

Concepts de base en génétique quantitative

Zulma G. Vitezica

zulma.vitezica@ensat.fr

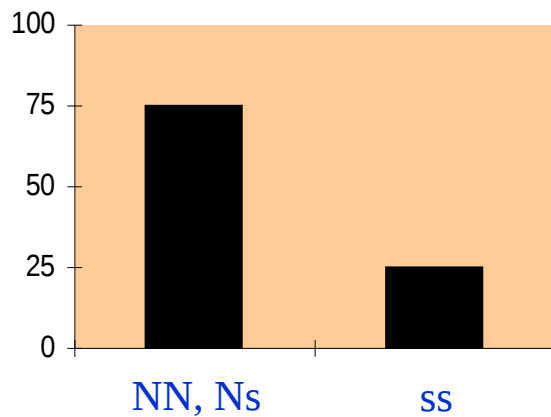
Département de Sciences Animales

Plan du cours ...

1. Qu'est-ce qu'un caractère quantitatif ?
2. Déterminisme génétique des caractères quantitatifs
3. Modèle « polygénique infinitésimal »
4. Décomposition $P = G + E$
5. Notion de héritabilité d'un caractère
6. Notion de parenté et de consanguinité

Caractères QUALITATIFS

Présence / absence de l'allèle sensible
Maladie (malade / sain)

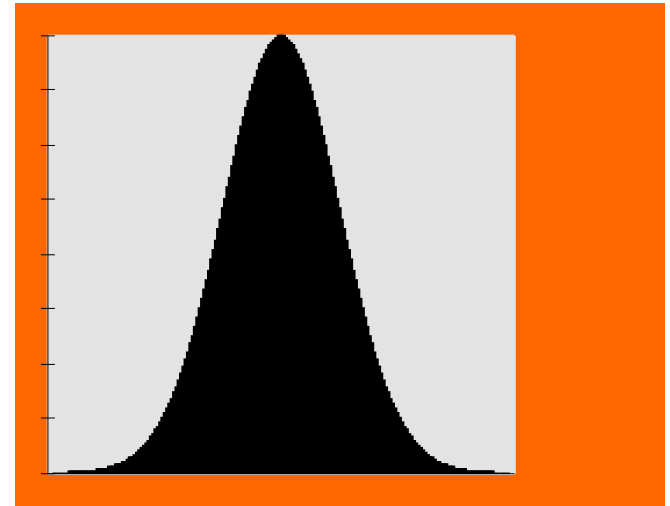


Exemple : Sensibilité à l'halothane chez les porcins. Locus autosomal, 2 allèles : **N** (normal), **s** (sensible). Individus non sensibles: NN, Ns.

Variation discontinue/binaire

Caractères QUANTITATIFS

Poids de grain par épi (g)
Teneur en protéines du lait (g/Kg)
Diamètre de fibre (micron)

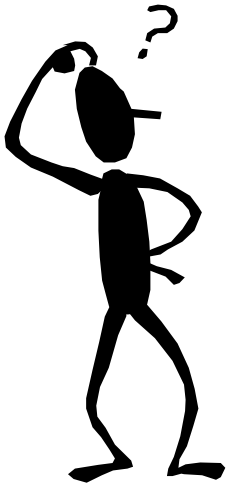


Variation continue

Qu'est-ce qu'un caractère quantitatif ?



Caractère qui se mesure de façon objective



Valeur observée, ou performance,
ou phénotype



Appréciation d'une aptitude

Exemples d'aptitudes pouvant, dans certaines populations, faire l'objet de programmes d'amélioration génétique

MORPHOLOGIE

Taille, état corporel

REPRODUCTION

Proliféricité, Fertilité, précocité sexuelle

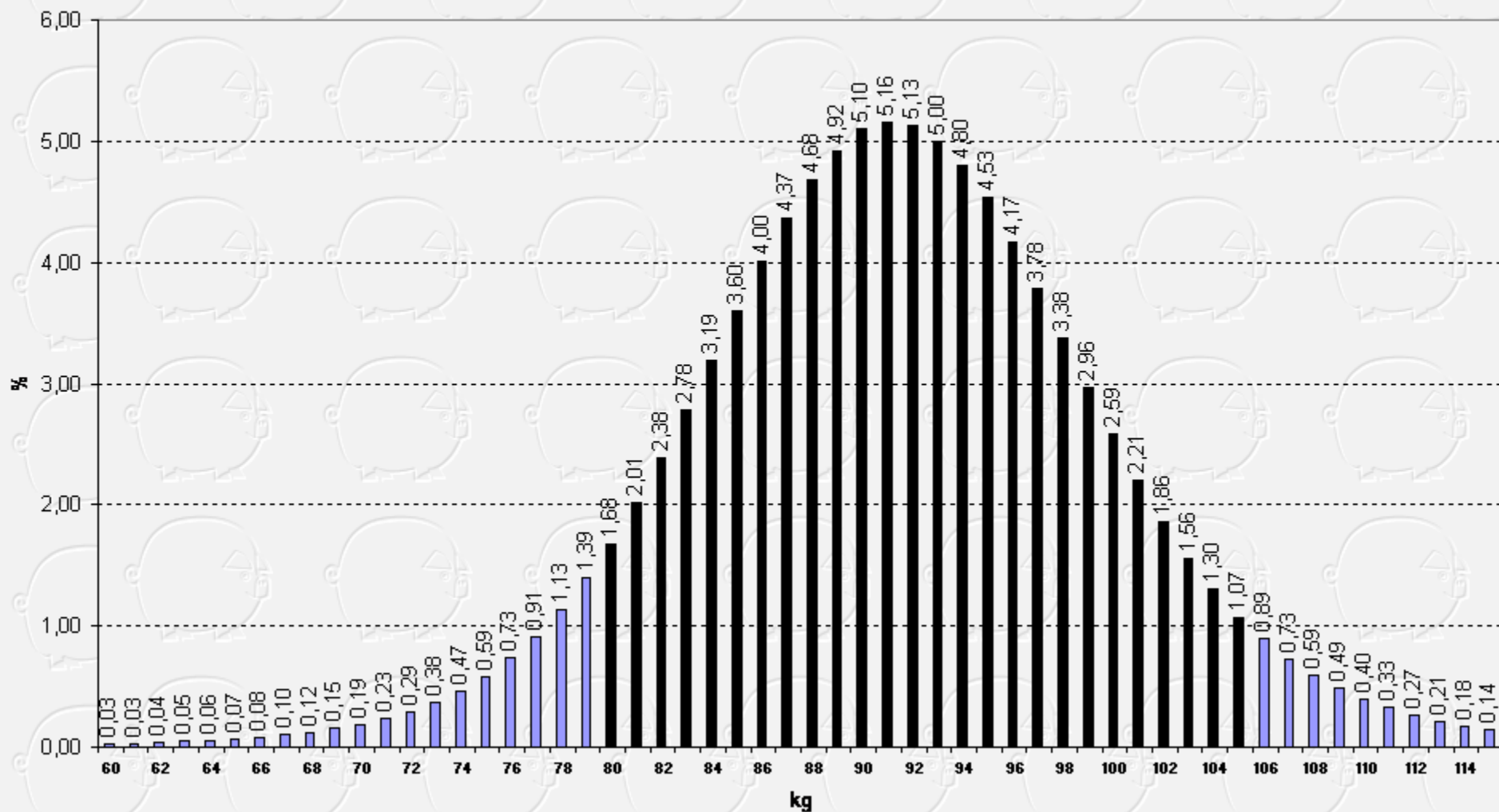
ADAPTATION AU CONDITIONS DU MILIEU

Élevage, Stress

PREDISPOSITION / RESISTANCE MALADIE

Taux d'anticorps, note d'atteinte

Répartition des porcs par poids en 2006



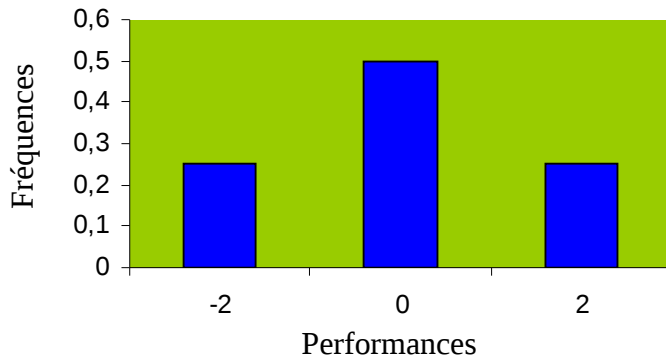
Source: Uniporc Ouest

1 gène, 1 couple d'allèles : **A et a**

$$\begin{aligned} f(A) &= 0,5 & \text{effet}(A) &= +1 \\ f(a) &= 0,5 & \text{effet}(a) &= -1 \end{aligned}$$

AA	+2	25%
Aa	0	50%
aa	-2	25%

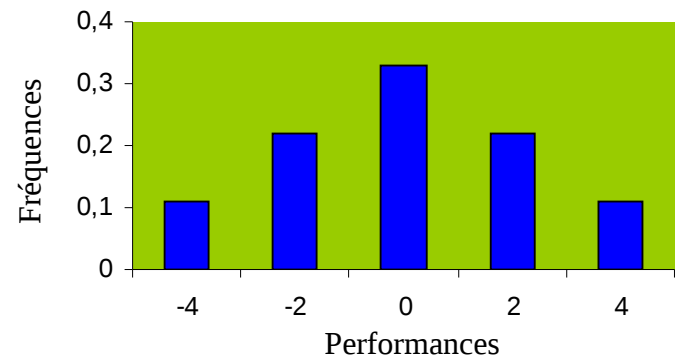
N = 1



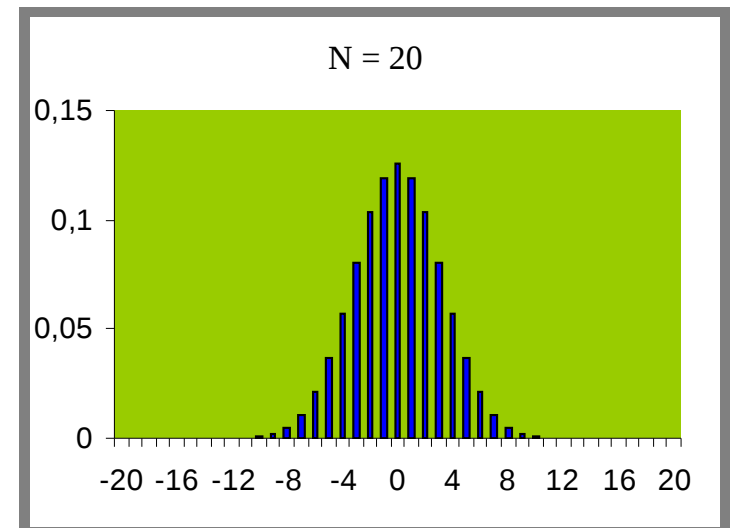
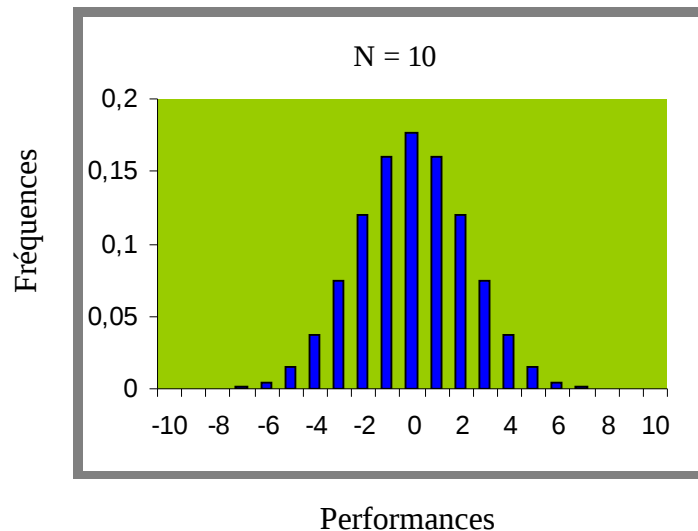
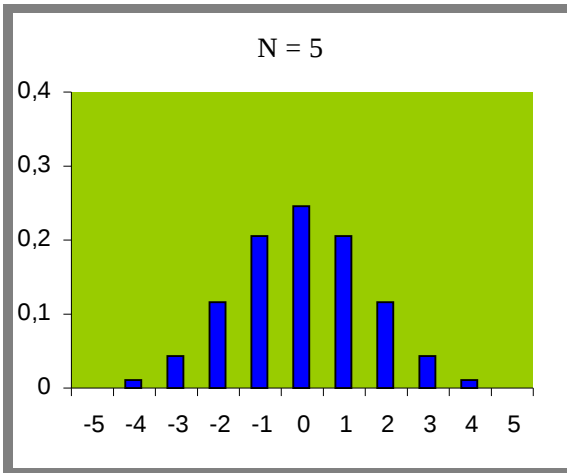
2 gènes, 2 couples d'alleles : **A, a et B, b**

AA, BB	+4
AA, Bb	+2
AA, bb	0
Aa, BB	+2
Aa, Bb	0
Aa, bb	-2
aa, BB	0
aa, Bb	-2
aa, bb	-4

N = 2



Distribution des performances
pour un caractère gouverné
par N couples de gènes



**La distribution de la variation
suit une loi normale**

Origine génétique de la variation quantitative

Théorème central limite

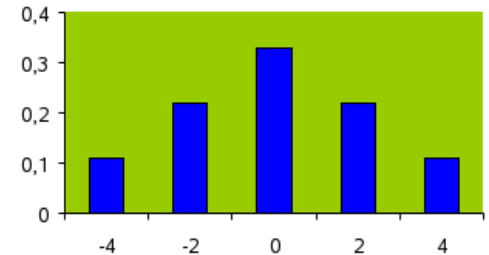
Soient X_1, X_2, \dots, X_N des variables aléatoires indépendantes, pas forcément de même loi, d'espérance μ et de variance σ^2 .

Soit Y la variable aléatoire définie comme la somme des N variables X_N .

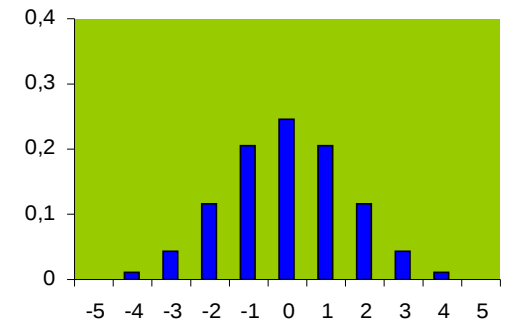
Si $N \rightarrow \infty$, alors Y tend vers une loi normale

**La loi normale est employée
comme référence en
génétique quantitative**

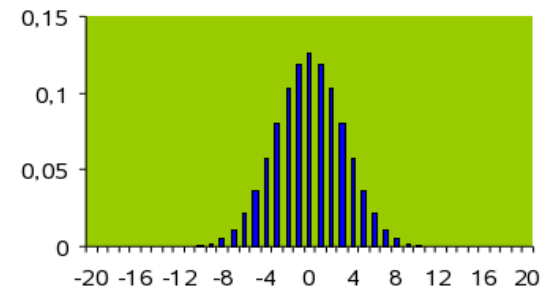
$N = 2$



$N = 5$



$N = 20$



Distribution de la longueur de la fleur (mm) dans un croisement entre deux variétés consanguines (P_1 et P_2) d'une espèce de tabac (East, 1916)

Les variations observées des caractères quantitatifs sont imputables ...

