



**HAL**  
open science

## Bases de la génétique quantitative

Nathalie Mandonnet, Michel Naves

► **To cite this version:**

Nathalie Mandonnet, Michel Naves. Bases de la génétique quantitative. Master Ecotrop (diversité génétique), 2013. hal-02803758

**HAL Id: hal-02803758**

**<https://hal.inrae.fr/hal-02803758v1>**

Submitted on 5 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Introduction à la Génétique quantitative

**Nathalie Mandonnet & Michel Naves**  
**INRA-URZ**



*UAG-Master ECOTROP*  
*Oct-nov 2013*

ALIMENTATION  
AGRICULTURE  
ENVIRONNEMENT





# Introduction

- ***Génétique des populations (GP) et génétique quantitative (GQ):***
  - *Deux cadres d'analyses de la diversité génétique d'espèces naturelles ou domestiquées*
- ***Deux disciplines proches et complémentaires, se distinguant par leur objet:***
  - ***GP:***
    - *structure et évolution des pop naturelles,*
    - *observation de caractères qualitatifs associés à des gènes identifiables.*
  - ***GQ :***
    - *s'intéresse souvent à évolution de pops soumises à l'action humaine,*
    - *traite caractères à variation continue, déterminés par nombreux gènes, la plupart non identifiables individuellement.*



# Introduction

- ***Sur le fond les facteurs de l'évolution des populations naturelles et « artificielles » ne diffèrent guère***
  - *Fractionnement en sous-populations,*
  - *sélection/dérive,*
  - *mutations,*
  - *migrations,*
  - *croisements*



# Introduction

- *En agriculture, l'homme modèle les animaux et les plantes à ses besoins.*
- *Selon des **processus analogues** à ceux par lesquels les espèces **s'adaptent à leur milieu***
  - *Seule l'échelle de temps des processus diffère*
  - *La **domestication** des plantes et des animaux a été le **modèle** sur lequel Darwin a conçu sa **théorie de la sélection naturelle***



# Introduction

- *La génétique quantitative s'intéresse aux **caractères quantitatifs (à variation continue)***
- *C'est la base de la **sélection artificielle raisonnée** (amélioration des plantes et des races animales)*
- ***Les traits de vie et la fitness** sont des caractères quantitatifs.*
- *La génétique quantitative permet donc aussi de **comprendre la sélection naturelle***



# Introduction

***Dans ce cours, présentation des principes et concepts de base de la génétique quantitative***

- *Comment on analyse des caractères à déterminisme complexe en génétique humaine ou en génétique et sélection animales ou végétales ?*
- *Exercices proposés à la fin de chaque chapitre : se familiariser avec les concepts : analyser des situations concrètes (se munir d'une calculatrice)*

# Introduction

## ***Quels caractères soumis à analyse génétique?***

### **Caractères discrets : nombre de modalités fini et souvent faible**

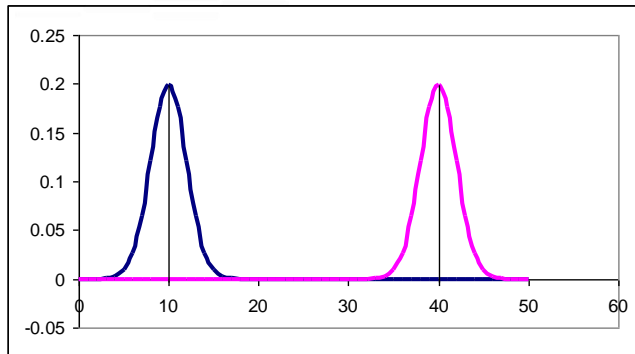
- *Ex : aspect lisse/ridé chez le pois, couleurs pétales de certaines fleurs, couleur pelage des bovins, groupes sanguins chez animaux ...*
- *Nature discontinue = relation simple entre phénotype et génotype*

### **Caractères continus : taille, poids corporel, quantité de lait produit, poids de 1000 grains, teneur en éléments nutritifs d'un fruit**

- *Interprétation génétique de la variation : plus complexe*
- *Souvent incapacité de distinguer et dénombrer gènes en ségrégation, responsables de la variation*
- *Association ambiguë entre un phénotype donné et un génotype particulier*



# Caractères qualitatifs

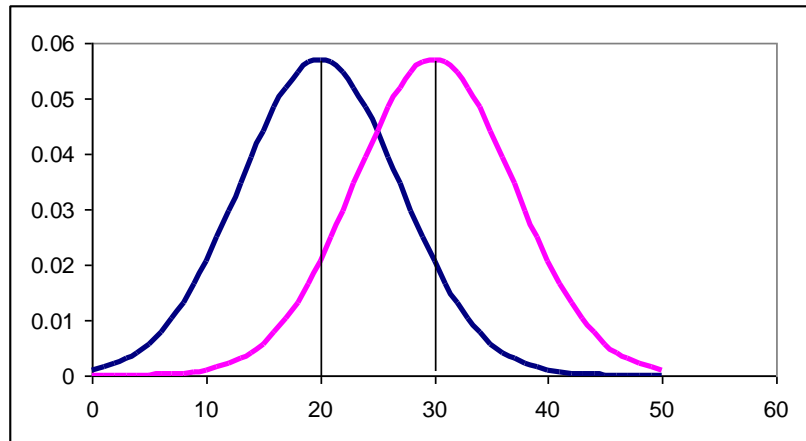


Graine		Fleur	Cosse		Tige	
Forme	Cotylédons	Couleur	Forme	Couleur	Emplacement	Taille
Gris & lisse	Jaune	Blanc	Plein	Jaune	Cosse axiale Fleur tout du long	Long (~3m)
Blanc & Ridé	Vert	Violet	Étroit	Vert	Cosse terminales Fleurs en haut	Court (~30 cm)
1	2	3	4	5	6	7

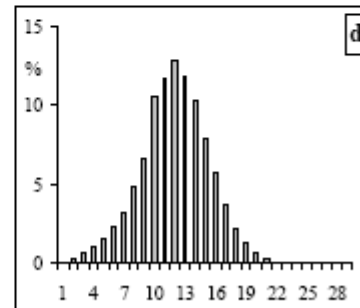
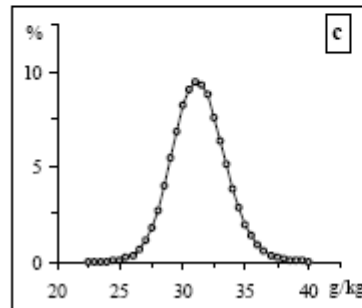
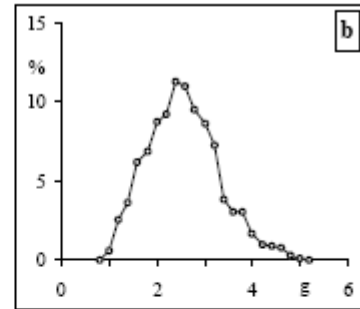
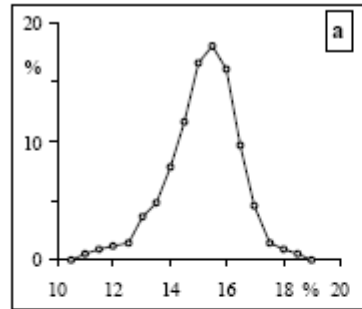
Génotypes distinguables par leur phénotype au locus

# Caractères quantitatifs

Génotypes non distinguables par l'observation de leur phénotype (car déterminé par loci +/- nombreux et dont expression influencée par environnement)



# Exemples de caractères quantitatifs



- a : Teneur en saccharose (%) dans la racine de betterave à sucre. Variété du début du XXème siècle. ( $N = 42\ 997$  ; Source : De Vries, 1909).
- b : Poids de grain (g) dans un épi de blé ( $N = 790$  ; Source : INA P-G, 2000).
- c : Teneur en protéines du lait de vache (g/kg), calculée sur l'ensemble de la lactation. Race Montbéliarde, contrôle de performances en ferme en 1988 ( $N = 251\ 705$  ; Source : FNOCL, 1989).
- d : Taille de portée (nombre de jeunes nés par portée) chez la truie. Race Large-White, contrôle de performances en ferme de 1990 à 2000. ( $N = 724\ 123$  ; Source : ITP-INRA, 2000).

# Introduction

***Controverse au début du XXe au sujet de l'hérédité des caractères quantitatifs :***

***Mutationnistes***  
***H. De Vries, W. Bateson***



***Biométriciens***  
***K. Pearson, W.F. Weldon***

***Rejet de l'application des lois de Mendel : continuité jugée incompatible avec ségrégation de facteurs Mendéliens discontinus***  
***Remise en cause de théorie de Darwin : caractères Quantitatifs ont subi une évolution graduelle sous action de la sélection naturelle***

***Soutien à la théorie de Darwin, Appréhension des caractères quantitatifs à travers l'analyse des liaisons statistiques entre phénotypes d'individus apparentés***



***Généticiens quantitatifs***



# Introduction

- **Objet de la GQ:**

Décrire la variabilité des caractères quantitatifs des plantes ou des animaux (Homme inclus) pour:

- donner des outils au sélectionneur
- mieux comprendre la biologie des organismes



# Introduction

## ***Applications essentielles de la GQ:***

- *Épidémiologie génétique chez l'homme :  
compréhension des maladies multifactorielles*
- *Amélioration des plantes et des animaux :  
compréhension de la domestication, élaboration  
de méthodes de sélection favorisant l'obtention de  
gains génétiques,*
- *Ecologie : compréhension des évolutions d'espèces  
naturelles sous des pressions anthropiques*

# Plan de cours

## 1. LA VARIABILITE AU SEIN DES POPULATIONS

- A. Description d'une population pour un caractère
- B. Description d'une population pour deux caractères
- Exercices

## II. HEREDITE ET MILIEU

- A. Définitions
- B. Mise en évidence des facteurs de variation génétiques
- C. Modélisation de l'action du génotype et du micro-milieu
- D. Action du macro-milieu et interaction génotype x milieu
- Exercices

## III. MODELISATION DE L'EFFET DES GENES SUR UN CARACTERE

- A. Le modèle génétique à un locus
- B. Généralisation à plusieurs locus du modèle génétique quantitatif
- C. Moyennes des variables  $G$ ,  $A$  et  $D$  au sein d'une population
- D. Intérêt prédictif de la notion de valeur génétique additive
- E. Décomposition de la variance génétique et paramètres génétiques
- Exercices


## IV. LA RESSEMBLANCE ENTRE APPARENTES

- A. Observations statistiques dans des populations
- B. Prédiction de la covariance entre apparentés
- C. Interprétation de la notion d'héritabilité et estimation
- D. Introduction à la notion de sélection
- Exercices

## V. EFFETS DES REGIMES DE REPRODUCTION

- A. Définition
- B. Effets de la consanguinité
- C. Hétérosis

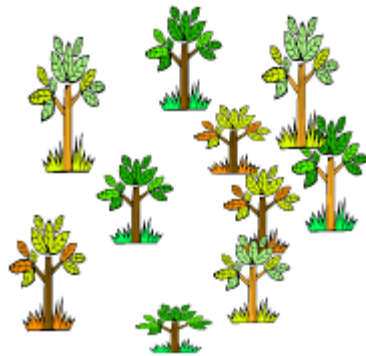
## VI. L'APPORT DES MARQUEURS MOLECULAIRES



# I. La variabilité de caractères quantitatifs au sein des populations

## A. Description d'une population pour un caractère

***Une population = ensemble plus ou moins hétérogène d'individus qui se reproduisent ensemble***



- *Descendants de croisements entre parents connus ou pas*
- *Populations artificielles : sélection récurrente*
- *Populations naturelles : sélection naturelle (ou sous la pression d'activités anthropiques)*

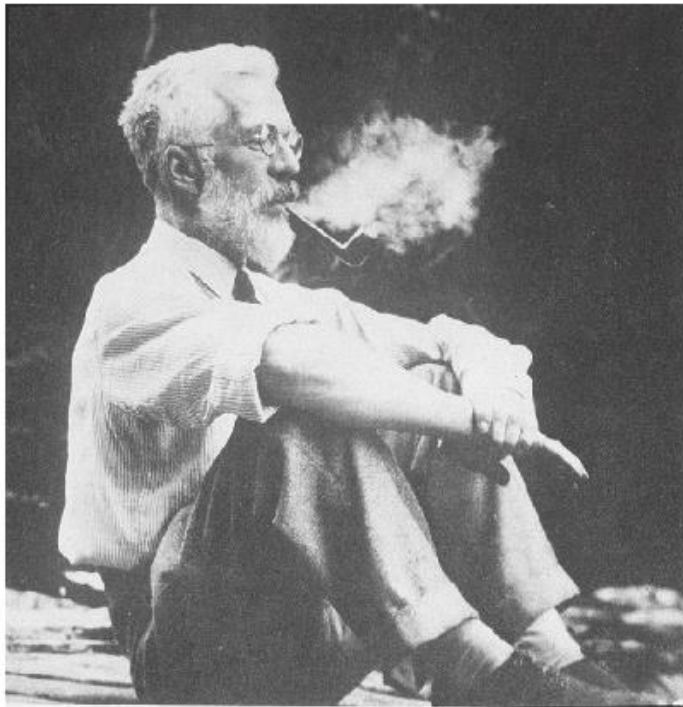
➤ *Phénotype = valeur d'un caractère mesurée chez un individu de la population*

➤ *Est déterminé par de nombreux gènes et par l'effet de l'environnement*



# Le déterminisme des caractères quantitatifs est Mendélien

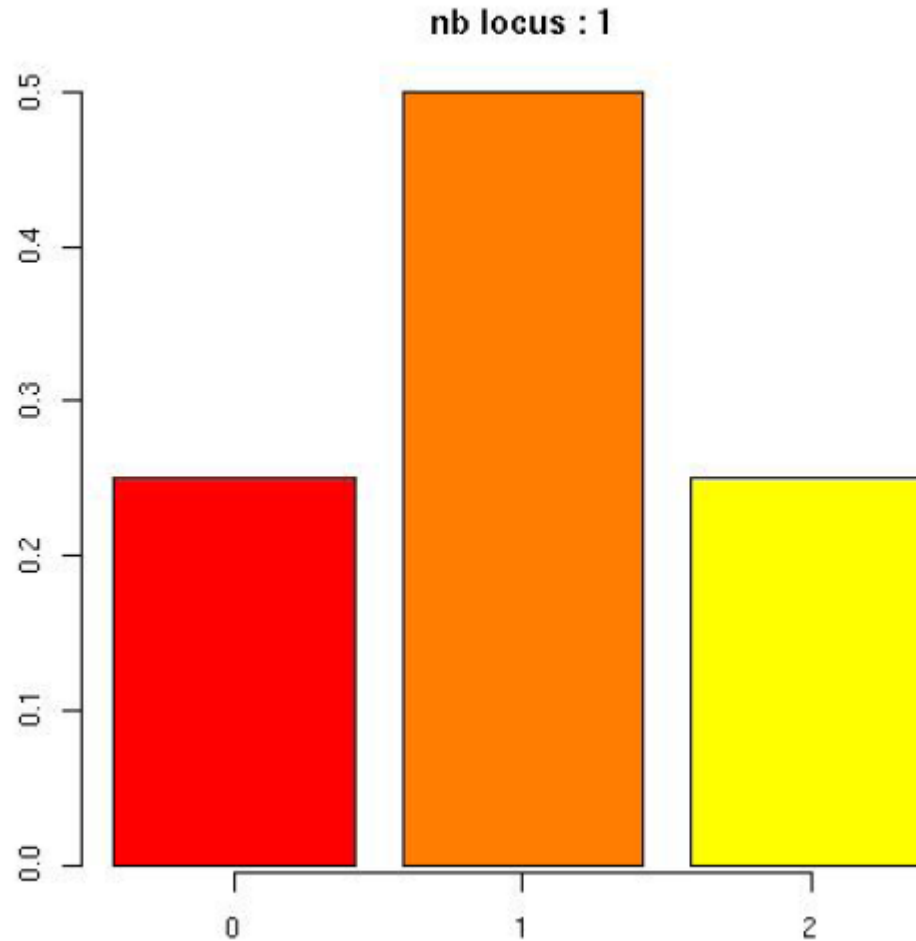
## Le modèle infinitésimal de Fisher (1918)



- Annoncé par Yule (1906), East (1910)
- Fisher, R. A. (1918). "The Correlation Between Relatives on the Supposition of Mendelian Inheritance." Philosophical Transactions of the Royal Society of Edinburgh **52**: 399-433.
- La ressemblance entre individus apparentés résulte du partage d'allèles en commun
- La variation observée pour les caractères quantitatifs résulte:
  - de l'action combinée d'un grand nombre de gènes (infini) à hérédité "mendélienne" individuelle
  - de l'influence de "l'environnement"

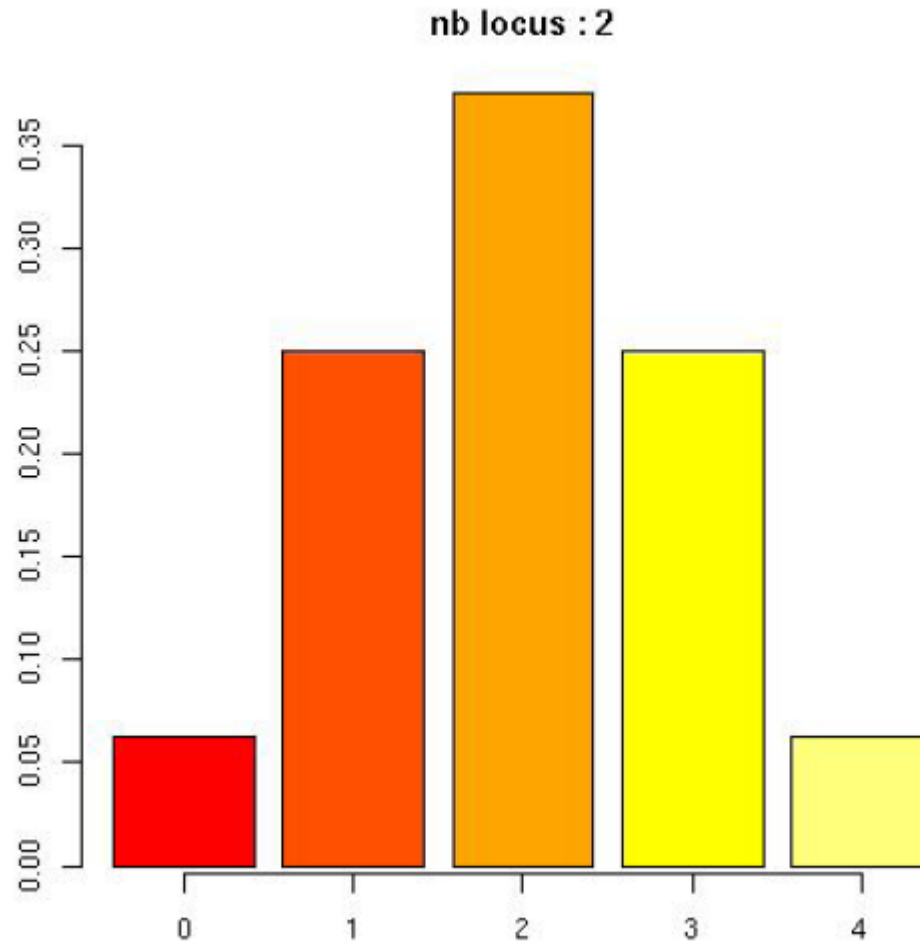
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



