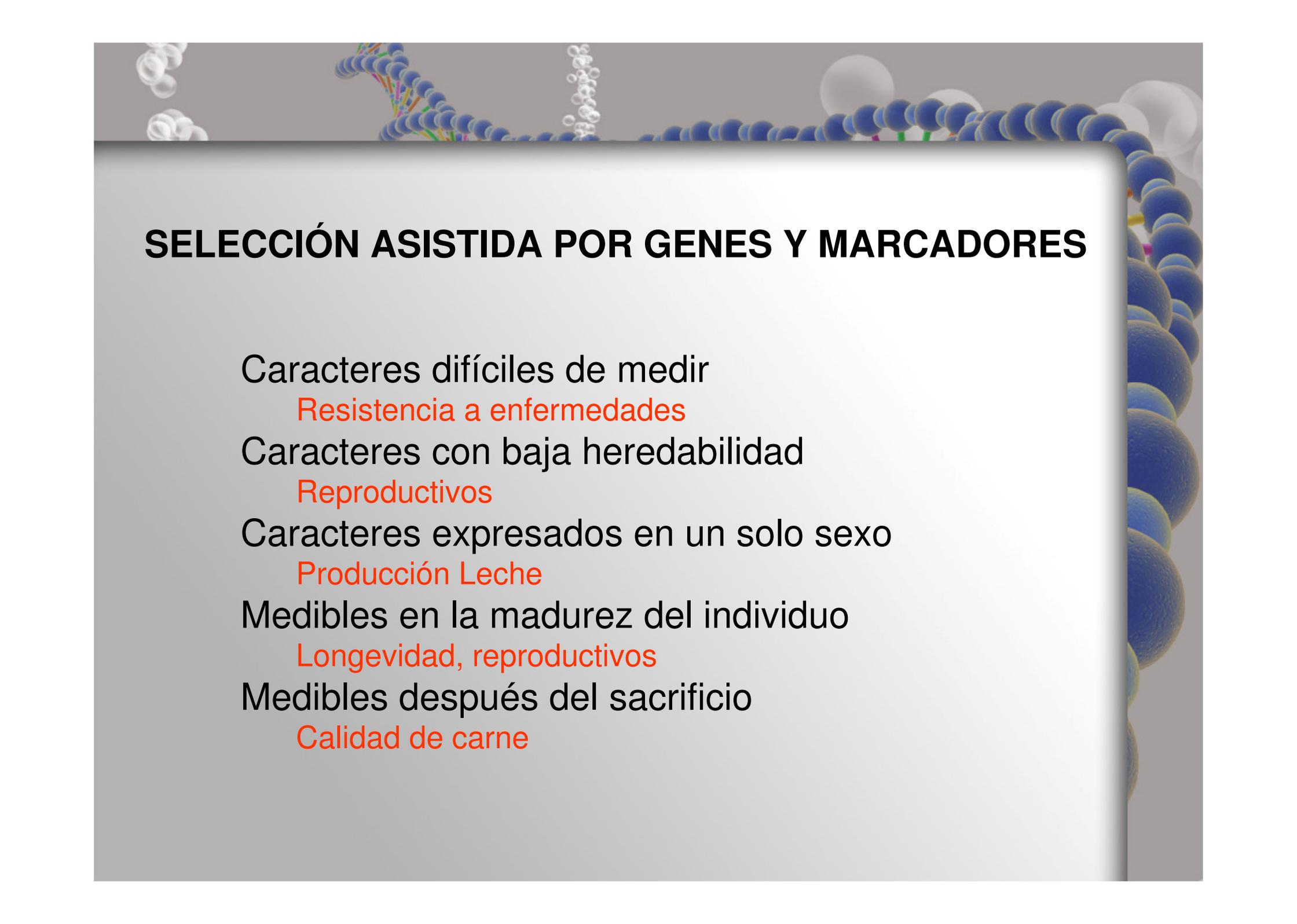
$$y_2 = b + a_2 + 0 + g_b + e_2$$

aa



Bb

$$y_3 = b + a_3 - g_a + 0 + e_3$$



SELECCIÓN ASISTIDA POR GENES Y MARCADORES

Caracteres difíciles de medir

Resistencia a enfermedades

Caracteres con baja heredabilidad

Reproductivos

Caracteres expresados en un solo sexo

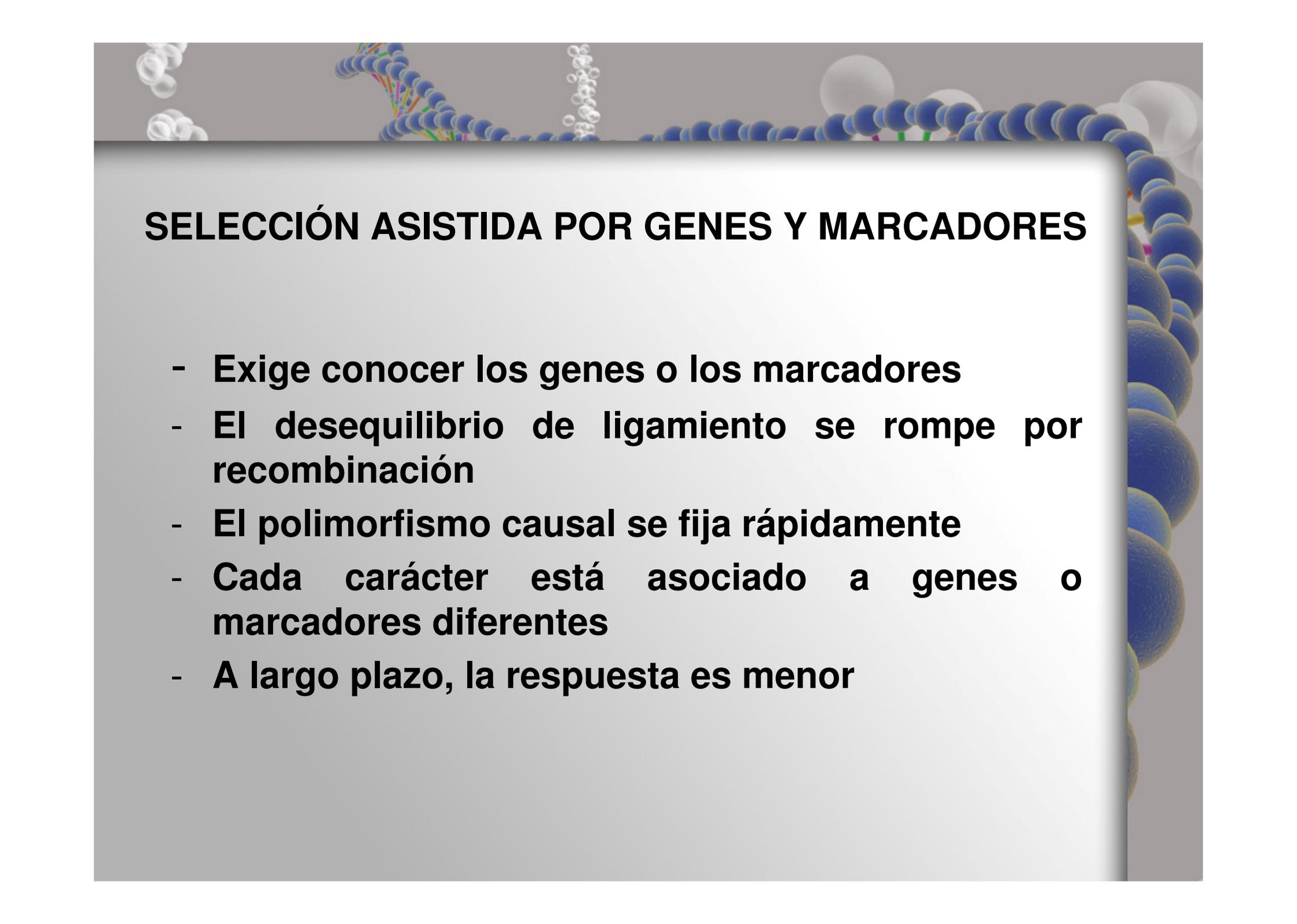
Producción Leche

Medibles en la madurez del individuo

Longevidad, reproductivos

Medibles después del sacrificio

Calidad de carne



SELECCIÓN ASISTIDA POR GENES Y MARCADORES

- **Exige conocer los genes o los marcadores**
- **El desequilibrio de ligamiento se rompe por recombinación**
- **El polimorfismo causal se fija rápidamente**
- **Cada carácter está asociado a genes o marcadores diferentes**
- **A largo plazo, la respuesta es menor**

SNP (Single Nucleotide Polymorphisms)

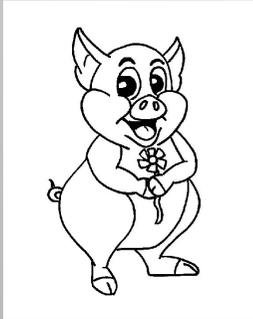
BovineSNP50 BeadChip. 50K SNP.

OvineSNP50 BeadChip. 50K SNP.

PorcineSNP60 BeadChip. 60K SNP.