➡ à des facteurs de milieu

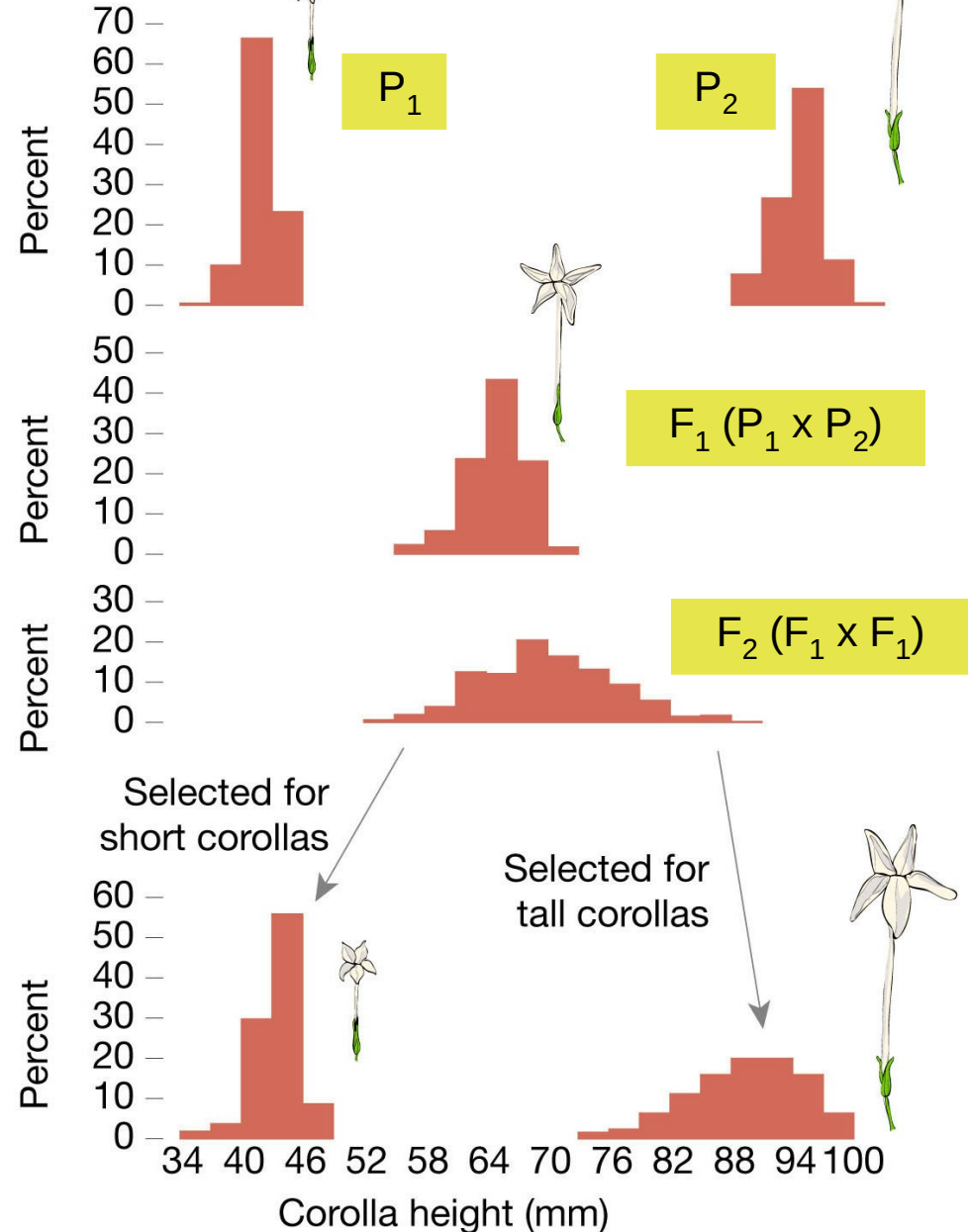
➡ aux ségrégations à plusieurs loci

Parents

F_1

F_2

F_5



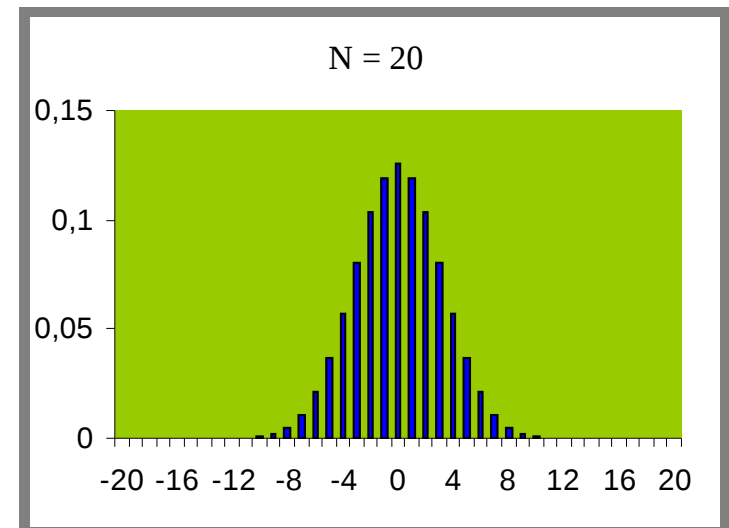
Déterminisme génétique des caractères quantitatifs

Origines de la continuité des caractères quantitatifs

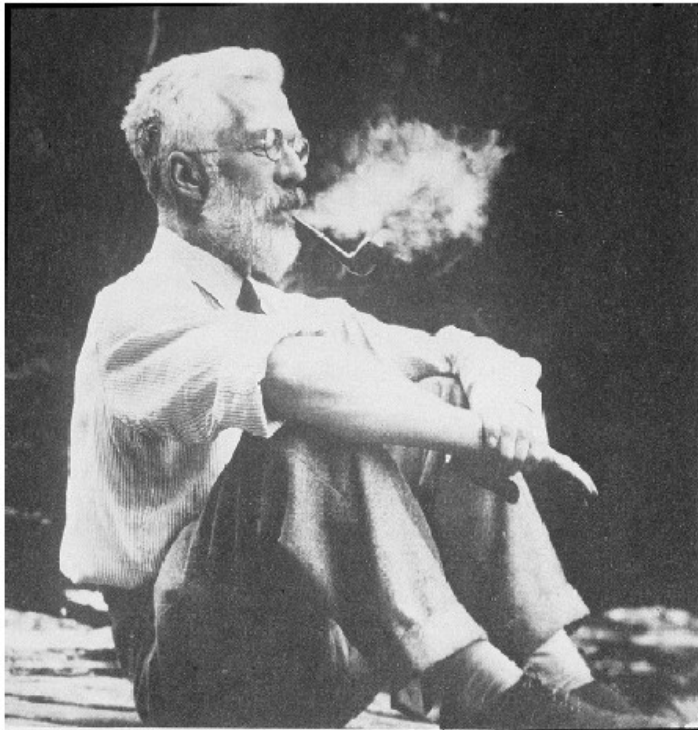
➔ Facteurs d'environnement

- l'alimentation
- fréquence de traites
- le mois du vêlage

➔ Intervention de nombreux gènes



Modèle Infinitésimal de Fisher (1918)



Fisher, R.A. (1918). 'The correlation between relatives on the supposition of Mendelian inheritance', Transactions of the Royal Society of Edinburgh 52: 399-433.

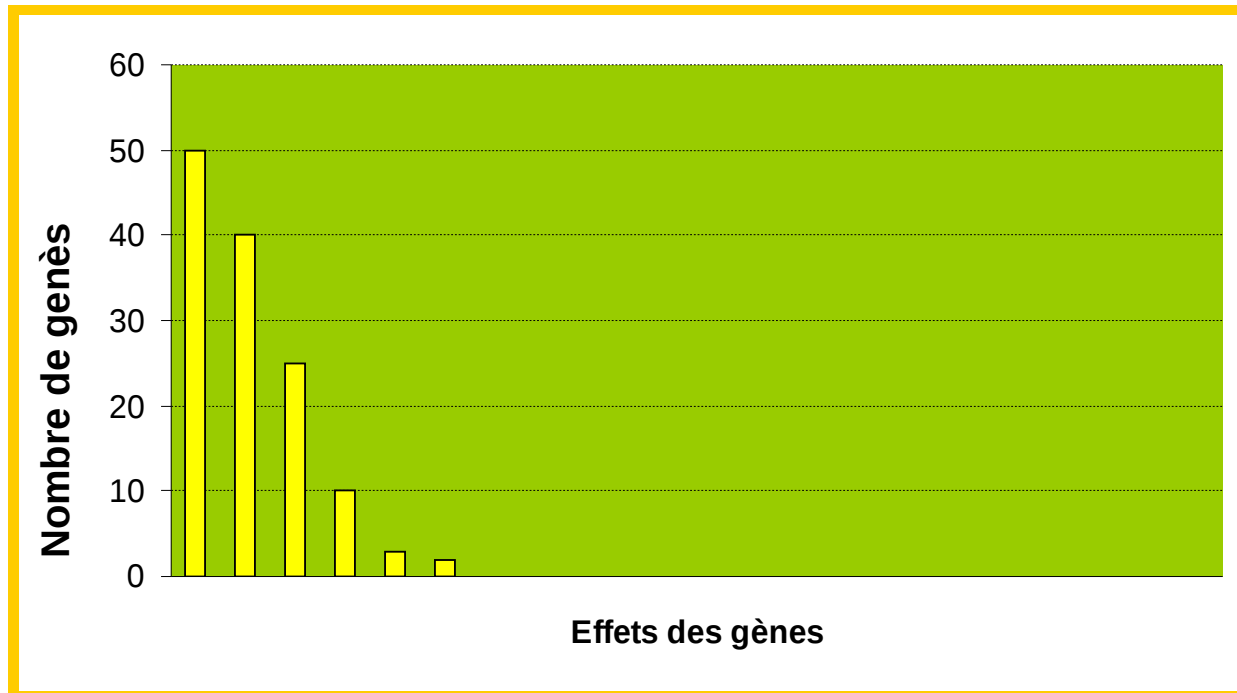
La **variation** observée pour les caractères quantitatifs est due à ...

- l'action combinée d'un grand nombre de gènes (infini) à hérédité "mendélienne" individuelle
- à l'effet de l'environnement

La **ressemblance entre individus** apparentés résulte du partage d'allèles en commun

Modèle Polygénique Infinitésimal

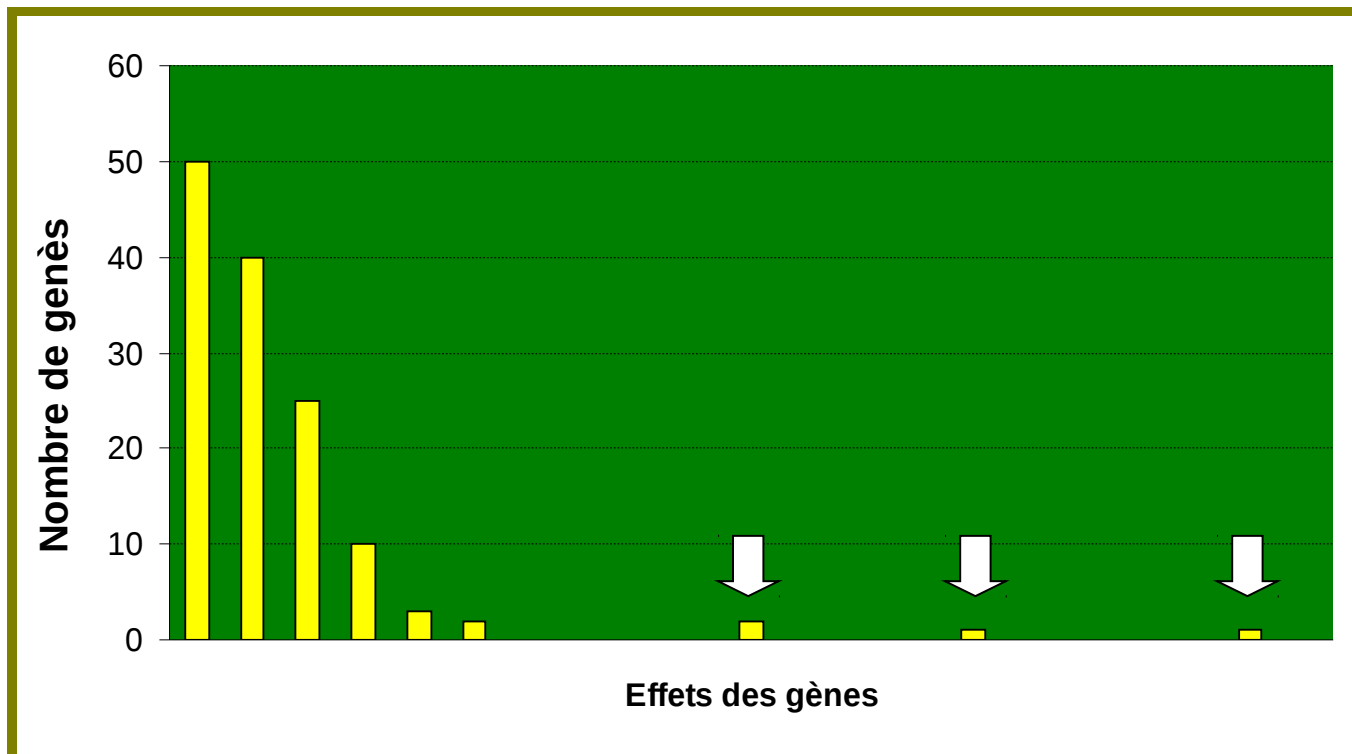
Le caractère est gouverné par une infinité (un très grand nombre) de gènes, chaque gène ayant un effet infinitésimal



Modèle Polygénique Infinitésimal ...

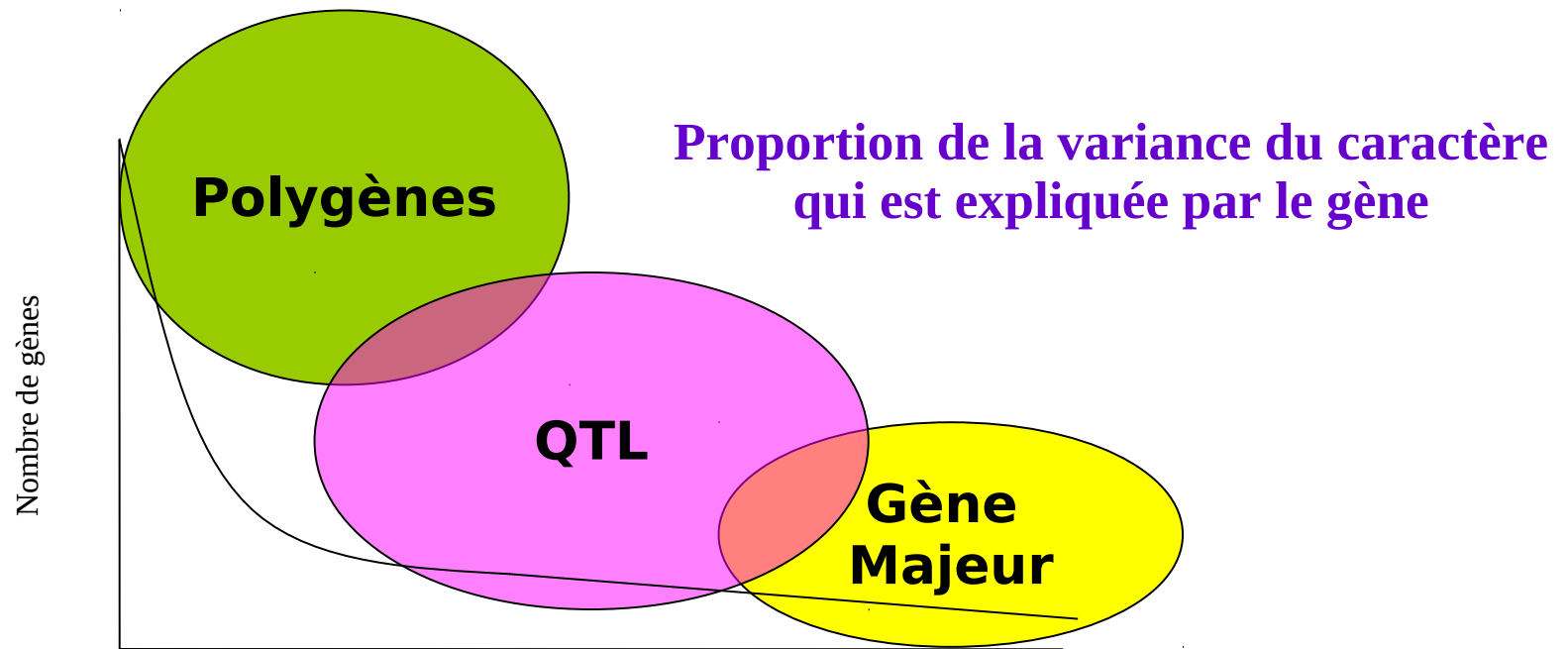
Pour certains caractères, on a pu mettre en évidence des **gènes à effet majeur**, ou des **QTLs**

QTL : Quantitative Trait Loci



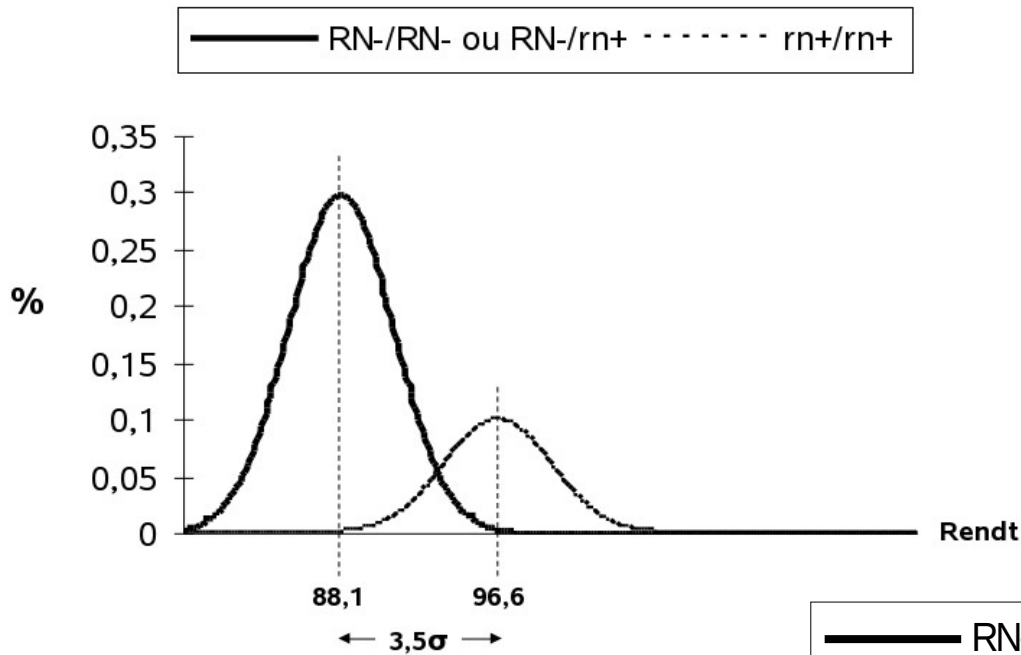
Modèle Polygénique Infinitésimal ...