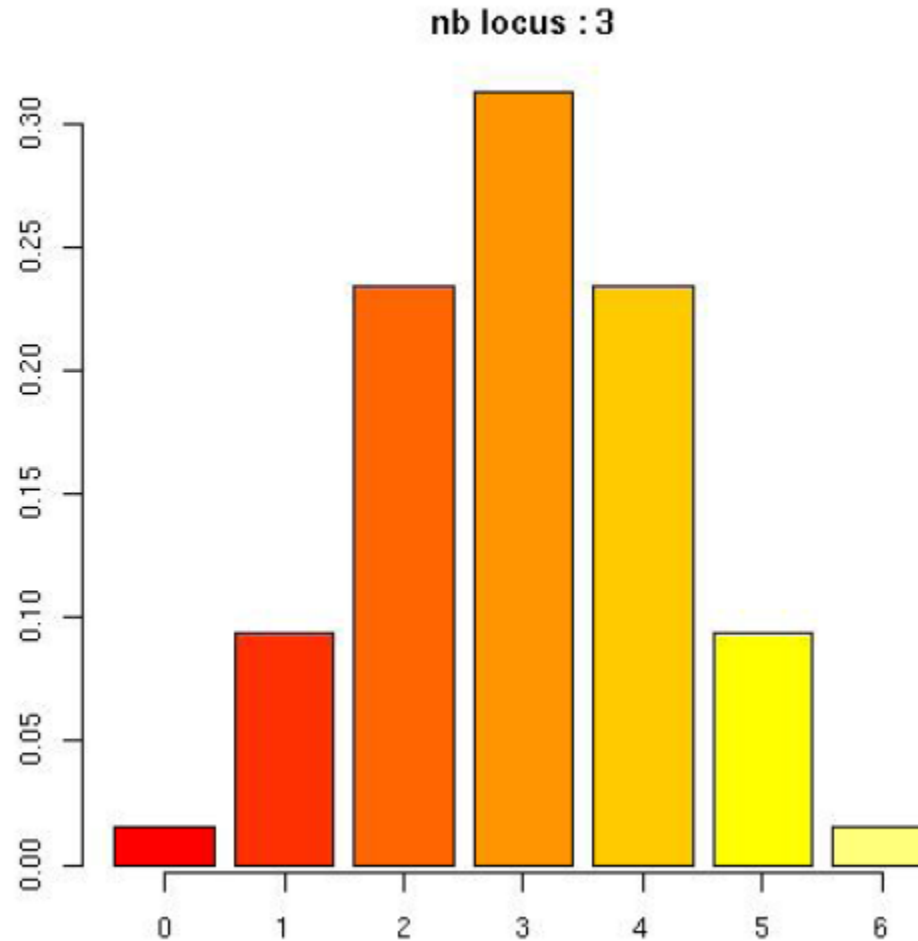
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



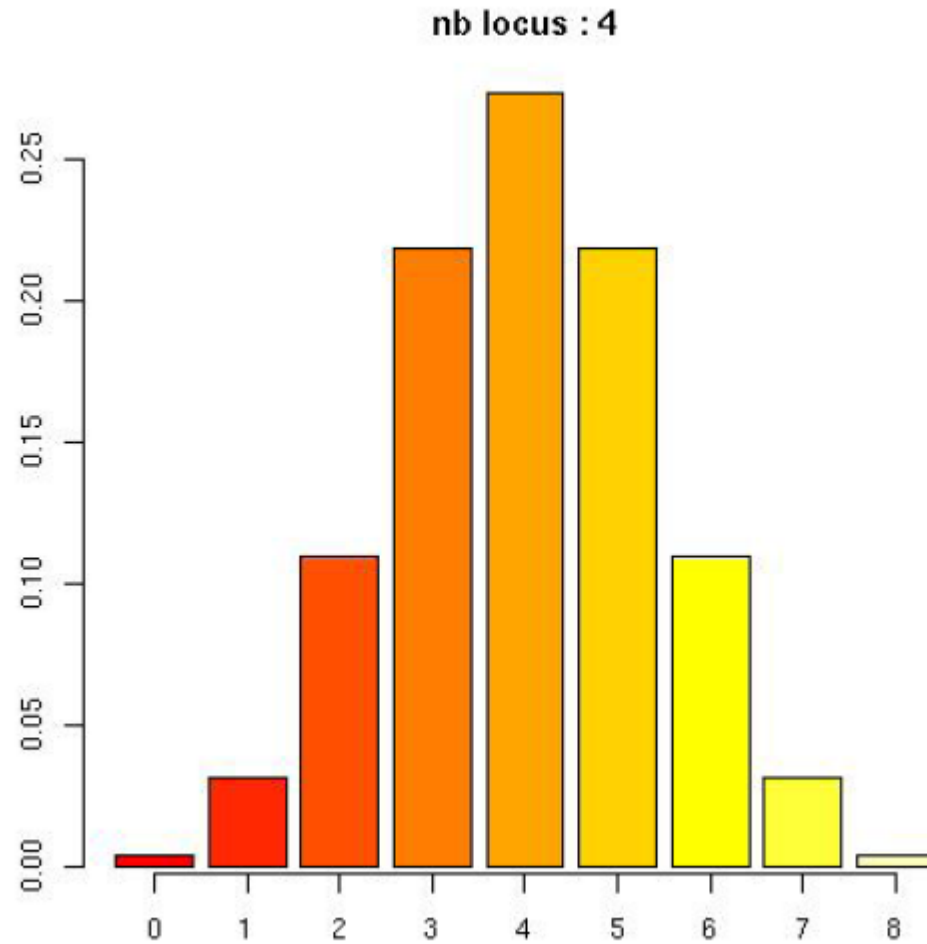
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



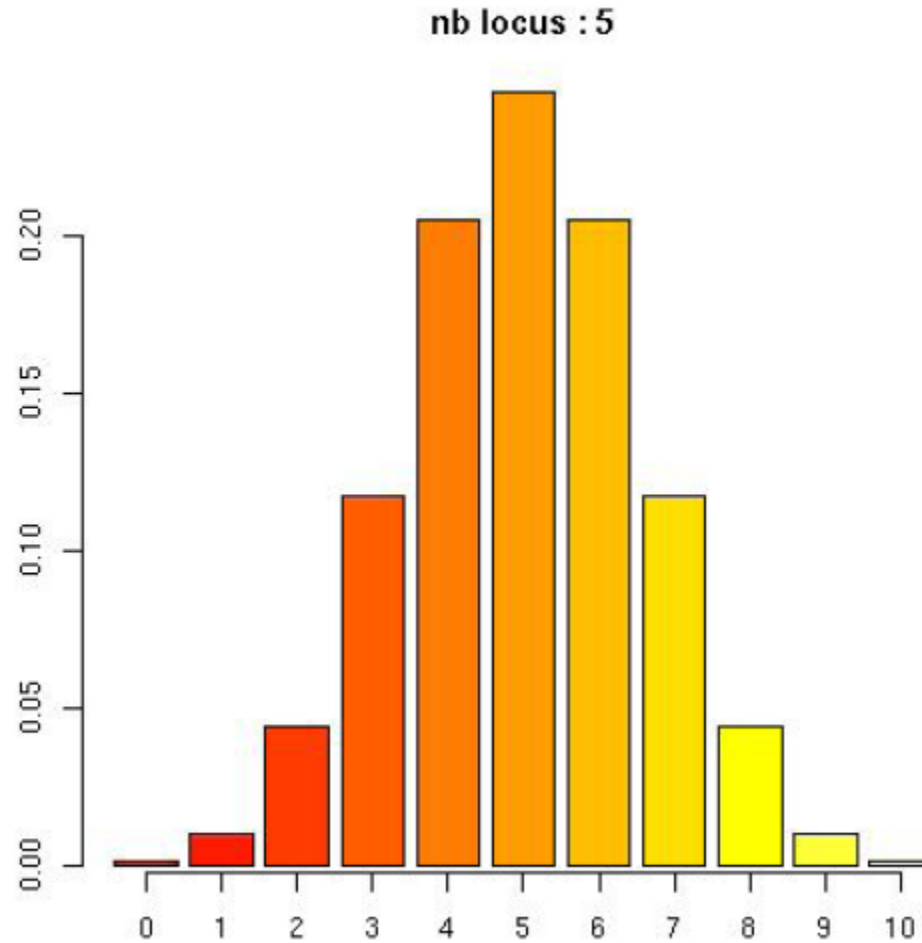
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



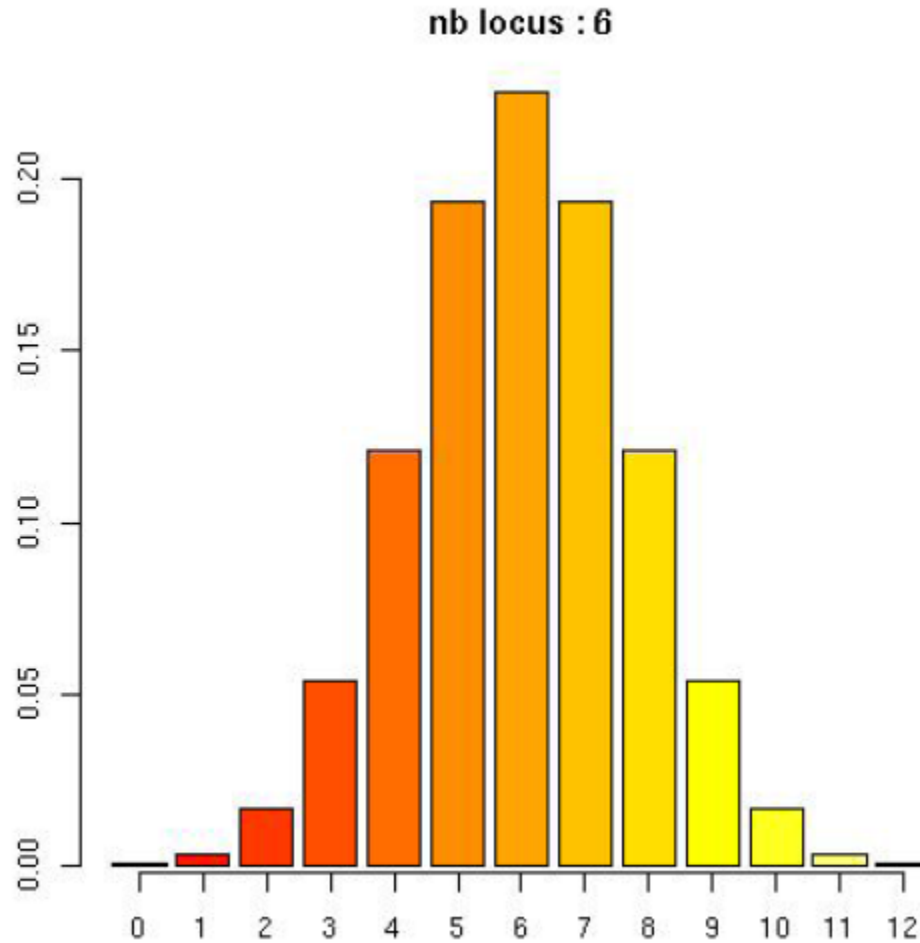
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



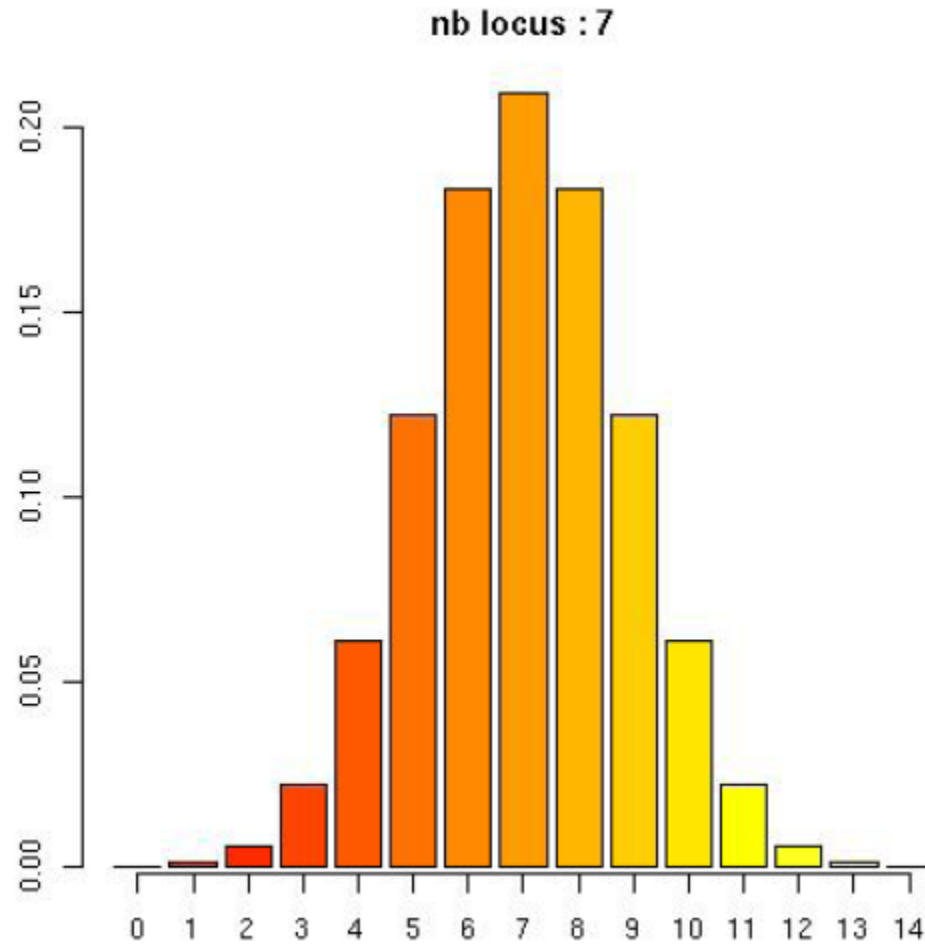
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



# Le modèle infinitésimal de Fisher (1918)

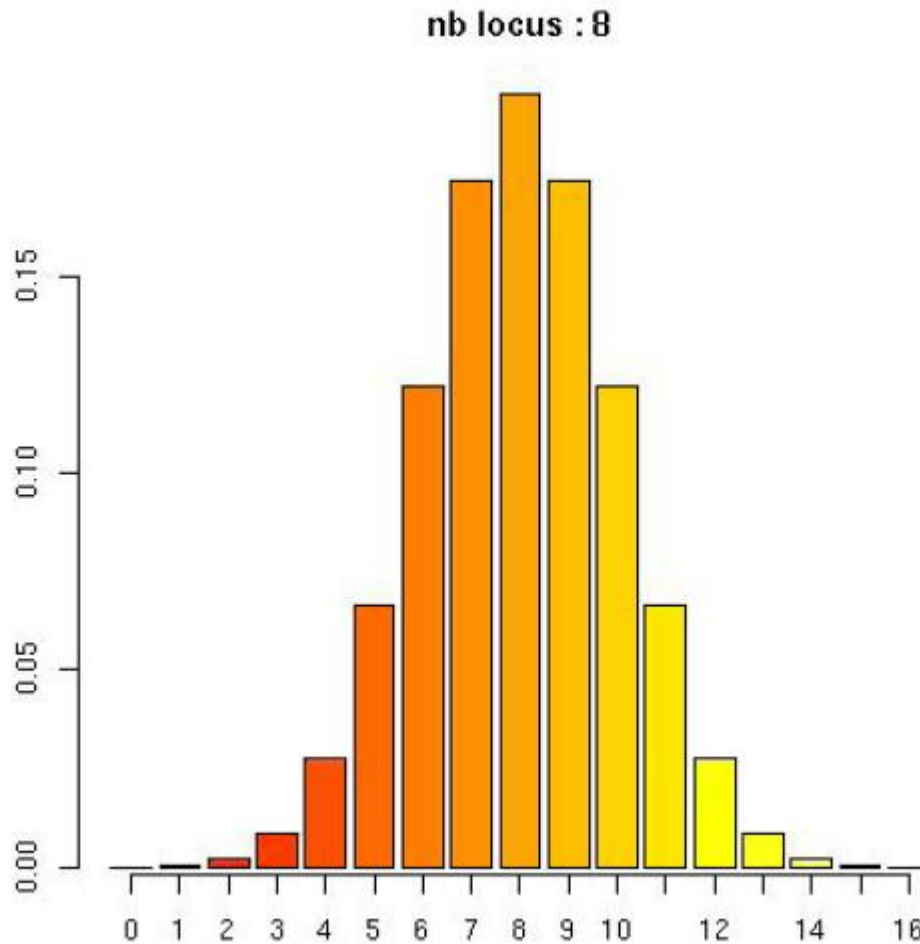
Locus bi-allélique, taille des effets alléliques identiques





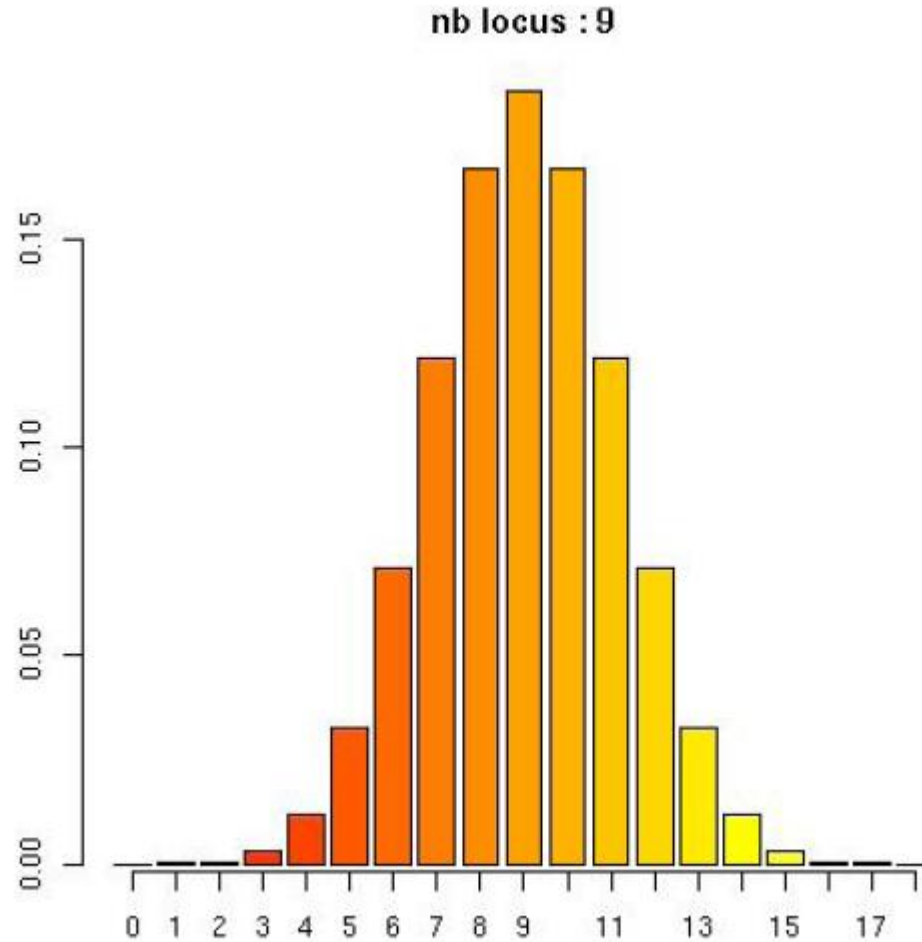
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