SELECCIÓN GENÓMICA



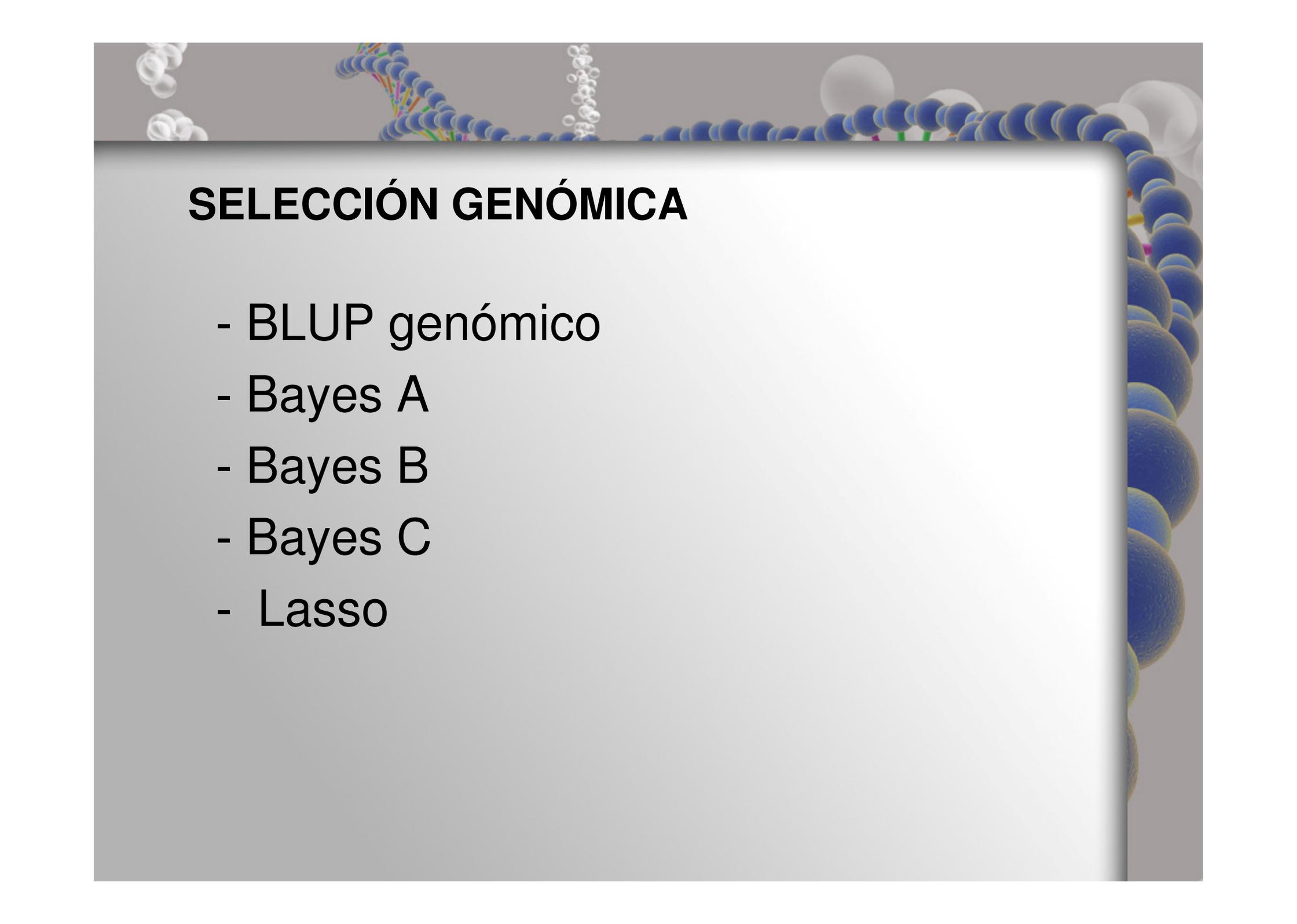
$$y_1 = b + \sum_i^{Ng} (G_{1i} = A_i A_i) g_i - (G_{1i} = a_i a_i) g_i + e_1$$



$$y_2 = b + \sum_i^{Ng} (G_{2i} = A_i A_i) g_i - (G_{2i} = a_i a_i) g_i + e_2$$



$$y_3 = b + \sum_i^{Ng} (G_{3i} = A_i A_i) g_i - (G_{3i} = a_i a_i) g_i + e_3$$



SELECCIÓN GENÓMICA

- BLUP genómico
- Bayes A
- Bayes B
- Bayes C
- Lasso

BLUP GENÓMICO

1. Todos los marcadores influyen en carácter
2. Todos los marcadores se asume que proceden de una distribución normal única.

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \mathbf{e}$$

$$\mathbf{g} \sim N(0, \sigma_g^2)$$

BAYES A

1. Todos los marcadores influyen en carácter
2. Cada marcador se asume que procede de una distribución normal particular

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \mathbf{e}$$

$$g_i \sim N(0, \sigma_{g_i}^2)$$

BAYES B

1. Un porcentaje prefijado de marcadores incluye en el carácter.
2. Cada marcador se asume que procede de una distribución normal particular

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \mathbf{e}$$

$$g_i \sim I(k = 1) \times N(0, \sigma_{g_i}^2) + I(k = 0) \times 0$$

BAYES C

1. Un porcentaje no prefijado de marcadores incluye en el carácter.
2. Este porcentaje es estimado.
3. Cada marcador se asume que procede de una distribución normal particular

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \mathbf{e}$$

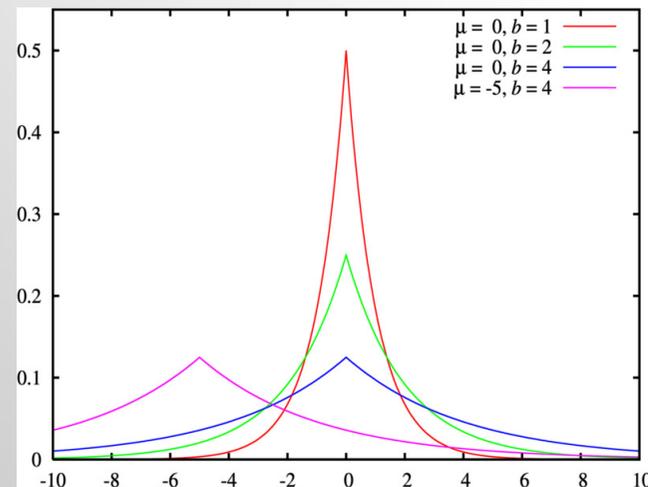
$$g_i \sim I(k = 1) \times N(0, \sigma_{g_i}^2) + I(k = 0) \times 0$$

BAYESIAN LASSO

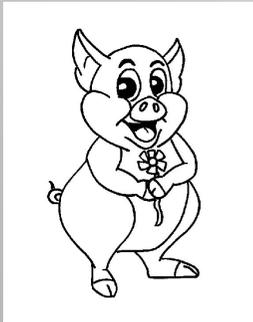
1. Todos o parte de los marcadores influyen en carácter
2. Los marcadores se asume que proceden de una distribución exponencial doble (Laplace)

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \mathbf{e}$$

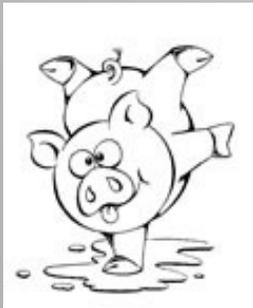
$$f(g_i) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right)$$



SELECCIÓN GENÓMICA: MÉTODOS NO PARAMÉTRICOS



$$y_1 = b + f(\mathbf{G}_1) + e_1$$



$$y_2 = b + f(\mathbf{G}_2) + e_2$$



$$y_3 = b + f(\mathbf{G}_3) + e_3$$

POBLACIÓN

GENOTIPADOS
FENOTIPADOS

GENOTIPADOS
NO FENOTIPADOS

NO GENOTIPADOS
FENOTIPADOS

NO GENOTIPADOS
NO FENOTIPADOS



MODELO LINEAL MIXTO: BLUP UNIVARIANTE

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

MODELO LINEAL MIXTO: BLUP UNIVARIANTE GENÓMICO

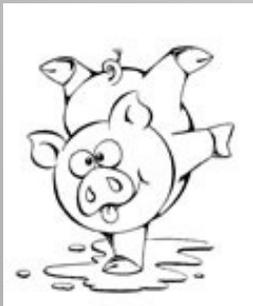
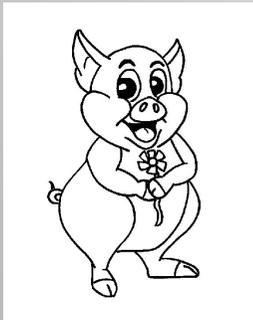
$$y = \mathbf{Xb} + \mathbf{Za} + e \quad \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{G}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

$$\mathbf{a} = \mathbf{Wg}$$

$$\text{Var}(a) = \mathbf{W}'\mathbf{W}\sigma_g^2$$

$$\text{Var}(a) = \mathbf{W}'\mathbf{W}\sigma_g^2 / k = \mathbf{G}\sigma_a^2$$

BLUP UNICARACTER



Parentesco genómico

$$y_1 = b + a_1 + e_1$$
$$y_2 = b + a_2 + e_2$$
$$y_3 = b + a_3 + e_3$$

The diagram illustrates the relationship between the additive genetic effects a_1 , a_2 , and a_3 . Arrows indicate that a_1 is the parent of a_2 and a_3 , and a_2 is the parent of a_3 . The label "Parentesco genómico" (Genomic relationship) points to the space between the equations, indicating the genetic relationship between the individuals.