La distinction entre gènes majeurs et polygènes repose sur l'importance de leurs effets.

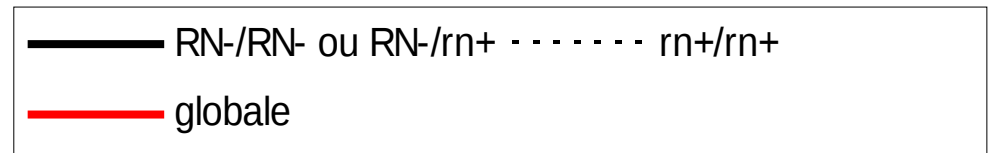


Effets de gènes

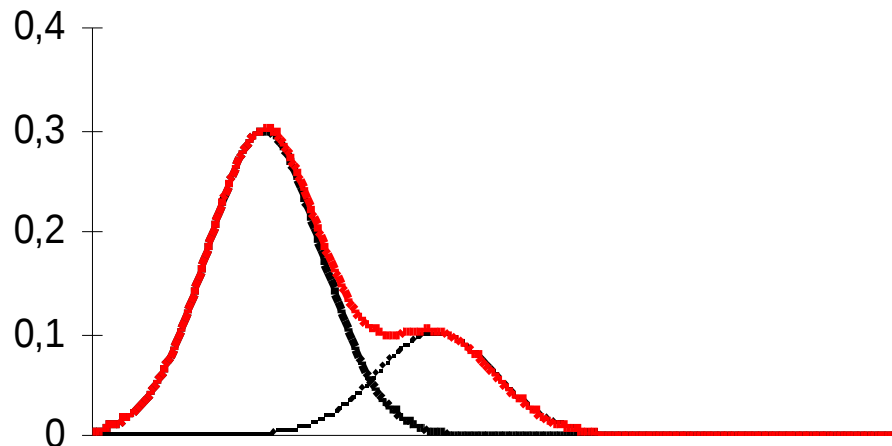
QTL : Quantitative Trait Loci



**Gène RN à effet majeur
sur la qualité de la
viande de porc**

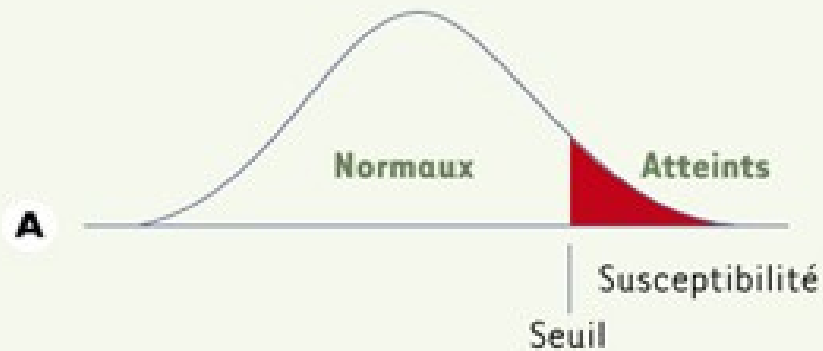


L'existence d'un gène
majeur n'exclut pas celle
d'autres gènes induisant
des variations pour le
caractère étudié

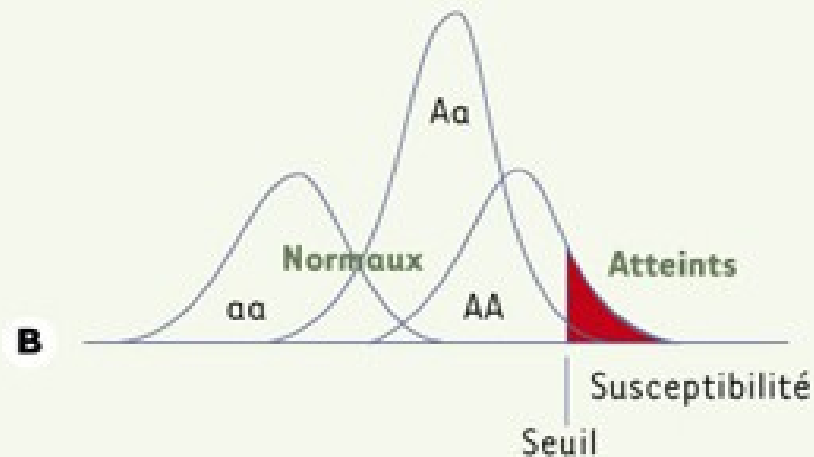


La distribution globale est bimodale

Maladies chez l'Homme



On suppose que la susceptibilité à la maladie est sous la dépendance de nombreux gènes et de facteurs de milieu



La susceptibilité est sous le contrôle d'un gène majeur, dont l'action est toutefois modulée par un système polygénique et des facteurs de milieu.

Modèle Polygénique Infinitésimal ...

La valeur phénotypique (P) s'exprime comme ...

$$P = G + E$$

avec la valeur génétique (G) et l'effet de l'environnement (E).

En l'absence d'interaction génotype x environnement, on a :

$$\text{Var}(P) = \text{Var}(G) + \text{Var}(E) \qquad \text{Cov}(G, E) = 0$$

Modèle Polygénique Infinitésimal

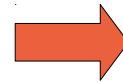
La *valeur phénotypique* (P) s'exprime comme ...

$$P = G + E$$

avec la *valeur génétique* (G) et l'environnement (E , *milieu non contrôlé*).

G de l'individu représente l'effet moyen du génotype sur un caractère donné.

On va étudier **les différents allèles** d'un même locus.



Modèle à un locus

Généralisation du modèle à plusieurs loci

Modèle à un locus

Expression de la valeur génétique

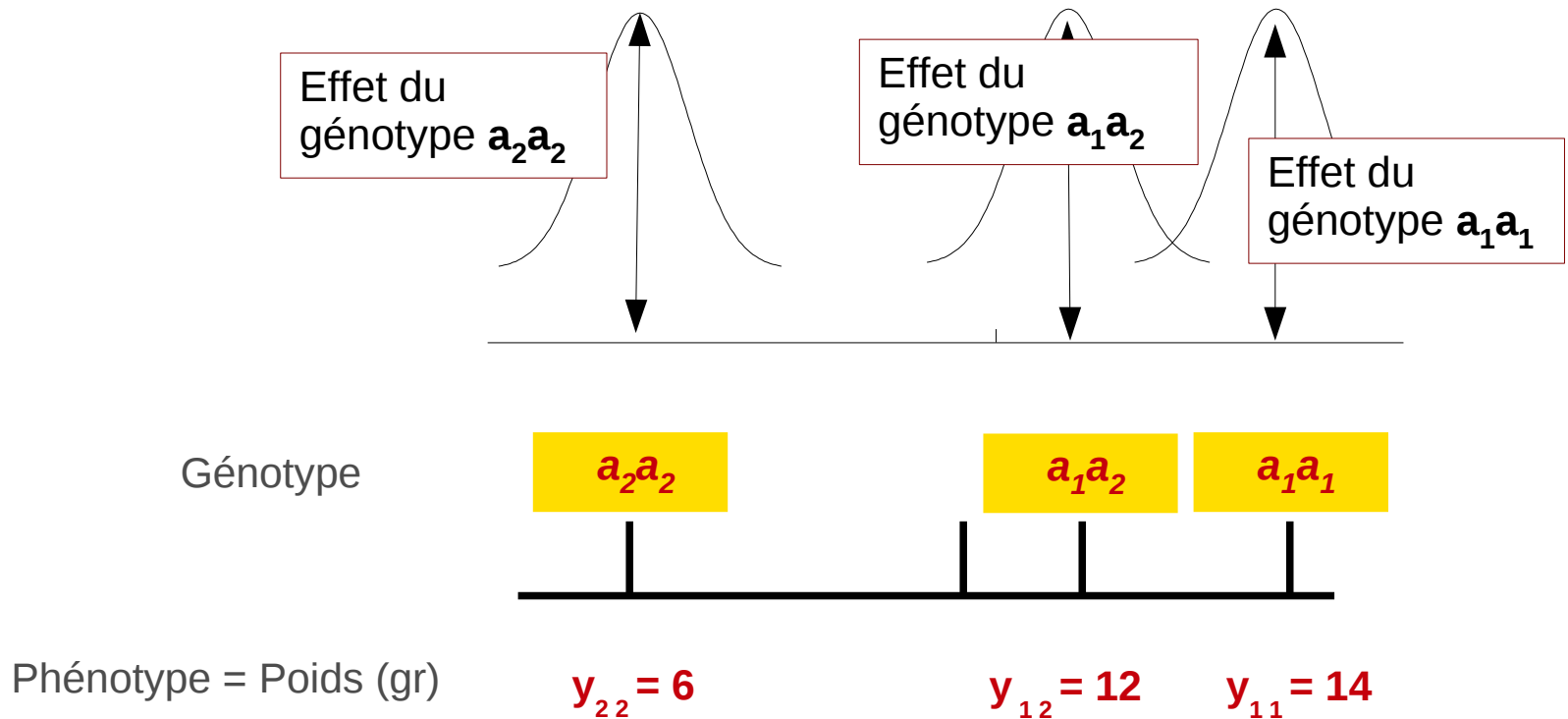
Considérons donc **un locus** avec **deux allèles** a_1 et a_2 en EHW.



On suppose que l'on peut identifier le génotype de chaque individu au locus considéré et donc calculer (sans erreurs) **la moyenne du poids à 6 semaines** (en grammes) pour chacun de génotypes sur un grand nombre de souris (King, 1950).

Modèle à un locus

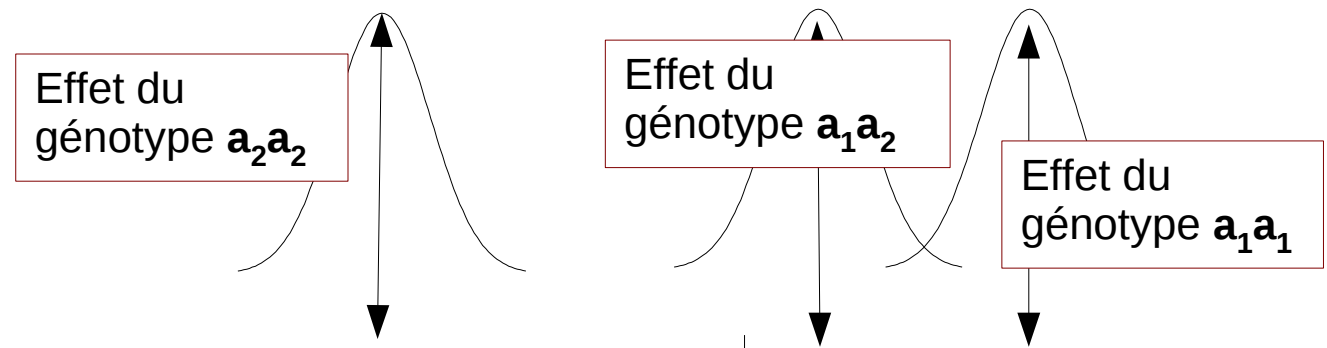
Expression de la valeur génétique



La moyenne du poids à 6 semaines

Modèle à un locus

Expression de la valeur génétique

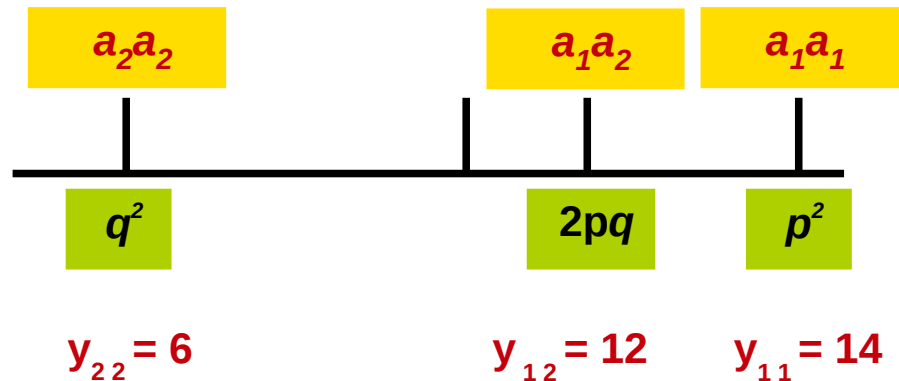


En EHW

Génotype

Pr(génotype) = Fréquence.

Phénotype = Poids (gr)



Modèle à un locus

Expression de la valeur génétique

Génotype	Fréq.	P	Fréq. x P
a_1a_1	p^2	y_{11}	$p^2 y_{11}$
a_1a_2	$2pq$	y_{12}	$2pq y_{12}$
a_2a_2	q^2	y_{22}	$q^2 y_{22}$

La moyenne de la population dépend donc des fréquences p et q

$$\mu = p^2 y_{11} + 2pq y_{12} + q^2 y_{22}$$

Exemple de calcul de la moyenne :



$$p=0,5$$

$$\mu = 0,25 \times 14 + 2 \times 0,25 \times 12 + 0,25 \times 6 = 11,00 \text{ g}$$

$$p=0,9$$

$$\mu = 0,81 \times 14 + 2 \times 0,09 \times 12 + 0,01 \times 6 = 13,56 \text{ g}$$

Modèle à un locus

Expression de la valeur génétique (G)

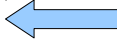


Génotype	Fréq.	P	G
a_1a_1	p^2	y_{11}	$y_{11} - \mu$
a_1a_2	$2pq$	y_{12}	$y_{12} - \mu$
a_2a_2	q^2	y_{22}	$y_{22} - \mu$

μ : la moyenne de la population

La **valeur génétique (G)** est une notion relative, elle dépend de la moyenne de la population dans laquelle on se trouve

Modèle à un locus

Expression de la valeur génétique (G)

Génotype	Fréq.	P	G
a_1a_1	p^2	y_{11}	$y_{11} - \mu$  G_{11}
a_1a_2	$2pq$	y_{12}	$y_{12} - \mu$  G_{12}
a_2a_2	q^2	y_{22}	$y_{22} - \mu$  G_{22}

Valeur génétique (G, en grammes) des 3 génotypes :

Fréquence du génome - moyenne des fréquences

$$G_{11} = (14 - 11) = 3$$

$$G_{12} = (12 - 11) = 1$$

$$G_{22} = (6 - 11) = -5$$

G_{11}	G_{12}	G_{22}
a_1a_1	a_1a_2	a_2a_2
3	1	-5

Modèle à un locus

Effet moyen d'un allèle (α_j)

De son génotype de valeur G , un individu ne transmet à sa descendance que la moitié des allèles



L'effet moyen d'un allèle

Père

a_1a_1

X

a_1a_2

Mère



$a_?a_?$

Modèle à un locus

Effet moyen d'un allèle (α_i)

L'effet moyen d'un allèle est désigné « effet additif » d'un allèle

$$f(a_1) = p = 0.5 \text{ et } f(a_2) = 1-p = q = 0.5$$

$$\alpha_1 = \alpha a_1 = f_{a_1} G_{a_1 a_1} + f_{a_2} G_{a_1 a_2} = p G_{a_1 a_1} + q G_{a_1 a_2}$$

$$\alpha_1 = 0.5 (3) + 0.5 (1) = 2$$

$$\alpha_2 = \alpha a_2 = f_{a_1} G_{a_1 a_2} + f_{a_2} G_{a_2 a_2} = p G_{a_1 a_2} + q G_{a_2 a_2}$$

$$\alpha_2 = 0.5 (1) + 0.5 (-5) = -2$$

Modèle à un locus

Effet moyen d'un allèle (α_i)

L'effet moyen d'un allèle est désigné « effet additif » d'un allèle

$$\alpha_i = \sum_{j=1}^n p_j G_{ij}$$

L'effet moyen de la substitution d'un allèle (α)

$$\alpha = \alpha_1 - \alpha_2$$

Modèle à un locus

La valeur génétique (G) ...

... en un locus est l'effet de la combinaison de deux allèles.

Il est ainsi naturel de rapprocher la valeur génétique et la somme des effets moyens de deux allèles qui composent ce génotype.

**G = effet moyen de l'allèle paternel + effet moyen de l'allèle maternel
+ effet d'interaction entre les 2 allèles**

Soit un individu de génotype $\mathbf{a_i a_j}$ (G_{ij} , i et j quelconques) ...,

$$G_{ij} = \alpha_i + \alpha_j + \delta_{ij}$$

Modèle à un locus

La valeur génétique (G) ...

Soit un individu de génotype $a_i a_j$ (G_{ij} , i et j quelconques) ...,

$$G_{ij} = \alpha_i + \alpha_j + \delta_{ij}$$

La **valeur génétique additive (A)** d'un individu est la somme des effets moyens des allèles qu'il possède.