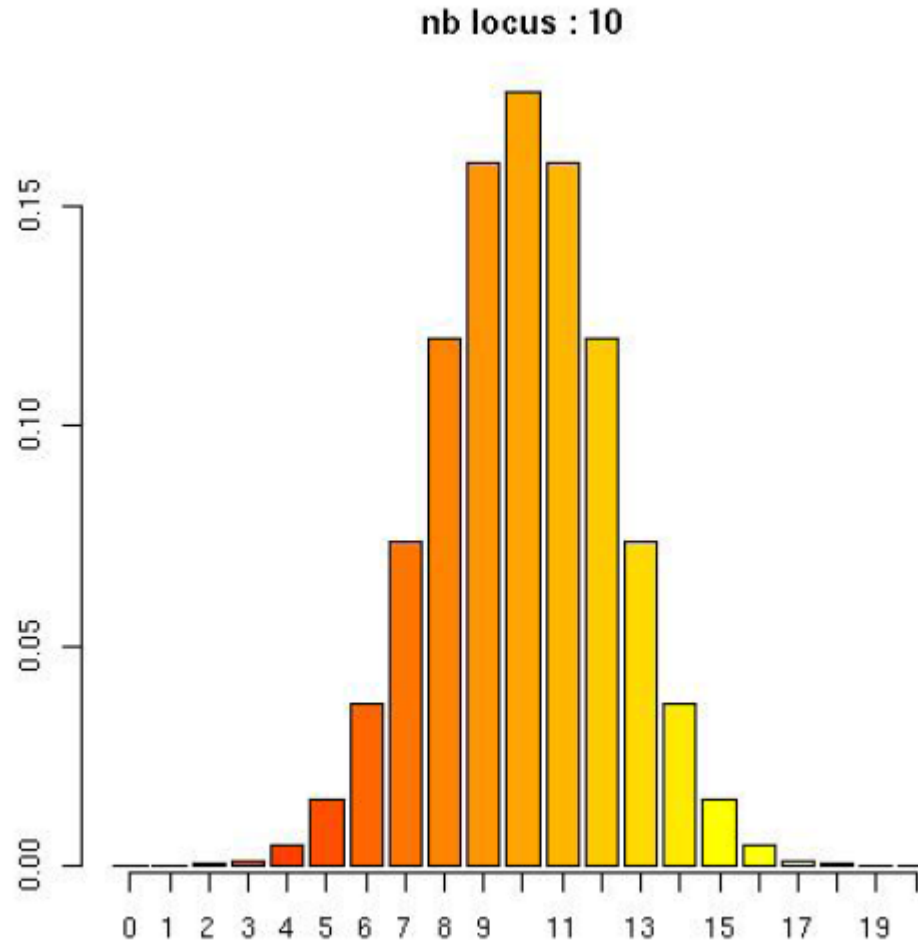
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



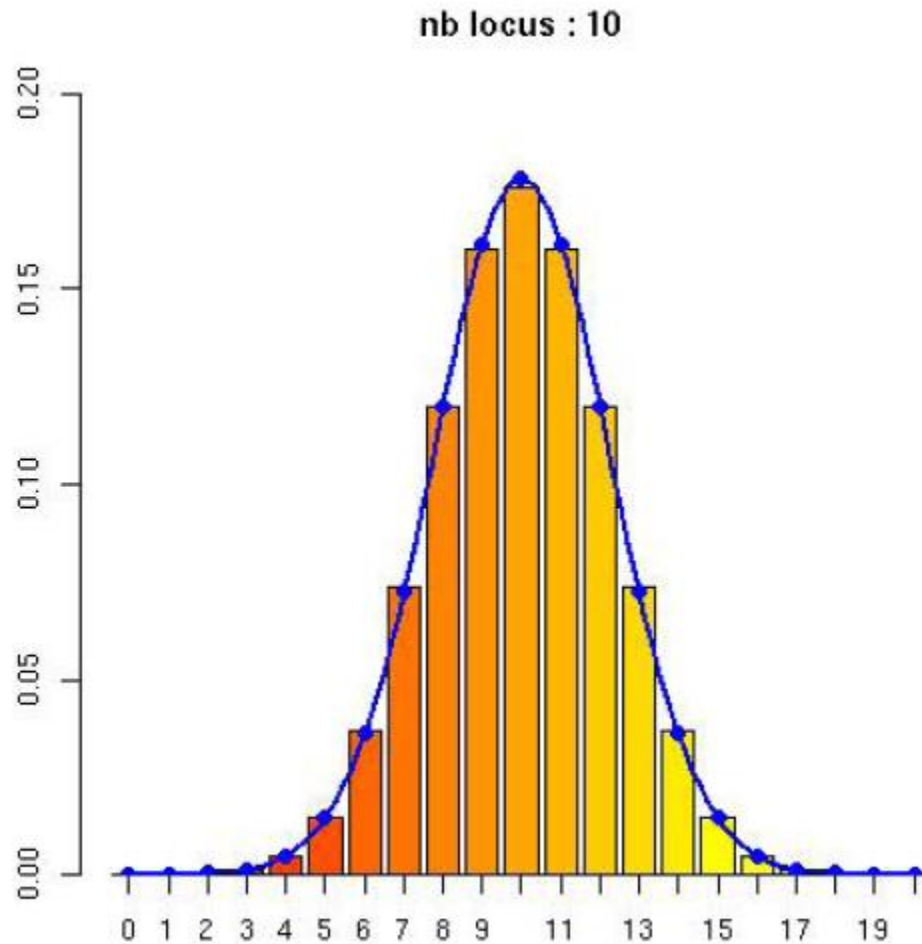
# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



# Le modèle infinitésimal de Fisher (1918)

Locus bi-allélique, taille des effets alléliques identiques



# Origine de la variation quantitative

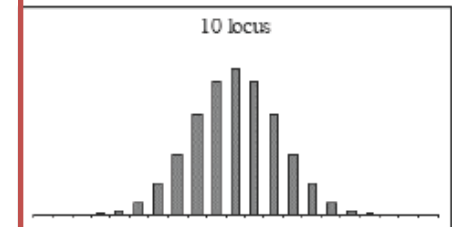
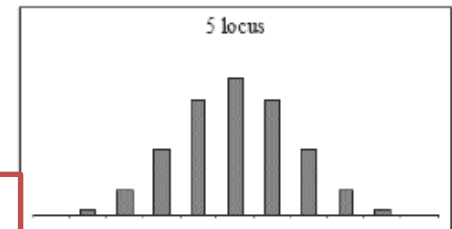
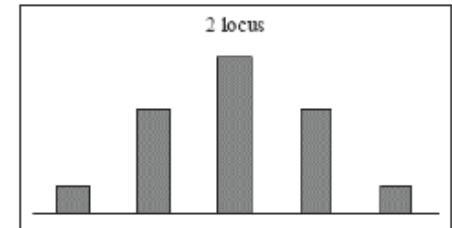
## Théorème central limite

Soient  $X_1, X_2, \dots, X_n$  des variables aléatoires indépendantes, pas forcément de même loi, d'espérance  $\mu$  et de variance  $\sigma_i^2$

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

Autrement dit, si une variable  $Y$  est la résultante d'un grand nombre de causes, petites, à effet additif, cette variable suit une loi normale. C'est à cause de cette interprétation que **la loi normale est très souvent employée comme référence, en particulier en génétique quantitative**



# L'espérance phénotypique

- L'espérance mathématique d'une variable aléatoire est sa moyenne **théorique**
- **Exemple :**
  - Soit une population  $\Omega$  d'effectif  $N$  et de moyenne  $\mu$  pour une caractéristique  $x$  mesurée sur les individus
  - Pour tout individu on a  $E(x) = \mu$
  - Pour tout échantillon de taille  $n$  tiré au hasard dans la population on a :

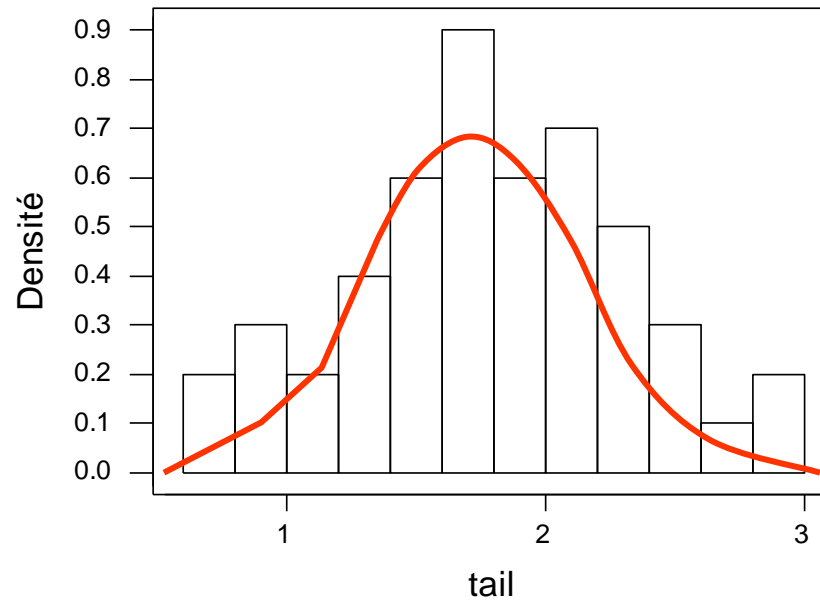
$$E(\bar{x}) = \mu$$

# La variabilité phénotypique

- Nous avons à définir la **variabilité phénotypique**.
- On la mesurera par la **variance** du trait considéré ou par **son écart-type**, qui est égal à la racine carrée de la variance et s'exprime en unité du caractère (dispersion moyenne autour du caractère = champ de variation du caractère considéré)

# Variance phénotypique

- Exemple : considérons la taille en mm d'un échantillon de femelles de l'hyménoptère parasitoïde *Leptomastyx dactylopi*





# Mesure de la variation phénotypique

- Par la variance totale:
  - **Empirique** : échantillon de la population


$$V_P = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

- - de la population

$$V_P = E(x_i - \mu)^2$$

$$V_P = E(x_i)^2 - \mu^2$$

$$V_P = E(x_i)^2 - E^2(x_i)$$

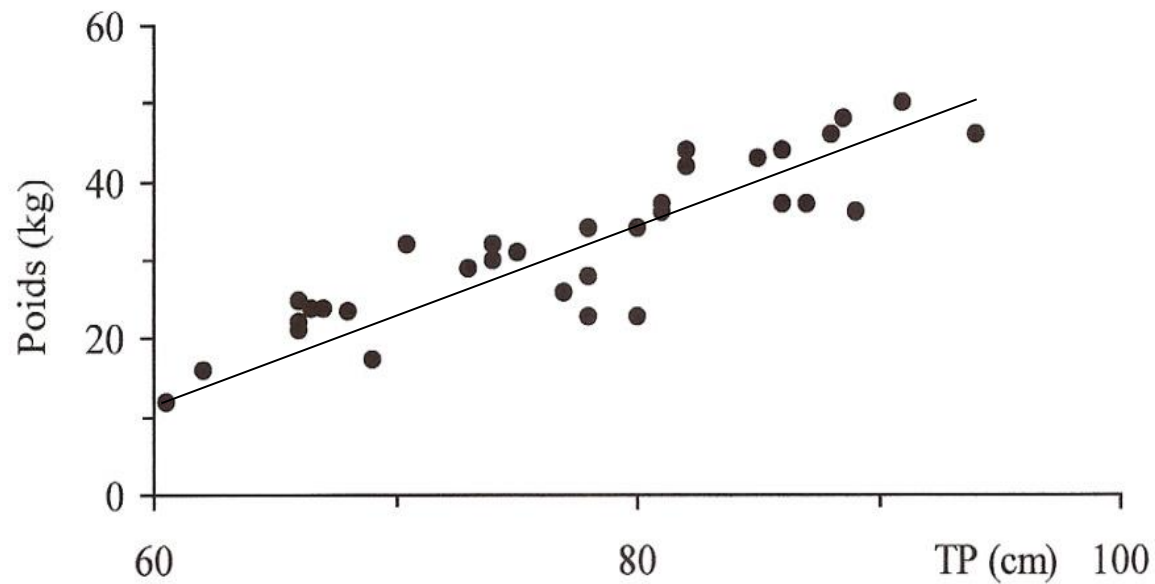


# I. La variabilité de caractères quantitatifs au sein des populations

B. Description d'une population pour deux  
caractères

# Distribution à deux dimensions

Relation entre le tour de poitrine (TP) et le poids dans un échantillon de 35 mâles de bouquetin des Alpes (*Capra ibex ibex*). Source: Toïgo, 1998



# Corrélation et covariance

Coefficient de **corrélation** (de Pearson) = quantifie association entre deux variables X et Y

= rapport de la covariance entre les deux variables X et Y au produit de leur écart-types

$$\frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Covariance =  $S_{XY}$  = espérance du produit des 2 variables centrées sur leur moyenne

Corrélation = sans dimension = varie de -1 à +1  
= degré de liaison linéaire entre 2 variables

$$\frac{s_{XY}}{s_X s_Y} = \frac{[(\sum XY) - (\sum X)(\sum Y)/n]}{\sqrt{[(\sum X^2) - ((\sum X)^2)/n][(\sum Y^2) - ((\sum Y)^2)/n]}}$$



## Corrélation et covariance

**corrélation** (de Pearson) positive = valeurs élevées de  $Y$  préférentiellement associées à des valeurs élevées de  $X$

**corrélation** (de Pearson) négative = valeurs élevées de  $Y$  préférentiellement associées à des valeurs faibles de  $X$

Corrélation élevée en valeur absolue = association rigoureuse  
= nuage de points resserré autour de la droite

A contrario : valeur proche de zéro = liaison linéaire imparfaite

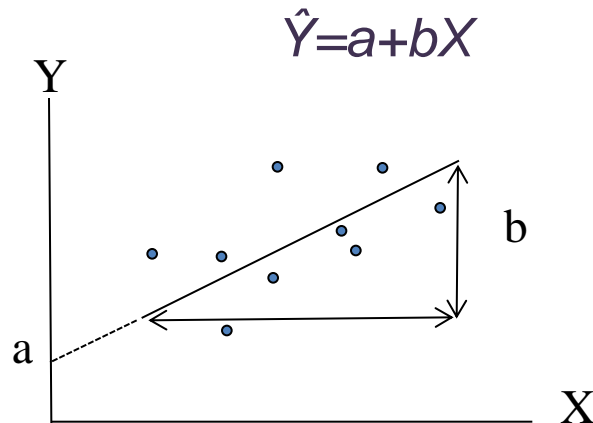
Corrélation nulle = indépendance linéaire de  $X$  et  $Y$

# Régression linéaire

Existence corrélation linéaire entre X et Y

= prédiction de la valeur d'une variable à partir de l'autre variable

En effet, connaissant X, il est possible de prédire Y par  $\hat{Y}$ , selon l'équation suivante



Coefficient de régression =  $b$  = pente droite = quantité moyenne de variation de Y quand X croit d'une unité

Coefficients a et b de la droite de régression = déterminés en minimisant la variance d'erreur entre valeurs prédites par équation ( $\hat{Y}$ ) et valeur réelles de Y (*méthode des moindres carrés*)



# Régression linéaire

Plus la corrélation est élevée entre les variables  $X$  et  $Y$ , plus la prédiction de l'une par rapport à l'autre est précise (plus  $Y$  est proche de  $\hat{Y}$ )

Avantage de la technique de régression linéaire = prédire valeur d'un individu pour un caractère alors que sa mesure est difficile

Ex : Bouquetin des Alpes : prédire poids des animaux en mesurant leur croissance afin de permettre suivi des populations

Régression linéaire = abondamment utilisée en génétique quantitative

= prédire valeur génétique d'un caractère donné, non observable directement à partir de phénotypes que l'on mesure sur une population