BLUP UNIVARIANTE GENÓMICO

SUBPOB 1
NO GENOTIPADOS

SUBPOB 2
GENOTIPADOS

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{H}^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{G}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{G} \\ \mathbf{G}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{G} \end{bmatrix}$$

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$



BLUP MULTIVARIANTE GENÓMICO

- Puede considerar relaciones entre caracteres
- BLUP Multivariante
- Incluye parentesco genómico y genealógico
- Utiliza toda la información disponible
- Aprovecha la estructura de selección disponible
- **CLAVE: Elección de individuos a genotipar**

RESPUESTA A LA SELECCIÓN

$$R = \frac{i \times r(a, \hat{a}) \times \sigma_a}{L}$$

- Intensidad:** Permite seleccionar animales no valorados
- Precisión:** Incremente la precisión de los predictores
- Intervalo Generacional:** Permite seleccionar animales más jóvenes

RESPUESTA A LA SELECCIÓN

$$R = \frac{i \times r(a, \hat{a}) \times \sigma_a}{L}$$

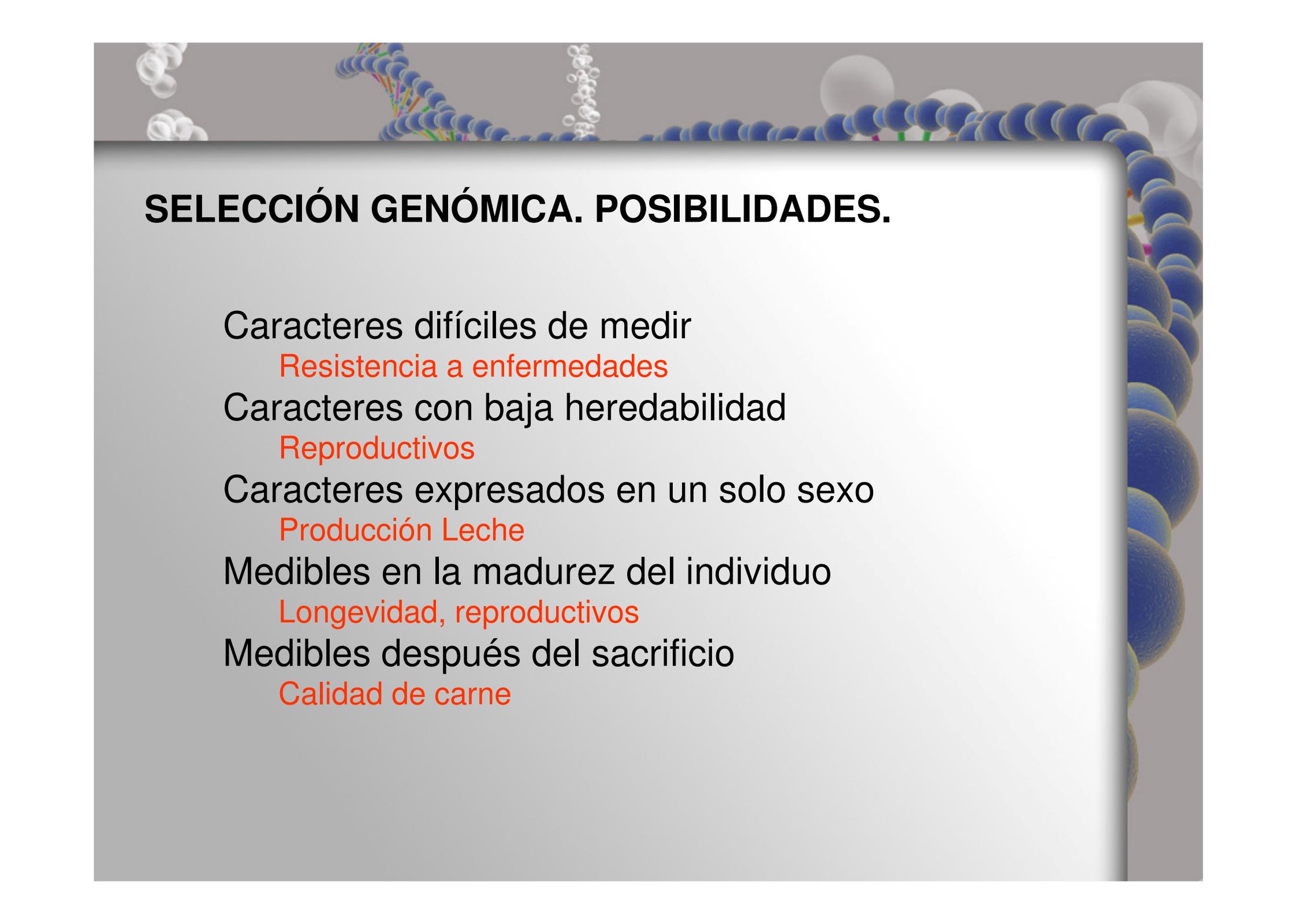
- Intensidad:** Permite seleccionar animales no valorados
- Precisión:** Incremente la precisión de los predictores
- Intervalo Generacional:** Permite seleccionar animales más jóvenes

RESULTADOS PRECISIÓN ($h^2=0.20$)

			SEL. GEN.	
	IND	BLUP	SNP	QTL
IND	1	28%	68%	78%
BLUP		1	31%	39%
SNP			1	9%
QTL				1

RESULTADOS PRECISIÓN ($h^2=0.40$)

			SEL. GEN.	
	IND	BLUP	SNP	QTL
IND	1	14%	37%	50%
BLUP		1	20%	31%
SNP			1	6%
QTL				1



SELECCIÓN GENÓMICA. POSIBILIDADES.