L'effet d'interaction est appelé **valeur ou résidu de dominance (D)**

$$G = A + D$$

$$P = G + E = A + D + E$$

Modèle à un locus

La valeur génétique additive (A) ...

Génotype Valeur génétique additive

a_1a_1

$$2\alpha_1 = 2q\alpha$$

a_1a_2

$$\alpha_1 + \alpha_2 = (q-p)\alpha$$

a_2a_2

$$2\alpha_2 = -2p\alpha$$

$$G = A + D$$



$$G_{ij} = \alpha_i + \alpha_j + \delta_{ij}$$

p= 0,5

Génotype	a_1a_1	a_1a_2	a_2a_2
G	+ 3	+ 1	-5
A	+ 4	0	-4
D	- 1	+ 1	-1

Modèle Polygénique Infinitésimal ...

Moyenne de valeurs au sein d'une population

Génotype Valeur génétique additive

$a_1 a_1$

$$2\alpha_1 = 2q\alpha$$

$a_1 a_2$

$$\alpha_1 + \alpha_2 = (q-p)\alpha$$

$a_2 a_2$

$$2\alpha_2 = -2p\alpha$$

L'espérance de la valeur génétique additive (A) dans une population est nulle


$$E(A) = 2 p^2 q \alpha + 2 p q \alpha (q-p) - 2 q^2 p \alpha = 2 p q \alpha (p + q - p - q) = 0$$

La variance des valeurs génétiques additives ($Var(A)$) est égale à l'espérance de son carré

$$Var(A) = 4 p^2 q^2 \alpha^2 + 2 p q \alpha^2 (q-p)^2 + 4 q^2 p^2 \alpha^2 = 2 p q \alpha^2$$

Prédiction de la valeur génétique additive d'un descendant

connaissant un seul parent

Génotype	G	A	Valeur attendue de la descendance
a_1a_1	$2 \alpha_1 + \delta_{11}$	$2 \alpha_1$	α_1
 a_1a_2	$\alpha_1 + \alpha_2 + \delta_{12}$	$\alpha_1 + \alpha_2$	$\frac{1}{2} \alpha_1 + \frac{1}{2} \alpha_2$
a_2a_2	$2 \alpha_2 + \delta_{22}$	$2 \alpha_2$	α_2

La valeur génétique additive (A) représente la fraction de la valeur génétique (G) dont on peut facilement prédire la transmission de parent à descendant.



Un parent transmet, en espérance, la moitié de sa valeur génétique additive

Modèle à un locus

Prédiction de la valeur génétique additive d'un descendant

connaissant les deux parents

Valeur génétique additive d'un descendant connaissant celles de ses parents

Aléa de méiose : d'un descendant à l'autre, ce ne sont pas exactement les mêmes gènes qui sont transmis par les parents

$$A_{desc} = \frac{1}{2} A_{père} + \frac{1}{2} A_{mère} + w$$

idem mère

$\frac{1}{2}$ **Valeur génétique additive du père**
(le père transmet la moitié de ses gènes)

Modèle Polygénique Infinitésimal

Modèle avec plusieurs loci

Modèle avec additivité d'un locus à l'autre

Pour / loci :

$$G = \sum_i G_{l_i}$$

$$A = \sum_i (\alpha_{l_i} + \alpha_{l_j})$$

$$D = \sum_i \delta_{l_i}$$

On généralise à plusieurs loci la décomposition vue pour un locus :

$$G = A + D$$

Modèle Polygénique Infinitésimal

Modèle avec plusieurs loci

$$P = G + E$$

$$P = A + D + I + E$$

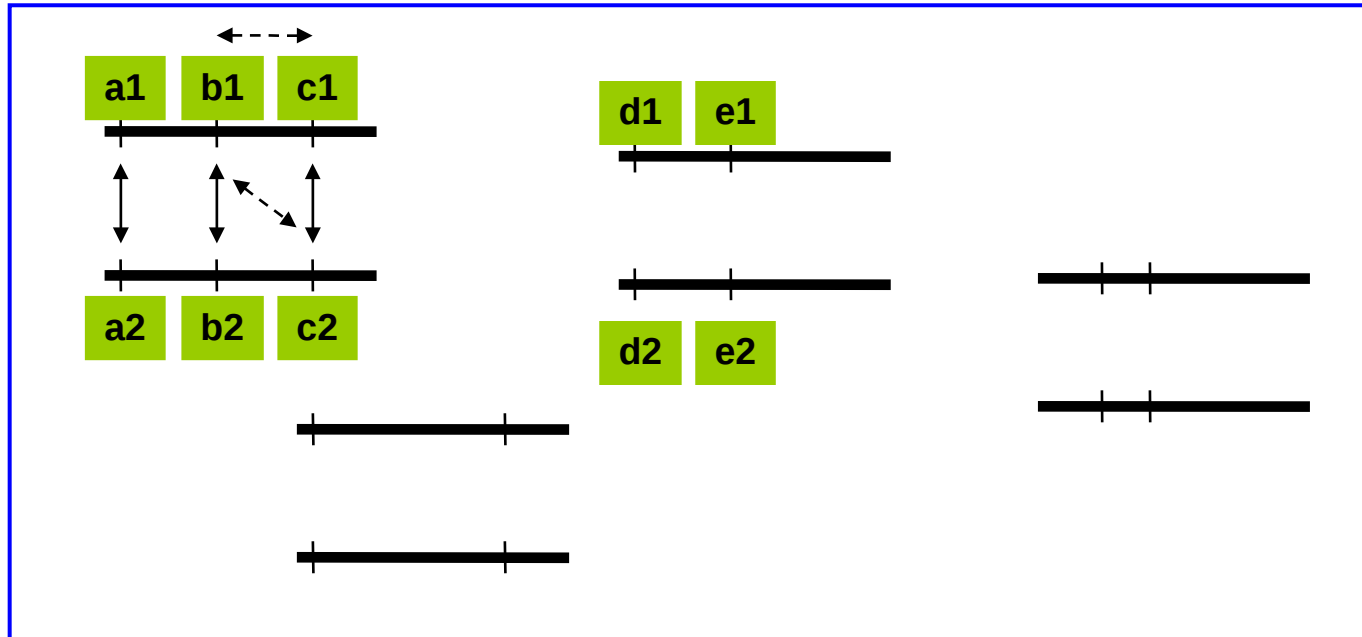
Avec : **A** = Valeur génétique additive
D = Valeur de dominance
I = Epistasie

A = Σ effet moyen de chaque gène

D = Σ effet produit par l'interaction des gènes présents au même locus

I = Σ effet produit par l'interaction des gènes présents à des loci différents

Modèle Polygénique Infinitésimal

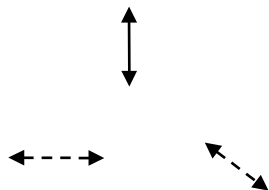


a1, a2, b1, b2, ...etc...

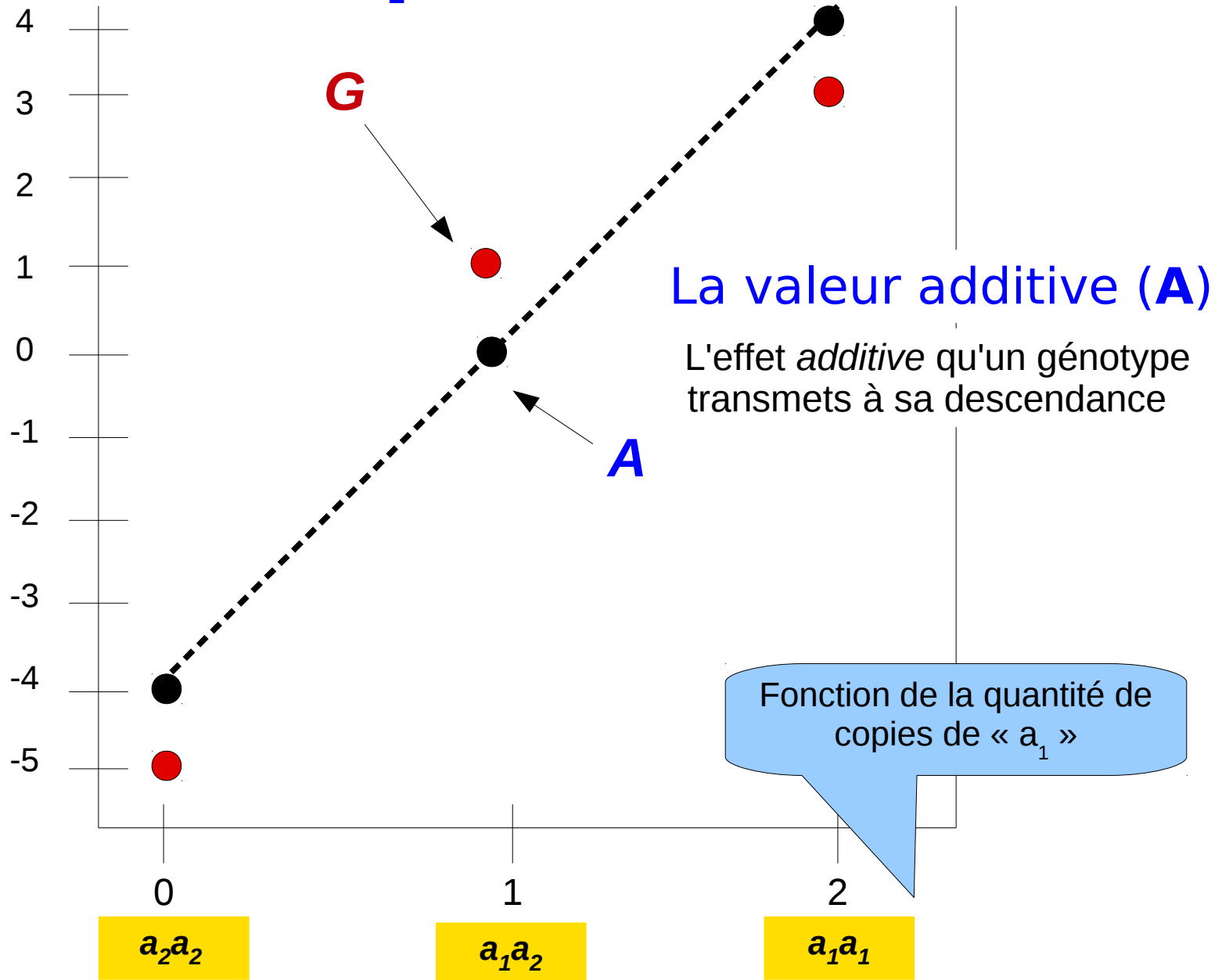
contribuent à **A** (additive)

contribuent à **D** (dominance)

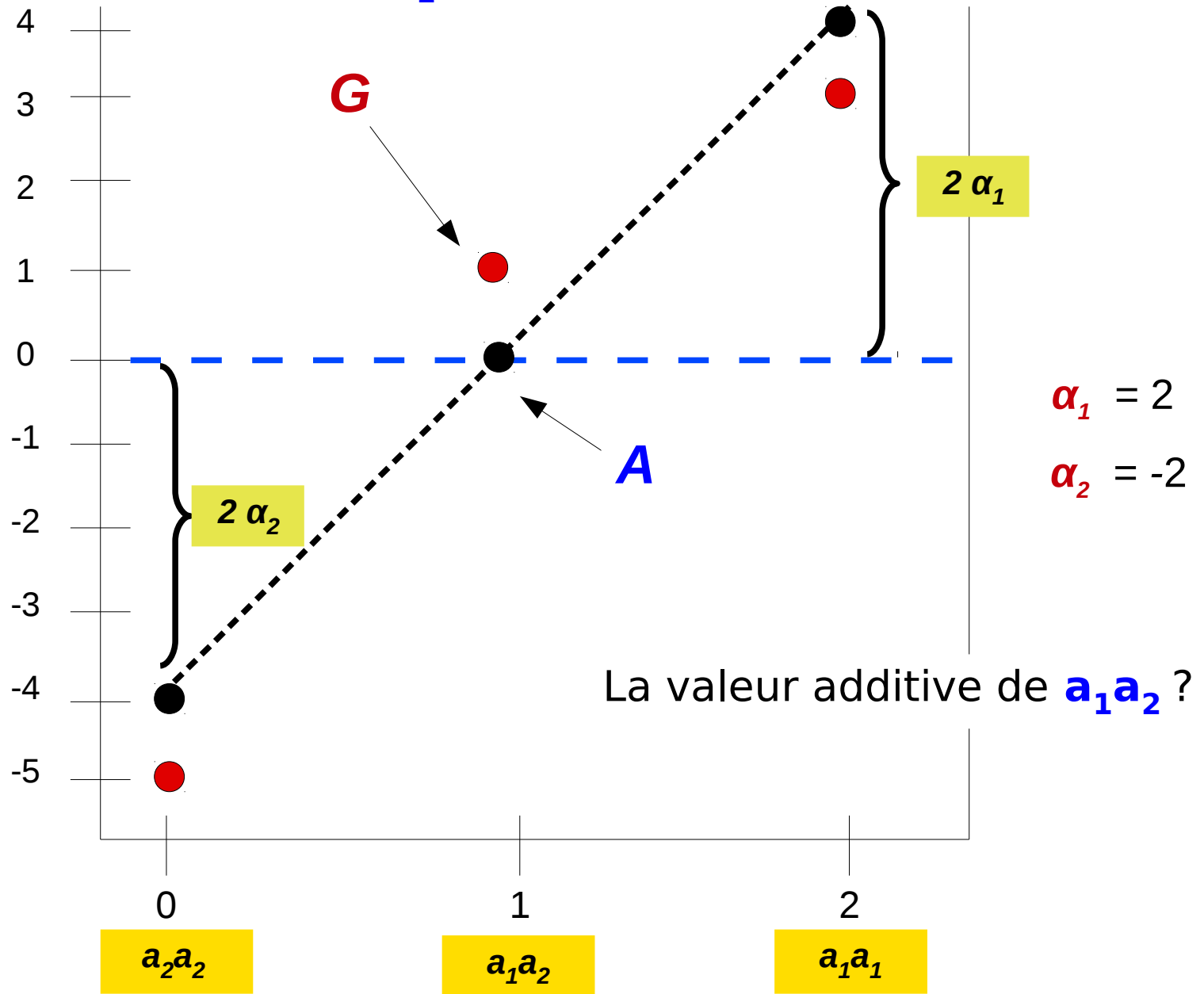
contribuent à **I** (épistasie)



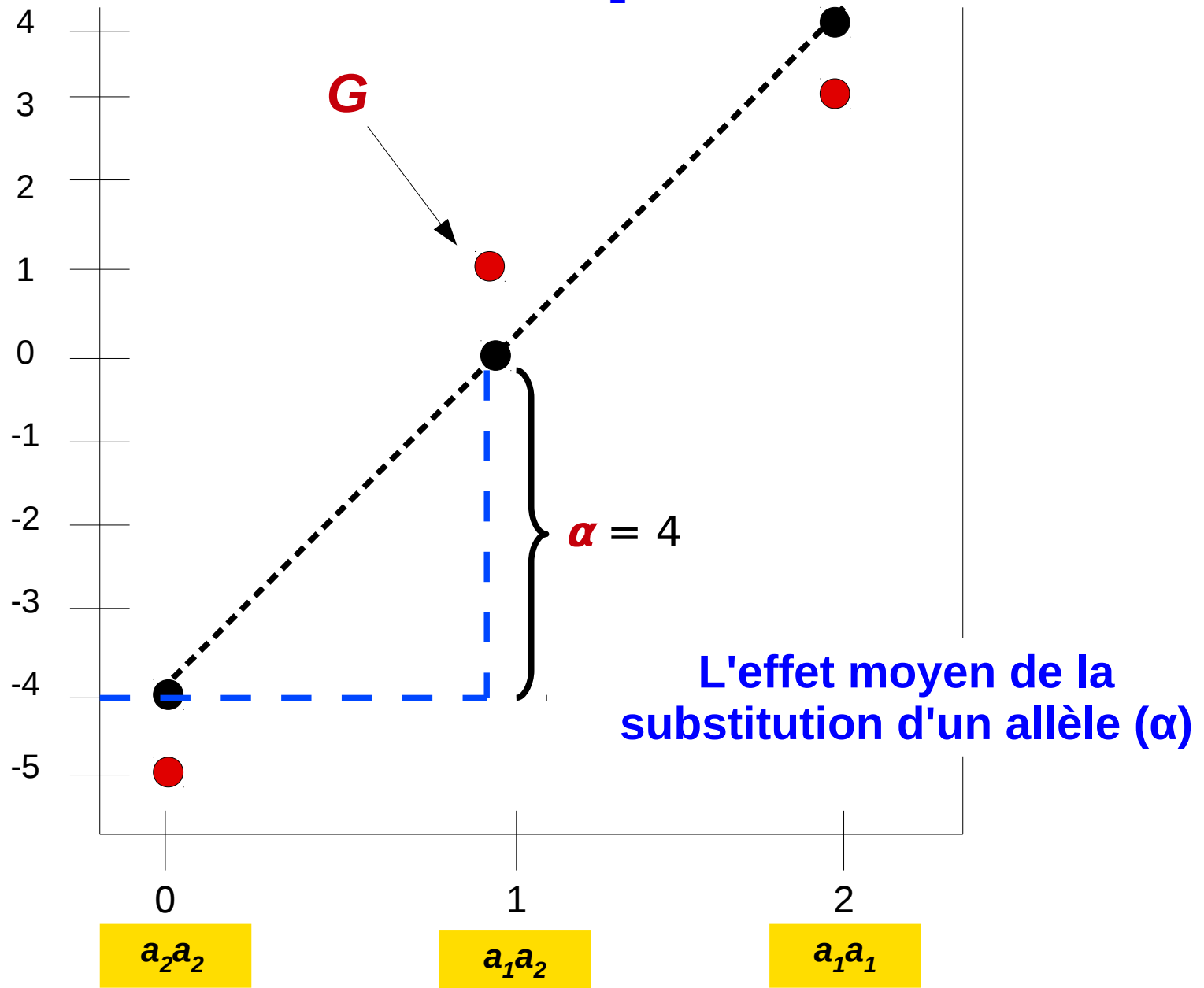
Régression linéaire de la valeur génétique (**G**) sur le nombre d'allèles **a₁** dans le génotype