# Mémento de calculs statistiques

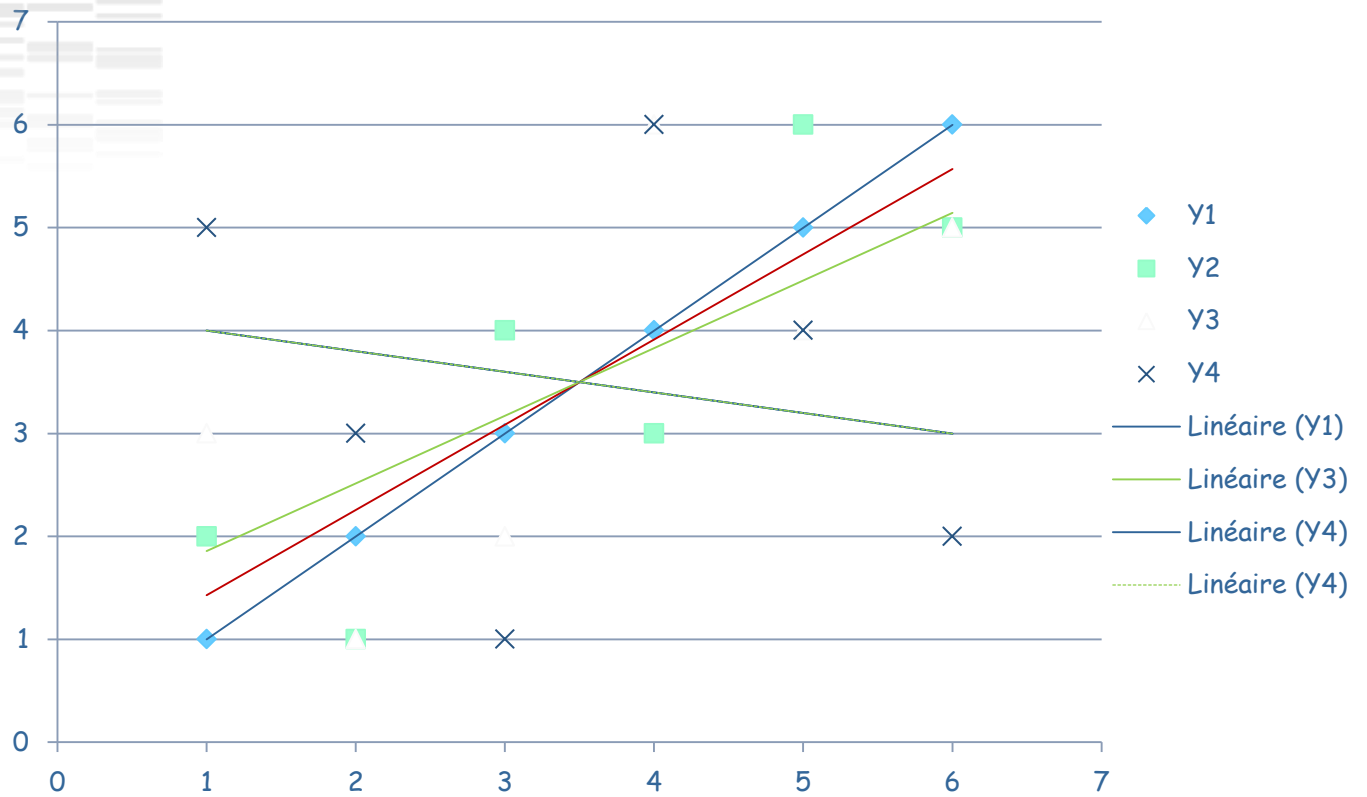
Paramètre	Population de taille infinie		Estimateur sur un échantillon de taille $n$		Champ de variation
	Symbole	Définition	Symbole	Formule de calcul	
Moyenne	$\mu, \mu_X$	$E(X)$	$\bar{X}, \hat{\mu}$	$\frac{1}{n} \sum X$	$-\infty, +\infty$
Variance	$\sigma^2, \sigma_X^2$ $V_X, \text{Var}(X)$	$E([X - E(X)]^2)$ $= E(X^2) - [E(X)]^2$	$s_X^2, \hat{\sigma}_X^2$ $\hat{V}_X$	$\frac{1}{n-1} [(\sum X^2) - (\sum X)^2/n]$ $= \frac{1}{n-1} [(\sum X^2) - n(\bar{X})^2]$	$0, +\infty$
Covariance	$\sigma_{XY}$ $\text{Cov}(X, Y)$	$E([X - E(X)][Y - E(Y)])$ $= E(XY) - E(X)E(Y)$	$s_{XY}$	$\frac{1}{n-1} [(\sum XY) - (\sum X)(\sum Y)/n]$	$-\infty, +\infty$
Corrélation	$\rho_{XY}$ $r(X, Y)$	$\frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$	$r(X, Y)$ $\hat{r}(X, Y)$	$\frac{s_{XY}}{s_X s_Y} = \frac{[(\sum XY) - (\sum X)(\sum Y)/n]}{\sqrt{[(\sum X^2) - (\sum X)^2/n][(\sum Y^2) - (\sum Y)^2/n]}}$	$-1, +1$
Régression	$\beta_{Y/X}$ $b_{Y/X}$	$\frac{\text{Cov}(X, Y)}{\sigma_X^2}$	$b_{Y/X}$ $\hat{b}_{Y/X}$	$\frac{s_{XY}}{s_X^2} = \frac{[(\sum XY) - (\sum X)(\sum Y)/n]}{[(\sum X^2) - (\sum X)^2/n]}$	$-\infty, +\infty$





# Exercices

- N°1:
  - Variable Poids: 32,3 kg +/- 9,86 kg
  - Variable TP: 77,4 cm +/- 8,79 cm
  - Coef. corrélation: 0,88
  - Coef. de régression de Y sur X: 0,98
- N°2:



							moyenne	écart-type	covariance(X,Y)	coef.regY/X
<b>X</b>	1	2	3	4	5	6	3,5	1,71		
<b>Y1</b>	1	2	3	4	5	6	3,5	1,71	2,92	<b>1,00</b>
<b>Y2</b>	2	1	4	3	6	5	3,5	1,71	2,42	<b>0,83</b>
<b>Y3</b>	3	1	2	6	4	5	3,5	1,71	1,92	<b>0,66</b>
<b>Y4</b>	5	3	1	6	4	2	3,5	1,71	-0,58	<b>-0,20</b>



# II. Hérité et milieu

## A. définitions



# Modèle de la GQ

$$P = G + E$$



# Le phénotype P

**Le phénotype** = manière dont un organisme nous apparaît, pour un niveau d'observation donné. En génétique quantitative, les caractères que l'on étudie font l'objet d'une mesure.

**Valeur phénotypique** = mesure effectuée sur un individu.

Dans le milieu professionnel de la sélection, animale ou végétale, on parle également de **performance**.

En génétique végétale, l'entité sur laquelle on réalise une mesure peut être un mélange d'individus, par exemple l'ensemble des individus cultivés sur une même parcelle.

Dans ce cas, la valeur phénotypique = une moyenne de valeurs phénotypiques individuelles

# Le génotype G

**Le génotype** d'un individu en un locus = ensemble des allèles qu'il possède à ce locus (deux allèles homologues pour un individu diploïde).

Le **génotype en plusieurs locus** = ensemble des génotypes à chacun des locus.

Si l'on considère l'ensemble du génome et l'ensemble des caractéristiques phénotypiques des organismes, aucun individu ne peut être strictement identique à un autre, hormis le cas d'organismes obtenus par multiplication végétative ou le cas des vrais jumeaux.

En pratique, les termes "génotype" et "phénotype" sont employés dans un sens restreint, c'est-à-dire pour les seuls caractères qui nous intéressent et les gènes qui les influencent.

*Sauf indication contraire, les termes «génotype» et «phénotype» sont employés par rapport à un ensemble défini de caractères.*

# Le milieu E

## *Ensemble des causes non génétiques de variation*

- (1) l'environnement dans lequel vit (ou a vécu) l'individu observé,
- (2) certains états physiologiques qui lui sont propres
- (3) l'observateur lui-même.

(1) Production végétale : l'année (influence du climat), la parcelle (influence des conditions topographiques et de sol), les doses d'engrais appliquées aux différents stades du développement de la plante, les traitements phytosanitaires effectués, les conditions de récolte, etc...

En production animale : l'année (influence du climat sur les ressources fourragères), l'élevage (influence des potentialités agricoles de l'exploitation, du mode de conduite des animaux et de la technicité de l'éleveur), la saison de mise-bas en cas de production laitière ou de croissance des jeunes, etc...

(2) âge concerne les cultures permanentes et les animaux : rendement en fruits d'un cep de vigne ou d'un arbre fruitier varie en fonction de son âge; cycle et stade de végétation des plantes fourragères, stade de lactation des femelles laitières, etc ...

(3) influence de l'observateur dans l'application du protocole de mesure, précision de ses instruments de mesure, erreurs de mesure etc ...

# Facteurs de milieux contrôlés et non contrôlés

## *Facteurs de milieu contrôlés*

**Facteurs identifiés et mesurés** = dont on sait qu'ils ont une influence sur le phénotype = facteurs de **macro-milieu**

= propres à plusieurs individus élevés ensemble

Ex : fumure azotée d'une culture, valeur agronomique d'une parcelle, rang de lactation d'un animal,

= s'appliquent en commun à tous les individus se trouvant dans une même catégorie de milieu

## *Facteurs de milieu non contrôlés*

**Que l'on ne maîtrise pas** = échappent à l'observateur

Ex : hétérogénéité intra-parcellaire

= induisent chacun des faibles variations, résultant de micro phénomènes locaux

= s'applique de manière différente à chaque individu

= facteurs de **micro-milieu**





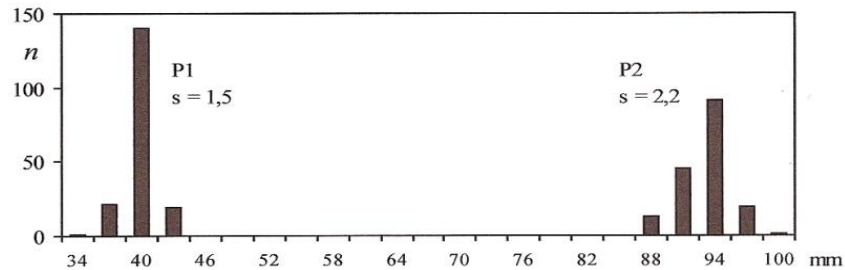
## II. Hérité et milieu

### B. Mise en évidence des facteurs de variation génétique

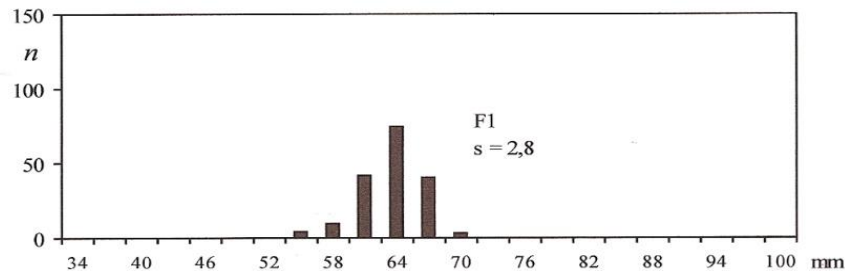
# L'hypothèse multi-factorielle

## Expérience de East 1916

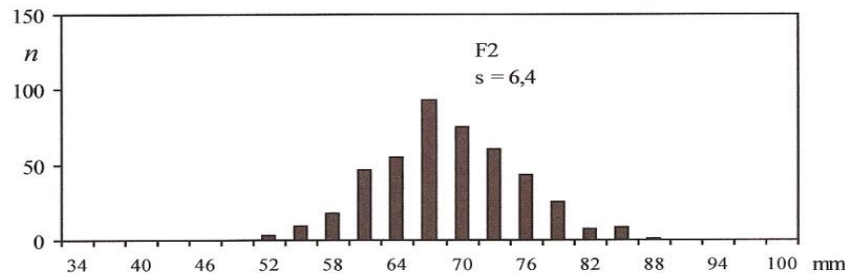
Croisement entre 2 lignées consanguines de tabac (P1 et P2)  
différents pour la longueur de leurs fleurs



Autofécondation à l'intérieur des lignées P1 et P2 : individus génétiquement identiques.



De même : individus de la première génération (F1) = génétiquement identiques  
**Variance observée = d'ordre environnementale**



L'accroissement substantiel de variance en F2 (F1xF1) évoque une ségrégation polygénique.  
Si longueur fleur déterminée par 1 locus, la F2 serait un mélange distributions P1, F1 et P2 dans proportions mendéliennes (1/4, 1/2, 1/4)

# L'hypothèse multi-factorielle

## Expérience de East 1916

Cette expérience fut poursuivie par l'analyse de la réponse à la sélection en F3 (F2 x F2).

L'ensemble de ces résultats a permis de suggérer que les lignées parentales différaient non pas à un locus, mais pour plusieurs ayant une influence sur le caractère mesuré.

Distributions globales observées s'accordent bien avec l'hypothèse d'une distribution normale en plusieurs locus indépendants les uns des autres.

***Les variations observées des caractères quantitatifs sont imputables***

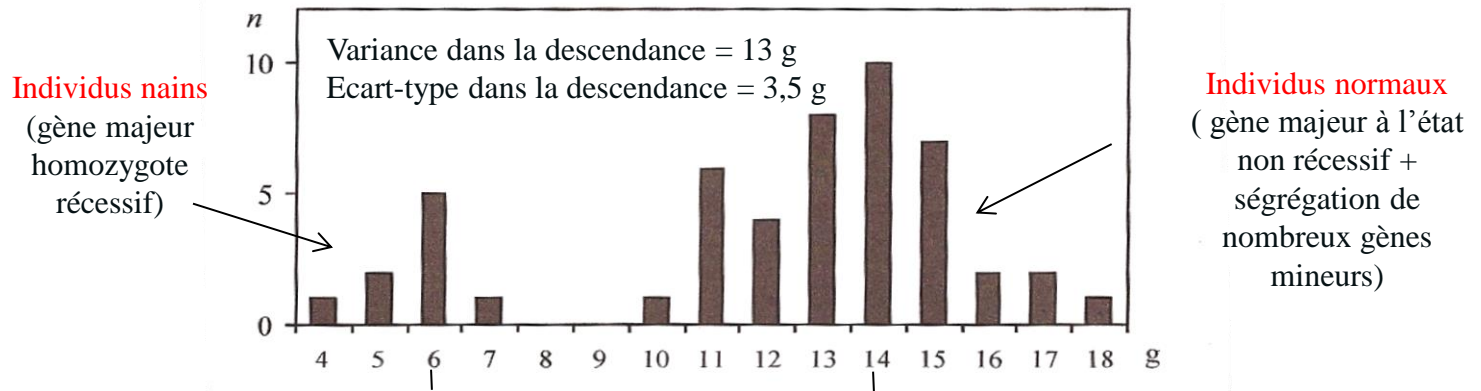
**(1) aux ségrégations à plusieurs locus  
et (2) à des facteurs de milieu.**

# Gènes majeurs et gènes mineurs (polygènes)

Gène majeur dans une population = gène qui détermine une différence phénotypique importante entre individus portant différents allèles = gène induisant des différences grandes au regard de la variation phénotypique totale

Exemple : King (1950) *Pygmy, a dwarfing gene in the house mouse*.  
J. Hered. 41: 249–252.

Poids à 60 jours (en g) de 50 souris mâles issues d'une lignée sélectionnée pour un faible poids et contenant un gène majeur de nanisme récessif



Différence moyenne entre normaux – nains = 8 g soit environ 2,2 unités d'écart-types

Equivalait à un gène qui induirait chez l'homme un écart d'une quinzaine de cm pour la taille

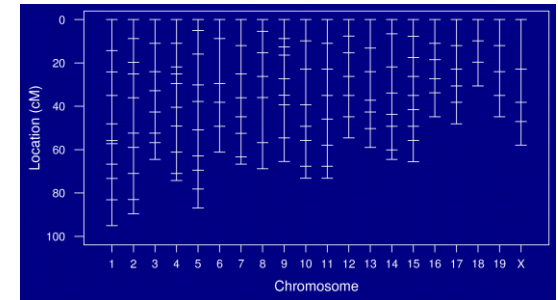
**Gène majeur** = provoque différences phénotypiques de grande ampleur

**Polygènes** = responsables de la variabilité autour de la moyenne (avec env.)

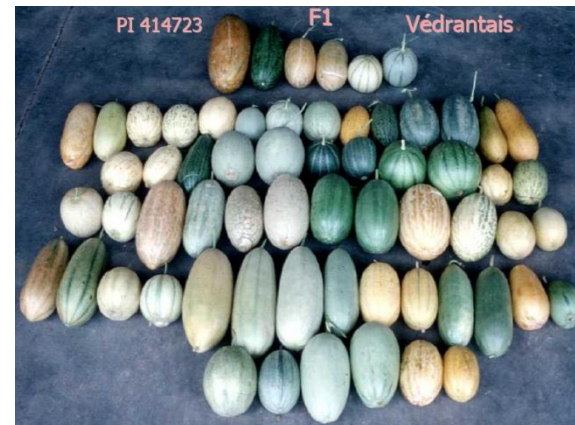
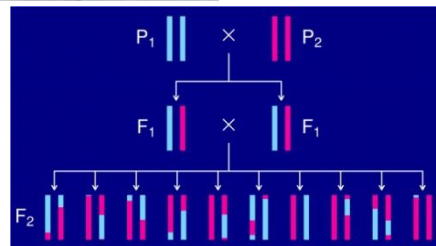
# Gènes majeurs et mineurs

Nombreux gènes majeurs et mineurs contrôlant divers caractères mis en évidence dans de très nombreux travaux de cartographie génétique de plantes cultivées ou d'animaux d'élevage

Cartographie génétique = balisage du génome avec des marqueurs moléculaires neutres (RFLP, AFLP, SSR, DArT, SNP, etc ...) ou ciblés par l'étude de populations en ségrégation



Localisation de gènes impliqués dans des variations de caractères quantitatifs en étudiant des populations en ségrégation issues de croisements entre génotypes contrastés phénotypiquement.



# Les gènes majeurs et mineurs sont appelés *des QTL = Quantitative Trait Loci*

**QTL** : locus impliqué dans la variation d'un caractère quantitatif

Un QTL qui ségrège (dont les allèles sont répartis plus ou moins aléatoirement) dans une population explique une part de la variation du caractère étudié dans la population (si petite = gène mineur, si grande = gène majeur)

Tout locus au niveau duquel une substitution allélique modifie la valeur du caractère est un QTL qui contribue à l'expression du caractère.

**Un QTL est par définition polymorphe !**

[si le locus étudié est monomorphe dans la population (1 seul allèle et tous les individus homozygotes), il ne peut pas influencer la variation phénotypique de la population]



# La détection de QTLs

Nécessité d'observer une population en ségrégation d'individus et d'étudier le polymorphisme de l'ADN de chaque individu avec des marqueurs

**But :** comprendre la nature de la variation quantitative ?

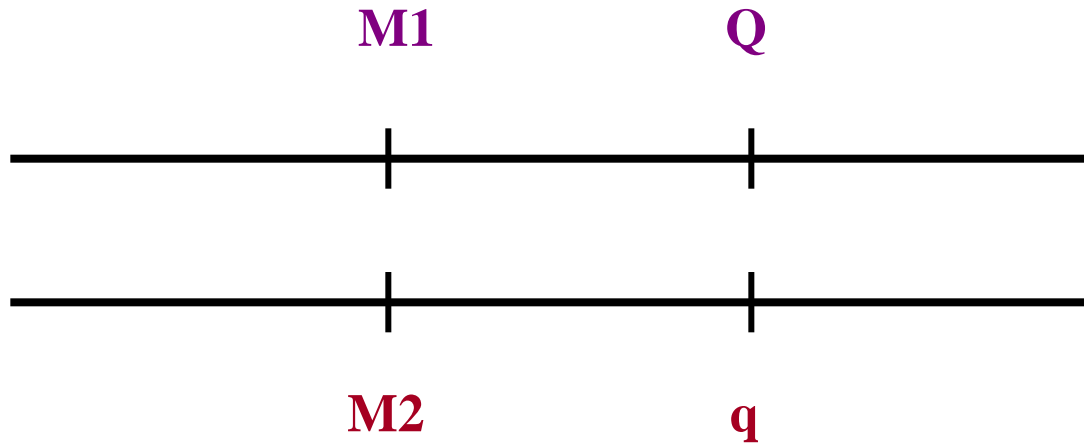
Inventaire des gènes en cause (QTL), de leurs allèles, de leurs effets

- combien de gènes décelables gouvernent le caractère étudié ?
- localisation des gènes sur quels chromosomes ? dans quelles régions ?
- contribution de chaque gène à la variation du caractère dans la population ?

