

# Cours 2 – Le Modèle génétique et le cas biallélique

Jemay Salomon

UMR GQE Le Moulon

Université Paris-Saclay, INRAE, CNRS, AgroParisTech

# Outline

Rappel

Le modèle à un locus

Généralisation du modèle

Conséquences pour la pratique de l'amélioration génétique

Activité - 30 min

# Outline

## Rappel

Le modèle à un locus

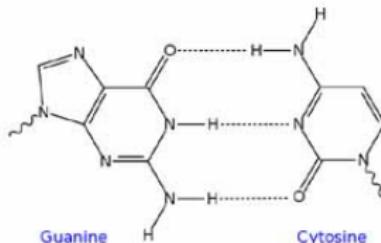
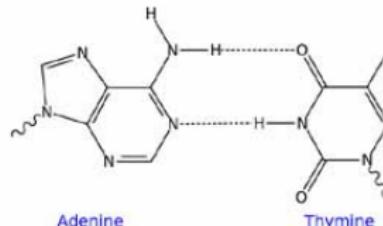
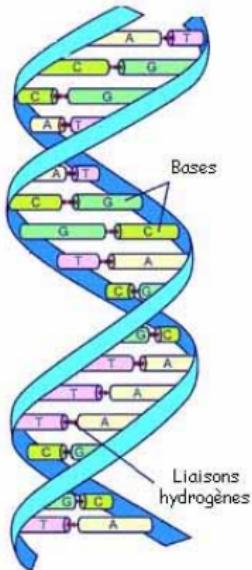
Généralisation du modèle

Conséquences pour la pratique de l'amélioration génétique

Activité - 30 min

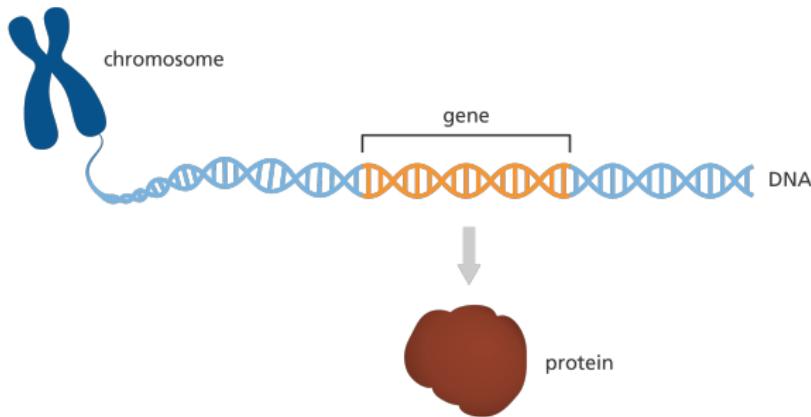
# Rappel des concepts génétiques

ADN : support de l'information génétique.