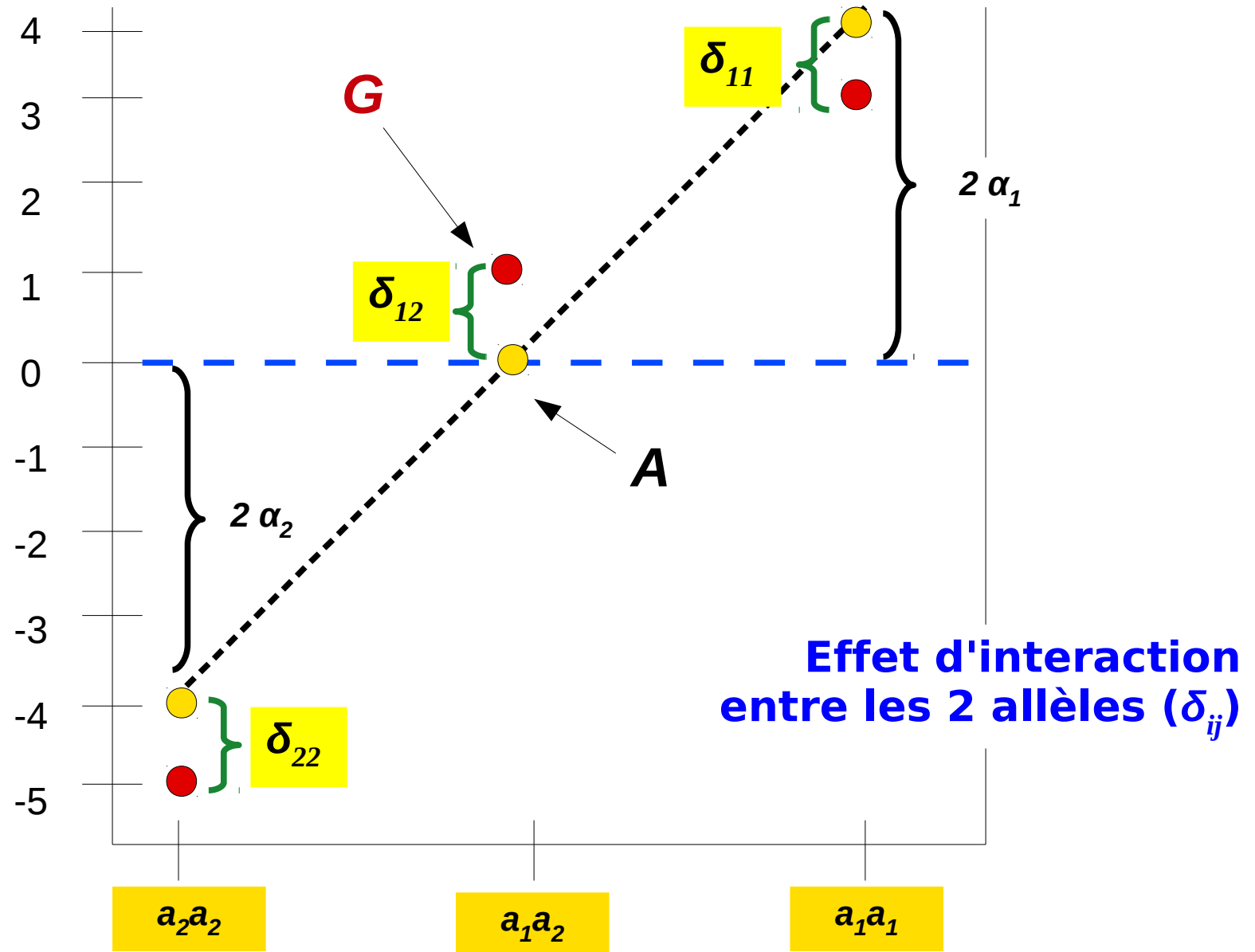


Régression linéaire de la valeur génétique (**G**) sur le nombre d'allèles **a₁** dans le génotype



α est le **coefficient de régression** de la valeur génétique (**G**) sur le nombre d'allèles a_1 dans le génotype





Modèle Polygénique Infinitésimal

$$P = G + E$$

En l'absence d'interaction génotype x environnement, on a :

$$\text{Cov}(G, E) = 0$$

$$\text{Var}(P) = \text{Var}(G) + \text{Var}(E)$$


$$\text{Var}(A)$$

 Notion d'héritabilité

Notion d'héritabilité

L'héritabilité au sens large

$$h^2 = \frac{V(G)}{V(P)}$$

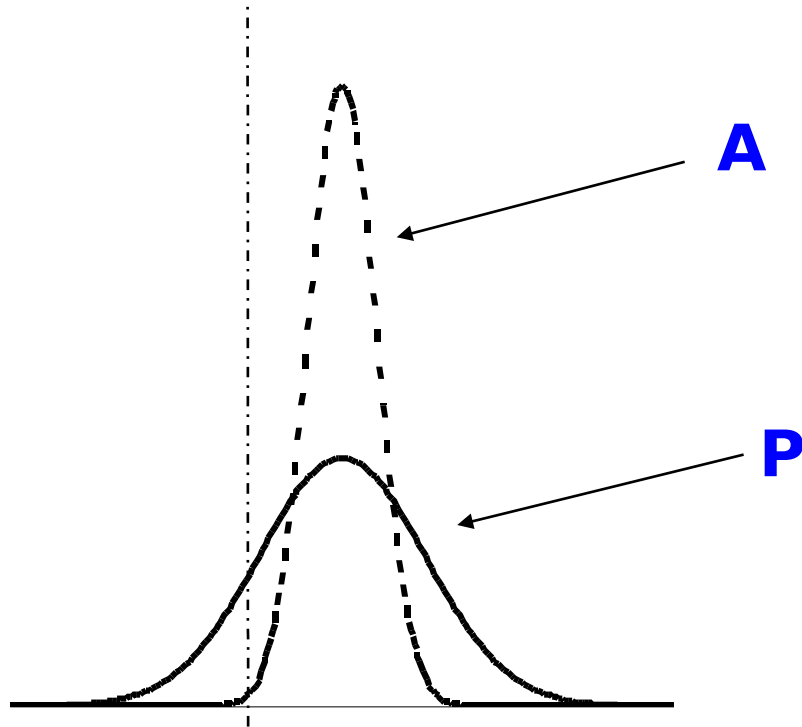
L'héritabilité au sens étroit



$$h^2 = \frac{V(A)}{V(P)}$$

L'héritabilité d'un caractère représente la proportion de la variance des productions pour ce caractère qui est de nature additive. Elle est comprise entre 0 et 1.

Notion d'héritabilité



La **SELECTION** exploite la variabilité génétique additive intra-population

$$h^2 = \frac{V(A)}{V(P)}$$

L'**héritabilité** permet de prédire si l'amélioration génétique par sélection sera efficace

Notion d'héritabilité

L'héritabilité au sens étroit

$$h^2 = \frac{V(A)}{V(P)}$$

Comme toute composante de la variance, l'héritabilité est un **paramètre spécifique du caractère** étudié et de la **population observée** et de son milieu.

Caractères fortement héritable

Caractères moyennement héritable

Caractères peu héritable

$$h^2 > 0,4$$

$$0,2 < h^2 < 0,4$$

$$h^2 < 0,20$$

Héritabilités : tendances générales

0

0,10

Reproduction (fertilité, prolificité ...)

0,20

Viabilité, résistance aux maladies

0,30

0,40

**Croissance (GMQ, Poids Age Type ...),
Efficacité alimentaire (IC)**

0,50

Quantité de produits (QL, QMG, QMP)

0,60

0,70

**Composition des produits
(% muscle, TVM, TB, TP, surface noix ...)**

0,80

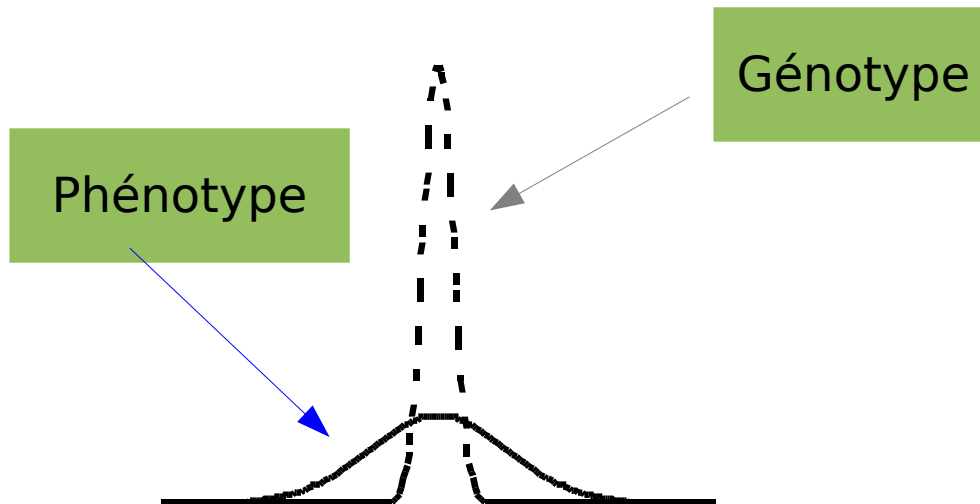
Possibilités de sélection ?

Caractère	μ (moyenne phénotypique)	h^2 (héritabilité)
Prolificité	1,50 pts	0,10
GMQ ferme	250 g	0,20
Rendement	45 à 50 %	0,30
Conformation	R- à U	0,20 à 0,30
Gras dorsal	4 à 5 mm	0,25 à 0,30
Surface noix côtelette	13 cm ²	0,50

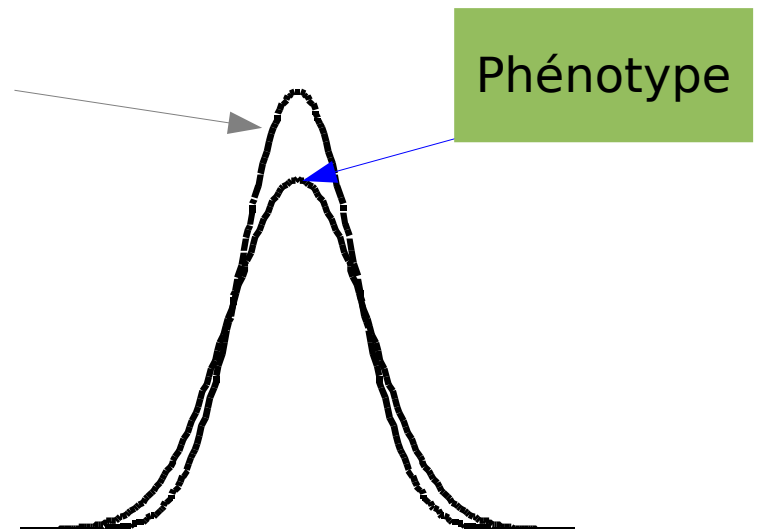
D'après Rattalino, 2008.

Notion d'héritabilité

$$h^2 = 0,1$$



$$h^2 = 0,8$$

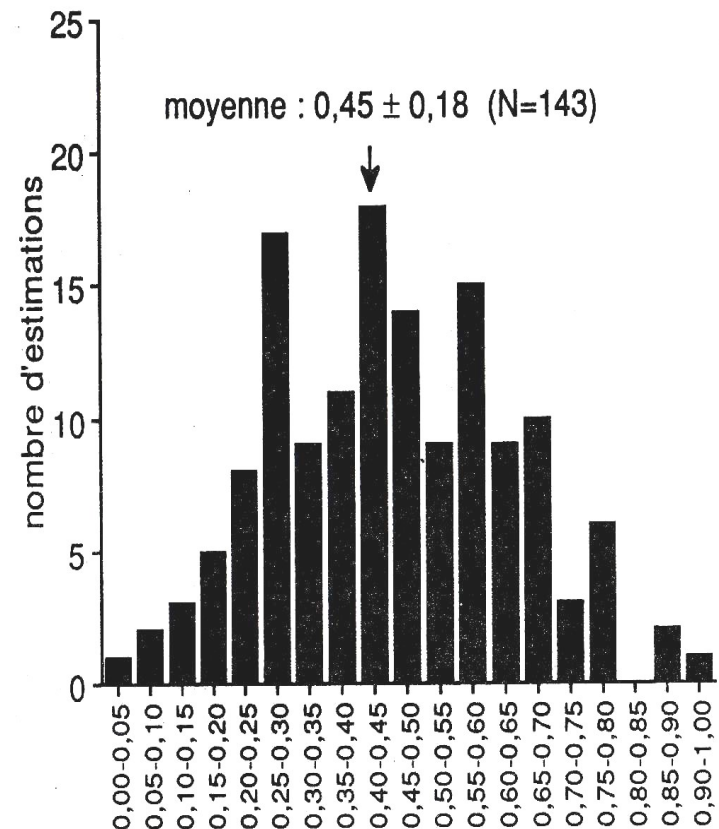


Notion d'héritabilité

L'héritabilité est un paramètre susceptible de varier suivant :

- la population considérée
(l'échantillon de cette population)
- le milieu de contrôle
- la période de contrôle
- la méthode d'estimation
- ...

a - hérabilité de l'épaisseur de lard
mesurée aux ultrasons



Notion de corrélation génétique

La corrélation génétique additive

Elle est la corrélation entre les valeurs génétique additives du même individu pour deux caractères différentes

En génétique animale, l'usage a fait que quand on parle de corrélation génétique sans autre précision, il s'agit toujours de la corrélation génétique additive.

Soit 2 caractères (1 et 2)
mesurés sur un grand
nombre d'individus ...

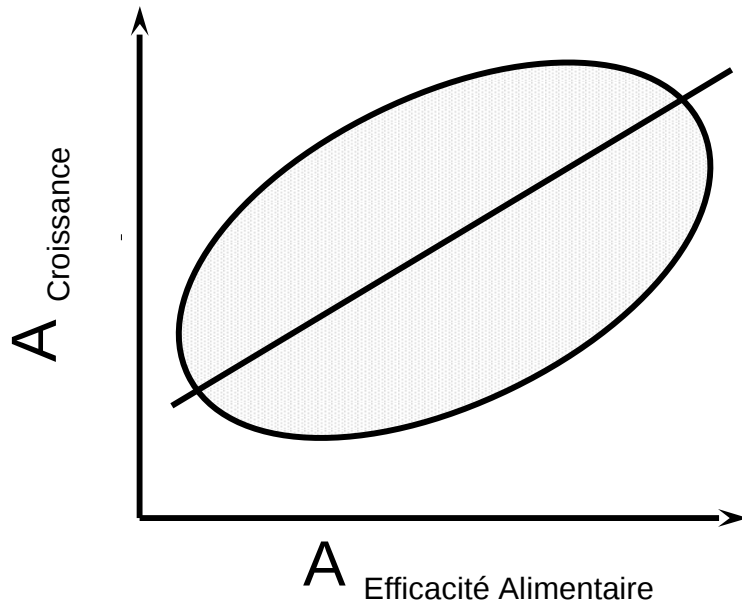
$$r_a = \frac{\text{Cov}(\mathbf{A}_1, \mathbf{A}_2)}{\sigma_{A_1} \sigma_{A_2}}$$

Le coefficient de corrélation mesure le degré de liaison linéaire entre deux variables X et Y. Ce coefficient se calcule comme le rapport de la covariance entre les deux variables et leurs écarts-types respectifs.