**Méthode :** recherche d'associations statistiques entre la ségrégation de chaque marqueur et celle du phénotype dans la population

Si association significative entre un marqueur et le caractère :  
existence d'un QTL à proximité du marqueur

# La détection de QTLs

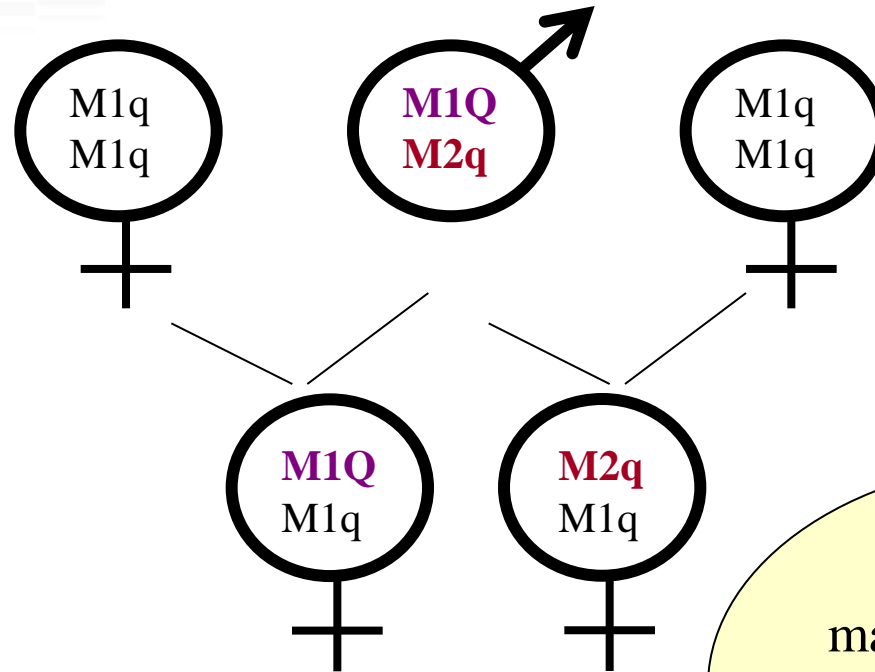


**Marqueur génétique** = Fragment d'ADN, correspondant à un loci, pour lequel il existe plusieurs formes ou allèles

**Quantitative trait loci** = Zone chromosomique influençant significativement le caractère

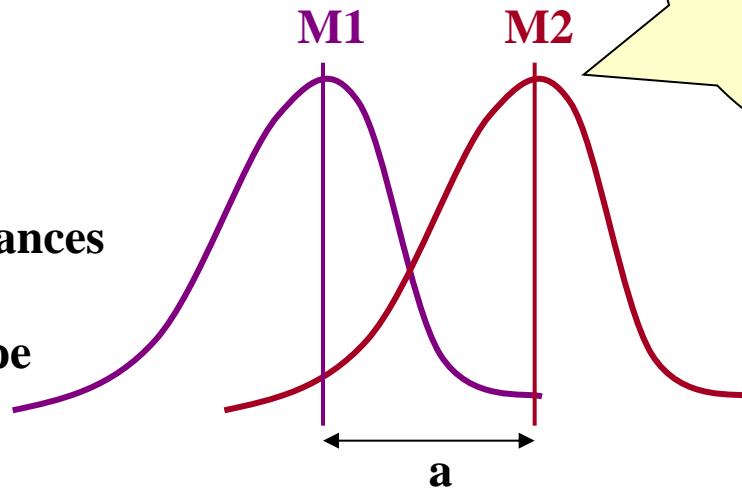


# La détection de QTLs



Le génotype au marqueur apporte de l'information sur la valeur génétique des individus

Distribution des performances des descendants en fonction du génotype Marqueur du père



# Une littérature pléthorique d'études QTL

Ex: revue de Rieseberg H et al. (2002) . *Directional selection is the primary cause of phenotypic diversification*. PNAS vol. 99 no. 19 : 12242–12245

Kingdom	Species	Cross	Mating system	Trait	Trait type	Total QTLs
Animal	<i>Acyrtosiphon pisum pisum</i>	Intra	Outcross	Alfalfa race - fecundity on clover	Life history	4
Animal	<i>Acyrtosiphon pisum pisum</i>	Intra	Outcross	Clover race - acceptance of clover	Life history	1
Animal	<i>Caenorhabditis elegans</i>	Intra	Self	Body length	Morphology	2
Animal	<i>Caenorhabditis elegans</i>	Intra	Self	Fertility	Life history	3
Animal	<i>Caenorhabditis elegans</i>	Intra	Self	Mean lifespan	Life history	1
Animal	<i>Caenorhabditis elegans</i>	Intra	Self	Total fertility	Life history	2
Animal	<i>Caenorhabditis elegans</i>	Intra	Self	Early fertility fraction	Life history	1
Animal	<i>Caenorhabditis elegans</i>	Intra	Self	Population growth	Life history	4
Animal	<i>Caenorhabditis elegans</i>	Intra	Self	Time of onset of egg laying	Life history	3
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Reproductive success	Life history	2
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Ovariole number	Life history	2
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Cuticular hydrocarbons	Physiology	1
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Sex comb tooth number	Morphology	2
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Lifespan male	Life history	5
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Lifespan female	Life history	5
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Male lifespan	Life history	23
Animal	<i>Drosophila melanogaster</i>	Intra	Outcross	Female lifespan	Life history	23
Animal	<i>Mus musculus</i>	Intra	Outcross	Body weight	Morphology	11
Animal	<i>Mus musculus</i>	Intra	Outcross	Aggressive behavior	Behavior	2
Animal	<i>Oncorhynchus mykiss</i>	Intra	Outcross	Time to hatch	Life history	3
Animal	<i>Oncorhynchus mykiss</i>	Intra	Outcross	Embryonic length	Morphology	2
Animal	<i>Oncorhynchus mykiss</i>	Intra	Outcross	Weight at swim-up	Morphology	2
Animal	<i>Apis mellifera</i>	Intra	Outcross	Alarm pheromones	Life history	9
Animal	<i>Apis mellifera</i>	Intra	Outcross	Foraging behavior	Behavior	3
Animal	<i>Apis mellifera</i>	Intra	Outcross	Drone wing length	Morphology	5
Animal	<i>Apis mellifera</i>	Intra	Outcross	Worker wing length	Morphology	4
Animal	<i>Sus scrofa</i>	Intra	Outcross	Birth weight	Life history	3
Animal	<i>Sus scrofa</i>	Intra	Outcross	Growth rate to 30 kg	Life history	2
Animal	<i>Sus scrofa</i>	Intra	Outcross	Length small intestine	Morphology	9
Animal	<i>Sus scrofa</i>	Intra	Outcross	Abdominal fat	Morphology	7
Animal	<i>Sus scrofa</i>	Intra	Outcross	Average depth back fat	Morphology	7
Animal	<i>Sus scrofa</i>	Intra	Outcross	Femur measurements PC1	Morphology	2
Animal	<i>Sus scrofa</i>	Intra	Outcross	Osteochondrosis	Life history	2

# Existence d'une littérature pléthorique d'études QTL

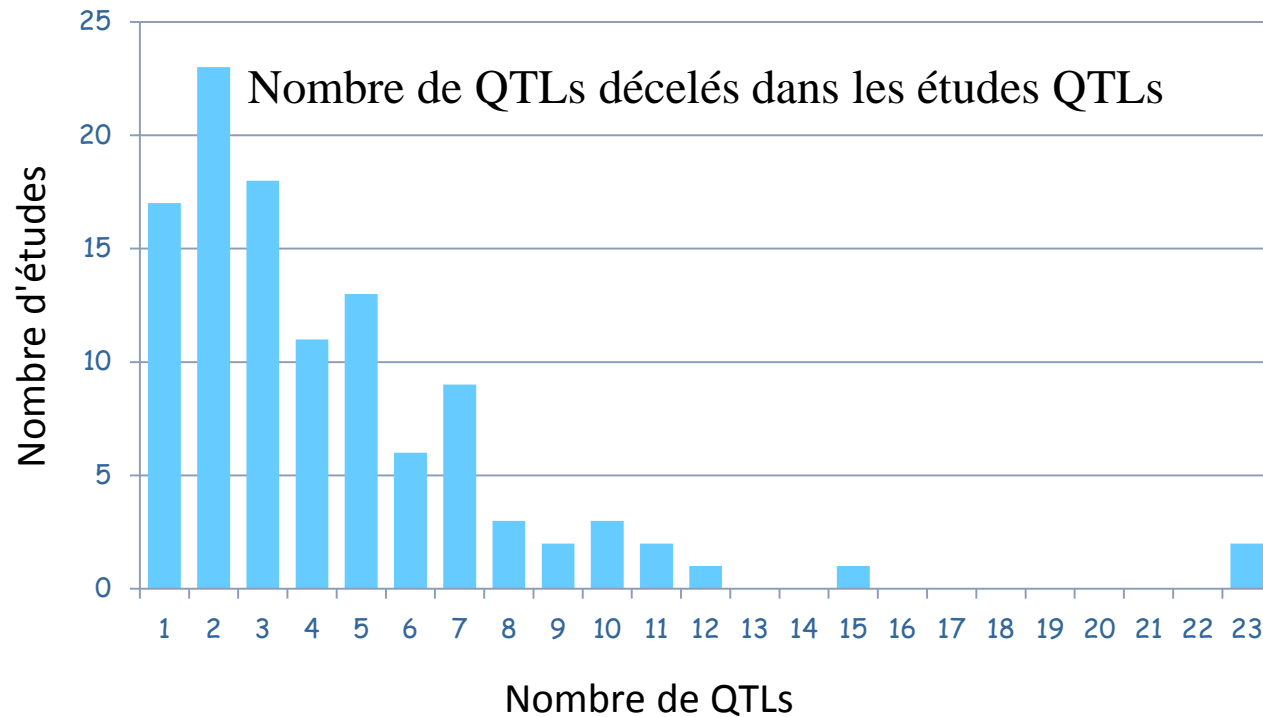
Kingdom	Species	Cross	Mating system	Trait	Trait type	Total QTLs
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Hypocotyl length: white light	Morphology	4
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Total leaf number	Morphology	4
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Largest leaf length	Morphology	7
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Plant height	Morphology	4
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Side shoot number	Morphology	4
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Flowering time	Life history	5
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Days to first flower	Life history	2
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	No. of nodes	Morphology	1
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Fruit no. in main infleur	Life history	3
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Seed weight	Morphology	10
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Seed length	Morphology	8
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Fruit length	Morphology	6
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Ovary length	Morphology	4
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Ovule no./fruit	Life history	4
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Petal length	Morphology	10
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Petal width	Morphology	2
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Sepal length	Morphology	5
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Sepal width	Morphology	3
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Short stamen length	Morphology	7
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Long stamen length	Morphology	4
Plant	<i>Arabidopsis thaliana</i>	Intra	Self	Pistil length	Morphology	5
Plant	<i>Mimulus cardinalis</i> x <i>M. lewisii</i>	Inter	Outcross	Anthocyanin conc.	Physiology	3
Plant	<i>Mimulus cardinalis</i> x <i>M. lewisii</i>	Inter	Outcross	Corolla width	Morphology	8
Plant	<i>Mimulus cardinalis</i> x <i>M. lewisii</i>	Inter	Outcross	Nectar volume	Physiology	3
Plant	<i>Mimulus cardinalis</i> x <i>M. lewisii</i>	Inter	Outcross	Stamen length	Morphology	7
Plant	<i>Mimulus cardinalis</i> x <i>M. lewisii</i>	Inter	Outcross	Pistil length	Morphology	7
Plant	<i>Pinus silvestris</i>	Intra	Outcross	Frost hardiness	Physiology	7
Plant	<i>Senecio vulgaris</i>	Intra	Outcross x	Days to bud formation	Life history	1
Plant	<i>Senecio vulgaris</i>	Intra	Outcross x	Plant height	Morphology	2
Plant	<i>Senecio vulgaris</i>	Intra	Outcross x	Leaf number	Morphology	2
Plant	<i>Senecio vulgaris</i>	Intra	Outcross x	Lateral branch number	Morphology	1
Plant	<i>Senecio vulgaris</i>	Intra	Outcross x	Outer involucral bracts	Morphology	1
Plant	<i>Senecio vulgaris</i>	Intra	Outcross x	Disc florets/capitulum	Morphology	1
Plant	<i>Zea mays</i> x <i>Z. mays</i>	Inter	Outcross	No. of tassel branches	Morphology	7
Plant	<i>Zea mays</i> x <i>Z. mays</i>	Inter	Outcross	Length of central spike	Morphology	2
Plant	<i>Zea mays</i> x <i>Z. mays</i>	Inter	Outcross	Length of branching space	Morphology	6
Plant	<i>Zea mays</i> x <i>Z. mays</i>	Inter	Outcross	Length of sessile spikelet	Morphology	8
Plant	<i>Zea mays</i> x <i>Z. mays</i>	Inter	Outcross	Width of rachis on central spike	Morphology	5
Plant	<i>Glycine max</i> x <i>G. soja</i>	Inter	Self	Pod dehiscence	Morphology	5
Plant	<i>Helianthus annuus</i>	Intra	Outcross x self	Days to flower	Life history	10
Plant	<i>Helianthus annuus</i>	Intra	Outcross x self	Stem diameter	Morphology	6

# Existence d'une littérature pléthorique d'études QTL

Kingdom	Species	Cross	Mating system	Trait	Trait type	Total QTLs
Plant	<i>Helianthus annuus</i>	Intra	Outcross x self	Leaf shape	Morphology	2
Plant	<i>Helianthus annuus</i>	Intra	Outcross x self	Leaf size	Morphology	4
Plant	<i>Helianthus annuus</i>	Intra	Outcross x self	Peduncle length	Morphology	3
Plant	<i>Helianthus annuus</i>	Intra	Outcross x self	No. of branches	Morphology	3
Plant	<i>Helianthus annuus</i>	Intra	Outcross x self	No. of heads	Morphology	5
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Soluble solids conc. - CA	Life history	3
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Wilting 6 h-GS-ch9	Morphology	3
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Inhibition ammonium uptake	Morphology	2
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Stigma exertion	Life history	1
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Bud morphology	Morphology	7
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Flower size	Morphology	3
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Number of flowers	Morphology	1
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Total yield	Morphology	12
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Fruit color	Morphology	15
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Firmness	Morphology	3
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Fruit weight	Morphology	3
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Cover	Morphology	6
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	% green yield	Morphology	11
Plant	<i>Lycopersicon esculentum</i>	Inter	Outcross x self	Maturity	Morphology	6
Plant	<i>Oryza japonica</i>	Intra	Self	Plant height	Morphology	2
Plant	<i>Oryza japonica</i>	Intra	Self	Heading date	Life history	3
Plant	<i>Oryza japonica</i>	Intra	Self	Tiller number	Morphology	2
Plant	<i>Oryza japonica</i>	Intra	Self	Leaf width	Morphology	1
Plant	<i>Oryza japonica</i>	Intra	Self	Panicle length	Morphology	1
Plant	<i>Oryza japonica</i>	Intra	Self	Spikelets/panicle	Morphology	1
Plant	<i>Oryza japonica</i>	Intra	Self	Primary branches/panicle	Morphology	2
Plant	<i>Oryza japonica</i>	Intra	Self	Secondary branches/panicle	Morphology	2
Plant	<i>Oryza japonica</i>	Intra	Self	Grain length	Morphology	2
Plant	<i>Oryza japonica</i>	Intra	Self	Grain width	Morphology	1
Plant	<i>Oryza japonica</i>	Intra	Self	Grain pilosity	Morphology	1
Plant	<i>Zea mays</i>	Intra	Outcross x self	Glume hardness	Morphology	5
Plant	<i>Zea mays</i>	Intra	Outcross x self	Internode length on primary lateral branch	Morphology	5
Plant	<i>Zea mays</i>	Intra	Outcross x self	Leaf length	Morphology	5
Plant	<i>Zea mays</i>	Intra	Outcross x self	Branch number in primary lateral	Morphology	5
Plant	<i>Zea mays</i>	Intra	Outcross x self	Plant height	Morphology	6
Plant	<i>Zea mays</i>	Intra	Outcross x self	No. of ears on lateral branch	Morphology	3
Plant	<i>Zea mays</i>	Intra	Outcross x self	No. of tillers	Morphology	1

# En général, il y a peu de QTLs décelés dans les études

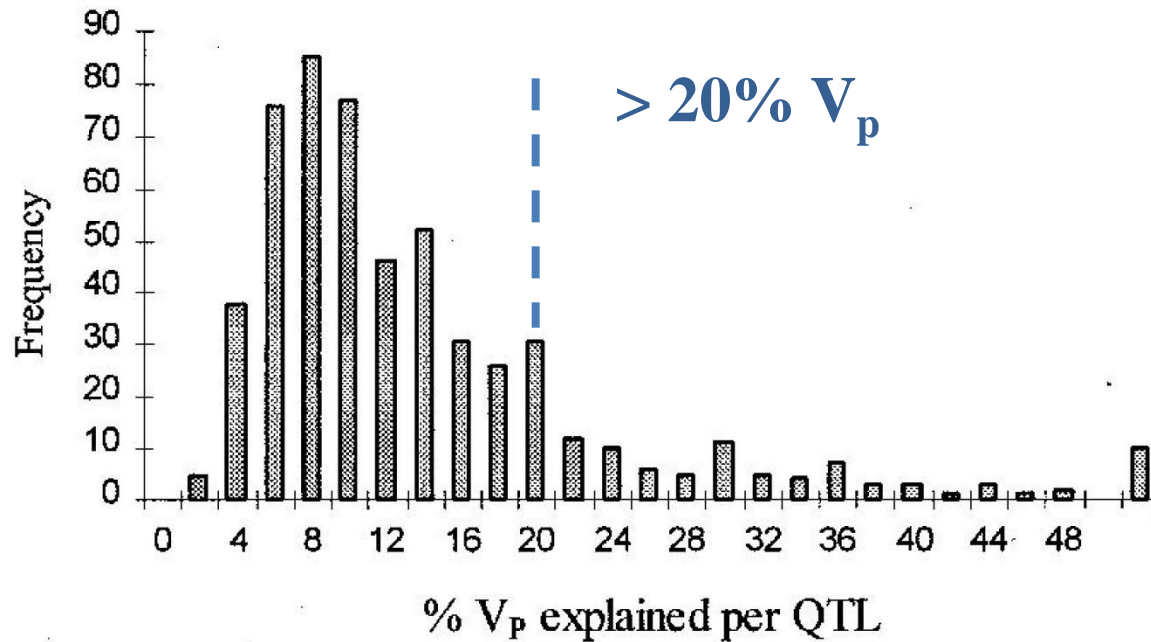
Revue de Rieseberg H et al. (2002) . *Directional selection is the primary cause of phenotypic diversification*. PNAS 99(19) : 12242–12245



# La détection de gènes majeurs dans les études QTLs est rare

Kearsey MJ, Farquhar AGL. 1998. QTL analysis in plants: where are we now? Heredity 80, 137–42.