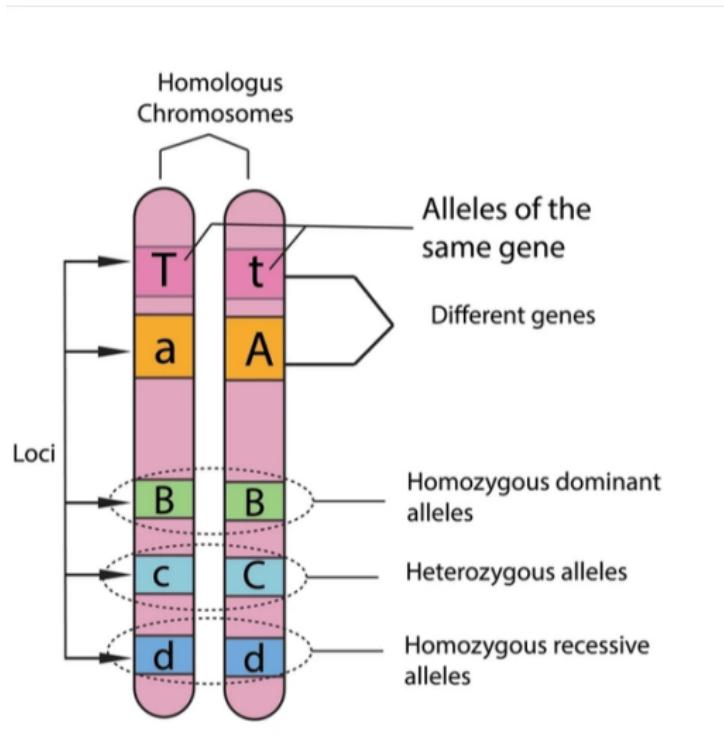
# Rappel des concepts génétiques

**Gène :** unité fonctionnelle d'hérédité.



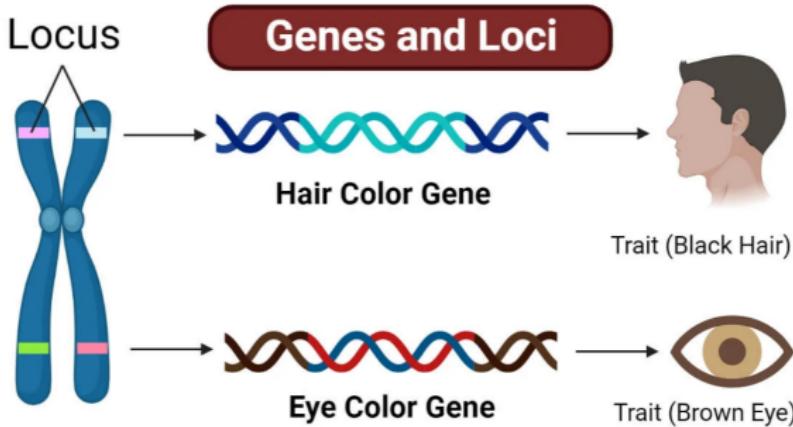
# Rappel des concepts génétiques

Allèle : version alternative d'un même gène.



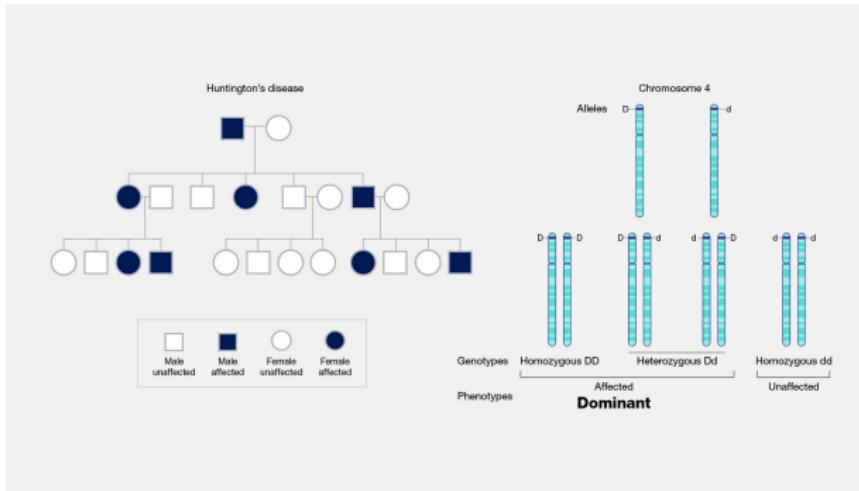
# Rappel des concepts génétiques

Locus : position d'un gène ou allèle sur le génome.



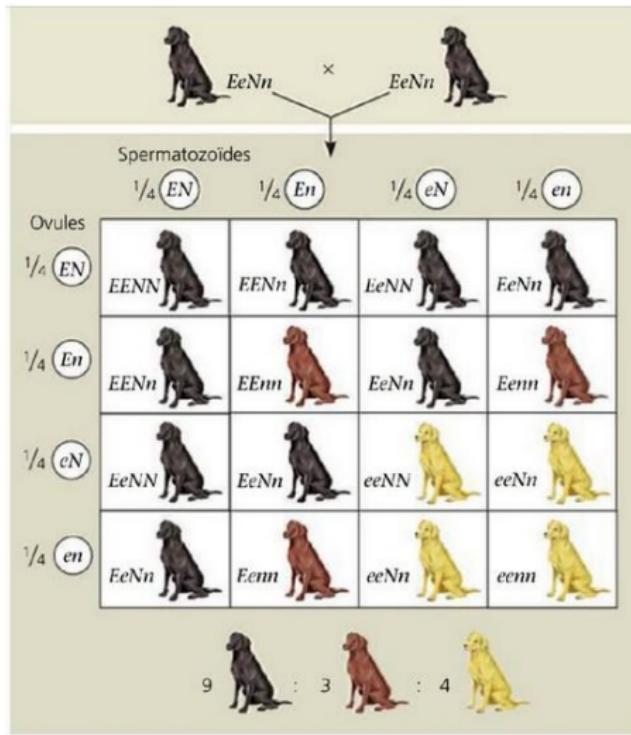
# Rappel des concepts génétiques

Dominance : interaction entre allèles d'un même gène.