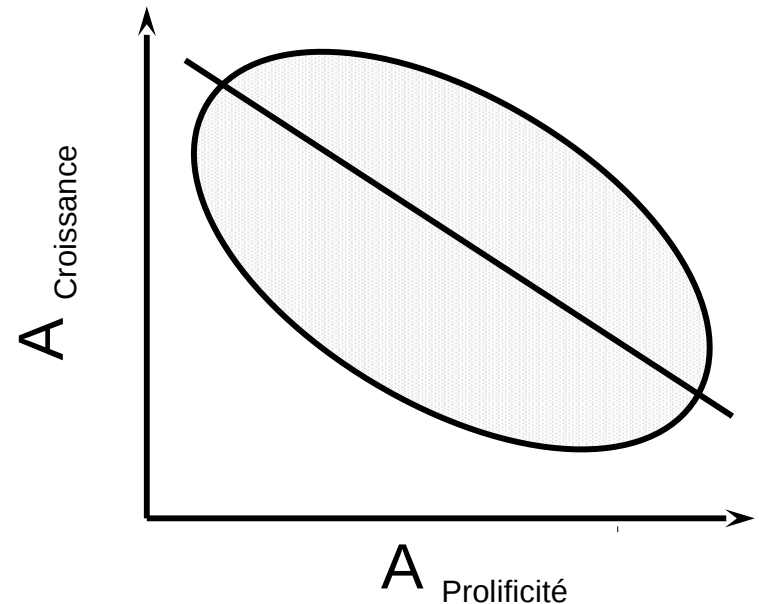
Notion de corrélation génétique

La corrélation génétique additive

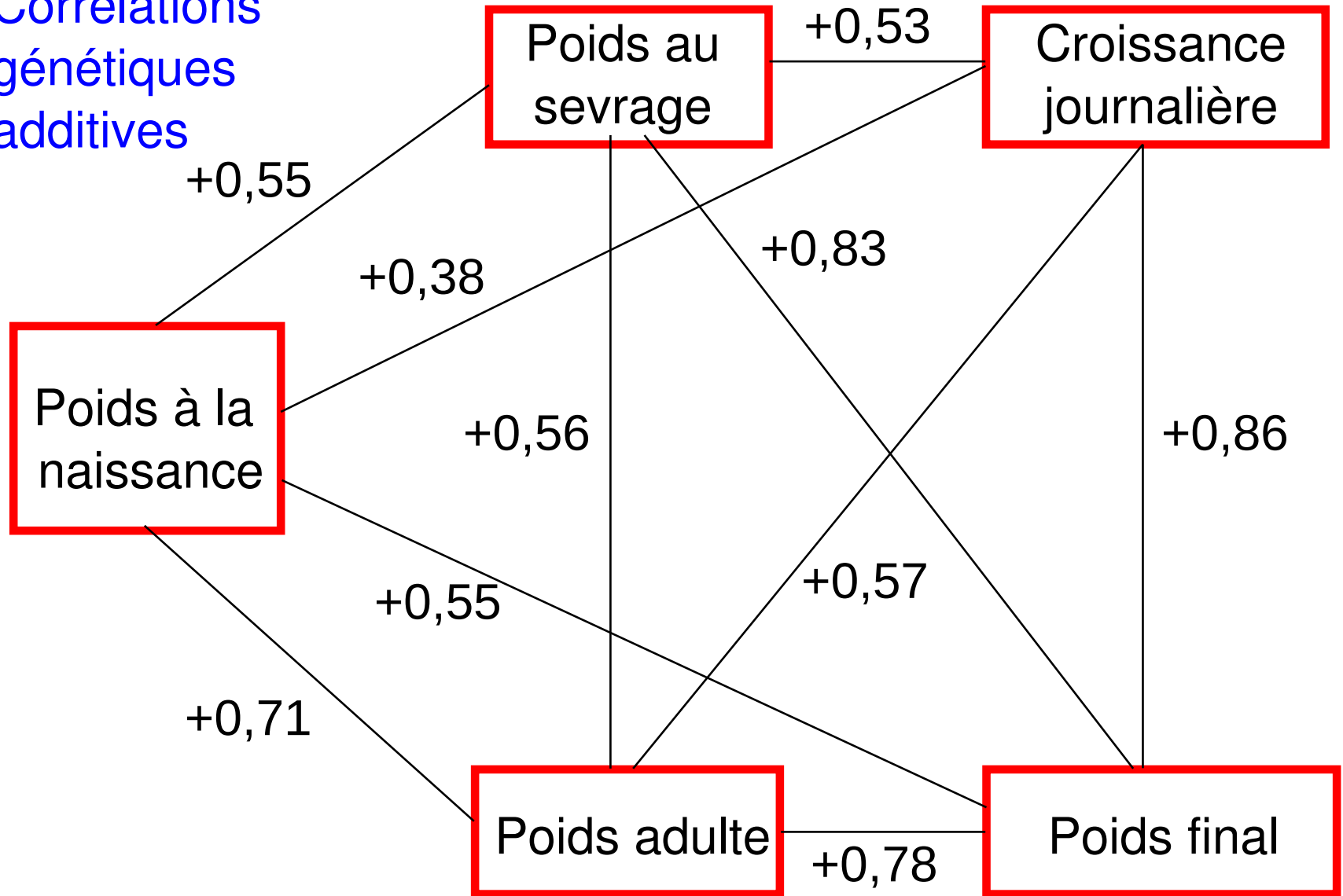
r_A Favorable

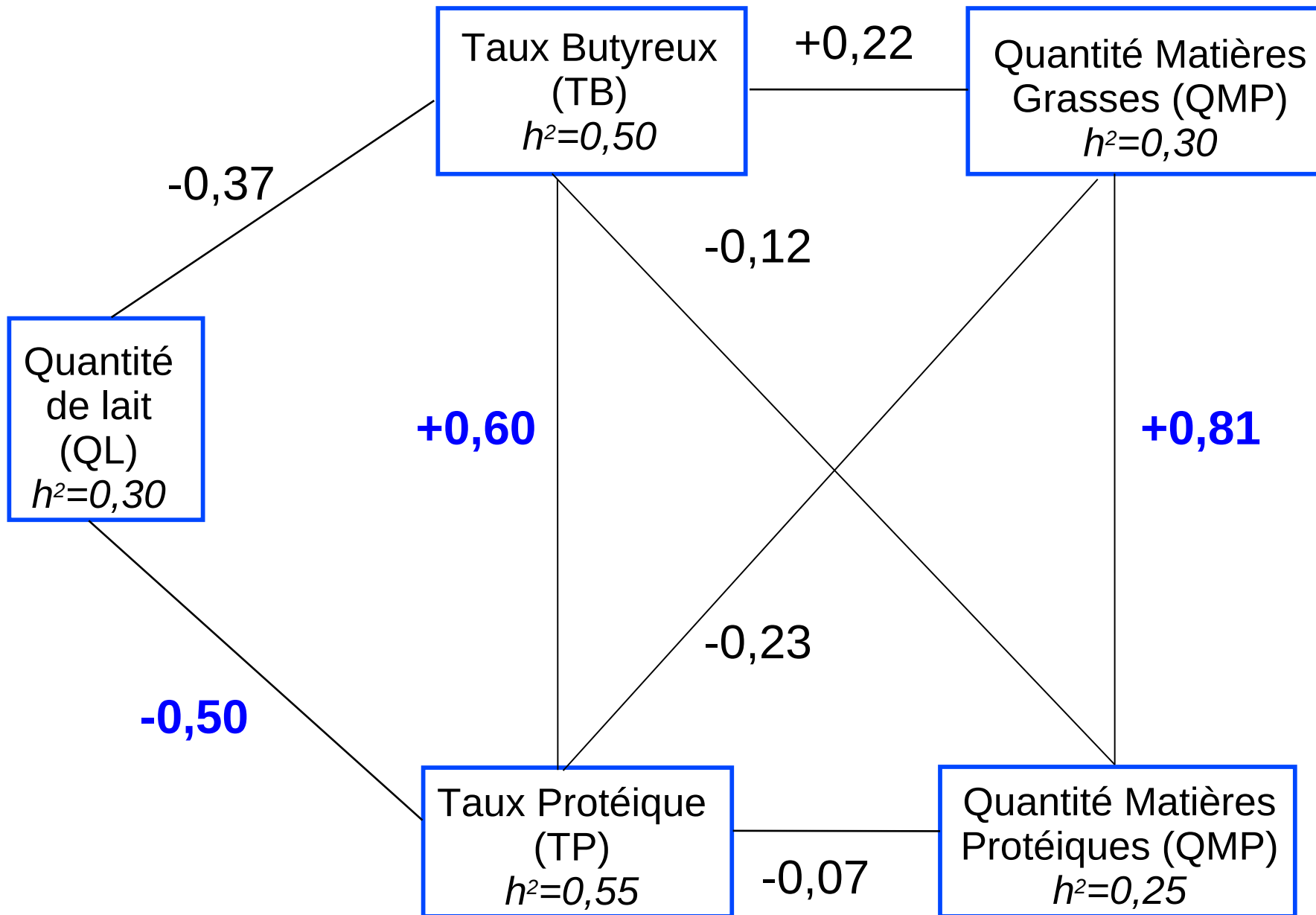


r_A Défavorable



Corrélations
génétiques
additives





Principe de l'amélioration génétique

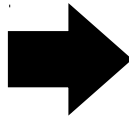
$$P = G + E$$

On veut améliorer **G**

$$G = A + D (+ I)$$

Améliorer A

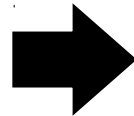
Exploiter la variabilité
génétique intra-race



SELECTION

Améliorer D

Exploiter la variabilité
génétique entre-races



CROISEMENT

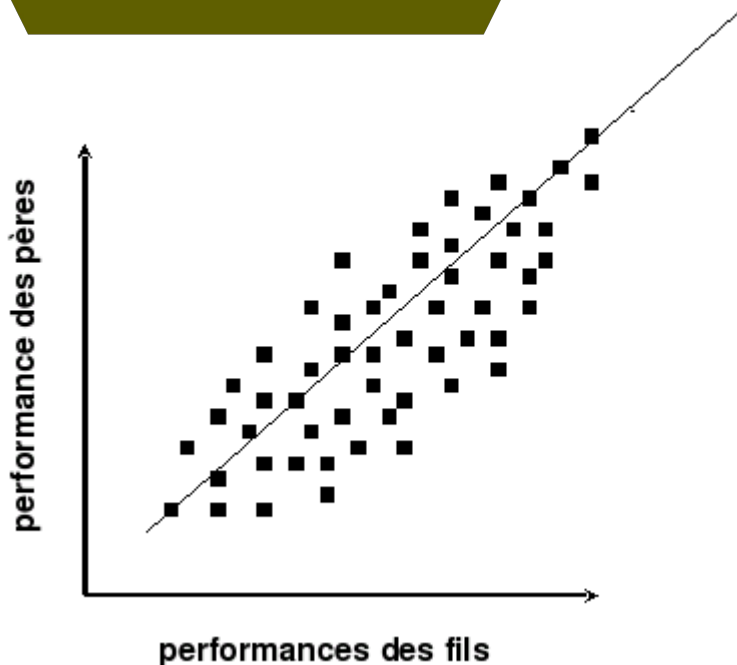
Ces deux stratégies sont
complémentaires !

La ressemblance entre apparentés

Covariance entre apparentés

$$h^2 = \frac{V(A)}{V(P)}$$

Deux individus qui partagent un ou deux parents en commun se ressemblent, sans leur être identiques



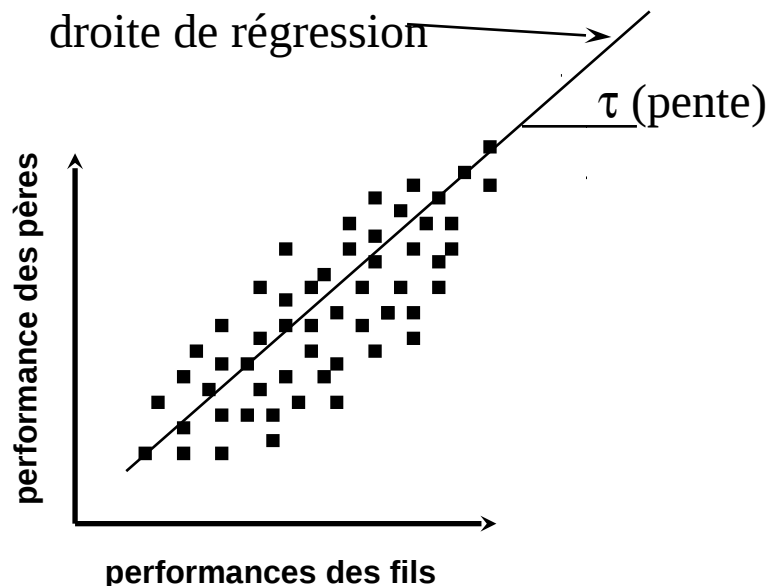
Le degré de ressemblance dépend du **lien de parenté** et du caractère étudié

Déterminer la partie de la **variance génétique additive** qui peut être estimée à partir de la ressemblance observée (phénotypique) entre individus apparentés

Paramètres génétiques

Interprétation de l'héritabilité

Soit P_i la valeur phénotypique d'un descendant, P_p et P_m celles de son père et sa mère. Les parents s'unissent au hasard. Il n'y a pas d'effet d'environnement commun entre parent et descendant. Considérons la valeur du “**parent moyen**” $P = (P_p + P_m) / 2$



$$\text{cov}(P_i, P) = \frac{1}{2} [\text{cov}(P_i, P_p) + \text{cov}(P_i, P_m)]$$

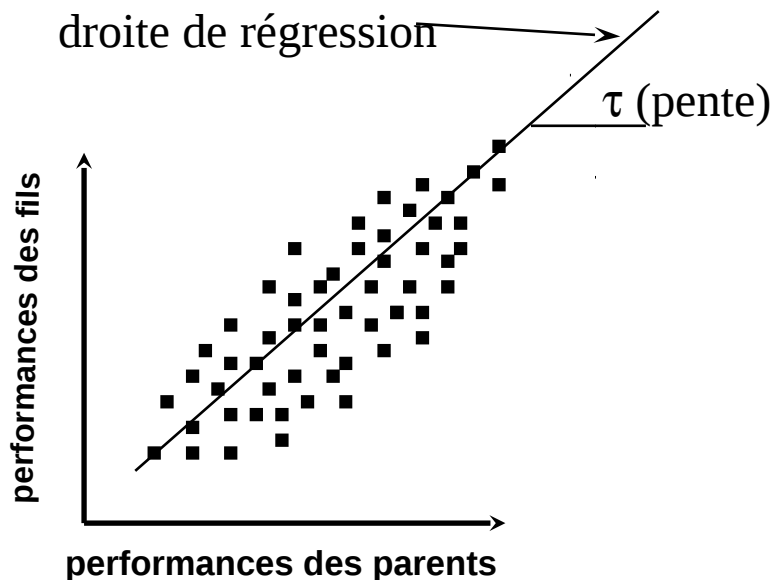
$$\text{cov}(P_i, P) = \frac{1}{2} [\frac{1}{2} V_A + \frac{1}{2} V_A] = \frac{1}{2} V_A$$

$$\text{Var}(P) = \frac{1}{4} \text{Var}(P_p) + \frac{1}{4} \text{Var}(P_m) = \frac{1}{2} V_P$$

Paramètres génétiques

Interprétation de l'héritabilité

On déduit la valeur du coefficient de régression du descendant sur son « parent moyen »



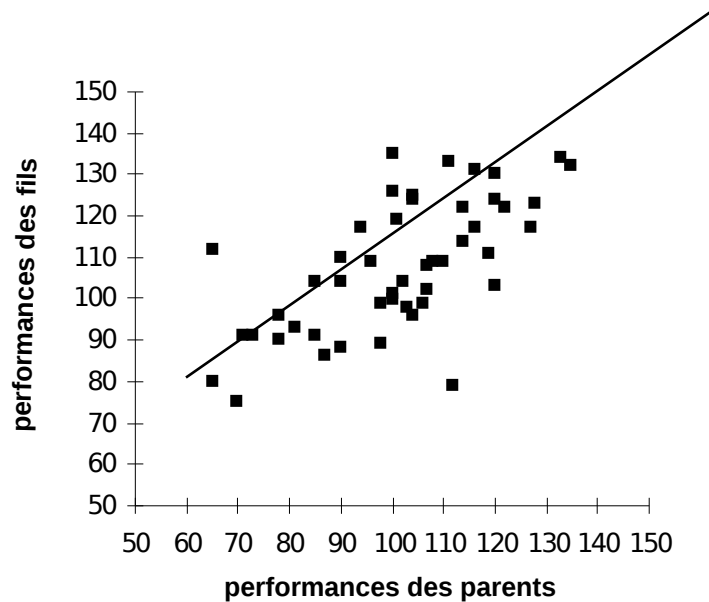
$$b_{Pi/P} = \frac{\text{cov}(P_i, P)}{\text{Var}(P)} = \frac{\frac{1}{2} V_A}{\frac{1}{2} V_P} = h^2$$

L'héritabilité au sens étroit s'interprète comme le coefficient de régression de la valeur phénotypique du descendant sur celle du parent moyen