*Revue de 176 études QTLs (expérience/caractère)*



# Les études QTLs ne permettent de déceler que la partie immergée de l'iceberg du déterminisme génétique

On décèle toujours les gènes majeurs ( $\% V_p > 20\%$ ) quand ils existent, mais ils sont rares !!

On décèle toujours de nombreux gènes de taille modeste ( $10\% < \% V_p < 20\%$ ) ou petite ( $4\% < \% V_p < 10\%$ )

L'ensemble des facteurs génétiques décelables explique très rarement plus de 50% de la variation de la population.

Il existe donc un grand nombre de gènes dont les effets ne sont pas décelables avec les pouvoirs de résolution usuels des études QTLs (taille de la population, densité du marquage moléculaire, type de population étudié etc ...)

**Les études QTL modernes ne démentent pas la conception génétique de Fischer (1918) de la nature des caractères quantitatifs = une somme d'un nombre très important de gènes d'une taille moyenne petite**

## Conséquence de l'hypothèse multifactorielle en matière de distribution

Pour une espèce diploïde, le nombre ( $N$ ) de génotypes différents en un locus augmente avec le nombre d'allèles ( $k$ ), selon la formule suivante:

$$N = k(k+1)/2$$

Pour  $L$  locus, chacun avec  $k$  allèles, le nombre de génotypes possibles devient:

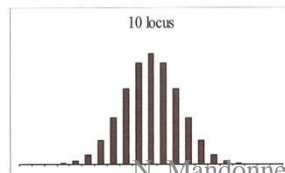
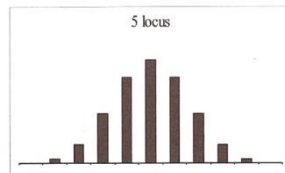
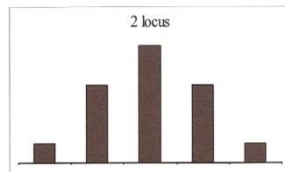
$$N = [k(k+1)/2]^L$$



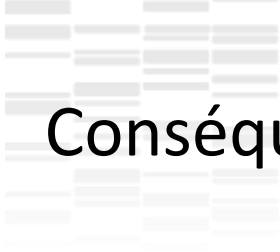
# Conséquence de l'hypothèse multifactorielle en matière de distribution

Nombre de géotypes possibles selon les nombres de locus et d'allèles par locus.

Nombre de locus	Nombre d'allèles par locus		
	2	3	5
1	3	6	15
2	9	36	225
3	27	216	3 375
5	243	7 776	759 375
10	59 049	60 466 176	576 650 390 500



Si l'on fait abstraction des effets du milieu, dès que le nombre de locus impliqués dépasse quelques unités, la distribution des valeurs prises par un caractère devient continue.



## Conséquence de l'hypothèse multifactorielle en matière de distribution

L'hypothèse multifactorielle concilie:

- La nature discontinue de l'hérédité mendélienne
- La continuité des caractères quantitatifs et leur distribution normale (le plus souvent)

Reste valide si l'on considère les variations dues au milieu mais

**Il faut mieux préciser les effets respectifs du génotype et du milieu, leur combinaison.**



## II. Hérité et milieu

### C. Modélisation de l'action du génotype et du micro-milieu

# Modélisation de l'action du génotype et du micro-milieu

Dans cette section, nous ne considérons pas les effets des facteurs de milieu contrôlés: toutes les variations de nature environnementale sont supposées résulter de facteurs de **micro-milieu** uniquement.

## Notion de valeur génétique d'un génotype

= moyenne phénotypique de nombreux individus de même génotype, exprimée en écart à la moyenne de la population

$$P_1 = G + E_1$$

$$P_2 = G + E_2$$

...

$$P_n = G + E_n$$

$$\overline{P_m} = G + \mu$$



$$G = P_m - \mu$$

# Valeur génétique d'un génotype

= moyenne phénotypique de nombreux individus de même génotype, exprimée en écart à la moyenne de la population

	Exemples	Particularités de la reproduction	Obtention de plusieurs exemplaires d'un même génotype	Notion de valeur génétique d'un génotype
<b>Génétique végétale</b>				
Plantes autogames	blé, tomate, haricot	autofécondation	rapide	concrète
Plantes allogames non multipliable végétativement	maïs	allofécondation	impossible	abstraite
Plantes à reproduction végétative	rosier, igname, fraisier	bouturage	immédiate	concrète
<b>Génétique animale</b>				
Petits animaux et insectes	souris, drosophile	Cycles rapides	rapide par croisement consanguins successifs	concrète
Gros animaux d'élevage	bovins	Cycles lents	impossible	abstraite

## Valeur génétique d'un génotype

représente **l'effet moyen du génotype sur un caractère** donné. Elle s'exprime en unités du caractère mesuré dans la population (*Pop*).

$$\mu = E(Pop)$$

En termes mathématiques, nous disons que la valeur génétique (*G*) d'un génotype est **l'espérance de la valeur phénotypique (*P*) conditionnée par le génotype et centrée sur la moyenne de la population**:

$$G = E( P | génotype ) - \mu$$

# Valeur phénotypique d'un individu

L'écart entre la valeur phénotypique ( $P$ ) d'un individu et sa valeur génétique ( $G$ ) est imputable aux conditions de micro-milieu dans lequel se trouve l'individu.

Par convention cet écart ( $E$ ) est désigné sous le terme d'écart environnemental.

**En définitive, la valeur phénotypique d'un individu  $i$  se décompose ainsi :**

$$P = \mu + G + E$$

## Valeur phénotypique ( $P_{ij}$ ) d'un individu ( $i$ ) observé ( $j$ ) fois dans la population

$$P_{ij} = \mu + G_i + E_{ij}$$

$E_{ij}$  est un résidu d'espérance nulle ( $E[E_{ij}] = 0$ ) pour tout individu ( $i$ ) (moyenne globale des résidus d'un très grand nombre de copies du même individu est nulle, compte tenu du caractère aléatoire des effets de **micro-milieu** subis par chaque copie qui ont tendance à se compenser).





# Héritabilité au sens large d'un caractère

Dans une population constituée d'individus de génotypes différents, la valeur génétique ( $G$ ) des génotypes est considérée comme une **variable aléatoire**, dont la distribution dépend des fréquences des génotypes.

On ne considère dans la variation environnementale que les facteurs de **micro-milieu** : les concepts que nous allons définir concernent les valeurs phénotypiques ( $P$ ) corrigées pour tous les effets systématiques des facteurs de milieu contrôlés.

La valeur environnementale ( $E$ ) est alors un résidu aléatoire, d'espérance nulle par construction, quel que soit le génotype.

# Héritabilité au sens large d'un caractère

Les variations phénotypiques observées entre individus pour un caractère quantitatif sont imputables à

- (1) des différences de génotype ( $G_i$ ) entre individus
- (2) des différences de conditions de milieu dans lesquelles sont placés les individus mesurés

$$P_{ij} = \mu + G_i + E_{ij}$$

$$\text{Var}(P_{ij}) = \text{Var}(\mu + G_i + E_{ij})$$

$$\text{Var}(P_{ij}) = \text{Var}(G_i + E_{ij}) = \text{Var}(G_i) + \text{Var}(E_{ij}) + \text{Covar}(G_i, E_{ij})$$

Or il existe une indépendance entre les variables aléatoires  $G_i$  et  $E_{ij}$  par la "randomisation", qui consiste à répartir les  $j$  individus  $i$  au hasard dans les différents micro-milieus :

$$\text{Covar}(G_i, E_{ij}) = 0$$

Donc :

$$\text{Var}(P_{ij}) = \text{Var}(G_i) + \text{Var}(E_{ij})$$

# Héritabilité au sens large d'un caractère

La variance phénotypique des individus d'une population est égale à la somme de leur variance génétique et de leur variance environnementale :

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

$$V_P = V_G + V_E$$

Selon la définition proposée par Lush (1937), on appelle **héritabilité au sens large**, notée  $h_{sl}^2$  également ( $H^2$ ) le rapport entre la variance génétique et la variance phénotypique :

$$h_{sl}^2 = \frac{V_G}{V_P} = \frac{V_G}{V_G + V_E}$$

## Héritabilité au sens large d'un caractère

$$h_{sl}^2 = \frac{V_G}{V_P} = \frac{V_G}{V_G + V_E}$$

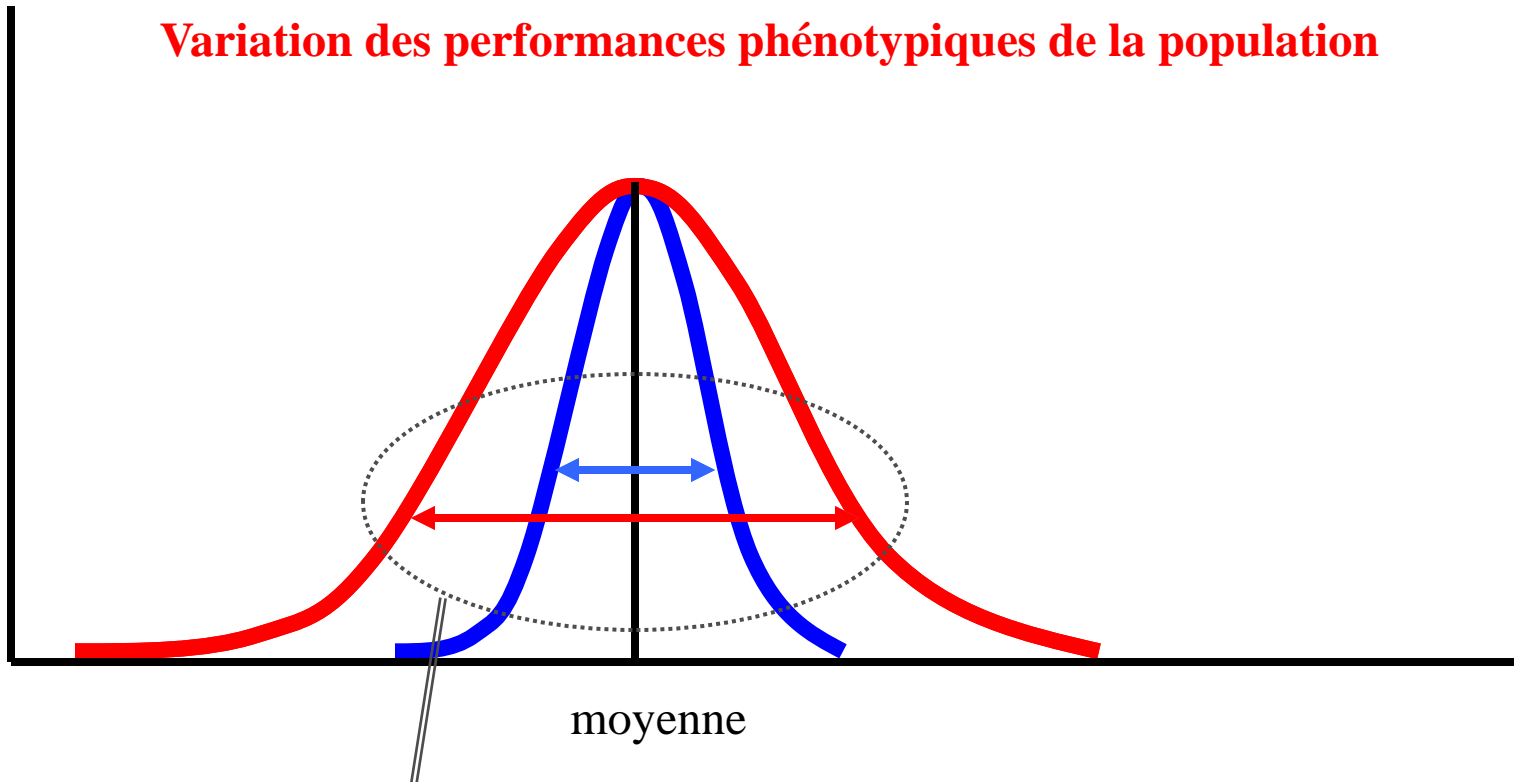
**L'héritabilité au sens large est donc une proportion: c'est la part de variance phénotypique qui est d'origine génétique (souvent exprimée en pourcentage).**

**L'héritabilité au sens large exprime le degré de confiance que l'on peut accorder au phénotype pour évaluer le génotype.**

# Héritabilité au sens large d'un caractère

Effets des gènes sur la performance de la population

Variation des performances phénotypiques de la population



L'héritabilité ( $H^2$ ) = part de la variance des performances expliquée par les gènes

# Héritabilité au sens large d'un caractère

Si l'on pouvait entretenir tous les individus d'une population strictement dans les mêmes conditions de micro-milieu (ce qui est une utopie car le résidu  $E$  n'est jamais nul, donc sa variance non plus), la variance phénotypique d'un caractère quelconque n'aurait qu'une seule composante, strictement génétique :

$h_{sl}^2$  serait alors égale à 1 (ou 100%).

A l'opposé, un caractère qui ne dépendrait presque pas du génotype des individus, mais serait beaucoup influencé par l'environnement aurait une héritabilité au sens large très faible.



## II. Hérité et milieu

### D. Action du macro-milieu et interaction génotype x milieu

# Additivité des effets des effets du macro-milieu et du génotype

## Comparaison de rendement de 2 variétés de blé

Macro-milieu 1 : 250 graines / m<sup>2</sup> au semis, 60 unités d'azote / ha

Macro-milieu 2 : 300 graines / m<sup>2</sup> au semis, 120 unités d'azote / ha

	Variété A	Variété B
Macro- milieu 1	75 quintaux/ha	80 quintaux/ha
Macro- milieu 2	85 quintaux/ha	90 quintaux/ha

$$P = \mu + m + G + E$$

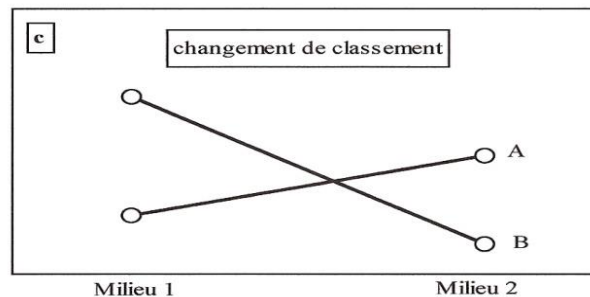
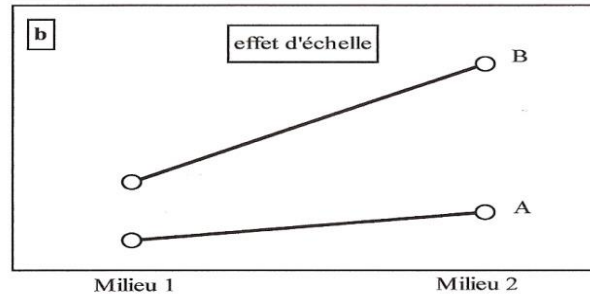
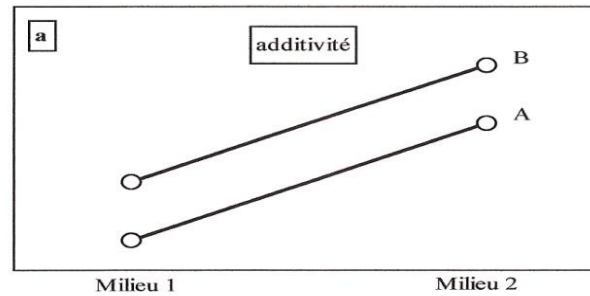
Additivité des effets du macro-milieu avec le génotype

= effet du milieu identique pour tous les génotypes

= le choix d'un génotype donné confère le même (dés)avantage quel que soit le milieu



# Interaction macro-milieu x génotype



**Il y a interaction génotype x milieu lorsque l'on s'écarte de la situation d'additivité**

**macro-milieu affecte différemment les génotypes**

**macro-milieu affecte différemment les génotypes**

$$P = \mu + m + G + m \times G + E$$



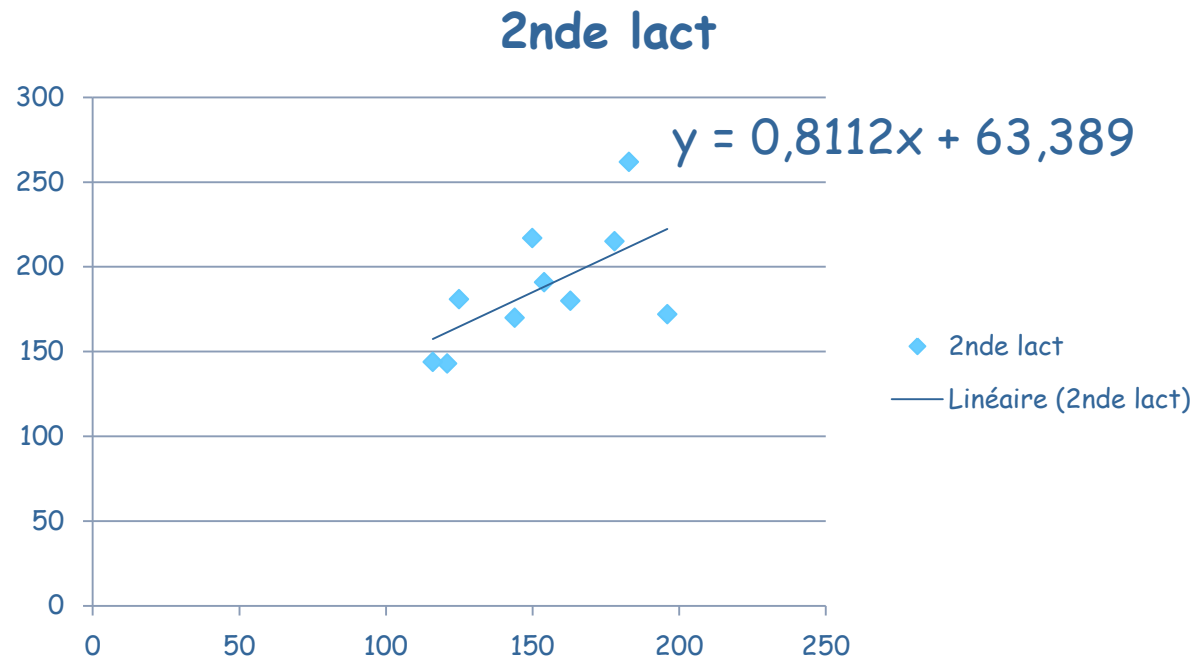
# Conséquence d'une interaction génotype x macro-milieu

Interactions = un changement important de classement entre génotypes (ou des différences entre génotypes) d'un milieu à l'autre

## Problèmes :

- dans un milieu donné, quel est le meilleur génotype (quel est le génotype qui répond le mieux aux objectifs de l'amélioration) ?
- dans quel milieu pratiquer la sélection ?