Caracteres difíciles de medir

Resistencia a enfermedades

Caracteres con baja heredabilidad

Reproductivos

Caracteres expresados en un solo sexo

Producción Leche

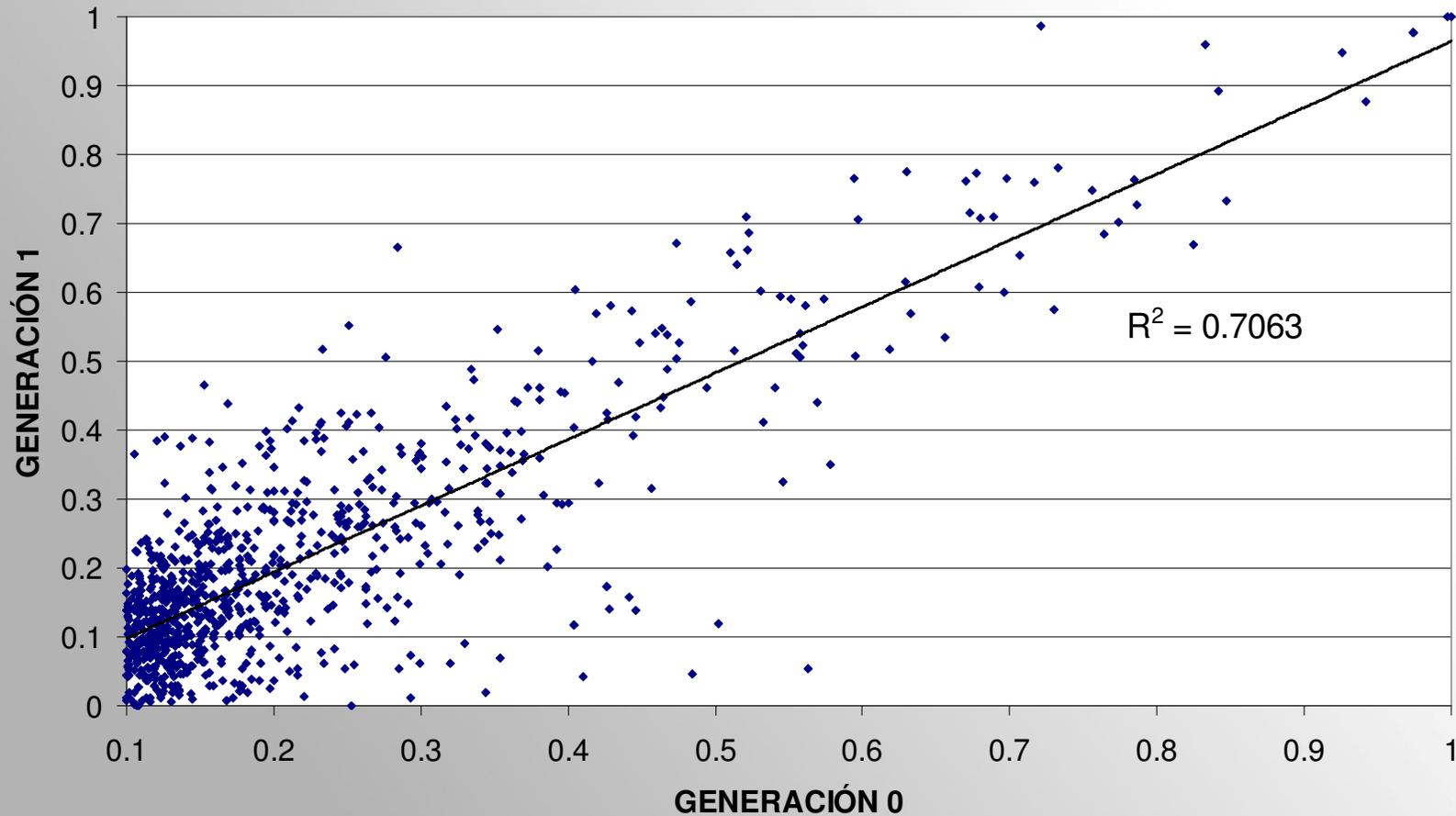
Medibles en la madurez del individuo

Longevidad, reproductivos

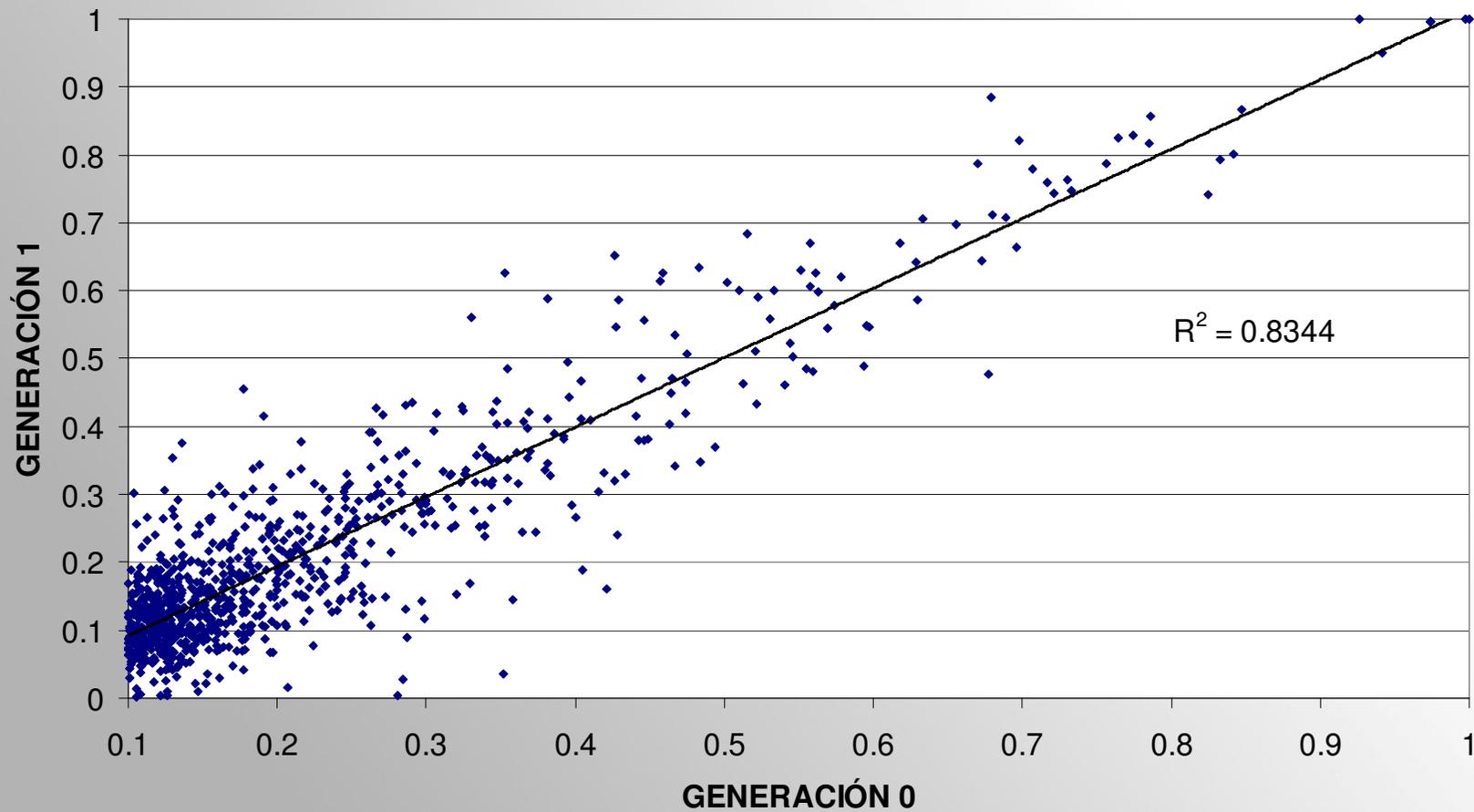
Medibles después del sacrificio

Calidad de carne

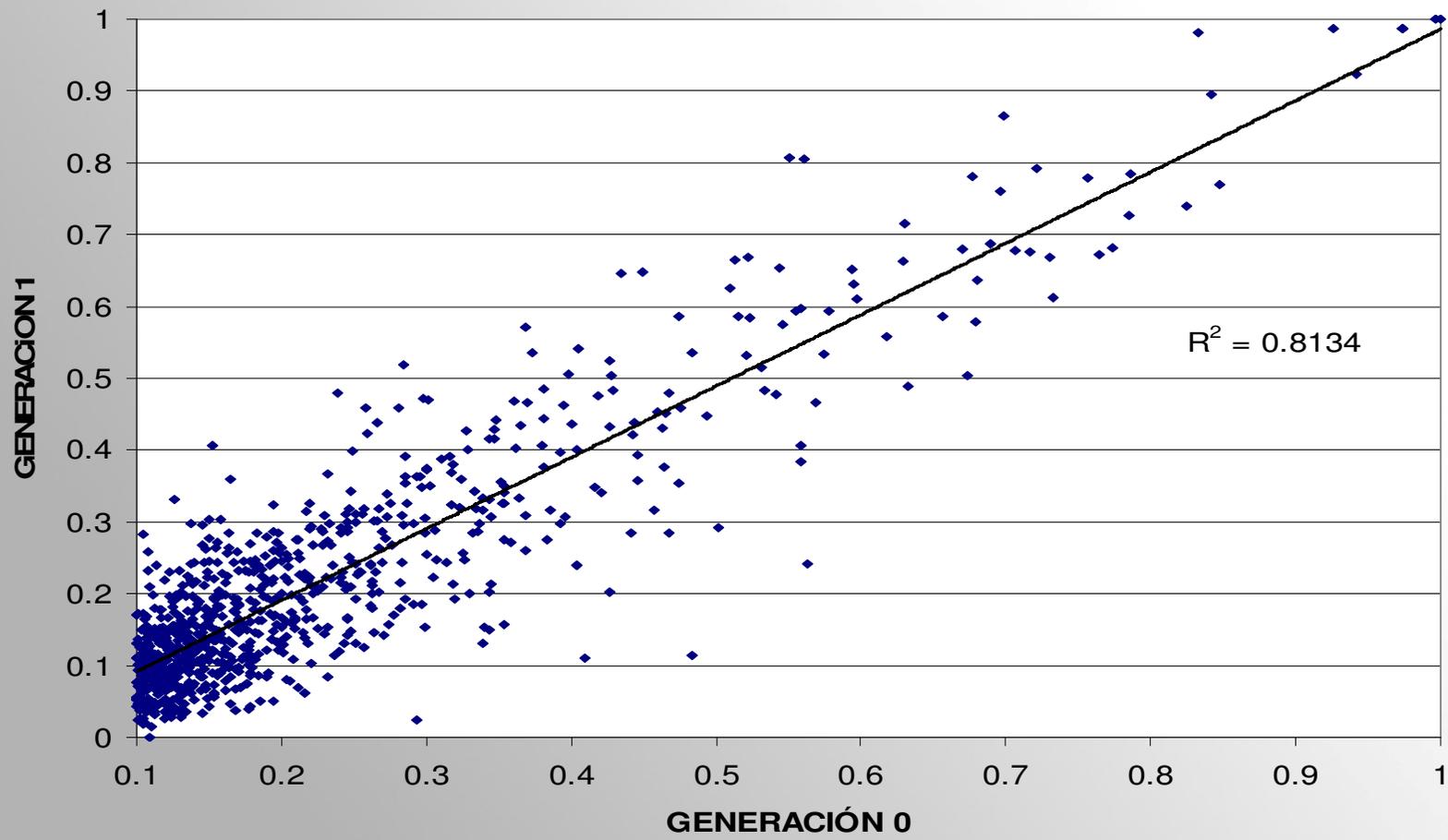
Desequilibrio de Ligamiento (r^2) Selección Genómica