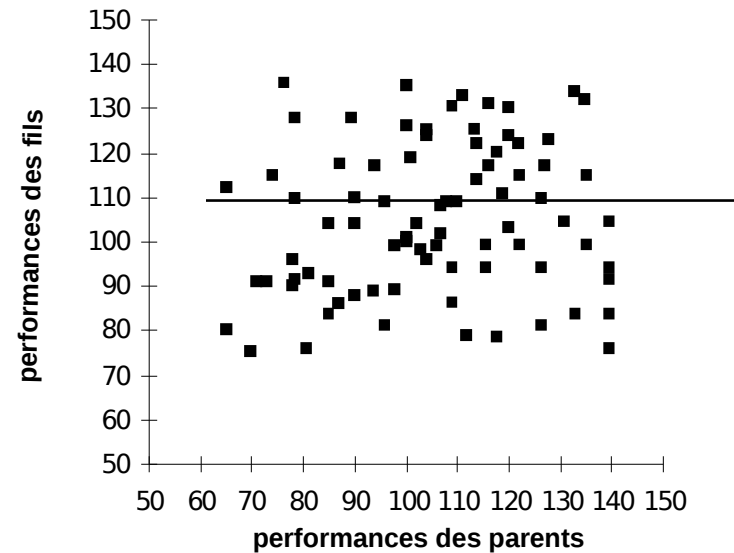
Paramètres génétiques

Estimation de l'héritabilité : régression parents-descendants

$$h^2 = 0.6$$



$$h^2 = 0.1$$



La ressemblance entre apparentés

Covariance entre apparentés

Soit 2 individus i et j appartenant à la même population et G et E sont non corrélés

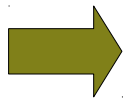
$$P_i = G_i + E_i \quad P_j = G_j + E_j$$

Lorsque l'on peut admettre l'absence d'effet d'environnement commun, l'expression de la covariance entre valeurs phénotypiques est

$$\text{cov}(P_i, P_j) = \text{cov}(G_i, G_j)$$

$$\text{cov}(G_i, G_j) = \text{cov}(A_i, A_j) + \text{cov}(D_i, D_j) = a_{ij} V(A) + d_{ij} V(D)$$

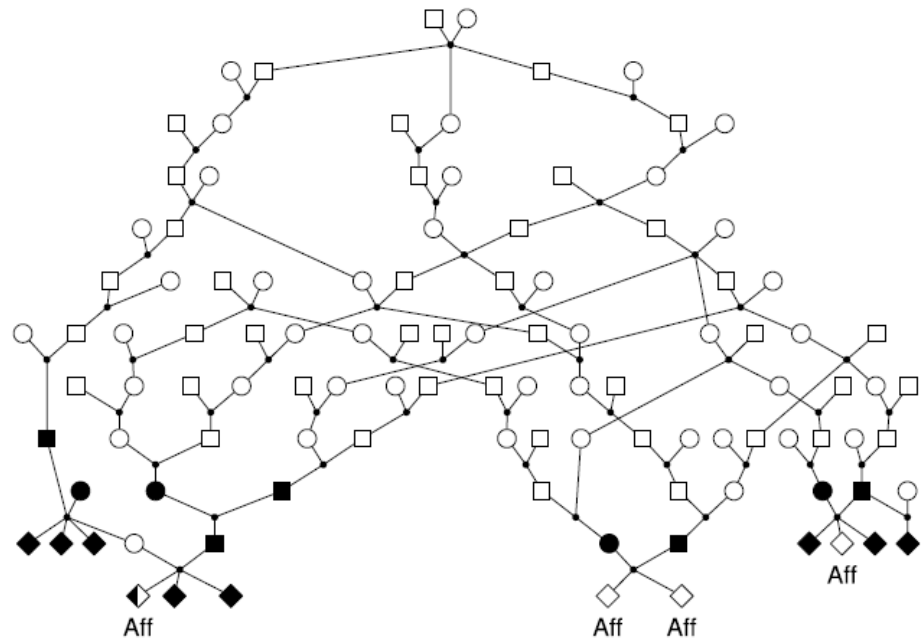
Coefficient
de relation



$$a_{ij} = 2 \phi_{ij} \quad (\phi_{ij} = \text{coefficient de parenté entre } i \text{ et } j)$$

(% gènes en commun)

Notion de parenté et de consanguinité



Loi d'équilibre de Hardy-Weinberg

1) Dans une population

- d'effectif infini,
- fermée (pas de migration),
- non soumise à sélection,
- dans laquelle il n'y a pas de mutation,

les fréquences géniques (p , q) restent constantes d'une génération à l'autre.

2) Si l'union des reproducteurs est panmictique, alors les fréquences génotypiques se déduisent automatiquement des fréquences géniques (p^2 , $2pq$, q^2)

1 caractère gouverné par un gène situé à 1 locus (A, a)

[t] $f(A) = p$ et $f(a) = q$

Grande taille, fermée
pas de mutation
pas de sélection

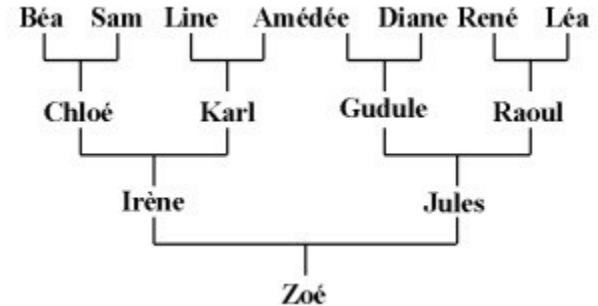
Panmixie

$f(AA)$	=	p^2
$f(Aa)$	=	$2pq$
$f(aa)$	=	q^2

[t+1] $f(A) = p$ et $f(a) = q$

Écarts à la panmixie

La consanguinité



Accouplements d'individus apparentés

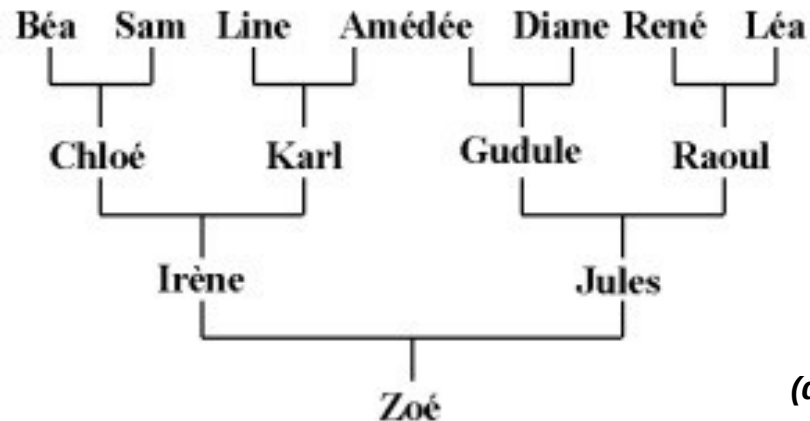
(1) Définitions

(2) Calcul des coefficients de parenté et de consanguinité

Écarts à la panmixie

La parenté et la consanguinité

(1) Définitions



(d'après Verrier et al., 2001)

Deux individus sont apparentés s'ils possèdent au moins un ancêtre commun
Irène et Jules sont apparentés

Un individu est consanguin si ses deux parents sont apparentés
Zoé est consanguine

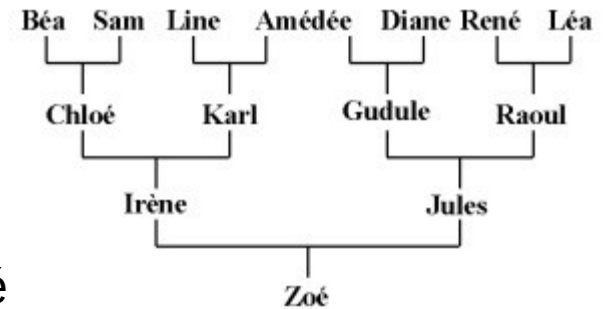
Écarts à la panmixie

(1) Définitions

Coefficients de parenté et de consanguinité

Le degré de parenté dépend du lien de parenté considéré (frère-soeur vs. deux cousins)

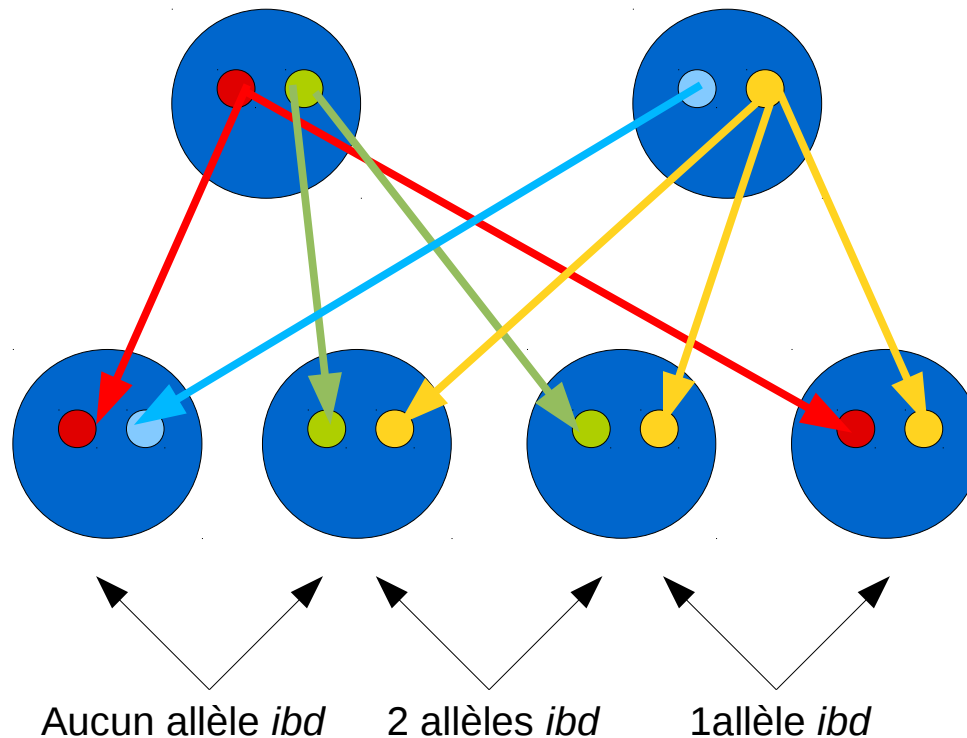
Le degré de consanguinité dépend de la parenté qui existe entre les 2 parents de l'individu concerné



(d'après Verrier et al., 2001)

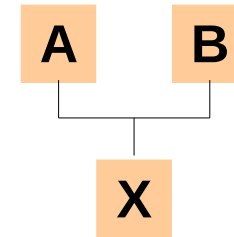
Pour quantifier ces phénomènes, on introduit les **coefficients de parenté et de consanguinité**.

On définit le **coefficient de parenté entre deux individus** A et B ($\Phi_{A,B}$) comme la probabilité que deux allèles tirés au hasard au même locus, l'un chez A et l'autre chez B, soient identiques par descendance¹



¹**Identiques par descendance** (*ibd : identical by descent*): deux copies d'un même allèle présent chez l'ancêtre commun

Écarts à la panmixie



Le coefficient de consanguinité d'un individu x (F_x) est égal à la probabilité pour que les deux allèles qui se trouvent à un locus donné soient identiques par descendance¹

Le coefficient de consanguinité d'un enfant est égal au coefficient de parenté entre ses deux parents:

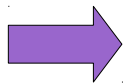
$$F_x = \Phi_{AB}$$

Écarts à la panmixie

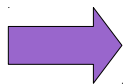
(2) Calcul des coefficients de parenté et de consanguinité

Son calcul suppose la connaissance de la **généalogie** et il est fondé sur la notion d'**identité des gènes**

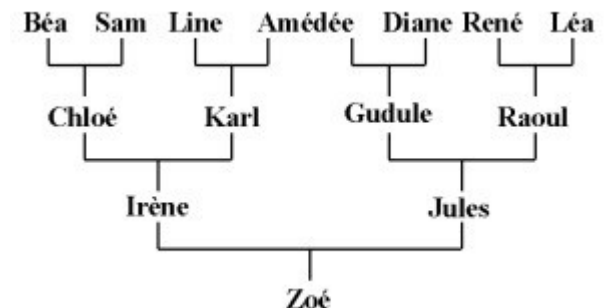
Deux méthodes de calcul peuvent être distinguées:



Méthode 'des chemins'



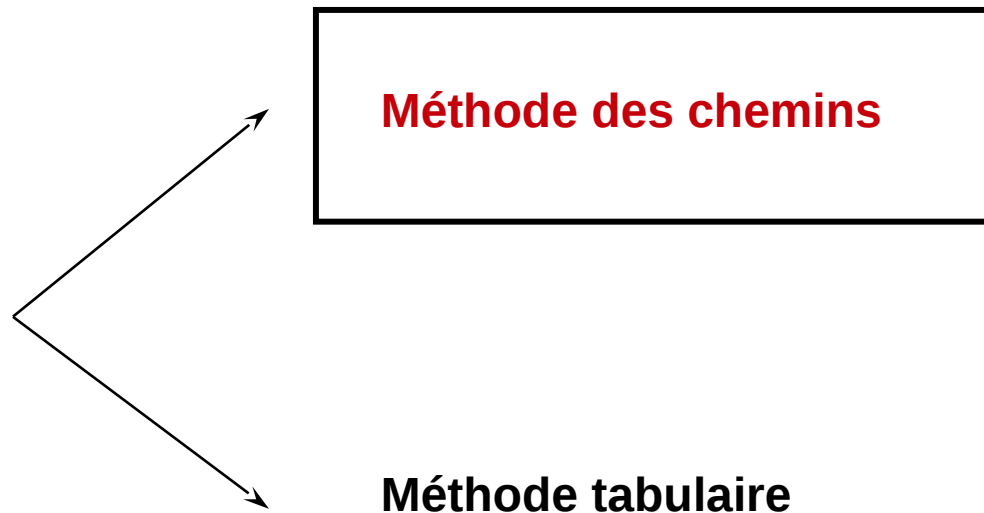
Méthode 'tabulaire'



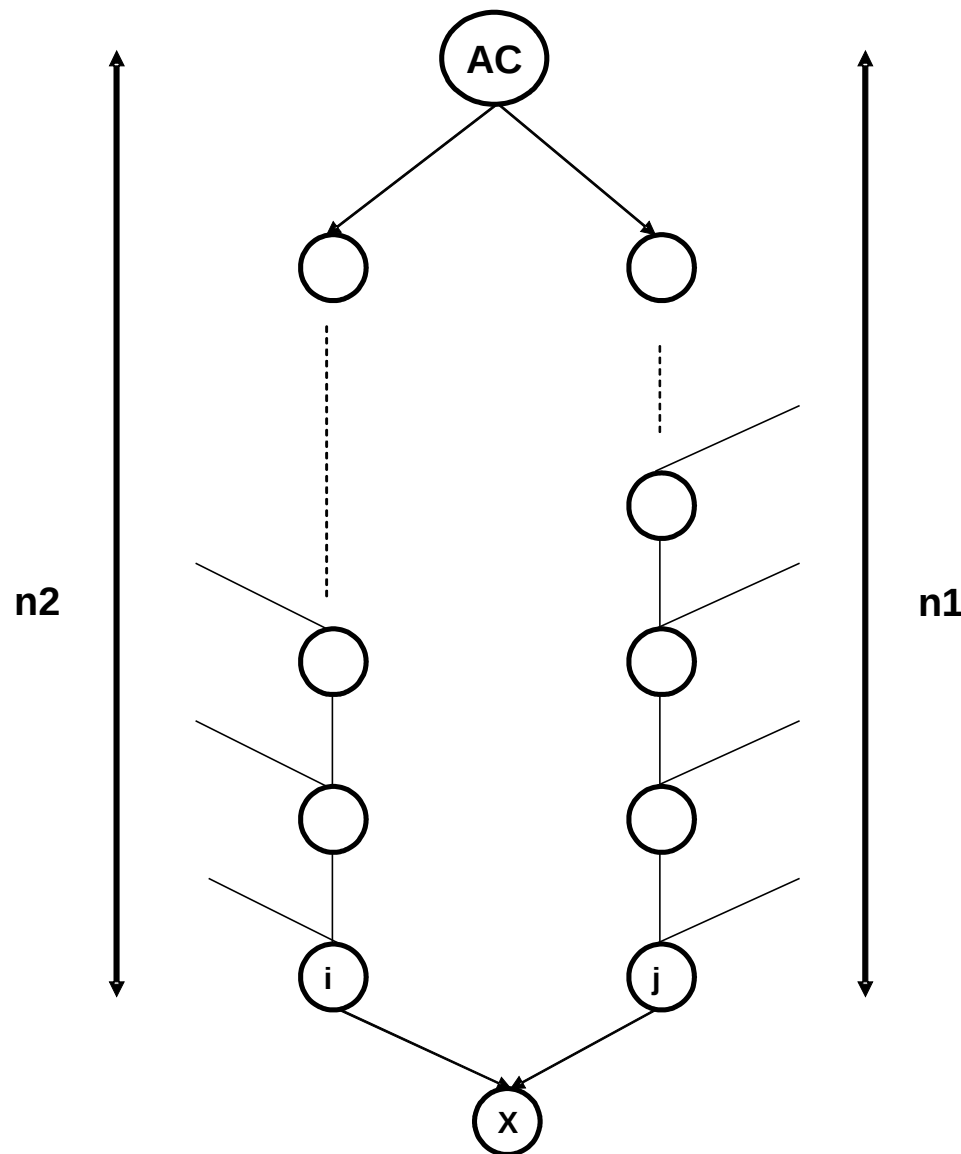
(d'après Verrier et al., 2001)