# Rappel des concepts génétiques

Épistasie : interaction entre gènes situés à des loci différents.



# Outline

Rappel

Le modèle à un locus

Généralisation du modèle

Conséquences pour la pratique de l'amélioration génétique

Activité - 30 min

# Décomposition de la valeur génétique (G)

**Population** : panmictique, taille infinie, reproduction sexuée.

**Un seul locus  $A$**  influence la valeur génétique.

Génotypes :  $A_iA_j$  avec  $A_iA_j = A_jA_i$ .

# Décomposition de la valeur génétique (G)

**Population** : panmictique, taille infinie, reproduction sexuée.

**Un seul locus  $A$**  influence la valeur génétique.

Génotypes :  $A_iA_j$  avec  $A_iA_j = A_jA_i$ .

- ▶  $P$  : Phénotype

# Décomposition de la valeur génétique (G)

**Population** : panmictique, taille infinie, reproduction sexuée.

**Un seul locus  $A$**  influence la valeur génétique.

Génotypes :  $A_iA_j$  avec  $A_iA_j = A_jA_i$ .

- ▶  $P$  : Phénotype
- ▶  $\mu$  : Moyenne de la population

# Décomposition de la valeur génétique (G)

**Population** : panmictique, taille infinie, reproduction sexuée.

**Un seul locus  $A$**  influence la valeur génétique.

Génotypes :  $A_iA_j$  avec  $A_iA_j = A_jA_i$ .

- ▶  $P$  : Phénotype
- ▶  $\mu$  : Moyenne de la population
- ▶  $G$  : Valeur génétique

# Décomposition de la valeur génétique (G)

**Population** : panmictique, taille infinie, reproduction sexuée.

**Un seul locus  $A$**  influence la valeur génétique.

Génotypes :  $A_iA_j$  avec  $A_iA_j = A_jA_i$ .

- ▶  $P$  : Phénotype
- ▶  $\mu$  : Moyenne de la population
- ▶  $G$  : Valeur génétique
- ▶  $E$  : Effet de l'environnement

# Décomposition de la valeur génétique (G)

**Population** : panmictique, taille infinie, reproduction sexuée.

**Un seul locus A influence la valeur génétique.**

Génotypes :  $A_iA_j$  avec  $A_iA_j = A_jA_i$ .

- ▶  $P$  : Phénotype
- ▶  $\mu$  : Moyenne de la population
- ▶  $G$  : Valeur génétique
- ▶  $E$  : Effet de l'environnement

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

# Décomposition de la valeur génétique (G)

Valeur génétique d'un génotype  $A_iA_j$  :

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

$$\begin{aligned} G = & \text{ effet moyen du gène paternel} & + \\ & \text{ effet moyen du gène maternel} & + \\ & \text{ effet d'interaction entre (dominance)} \end{aligned}$$

où :

# Décomposition de la valeur génétique (G)

Valeur génétique d'un génotype  $A_iA_j$  :

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

$$\begin{aligned} G = & \text{ effet moyen du gène paternel} & + \\ & \text{ effet moyen du gène maternel} & + \\ & \text{ effet d'interaction entre (dominance)} \end{aligned}$$

où :

- ▶  $\alpha_i$  : effet additif de l'allèle  $A_i$ ;

# Décomposition de la valeur génétique (G)

Valeur génétique d'un génotype  $A_iA_j$  :

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

$$\begin{aligned} G = & \text{ effet moyen du gène paternel} & + \\ & \text{ effet moyen du gène maternel} & + \\ & \text{ effet d'interaction entre (dominance)} \end{aligned}$$

où :

- ▶  $\alpha_i$  : effet additif de l'allèle  $A_i$ ;
- ▶  $\alpha_j$  : effet additif de l'allèle  $A_j$

# Décomposition de la valeur génétique (G)

Valeur génétique d'un génotype  $A_iA_j$  :

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

$$\begin{aligned} G = & \text{ effet moyen du gène paternel} & + \\ & \text{ effet moyen du gène maternel} & + \\ & \text{ effet d'interaction entre (dominance)} \end{aligned}$$

où :

- ▶  $\alpha_i$  : effet additif de l'allèle  $A_i$ ;
- ▶  $\alpha_j$  : effet additif de l'allèle  $A_j$ ;
- ▶  $\beta_{ij}$  : déviation due à la dominance (interaction  $A_iA_j$ )