Desequilibrio de Ligamiento (r^2) Sin selección

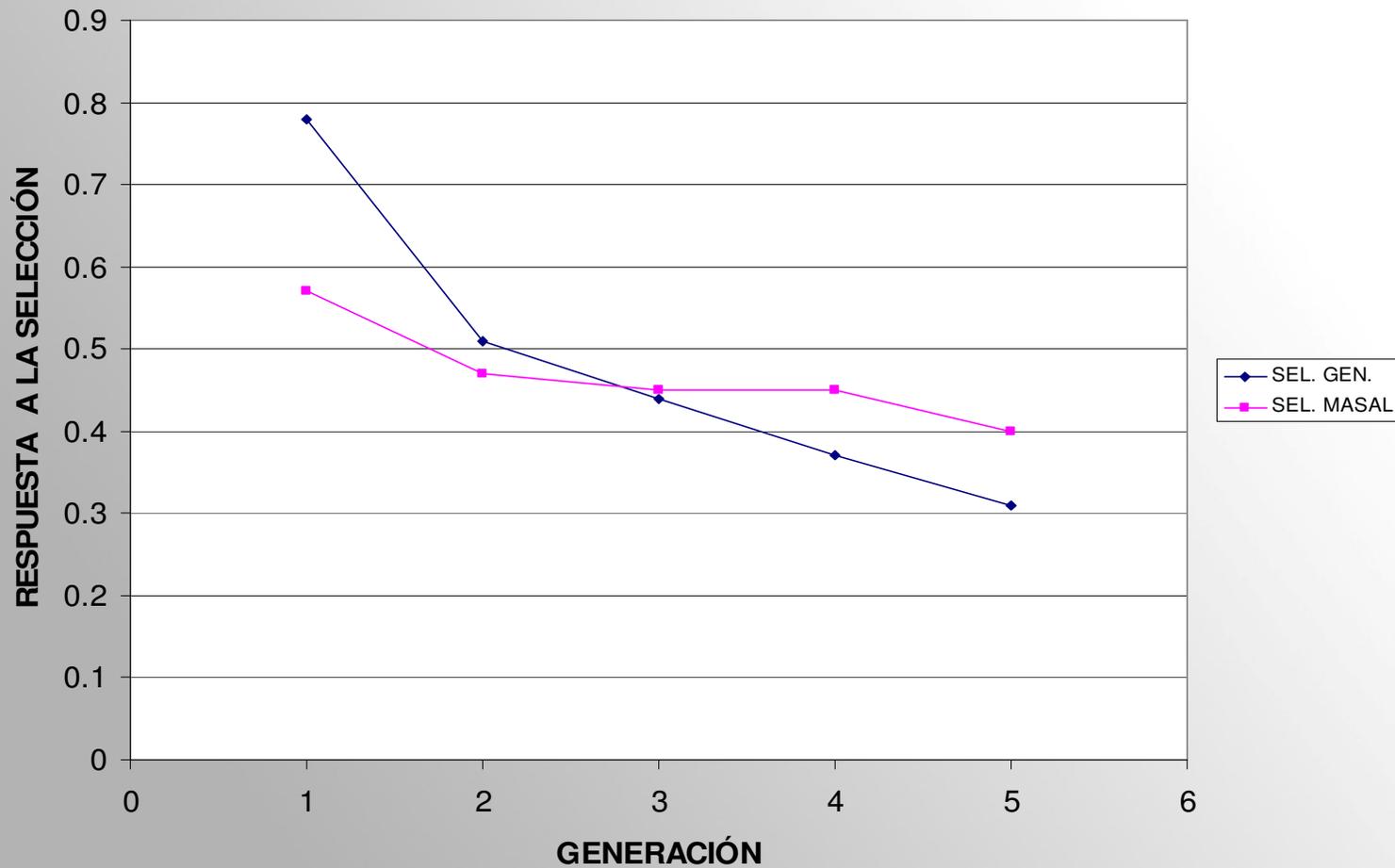


Desequilibrio de Ligamiento (r^2) Selección Individual

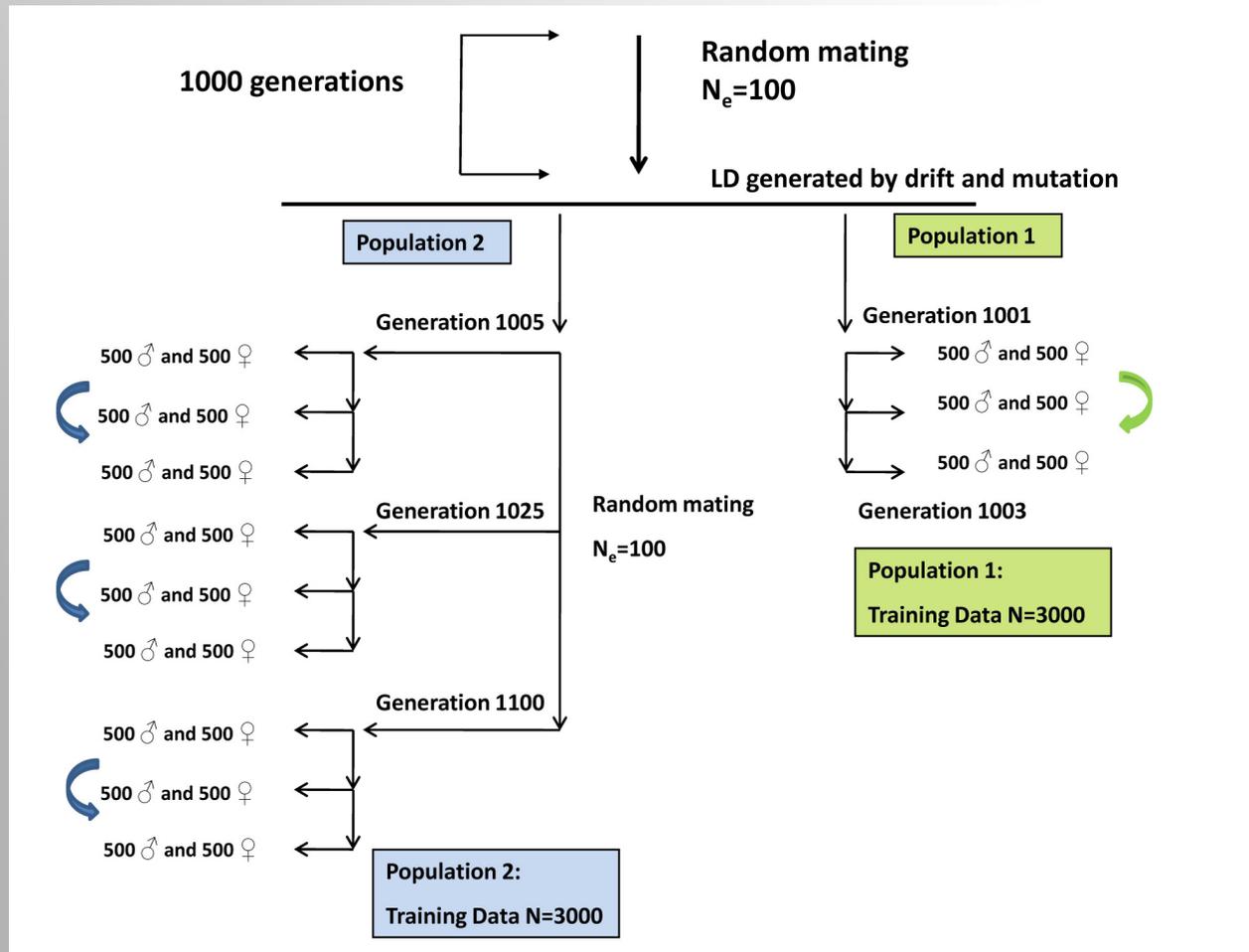


RESPUESTA VARIAS GENERACIONES

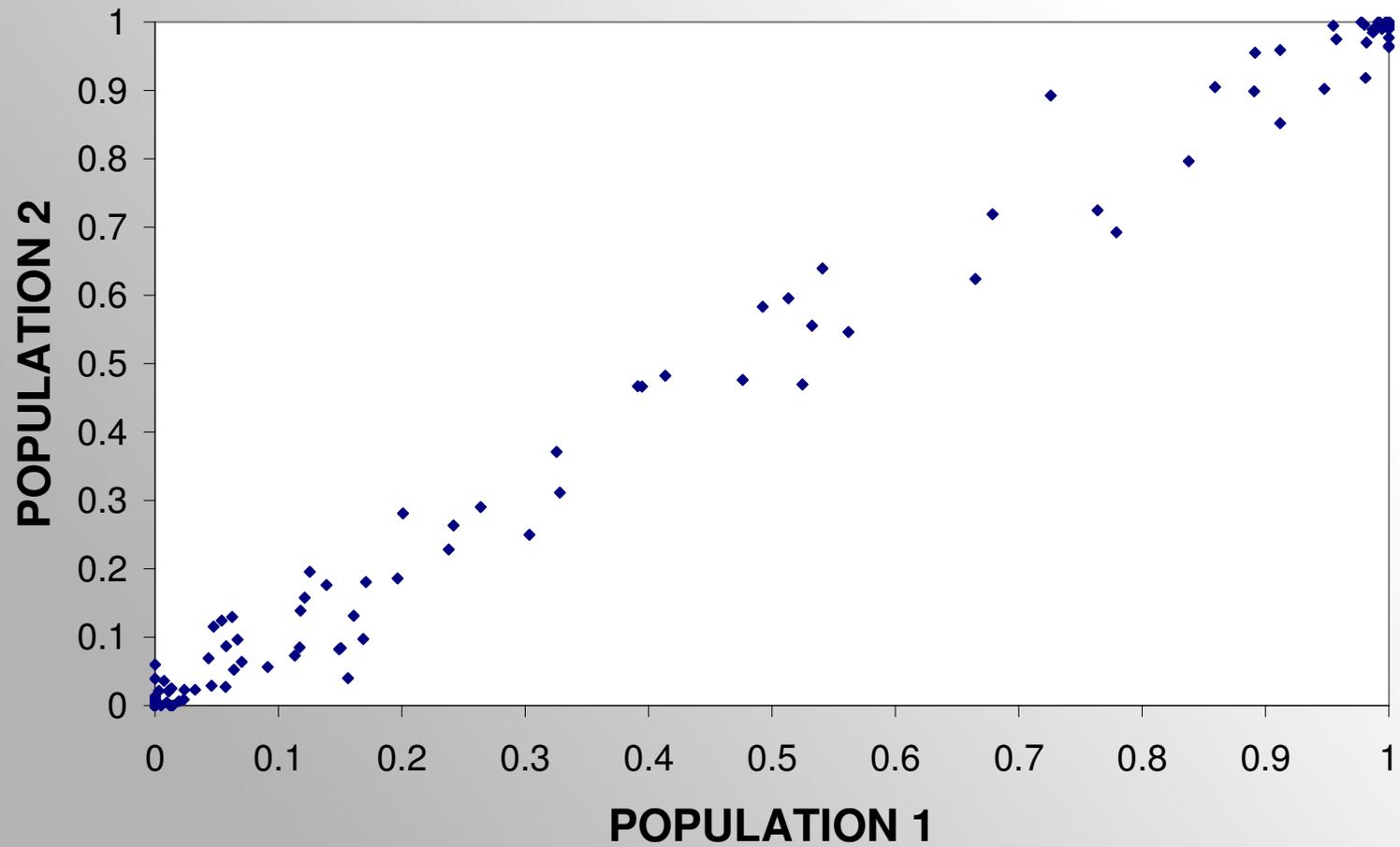
$h^2=0.40$



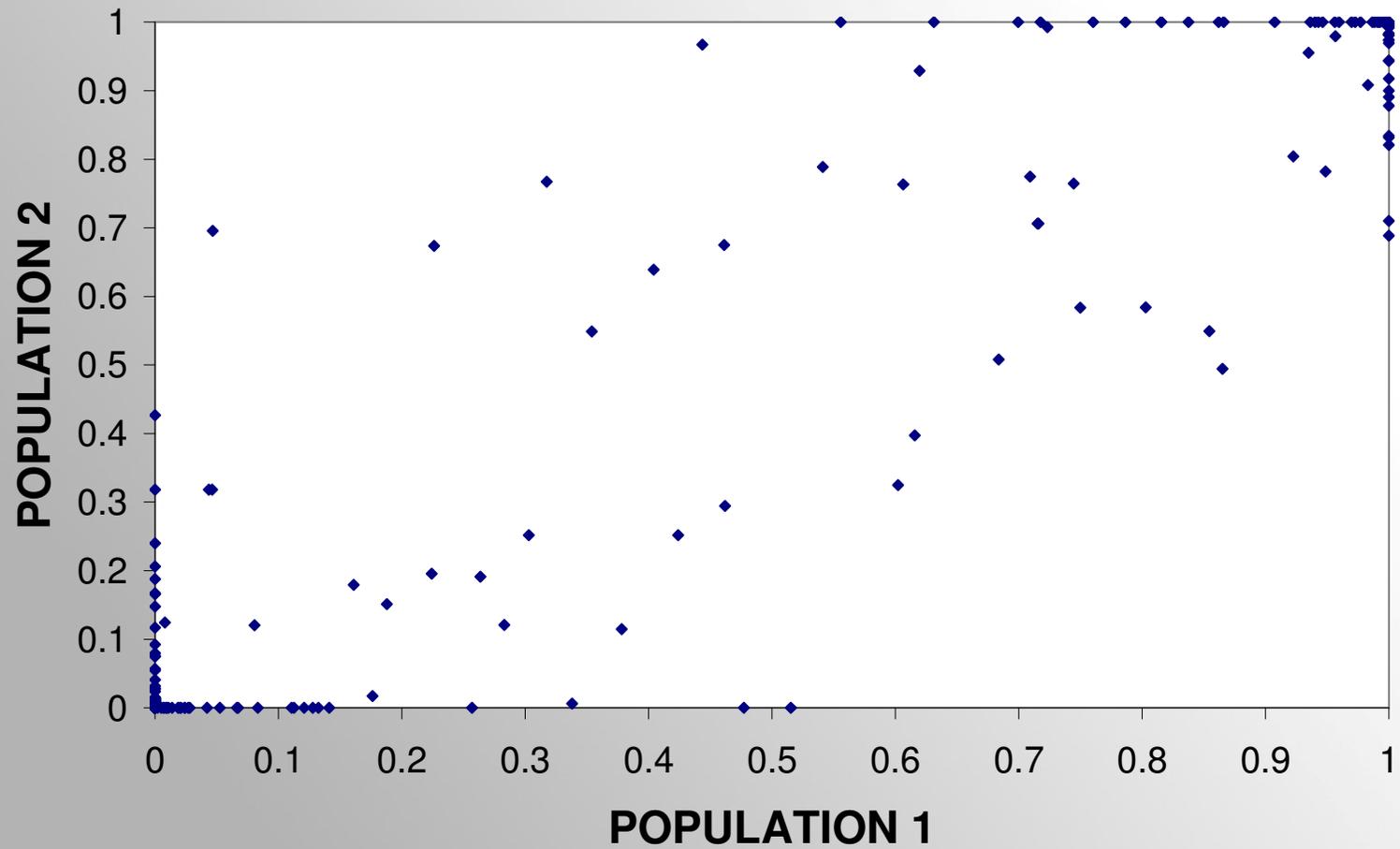
POBLACIONES MULTIPLES



FRECUENCIAS GÉNICAS QTL (5 GEN.)



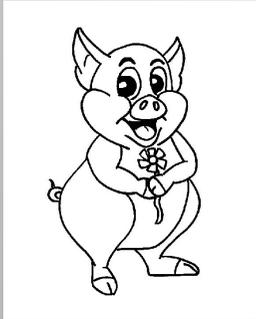
FRECUENCIAS GÉNICAS QTL (100 GEN.)



POBLACIONES MÚLTIPLES (PRECISIÓN)

h^2	G^*	Pob. Indiv.	Pob. Conjunta	Pob. Ponderada	Correlacion SNP
0.15	5	0.685 (0.015)	0.719 (0.014)	0.727 (0.013)	0.777 (0.028)
0.15	25	0.682 (0.013)	0.693 (0.013)	0.703 (0.013)	0.609 (0.049)
0.15	100	0.686 (0.013)	0.653 (0.014)	0.690 (0.013)	0.256 (0.044)
0.35	5	0.773 (0.011)	0.805 (0.010)	0.809 (0.010)	0.851 (0.033)
0.35	25	0.771 (0.011)	0.777 (0.011)	0.792 (0.010)	0.743 (0.038)
0.35	100	0.782 (0.008)	0.756 (0.015)	0.787 (0.008)	0.232 (0.040)

MODELO CON DOMINANCIA



$$y_1 = b + \sum_i^{Nm} I(G_{1i} = A_i A_i) g_i - I(G_{1i} = a_i a_i) g_i + \sum_i^{Nm} I(G_{1i} = A_i a_i) d_i + e_1$$



$$y_2 = b + \sum_i^{Nm} I(G_{2i} = A_i A_i) g_i - I(G_{2i} = a_i a_i) g_i + \sum_i^{Nm} I(G_{2i} = A_i a_i) d_i + e_2$$



$$y_3 = b + \sum_i^{Nm} I(G_{3i} = A_i A_i) g_i - I(G_{3i} = a_i a_i) g_i + \sum_i^{Nm} I(G_{3i} = A_i a_i) d_i + e_3$$