# Exercice n°3





# III. Modélisation de l'effet des gènes (des allèles) sur un caractère

## A. Le modèle à un locus

# Notion d'effet moyen d'un allèle

Jusqu'à présent, nous avons défini la valeur génétique  $G$  d'un génotype : effet moyen d'un génotype sur caractère :

$$G = E( P \mid \text{génotype} ) - \mu$$

Nécessité de s'intéresser aux effets des différents allèles d'un même locus :

1) Transmission que de la moitié des allèles d'un géniteur à sa descendance ( 1 allèle sur 2)

2) Nécessité de lier la GQ à la GP qui essentiellement fondée sur des considérations de fréquences alléliques

[Lien entre évolution de la moyenne d'un caractère quantitatif dans une population et l'évolution des fréq. alléliques (des gènes contrôlant le caractère) sous l'effet de la sélection]

3) Nécessité de décomposer la valeur génétique  $G$ , à un niveau allélique, pour introduire la notion et aborder le calcul de l'héritabilité au sens étroit

# Notion d'effet moyen d'un allèle

Soit un locus :  $n$  allèles  $A_i$  de fréquence  $p_i$   $1 \leq i \leq n$  :  $\sum p_i = 1$

Population de grande taille, panmictique (population fermée, absence mutations, sélection, unions au hasard)

Soit  $f_{ij}$  fréquence du génotype  $A_iA_j$ , avec un équilibre Hardy-Weinberg on a :

$f_{ij} = 2p_i p_j$  pour tout génotype hétérozygote  $A_iA_j$

$f_{ii} = p_i^2$  pour tout génotype homozygote  $A_iA_i$

Moyenne phénotypique  $\mu$  de la population pour le caractère mesuré

$$\mu = \sum_{i \leq j} f_{ij} \mu_{ij}$$

$\mu_{ij}$  : moyenne de tous les génotypes  $A_iA_j$

# Décomposition de la valeur génétique d'un génotype en fonction des effets de ses allèles

L'effet moyen ( $\alpha_i$ ) d'un allèle ( $A_i$ ) est l'espérance de toutes les valeurs génétiques des individus ayant reçu cet allèle ( $A_i$ ), l'autre allèle étant tiré au hasard.

Valeur génétique  $G_{ij}$  d'un génotype à un locus ( $A_i A_j$ )  
= effet moyen de l'allèle paternel ( $\alpha_i$ )  
+ effet moyen de l'allèle maternel ( $\alpha_j$ )  
+ effet d'interaction entre les 2 allèles ( $\delta_{ij}$ )

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

Valeur génétique au locus →  $G_{ij} = \alpha_i + \alpha_j + \delta_{ij}$  ← Valeur ou résidu de dominance au locus : écart entre valeur génétique et valeur génétique additive

Valeur génétique additive au locus

Sommation de cette décomposition raisonnée au niveau d'un locus *sur l'ensemble* des loci impliqués dans le contrôle du caractère :

$$G = A + D$$

Valeur génétique de l'individu →  $G = A + D$  ← Valeur de dominance

Valeur génétique additive de l'individu



# III. Modélisation de l'effet des gènes (des allèles) sur un caractère

## B. Généralisation à plusieurs locus



## Décomposition de la valeur génétique d'un génotype en considérant $l$ loci

Si on a ( $l$ ) loci impliqués dans le contrôle du caractère étudié,

La valeur génétique ( $G$ ) d'un individu peut s'écrire dans un **modèle additif** entre locus comme la somme des valeurs génétiques ( $G_{lij}$ ) à chacun d'eux :

$$G = \sum_l G_{lij}$$

La valeur génétique ( $G_{lij}$ ) étant l'espérance de la valeur phénotypique des individus porteurs des allèles  $ij$  au locus  $l$  considéré.

De la même façon, on peut généraliser la valeur génétique additive ( $A$ ) et la valeur de dominance ( $D$ ) d'un individu, précédemment définis à un locus, à l'ensemble des locus  $l$

$$A = \sum_l (\alpha_{li} + \alpha_{lj}) \quad ; \quad D = \sum_l \delta_{lij}$$

**On généralise ainsi à plusieurs locus, la décomposition de la valeur génétique d'un individu définie précédemment:  $G = A + D$**

## Décomposition de la valeur génétique d'un génotype en considérant $l$ loci

- $G = A + D$  encore une simplification de la réalité
- $G = A + D + E$  pour prendre en compte des ***interactions entre gènes*** situés à des locus différents: résidus d'***épistasie***
- En pratique, souvent négligés.

## Modèle quantitatif et complexité biologique

Les concepts introduits sont des concepts fondamentaux en génétique quantitative. **Les analyses de génétique quantitative et la pratique de l'amélioration génétique sont largement fondées sur le modèle polygénique.**

On peut se demander si cette approche statistique relativement simple est à même de rendre compte de la complexité du fonctionnement biologique des organismes (de l'expression des gènes jusqu'à un phénotype en passant par les synthèses protéiques, le métabolisme et la physiologie)

**De nombreux reproches ont d'ailleurs été formulés à l'égard de ce modèle dans les dernières décennies.**

Il ne faut cependant pas perdre de vue que **cette approche n'a aucune prétention explicative : son but est essentiellement opérationnel ou prédictif de l'évolution à plus ou moins court terme de la valeur de caractères quantitatifs au sein de populations soumises à la sélection naturelle ou artificielle.**

La décomposition suivant un modèle additif est fréquente en statistique et permet souvent des approximations de bonne qualité. **Ainsi, les insuffisances du modèle à l'échelon du fonctionnement biologique n'excluent pas un bon pouvoir prédictif à l'échelon macroscopique d'un phénotype quantitatif et de ses lois de transmission.**



# III. Modélisation de l'effet des gènes (des allèles) sur un caractère

E. Décomposition de la variance génétique

# Décomposition de la variance génétique

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

Valeur phénotypique → Moyenne de la population → Valeur génétique → Valeur génétique additive → Erreur ← Résidu de dominance ←

## Décomposition de la variance

[covariance nulles entre les variables  $A/D$ ,  $A/E$ , et  $D/E$ ]

$$V_P = V_G + V_E = V_A + V_D + V_E$$

La valeur génétique additive est la fraction de la valeur génétique dont on peut facilement prédire la transmission.

On définit l'**héritabilité au sens étroit d'un caractère ( $h^2$ )** comme le rapport entre la variance génétique additive et la variance phénotypique :

$$h_{ss}^2 = \frac{V_A}{V_P} = \frac{V_A}{V_A + V_D + V_E}$$

## L'héritabilité au sens étroit d'un caractère

$$h_{ss}^2 = \frac{V_A}{V_P} = \frac{V_A}{V_A + V_D + V_E}$$

Comme l'héritabilité au sens large ( $V_G/V_P$ ), l'héritabilité au sens étroit ( $V_A/V_P$ ) est un ratio compris entre 0 et 1

Un caractère dans une population est d'autant plus héritable que la variance génétique additive des loci qui gouverne le caractère, est élevée

L'héritabilité au sens large est mathématiquement supérieure à celle au sens étroit ( $h_{ss}^2 < h_{sl}^2$ )

L'héritabilité au sens étroit représente l'héritabilité au sens commun du terme, c'est-à-dire l'aptitude d'un caractère à se transmettre d'une génération à l'autre. Elle dépend principalement de la nature du caractère et de la population étudiée et du milieu.

L'héritabilité au sens large dépend uniquement de la « robustesse » du dispositif d'expérimentation, quelque soit par ailleurs la valeur de l'héritabilité au sens étroit.

En d'autre terme, il est possible d'obtenir des estimations d'héritabilité d'un caractère, très divergentes dans certaines populations ( $h_{ss}^2 \ll h_{sl}^2$ )

## Facteurs de variation de l'héritabilité au sens étroit

$$h_{ss}^2 = \frac{V_A}{V_P} = \frac{V_A}{V_A + V_D + V_E}$$

L'héritabilité dépend :

- (a) des fréquences alléliques dans la population, puisque  $V_A$  et  $V_D$  en dépendent;
- (b) de l'importance des variations dues aux effets de dominance;
- (c) de l'importance des variations dues au micro-milieu (on rappelle que l'on ne considère ici que les facteurs de milieu non contrôlés).

Comme toute composante de la variance, l'héritabilité au sens étroit est un paramètre spécifique du caractère étudié et de la population observée.

Toute évolution des fréquences alléliques fait évoluer les variances génétique additive et de dominance. L'héritabilité au sens étroit d'un caractère est susceptible d'évoluer sous l'effet de la sélection ou de la dérive génétique.

L'héritabilité au sens étroit d'un caractère est d'autant plus grande que les conditions de micro-milieu sont homogènes ( $V_E$  faible).

L'héritabilité au sens étroit d'un caractère a un sens essentiellement chez les animaux où les régimes de reproduction sont stables.

# Quelques exemples de valeurs d'héritabilité au sens étroit chez des espèces animales mentionnées dans la littérature

	$h^2$
<b>Bovins</b>	
Taux de réussite en 1ère insémination artificielle	0,00 - 0,05
Vitesse de croissance	0,25 - 0,35
Production laitière (traite)	0,25 - 0,35
Teneur en protéines du lait	0,45 - 0,55
<b>Ovins</b>	
Taille de portée	0,05 - 0,10
Taux de mortalité avant sevrage	0,00 - 0,05
Finesse de la laine	0,60 - 0,65
Production laitière (allaitement)	0,20 - 0,30
<b>Caprins</b>	
Production laitière (traite)	0,25 - 0,35
Teneur en protéines du lait	0,50 - 0,55
<b>Porc</b>	
Taille de portée	0,05 - 0,10
Indice de consommation	0,25 - 0,35
Épaisseur de lard dorsal	0,30 - 0,60
Taux de muscle dans la carcasse	0,50 - 0,60
<b>Cheval (sport et course)</b>	
Record sur 1 km entre 3 et 4 ans (trot)	0,30 - 0,40
Gain annuel en concours de saut d'obstacles	0,25 - 0,35
<b>Lapin</b>	
Taille de portée	0,00 - 0,20
Poids de la portée à l'abattage	0,10 - 0,30
<b>Poule</b>	
Taux d'éclosion	0,05 - 0,20
Age à la maturité sexuelle	0,20 - 0,40
Production d'œufs	0,15 - 0,30
Poids moyen de l'œuf	0,40 - 0,60
<b>Truite arc-en-ciel</b>	
Taux de mortalité des alevins	0,05 - 0,20
Poids à 2 ans	0,25 - 0,35
Teneur en gras de la carcasse	0,40 - 0,60



## Exercice 2

Deux lignées de haricot sont croisées entre elles. Dans la F1, la variance du poids des fèves mesurée est de 1,5.

La F1 est autofécondée, et dans la F2, la variance du poids des fèves est de 6,1.

Estimez l'héritabilité au sens large du poids des fèves dans cette expérience.

# Exercice 1

Dans un vaste troupeau de bestiaux, trois caractères dont la distribution est continue sont mesurés et leurs variances calculées comme suit :

	Caractères		
Variances	Longueur du jarret	Longueur du collier	Contenu en graisses
Variance phénotypique	310,2	730,4	106,0
Variance due au milieu	248,1	292,2	53,0
Variance génétique additive	46,5	73,0	42,4
Variance génétique due à la dominance	15,6	365,2	10,6

**0,15**

**0,10**

**0,40**


- 1) Calculez l'héritabilité au sens large et étroit pour chacun de ces caractères
- 2) Dans la population étudiée, quel caractère répondrait le mieux à la sélection ? Pourquoi ?

$$\mathbf{R = h^2 S}$$



# Résumé de la démarche

- GQ = exploitation de la variabilité des caractères
  - Décrite au sein d'une population et pour un caractère par  $V_p$  ou, au sein d'une population et entre caractères par corrélation/régression
  - Un modèle  $P = G + E$ 
    - $G$  = somme des effets de  $G_{\text{Majeurs}}$ ,  $G_{\text{mineurs}}$ , polygènes
    - $E$  = macromilieu (contrôlé) + micromilieu (nc)

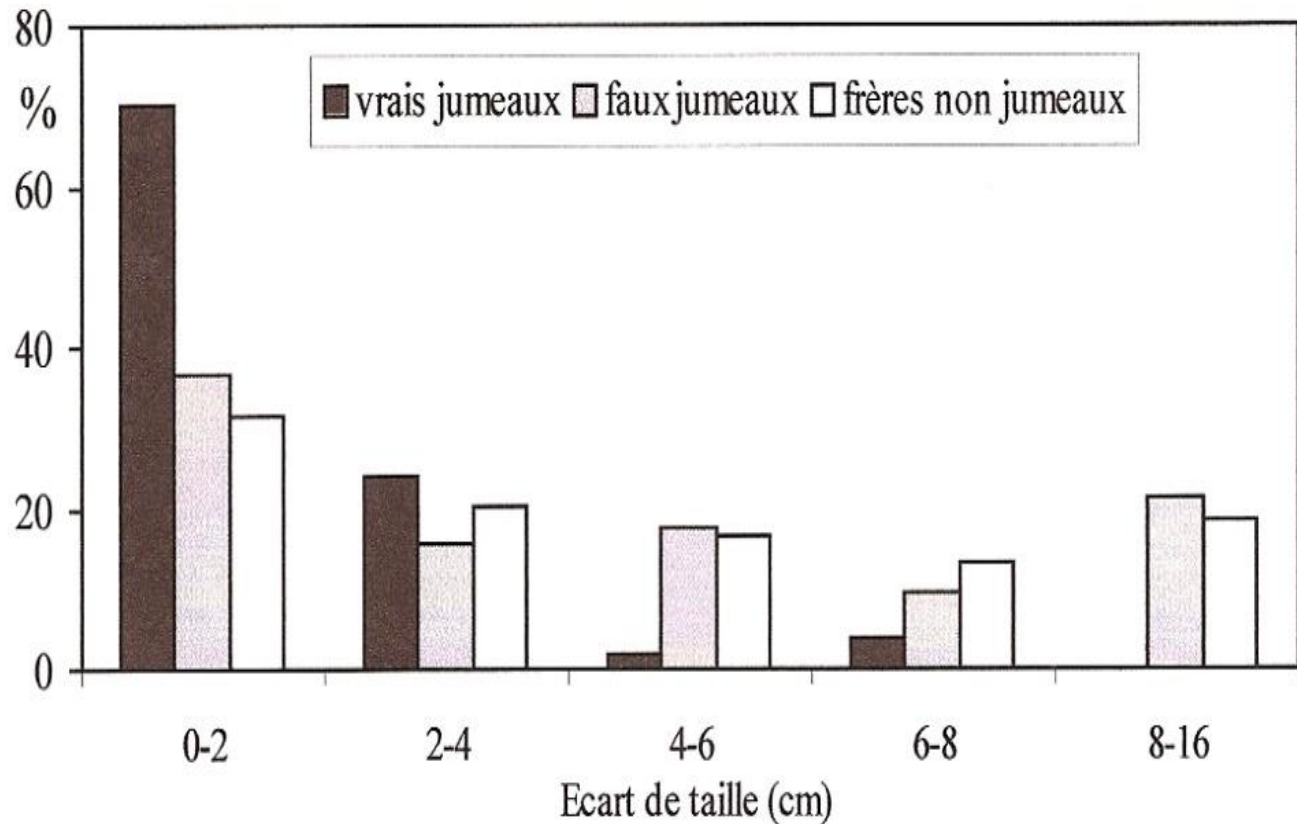


# IV. La ressemblance entre apparentés

## A. Observations statistiques au sein de populations

# Liaison entre membres d'une même fratrie

Fréquence (en %) des écarts de taille chez l'homme observés au sein de paires de vrais jumeaux (50 paires), de faux jumeaux (52 paires) et de pleins frères non jumeaux (52 paires), les membres d'une même paire étant élevés ensemble. (Source: Stem, 1960).



# Liaison entre membres d'une même fratrie

Coefficients de corrélation entre apparentés, pour différents caractères mesurés en ferme chez le porc (race Large-White).  $n$  = nombre de paires observées.  $r$  = valeur calculée de la corrélation

Caractère mesuré		non apparentés	demi frères-sœurs de père commun	pleins frères-sœurs de même portée
	$n$	725	717	656
pH du muscle semi-membraneux	$r$	0,013	0,032	0,085
	sign.	ns	ns	*
Age à 100 kg	$n$	3 836	2 770	2 732
	$r$	0,024	0,068	0,168
	sign.	ns	***	***
Epaisseur moyenne de lard dorsal à 100 kg	$n$	3 835	2 770	2 728
	$r$	0,009	0,107	0,238
	sign.	ns	***	***

- ▶ corrélation n'est pas significativement différente de zéro pour des individus non apparentés entre eux
- ▶ pour un lien de parenté donné, la corrélation dépend du caractère mesuré
- ▶ corrélation entre frères-sœurs de même portée plus que le double de la corrélation entre demi-frères-sœurs



# Liaison entre membres d'une même fratrie

Deux individus apparentés entre eux se ressemblent plus que deux individus choisis au hasard dans la population.