## Effet moyen associé à un allèle ( $\alpha_i$ )

L'effet moyen de l'allèle  $A_i$  (average effect, Fisher) est défini comme :

$$\alpha_i = E_j(G_{ij}) = \sum_j p_j G_{ij}$$

- ▶  $p_j$  : fréquence de l'allèle  $A_j$
- ▶ Panmixie :  $j$  est tiré indépendamment de  $i$
- ▶  $\alpha_i$  a par construction une **moyenne nulle** dans la population

## Effet de dominance ( $\beta_{ij}$ )

L'effet de dominance est défini comme l'interaction entre les deux allèles :

$$\beta_{ij} = G_{ij} - \alpha_i - \alpha_j$$

- ▶ Moyenne nulle :  $E_j(\beta_{ij}) = 0$
- ▶ Par construction,  $\alpha_i$  et  $\beta_{ij}$  sont **indépendants**
- ▶ Covariances nulles entre effets tirés indépendamment :

$$\text{cov}(\alpha_i, \beta_{ij}) = 0,$$

# Indépendance des effets alléliques

Sous panmixie, les deux allèles  $i$  et  $j$  d'un individu sont tirés indépendamment :

$$\text{cov}(\alpha_i, \alpha_j) = 0$$

De même :

$$E(\alpha_j) = E(\beta_{ij}) = 0$$

# Variance génétique

La variance génétique d'un génotype est :

$$V_G = V(G_{ij}) = V(\alpha_i) + V(\alpha_j) + V(\beta_{ij}) \quad (\text{covariances nulles})$$

$$V_G = 2V(\alpha) + V(\beta)$$

On définit :

$$V_A = 2V(\alpha) = 2E(\alpha_i^2) \quad (\text{variance additive})$$

$$V_D = V(\beta) = E(\beta_{ij}^2) \quad (\text{variance de dominance})$$

Donc :

$$V_G = V_A + V_D$$

## Héritabilité au sens strict

L'héritabilité au sens strict ( $h_{ss}^2$ ) est la proportion de la variance phénotypique due à la **variance génétique additive** :

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

où  $V_P$  est la variance phénotypique totale.

# Outline

Rappel

Le modèle à un locus

Généralisation du modèle

Conséquences pour la pratique de l'amélioration génétique

Activité - 30 min

## Généralisation à plusieurs locus

$$G_{I_{ij}} = \alpha_{I_i} + \alpha_{I_j} + \beta_{I_{ij}}$$

- ▶  $\alpha_{I_i}$  : effet additif de l'allèle  $i$  au locus  $I$
- ▶  $\alpha_{I_j}$  : effet additif de l'allèle  $j$  au locus  $I$
- ▶  $\beta_{I_{ij}}$  : déviation due à la dominance  $ij$  au locus  $I$

## Généralisation à plusieurs locus

- ▶ En l'absence d'épistasie et pour des loci en équilibre de liaison : **les variances s'additionnent** pour un caractère quantitatif.
- ▶ Variance additive totale :

$$V_A = 2 \sum_{\text{locus}} V(\alpha_{\text{loc}}) = 2 \sum_{\text{locus}} E(\alpha_{i,\text{loc}}^2)$$

- ▶ Variance de dominance totale :

$$V_D = \sum_{\text{locus}} V(\beta_{\text{loc}}) = \sum_{\text{locus}} E(\beta_{ij,\text{loc}}^2)$$

# Signification génétique des effets additifs

- ▶ **Effet moyen d'un allèle** : valeur moyenne d'un allèle en combinaison avec tous les autres.
- ▶ Dépend de l'effet biologique propre et des **fréquences alléliques**.
- ▶ Les variances **diffèrent selon la population** et au cours du temps.
- ▶ On peut utiliser un **modèle biologique** pour fixer les contributions des allèles/génotypes.