DISEÑO DE APAREAMIENTOS

	M1	M2	M3	M4	M5
H1					
H2					
H3					
H4					
H5					
H6					
H7					
H8					
H9					
H10					

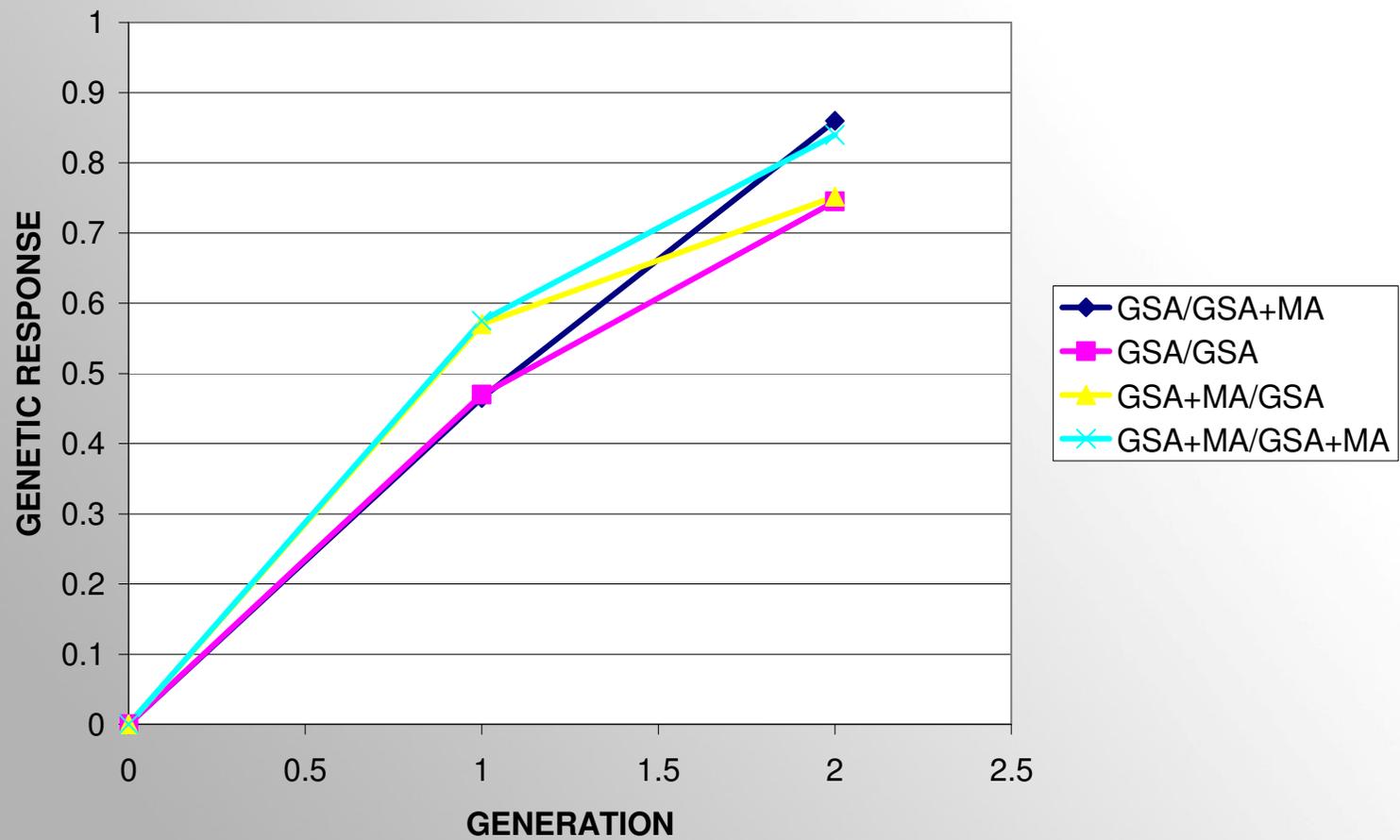
DISEÑO DE APAREAMIENTOS

	M1			M4	
H2					
H3					
H4					
H10					

DISEÑO DE APAREAMIENTOS

	M1			M4	
H2	x				
H3				x	
H4	x				
H10				x	

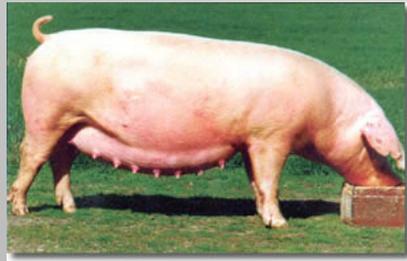
DISEÑO DE APAREAMIENTOS



COMPLEMENTARIDAD Y HETEROSIS

LW

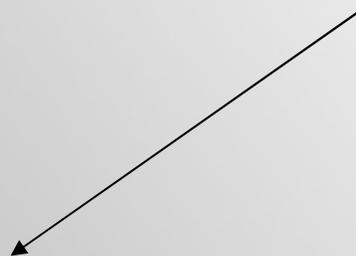
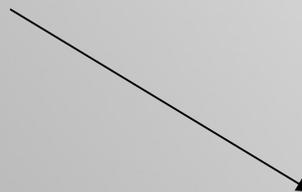
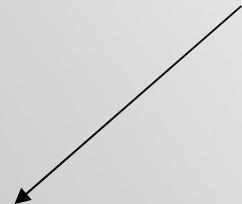
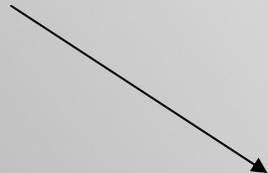
LD