Écarts à la panmixie

(2) Calcul des coefficients de parenté et de consanguinité



Méthode des chemins



$$\phi_{ij} = F_x = \sum_{AC} \left(\frac{1}{2} \right)^{n1+n2+1} (1 + F_{AC})$$

(1) Identifier tous les ancêtres communs à i et à j

(2) Calcule le coef. de consanguinité (F_{AC}) de chacun des ancêtres communs

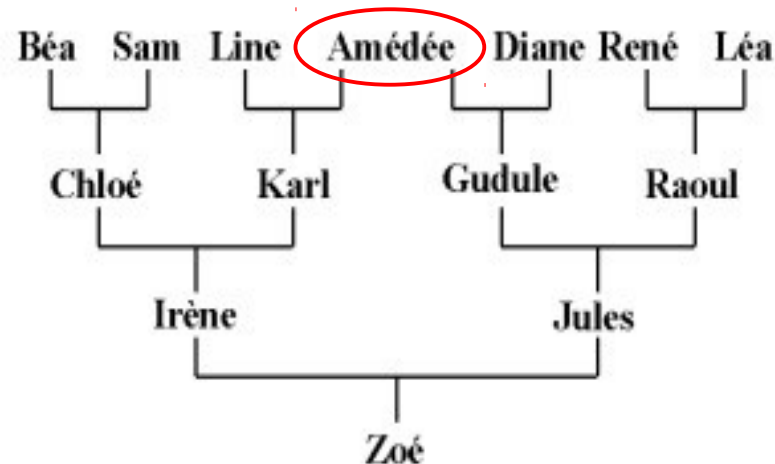
(3) Identifier pour chaque ancêtre toutes les chaînes de parenté reliant i à j en passant par lui même

(4) Pour chaque chaîne déterminer le nombre total de maillons entre l'ancêtre commun et i (n_1) et j (n_2), et calcules la probabilité

(5) Φ_{ij} s'obtient en sommant sur toutes les chaînes et tous les ancêtres, les probabilités calculées précédemment,

Calcul des coefficients de parenté et de consanguinité

Méthode des chemins

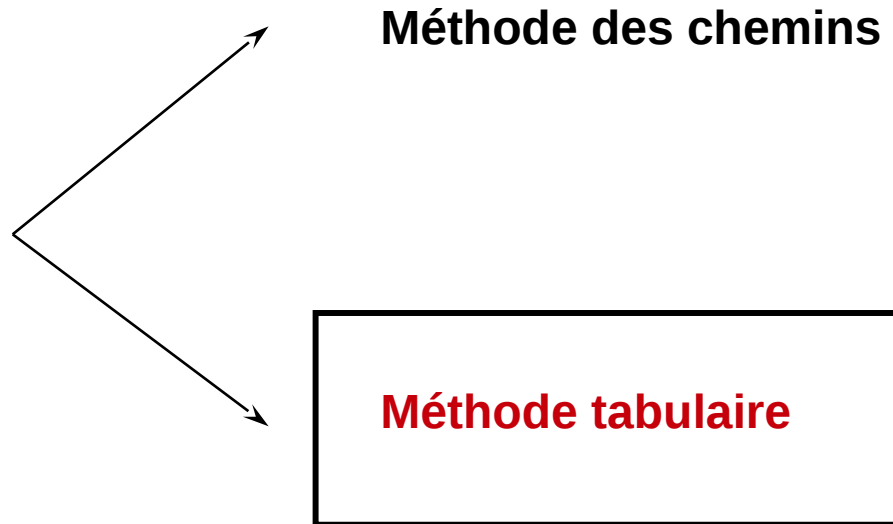


$$\phi_{Irène, Jules} = F_{Zoé} = \left(\frac{1}{2}\right)^5 (1 + F_{Amédée}) = \frac{1}{32} = 0,03125$$

Le coefficient de consanguinité de Zoé est 3,1%

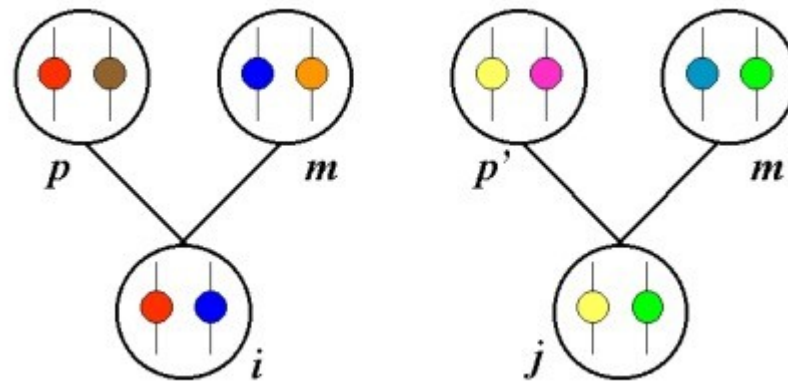
Écarts à la panmixie

(2) Calcul des coefficients de parenté et de consanguinité



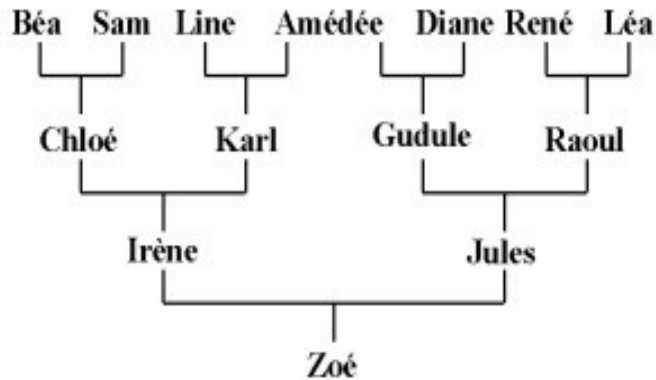
Calcul des coefficients de parenté et de consanguinité

Méthode tabulaire



$$\Phi_{ij} = \frac{1}{4} \Phi_{pp'} + \frac{1}{4} \Phi_{pm'} + \frac{1}{4} \Phi_{mp'} + \frac{1}{4} \Phi_{mm'}$$

Méthode tabulaire



Elle est bien adaptée à un traitement informatique des données (applications avec un grand nombre d'individus)

$$\Phi_{\text{Irène}, \text{Jules}} = \left(\frac{1}{4}\right) \Phi_{\text{Chloé}, \text{Gudule}} + \left(\frac{1}{4}\right) \Phi_{\text{Chloé}, \text{Raoul}} + \left(\frac{1}{4}\right) \Phi_{\text{Karl}, \text{Gudule}} + \left(\frac{1}{4}\right) \Phi_{\text{Karl}, \text{Raoul}}$$

$$\Phi_{\text{Irène}, \text{Jules}} = \left(\frac{1}{4}\right) \Phi_{\text{Karl}, \text{Gudule}} \quad \Phi_{\text{Chloé}, \text{Gudule}} = \Phi_{\text{Chloé}, \text{Raoul}} = \Phi_{\text{Karl}, \text{Raoul}} = 0$$

$$\Phi_{\text{Karl}, \text{Gudule}} = \left(\frac{1}{4}\right) \Phi_{\text{Line}, \text{Amédée}} + \left(\frac{1}{4}\right) \Phi_{\text{Line}, \text{Diane}} + \left(\frac{1}{4}\right) \Phi_{\text{Amédée}, \text{Amédée}} + \left(\frac{1}{4}\right) \Phi_{\text{Amédée}, \text{Diane}}$$

$$\Phi_{\text{Amédée}, \text{Amédée}} = \frac{1}{2} (1 + F_{\text{Amédée}}) = \frac{1}{2}$$

$$\Phi_{\text{Irène}, \text{Jules}} = \left(\frac{1}{4}\right) \Phi_{\text{Karl}, \text{Gudule}} = \frac{1}{4} \frac{1}{4} \Phi_{\text{Amédée}, \text{Amédée}} = \frac{1}{4} \frac{1}{4} \frac{1}{2} = \frac{1}{32}$$

Méthode tabulaire pour le calcul du coefficient de parenté

- 1) On trie les individus du plus vieux au plus jeune
- 2) On parcourt la généalogie.

Pour l'individu i ,

$$\Phi_{ii} = \frac{1}{2} + \frac{1}{2} \Phi_{\text{père}(i), \text{mère}(i)}$$

0 si inconnus

Pour les individus $j = (1 \text{ jusqu'à } i)$,

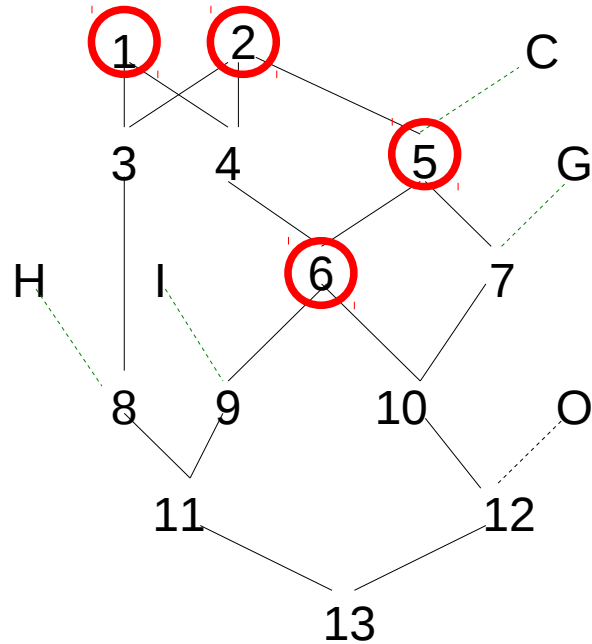
$$\Phi_{ij} = \frac{1}{2} (\Phi_{j, \text{père}(i)} + \Phi_{j, \text{mère}(i)})$$

0 si mère (père) inconnu

On remplit ainsi le tableau avec tous les Φ_{ij}

Méthode tabulaire pour le calcul de la relation additive

$(a_{ij} = 2 \Phi_{ij})$

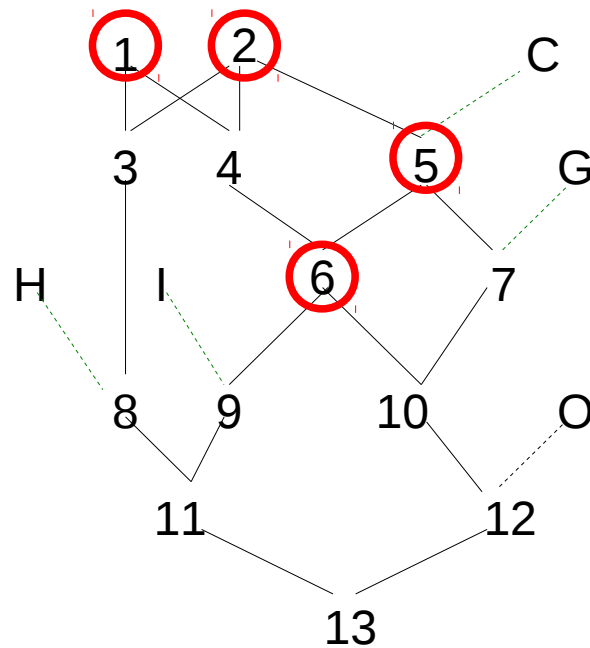


```
ped=matrix(c(
    1,0,0,
    2,0,0,
    3,1,2,
    4,1,2,
    5,2,0,
    6,4,5,
    7,5,0,
    8,3,0,
    9,6,0,
    10,6,7,
    11,8,9,
    12,10,0,
    13,11,12),
nrow=13,ncol=3,byrow=TRUE)
getA(ped)
```

Méthode tabulaire pour le calcul de la relation additive

$$(a_{ij} = 2 \Phi_{ij})$$

```
getA <- function(pedigree){  
  # A, tabular method  
  A=matrix(0,nrow(pedigree),nrow(pedigree))  
  f=rep(0,nrow(pedigree))  
  #read pedigree  
  for(i in 1: nrow(pedigree)){  
    A[i,i]=1  
    if(pedigree[i,2]!=0 && pedigree[i,3]!=0) A[i,i]=A[i,i]  
+0.5*A[pedigree[i,2],pedigree[i,3]]  
    f[i]=A[i,i]  
    for (j in 1:(i-1)){  
      if(pedigree[i,2]!=0) A[i,j]=A[i,j]  
+0.5*A[j,pedigree[i,2]]  
      if(pedigree[i,3]!=0) A[i,j]=A[i,j]  
+0.5*A[j,pedigree[i,3]]  
      A[j,i]=A[i,j]  
    }  
  }  
  list(A=A,f=f)  
}
```



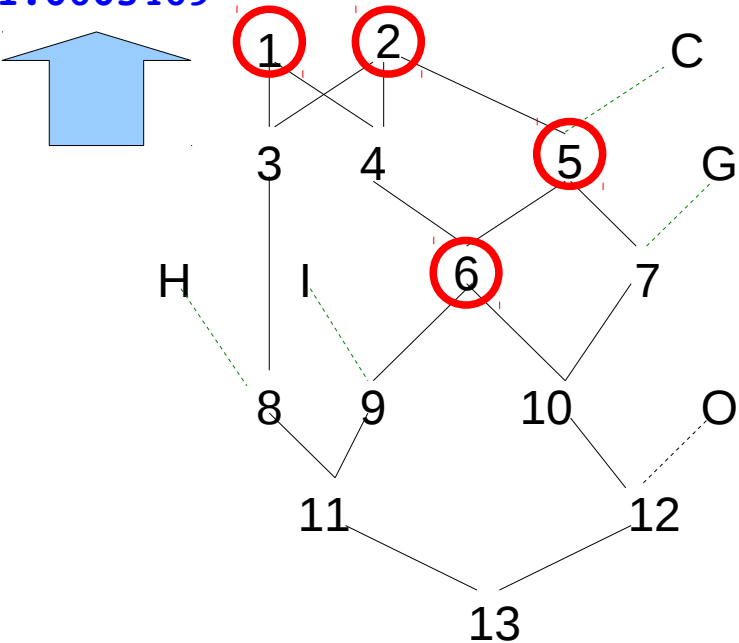
$$\text{Diag} = 1 + F_i$$

\$A

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]
[1,]	1.0000	0.00000	0.500000	0.500000	0.00000	0.2500000	0.000000	0.2500000
[2,]	0.0000	1.00000	0.500000	0.500000	0.50000	0.5000000	0.250000	0.2500000
[3,]	0.5000	0.50000	1.000000	0.500000	0.25000	0.3750000	0.125000	0.5000000
[4,]	0.5000	0.50000	0.500000	1.000000	0.25000	0.6250000	0.125000	0.2500000
[5,]	0.0000	0.50000	0.250000	0.250000	1.00000	0.6250000	0.500000	0.1250000
[6,]	0.2500	0.50000	0.375000	0.625000	0.62500	1.1250000	0.312500	0.1875000
[7,]	0.0000	0.25000	0.125000	0.125000	0.50000	0.3125000	1.000000	0.0625000
[8,]	0.2500	0.25000	0.500000	0.250000	0.12500	0.1875000	0.062500	1.0000000
[9,]	0.1250	0.25000	0.187500	0.312500	0.31250	0.5625000	0.156250	0.0937500
[10,]	0.1250	0.37500	0.250000	0.375000	0.56250	0.7187500	0.656250	0.1250000
[11,]	0.1875	0.25000	0.343750	0.281250	0.21875	0.3750000	0.109375	0.5468750
[12,]	0.0625	0.18750	0.125000	0.187500	0.28125	0.3593750	0.328125	0.0625000
[13,]	0.1250	0.21875	0.234375	0.234375	0.25000	0.3671875	0.218750	0.3046875

\$A

	[,9]	[,10]	[,11]	[,12]	[,13]
[1,]	0.1250000	0.1250000	0.1875000	0.0625000	0.1250000
[2,]	0.2500000	0.3750000	0.2500000	0.1875000	0.2187500
[3,]	0.1875000	0.2500000	0.3437500	0.1250000	0.2343750
[4,]	0.3125000	0.3750000	0.2812500	0.1875000	0.2343750
[5,]	0.3125000	0.5625000	0.2187500	0.2812500	0.2500000
[6,]	0.5625000	0.7187500	0.3750000	0.3593750	0.3671875
[7,]	0.1562500	0.6562500	0.1093750	0.3281250	0.2187500
[8,]	0.0937500	0.1250000	0.5468750	0.0625000	0.3046875
[9,]	1.0000000	0.3593750	0.5468750	0.1796875	0.3632812
[10,]	0.3593750	1.1562500	0.2421875	0.5781250	0.4101562
[11,]	0.5468750	0.2421875	1.0468750	0.1210938	0.5839844
[12,]	0.1796875	0.5781250	0.1210938	1.0000000	0.5605469
[13,]	0.3632812	0.4101562	0.5839844	0.5605469	1.0605469



$$\text{Diag} = 1 + F_i$$

A= Matrice de parenté