# Cas d'une population biallélique (Fisher)

- ▶ Population panmictique, 1 locus, 2 allèles : B et b, fréquences  $p$  et  $q$
- ▶ Trois génotypes possibles :

$$BB : p^2, \quad Bb : 2pq, \quad bb : q^2$$

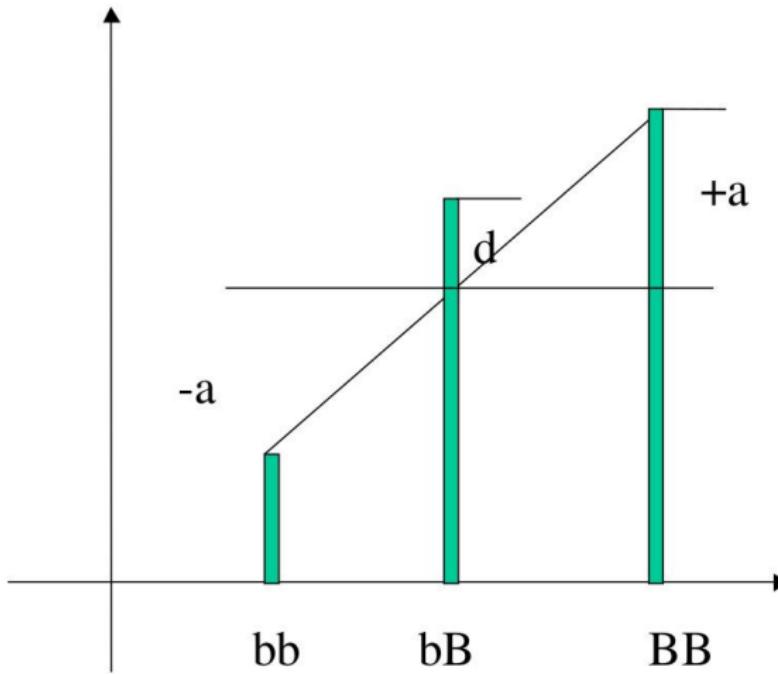
- ▶ Valeurs phénotypiques moyennes arbitraires mais fixes :  
 $Y_{BB}, Y_{Bb}, Y_{bb}$

# Cas d'une population biallélique (Fisher)

$$c = \frac{Y_{BB} + Y_{bb}}{2}, \quad a = \frac{Y_{BB} - Y_{bb}}{2}, \quad d = Y_{Bb} - c$$

- ▶  $d$  : paramètre lié à la **dominance** de B sur b
- ▶ Degré de dominance :  $d/a$ 
  - ▶  $d/a = 1$  : dominance complète
  - ▶  $0 < d/a < 1$  : dominance partielle
  - ▶  $d/a > 1$  : superdominance

## Cas d'une population biallélique (Fisher)



# Phénotype et valeur génétique

$$Y_{BB} = \mu + G_{BB} = \mu + 2\alpha_B + \beta_{BB}$$

$$Y_{Bb} = \mu + G_{Bb} = \mu + \alpha_B + \alpha_b + \beta_{Bb}$$

$$Y_{bb} = \mu + G_{bb} = \mu + 2\alpha_b + \beta_{bb}$$

$$\mu = p^2 Y_{BB} + 2pq Y_{Bb} + q^2 Y_{bb} = c + a(p - q) + 2pqd$$

## Effets moyens ( $\alpha_B, \alpha_b$ )

$$\alpha_B = pY_{BB} + qY_{Bb} - \mu = q\alpha, \quad \alpha_b = -p\alpha$$

- ▶ **Remarque :** L'effet moyen dépend des fréquences alléliques
- ▶ Même allèle, même effet biologique, peut avoir un **effet moyen différent** selon la population

## Termes de dominance ( $\beta_{ij}$ )

$$\beta_{BB} = -2q^2d, \quad \beta_{Bb} = 2pqd, \quad \beta_{bb} = -2p^2d$$

- ▶ Déviation entre valeurs **additives prédictives** et **valeurs réelles**
- ▶ Mesure de la dominance dans le modèle génétique