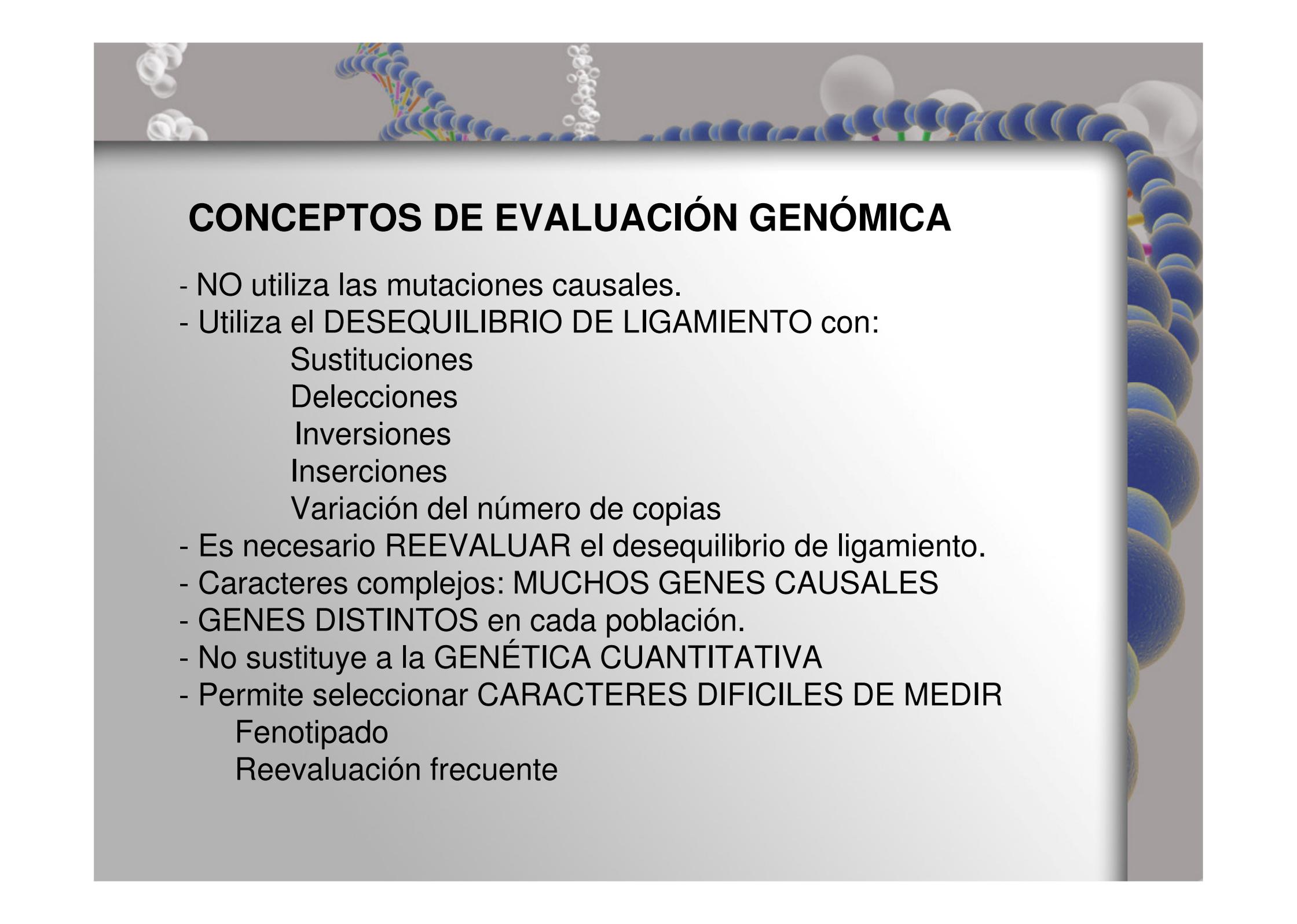


LD x LW

DUROC

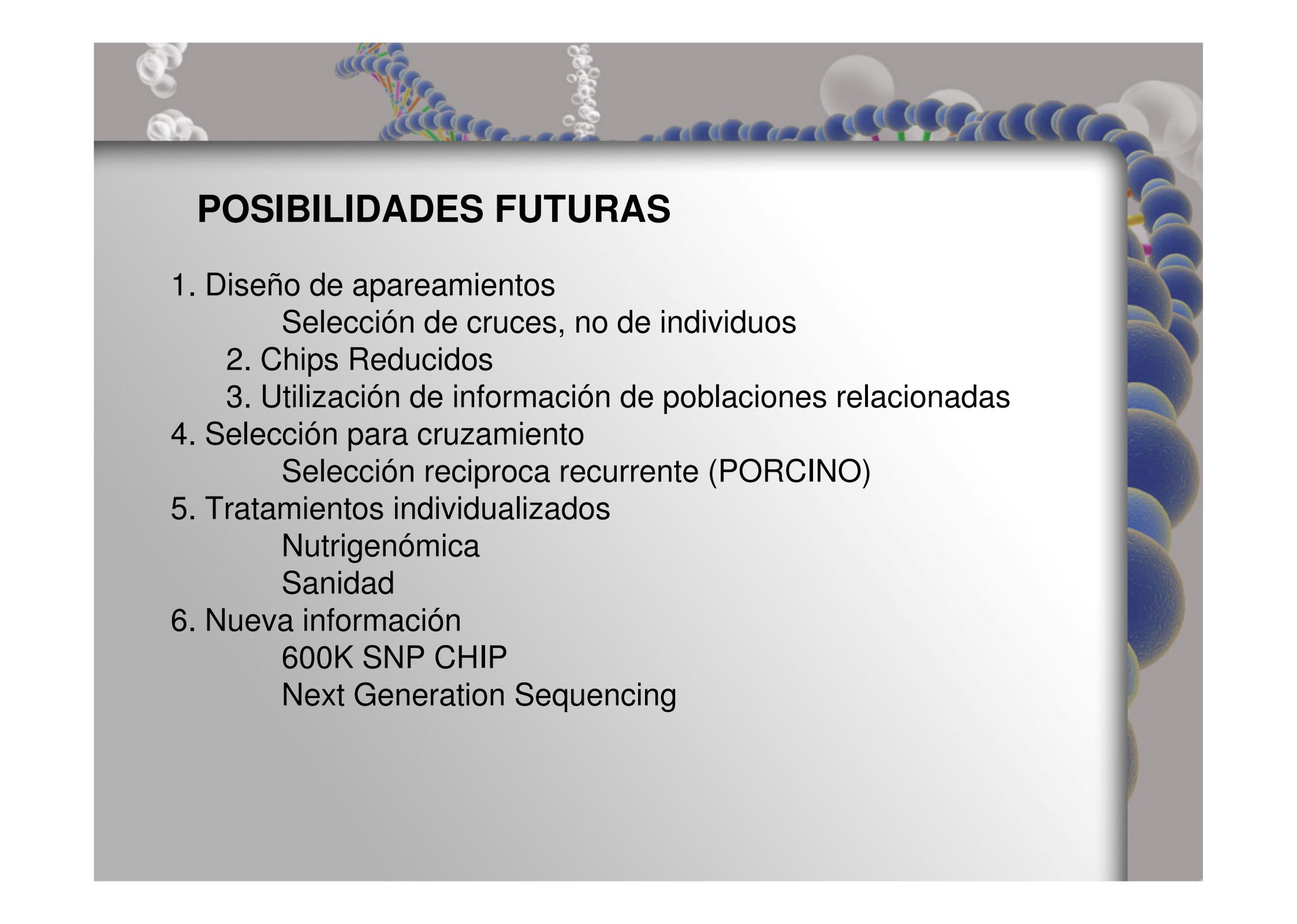
PRODUCTO FINAL





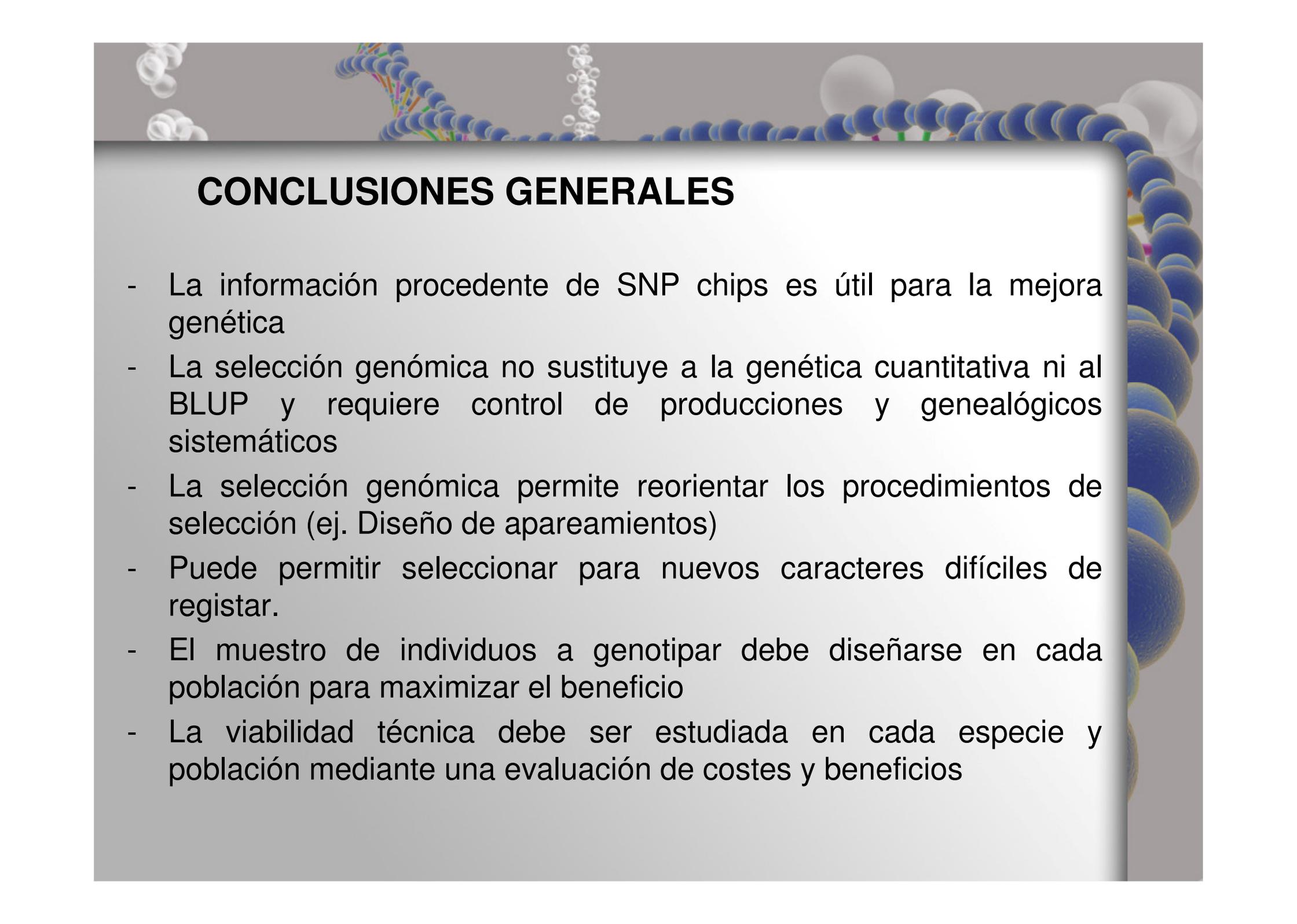
CONCEPTOS DE EVALUACIÓN GENÓMICA

- NO utiliza las mutaciones causales.
- Utiliza el DESEQUILIBRIO DE LIGAMIENTO con:
 - Sustituciones
 - Delecciones
 - Inversiones
 - Inserciones
 - Variación del número de copias
- Es necesario REEVALUAR el desequilibrio de ligamiento.
- Caracteres complejos: MUCHOS GENES CAUSALES
- GENES DISTINTOS en cada población.
- No sustituye a la GENÉTICA CUANTITATIVA
- Permite seleccionar CARACTERES DIFÍCILES DE MEDIR
 - Fenotipado
 - Reevaluación frecuente



POSIBILIDADES FUTURAS

1. Diseño de apareamientos
Selección de cruces, no de individuos
2. Chips Reducidos
3. Utilización de información de poblaciones relacionadas
4. Selección para cruzamiento
Selección recíproca recurrente (PORCINO)
5. Tratamientos individualizados
Nutrigenómica
Sanidad
6. Nueva información
600K SNP CHIP
Next Generation Sequencing



CONCLUSIONES GENERALES

- La información procedente de SNP chips es útil para la mejora genética
- La selección genómica no sustituye a la genética cuantitativa ni al BLUP y requiere control de producciones y genealógicos sistemáticos
- La selección genómica permite reorientar los procedimientos de selección (ej. Diseño de apareamientos)
- Puede permitir seleccionar para nuevos caracteres difíciles de registrar.
- El muestro de individuos a genotipar debe diseñarse en cada población para maximizar el beneficio
- La viabilidad técnica debe ser estudiada en cada especie y población mediante una evaluación de costes y beneficios