Le degré de ressemblance phénotypique, apprécié par une corrélation, dépend :

*(1) du lien de parenté considéré,*

*(2) d'éventuels facteurs d'environnement communs aux apparentés considérés et*

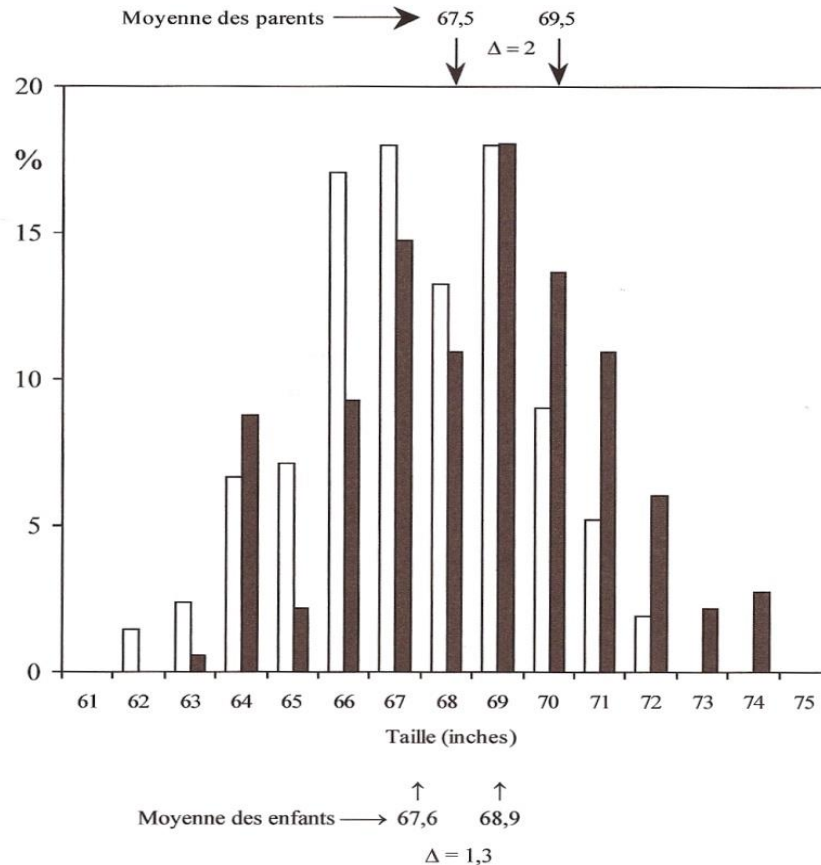
*(3) du caractère mesuré*

# Liaison entre parent-descendant

Distribution, chez l'homme, de la taille à l'âge adulte des enfants issus de parents répartis en deux catégories selon leur taille moyenne:

Blanc = 211 enfants issus de 33 couples de parents dont la taille moyenne est comprise entre 67 et 68 inches (170 – 173 cm).

Foncé = 183 enfants issus de 41 couples de parents dont la taille moyenne est comprise entre 69 et 70 inches. (175-178 cm).



Francis Galton (1886)

- ▶ larges variabilités intra-groupes comparables
- ▶ larges distributions chevaugantes
- ▶ Moyenne descendants issus de parents les plus grands supérieure à celles des parents les plus petits
- ▶ Ecart entre moyenne descendants inférieur à écart entre moyenne des 2 catégories de parents

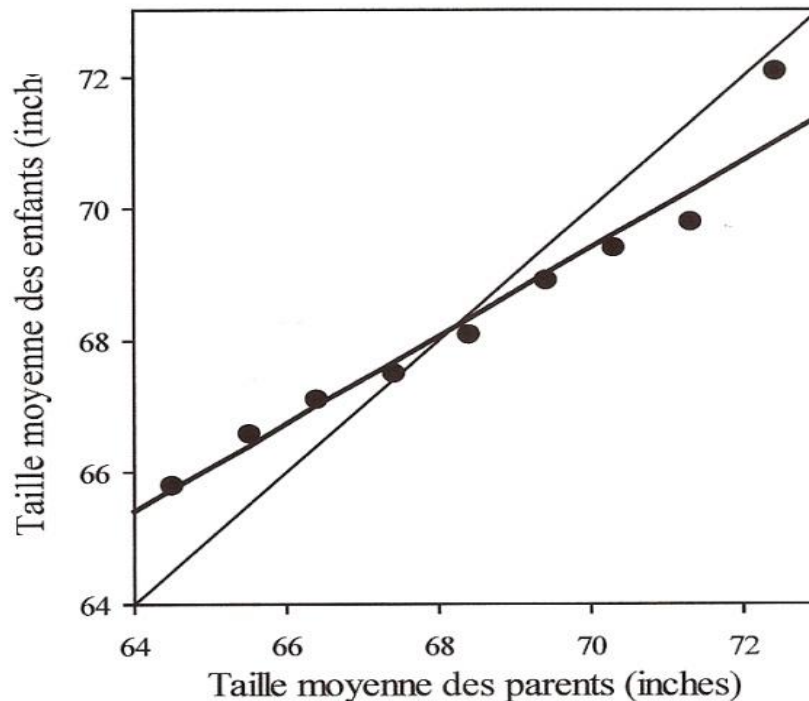


# Liaison entre parent-descendant

Liaison, chez l'homme, entre la taille à l'âge adulte des enfants (928 enfants) et la taille moyenne de leurs parents (205 couples de parents). D'après Galton 1886

« régression des descendants vers la moyenne »


les individus grands donnent des descendants qui, en moyenne, sont grands, mais moins grands qu'eux



Chaque point représente une moyenne de plusieurs familles

Les écarts entre parents se retrouvent en partie seulement entre leurs descendances

Ici coefficient de régression Parent-descendant =  $2/3$



# IV. La ressemblance entre apparentés

## B. Prédiction de la covariance entre apparentés

# Covariance entre individus

Soient deux individus  $i$  et  $j$  appartenant à la même population

$$P_i = \mu + G_i + E_i \quad \text{et} \quad P_j = \mu + G_j + E_j$$

répartition aléatoire des différents génotypes dans les différentes conditions de micro-milieu, les variables  $G$  et  $E$  sont non corrélées

$$\text{Cov}(P_i, P_j) = \text{Cov}(G_i, G_j) + \text{Cov}(E_i, E_j)$$

La covariance phénotypique entre deux individus a deux composantes:

- Une composante génétique, si les individus ont certains gènes en commun.
- Une composante environnementale, si les individus partagent un certain environnement commun.

En l'absence d'effet d'environnement commun

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



## Covariance entre individus: Des exemples

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


Covariance entre non apparentés = 0

Covariance entre demi-frères-sœurs =  $1/4 V_A$

Covariance entre pleins-frères =  $1/2 V_A + 1/4 V_D$

Covariance entre vrais jumeaux =  $V_A + V_D = V_G$

Covariance entre parent-descendants =  $1/2 V_A$



# IV. La ressemblance entre apparentés

## C. Interprétation de la notion d'héritabilité

## Corrélation entre 2 apparentés $i$ et $j$

$$r = \frac{Cov(P_i, P_j)}{V_P} = \frac{Cov(G_i, G_j) + Cov(E_i, E_j)}{V_P}$$

Si l'on peut admettre l'absence d'effet d'environnement commun

$$Cov(E_i, E_j) = 0$$

demi-frères-sœurs :  $r = \frac{\frac{1}{4}V_A}{V_P} = \frac{1}{4}h^2$

pleins frères-sœurs :  $r = \frac{\frac{1}{2}V_A + \frac{1}{4}V_D}{V_P} = \frac{1}{2}h^2 + \frac{1}{4}\frac{V_D}{V_P}$

vrais jumeaux :  $r = \frac{V_A + V_D}{V_P} = h_{sl}^2$

La valeur attendue du coefficient de corrélation entre valeurs phénotypiques d'une même fratrie s'exprime en fonction de paramètres génétiques,  $(V_A, V_D, h^2)$

## Corrélation entre 2 apparentés $i$ et $j$

demi-frères-sœurs :  $r = \frac{\frac{1}{4}V_A}{V_P} = \frac{1}{4}h^2$

pleins frères-sœurs :  $r = \frac{\frac{1}{2}V_A + \frac{1}{4}V_D}{V_P} = \frac{1}{2}h^2 + \frac{1}{4}\frac{V_D}{V_P}$

vrais jumeaux:  $r = \frac{V_A + V_D}{V_P} = h_{sl}^2$

(i) avec l'accroissement de la corrélation observée entre valeurs phénotypiques lorsque le lien de parenté est plus étroit

(ii) corrélation entre pleins frères-soeurs vaut plus du double de la corrélation entre demi-frères-soeurs

(iii) pour un lien de parenté donné, la corrélation dépend du caractère observé, puisque les paramètres génétiques ( $V_A$ ,  $V_D$ ,  $h^2$ ) sont spécifiques d'un caractère donné

# Estimation de l'héritabilité

demi-frères-sœurs :  $r = \frac{\frac{1}{4}V_A}{V_P} = \frac{1}{4}h^2$

pleins frères-sœurs :  $r = \frac{\frac{1}{2}V_A + \frac{1}{4}V_D}{V_P} = \frac{1}{2}h^2 + \frac{1}{4}\frac{V_D}{V_P}$

vrais jumeaux:  $r = \frac{V_A + V_D}{V_P} = h_{sl}^2$

héritabilité au sens étroit


$$\widehat{h^2} = 4 \times \text{corrélation entre demi - frères - soeurs}$$

l'héritabilité au sens large

$$\widehat{h_{sl}^2} = \text{corrélation entre vrais jumeaux ou entre individus d'un clone}$$

On exploite les ressemblances entre apparentés pour estimer l'h<sup>2</sup>.





# IV. La ressemblance entre apparentés

D. Introduction de la notion de sélection

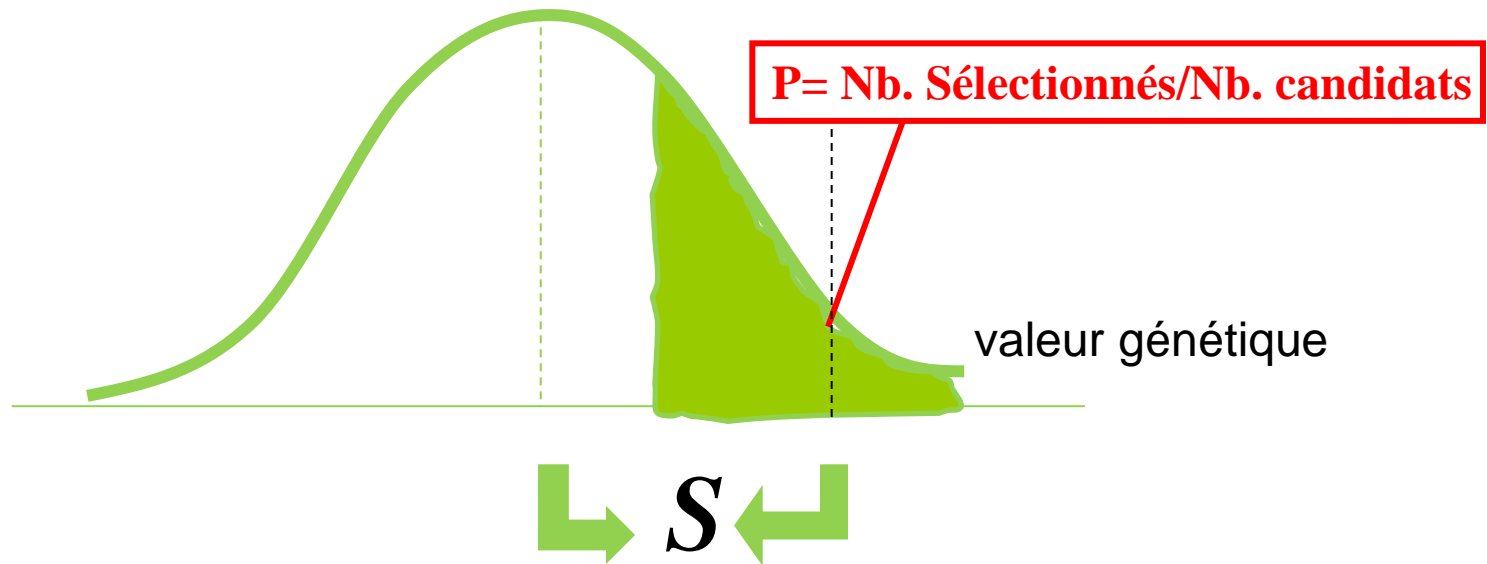


# Effets de la sélection dirigée

- Trois concepts clés : S, R, I
  - S : Différentiel de sélection
    - Moyenne des parents sélectionnés – moyenne de la population
  - R : Réponse à la sélection
    - Moyenne des descendants – moyenne de la population de la génération précédente
  - I : Intensité de la sélection
    - Différentiel de sélection/ écart type phénotypique de la population des parents

# Illustration graphique pour la sélection massale

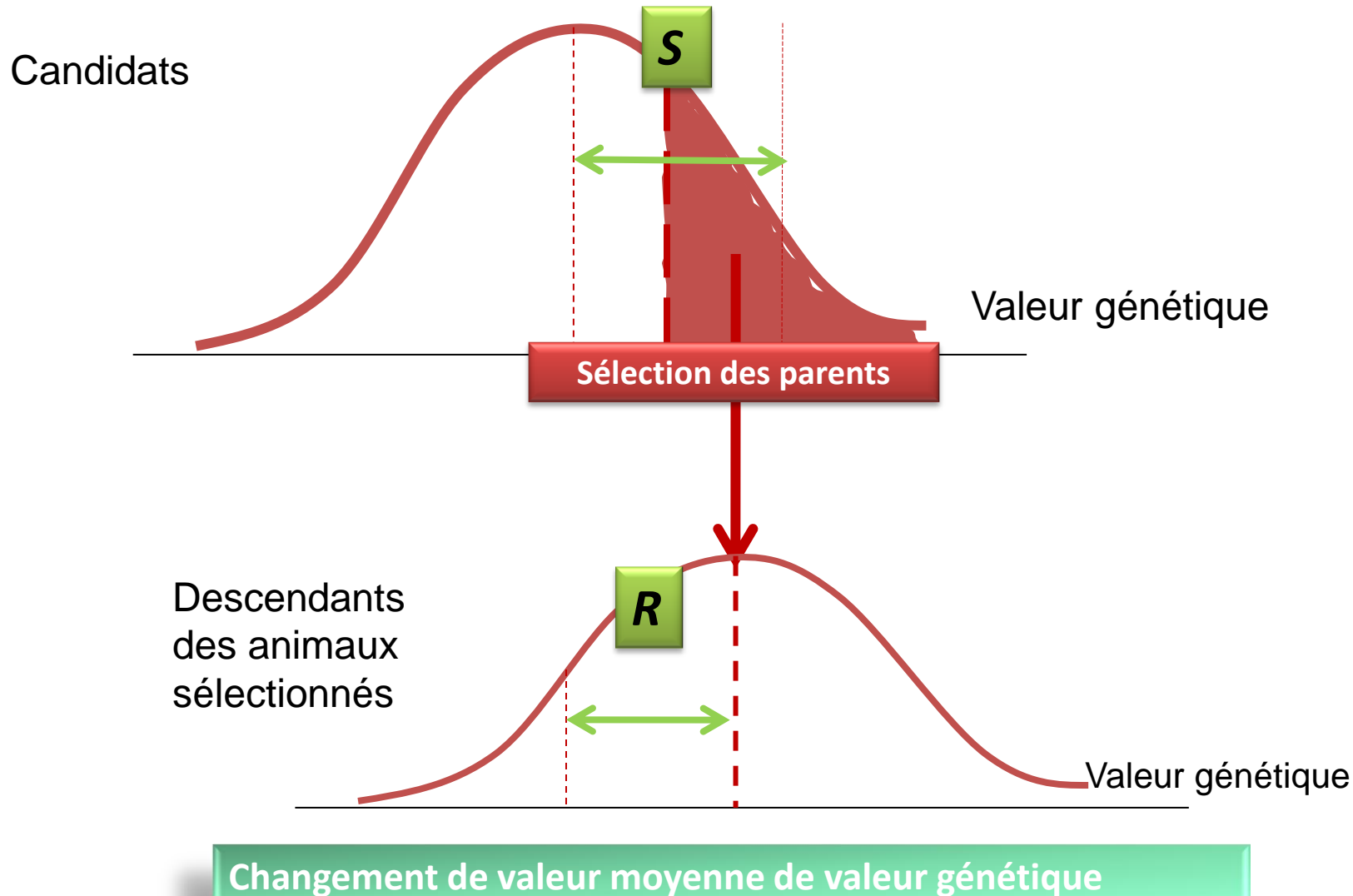
## La sévérité du choix



Intensité de sélection:

$i =$  différentielle standardisée  $= S / \text{écart-type génétique}$

# D'une génération à l'autre



# Résumé des conséquences du modèle: si « $P = G$ »

1. Les distributions phénotypiques sont normales
2. Le coefficient de corrélation phénotypique entre apparentés est égal à la proportion de gènes identiques par descendance
3. La consanguinité produira une population génotypiquement et phénotypiquement homogène

Alors la réponse à la sélection  $R$  devrait être égale au différentiel de sélection  $S$

- Il y a donc des faits qui ne cadrent pas avec le modèle → Recherche d'un modèle plus réaliste. Notamment  $R < S$

# Relation entre R et S

- On prend en compte l'environnement pour être plus réaliste:
  - $Y = \mu + G + E$
  - $V_P = V_G + V_E$
- Explique que les lignées consanguines restent variables :  $V_G$  disparaît mais  $V_E$  demeure
  - Mais n'explique pas que parfois la variance des lignées consanguines peut être supérieure à celles de lignées hétérogames (il faudra prendre en compte la non additivité)



# Relation entre R et S

La relation est alors:

$$R = h^2S$$

- $h^2$  est l'***héritabilité*** comprise entre 0 et 1

# Plan de cours

## 1. LA VARIABILITE AU SEIN DES POPULATIONS

- A. Description d'une population pour un caractère
- B. Description d'une population pour deux caractères
- Exercices

## II. HEREDITE ET MILIEU

- A. Définitions
- B. Mise en évidence des facteurs de variation génétiques
- C. Modélisation de l'action du génotype et du micro-milieu
- D. Action du macro-milieu et interaction génotype x milieu
- Exercices

## III. MODELISATION DE L'EFFET DES GENES SUR UN CARACTERE

- A. Le modèle génétique à un locus
- B. Généralisation à plusieurs locus du modèle génétique quantitatif
- C. Moyennes des variables  $G$ ,  $A$  et  $D$  au sein d'une population
- D. Intérêt prédictif de la notion de valeur génétique additive
- E. Décomposition de la variance génétique et paramètres génétiques
- Exercices

## IV. LA RESSEMBLANCE ENTRE APPARENTES

- A. Observations statistiques dans des populations
- B. Prédiction de la covariance entre apparentés
- C. Interprétation de la notion d'héritabilité et estimation
- D. Introduction à la notion de sélection
- Exercices

## V. EFFETS DES REGIMES DE REPRODUCTION

- A. Définition
- B. Effets de la consanguinité
- C. Hétérosis

## VI. L'APPORT DES MARQUEURS MOLECULAIRES



# La démarche en résumé

GQ = exploitation de la variabilité des caractères à des fins de sélection

- Décrite au sein d'une population et pour un caractère par  $V_p$  ou, au sein d'une population et entre caractères par corrélation/régression
- Un modèle  $P = G + E$ 
  - $G$  = somme des effets de QTL ( $G_{Majeurs}$ ,  $G_{mineurs}$ ), polygènes
  - $E$  = macromilieu (contrôlé) + micromilieu (nc)
  - Interactions  $G^*E$
  - Héritabilité au sens large ( $H^2$ )

# Résumé de la démarche

- Prise en compte des effets aux loci, on complique:
  - $P = (A + D + E) + \text{Environnement}$
  - $h^2$  héritabilité au sens étroit (domaine animal)
- La ressemblance entre apparentés (corrélation entre individus) permet d'estimer  $h^2$ 
  - $h^2$  permet de prédire la réponse à la sélection dirigée
  - $R = h^2 S$
- Différentes notions qui sont les outils de base de l'améliorateur