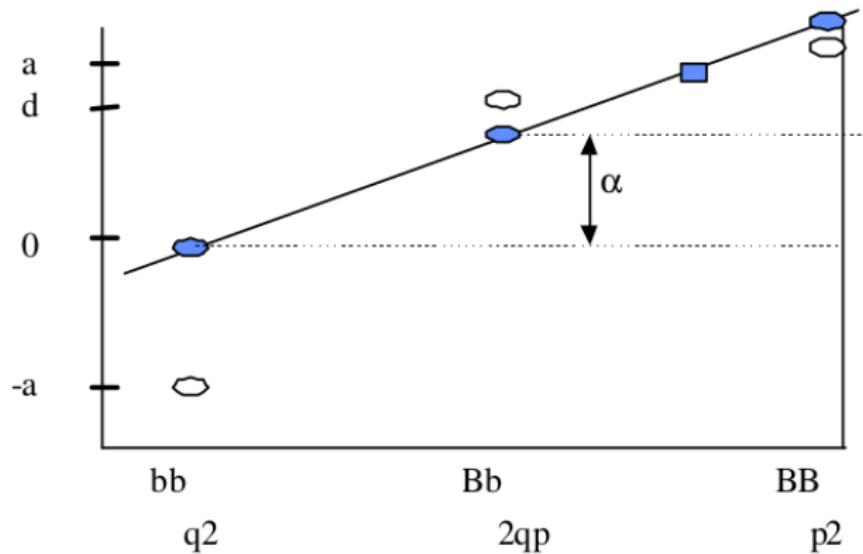
# Régression géno-phénotypique

- ▶ Dose d'allèle B : 0 (bb), 1 (Bb), 2 (BB)
- ▶ Approximation additive :

$$\text{phénotype prédit} = 2\alpha \times \text{dose}$$

- ▶ **Pente de la régression** =  $\alpha = \alpha_B - \alpha_b$
- ▶ La droite de régression passe près des génotypes les plus fréquents

# Régression géno-phénotypique



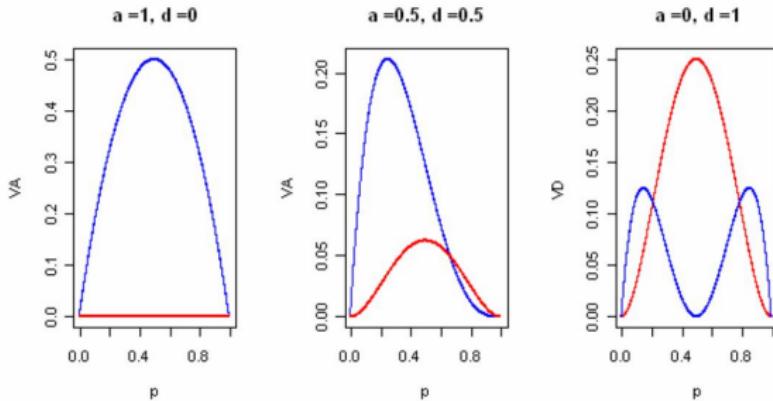
# Composantes de la variance

$$\sigma_A^2 = 2E(\alpha_i^2) = 2(p\alpha_B^2 + q\alpha_b^2) = 2pq\alpha^2$$

$$\sigma_D^2 = E(\beta_{ij}^2) = 4p^2q^2d^2$$

# Composantes de la variance

Ci-dessous quelques exemples de changement des différentes composantes de la variance en fonction des fréquences alléliques



- ▶ La variance additive est en bleu, la variance de dominance est en rouge.

## Interprétation et implications

- ▶ Variance additive maximale lorsque les fréquences sont équilibrées
- ▶ Superdominance ou dominance complète : variance peut s'annuler
- ▶ Relation **modèle biologique vs modèle statistique** n'est pas triviale
- ▶ Importance pour la **cartographie génétique, QTL, et prédition de valeur génomique**

# Outline

Rappel

Le modèle à un locus

Généralisation du modèle

**Conséquences pour la pratique de l'amélioration génétique**

Activité - 30 min

# Conséquences pour la pratique de l'amélioration génétique

## Conséquences pour la pratique de l'amélioration génétique

- ▶ Sélection sur la valeur génétique additive : Rechercher les individus ayant la meilleure valeur additive

## Conséquences pour la pratique de l'amélioration génétique

- ▶ **Sélection sur la valeur génétique additive** : Rechercher les individus ayant la meilleure valeur additive
- ▶ **Sélection sur les effets d'interaction** : Rechercher les couples dont la descendance présentera des effets de dominance ou d'interaction maximaux

## Conséquences pour la pratique de l'amélioration génétique

- ▶ **Sélection sur la valeur génétique additive** : Rechercher les individus ayant la meilleure valeur additive
- ▶ **Sélection sur les effets d'interaction** : Rechercher les couples dont la descendance présentera des effets de dominance ou d'interaction maximaux
- ▶ **Complémentarité des deux voies** : Dans la pratique, ces deux approches se complètent. Exemple :
  1. Sélection des parents à forte valeur additive
  2. Recherche des meilleurs couples/parents parmi les géniteurs retenus

# Outline

Rappel

Le modèle à un locus

Généralisation du modèle

Conséquences pour la pratique de l'amélioration génétique

Activité - 30 min

# Activité

- Code R