

Réponse à la sélection

Jemay Salomon

UMR GQE Le Moulon

Université Paris-Saclay, INRAE, CNRS, AgroParisTech

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du sélectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Pourquoi étudier la réponse à la sélection ?

Sélection artificielle et évolution

- ▶ En amélioration des plantes et des animaux, la sélection vise à
 - ▶ améliorer la valeur moyenne d'une population,
 - ▶ orienter l'évolution des caractères dans une direction désirée,
 - ▶ en choisissant les parents de la génération suivante,
 - ▶ et en contrôlant les schémas de reproduction (autofécondation, allofécondation).
- ▶ En populations naturelles ou en pré-breeding, l'objectif est de
 - ▶ comprendre comment la sélection modifie les caractères,
 - ▶ décrire les trajectoires évolutives,
 - ▶ identifier les traits sur lesquels la sélection agit le plus fortement.

Lynch & Walsh (1998), Gallais (1990)

Prédire l'effet de la sélection

Un enjeu central en génétique quantitative

- ▶ Un problème fondamental est de **prédir la valeur des descendants** issus des individus sélectionnés.
- ▶ À partir du cadre statistique vu dans les cours précédents, on cherche à prédire
 - ▶ le changement de moyenne phénotypique,
 - ▶ la réponse de la population à la sélection.
- ▶ Ces prédictions permettent :
 - ▶ d'optimiser les schémas de sélection en amélioration,
 - ▶ de comprendre les facteurs évolutifs gouvernant l'évolution des caractères,
 - ▶ y compris l'évolution de la fitness elle-même.

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Modélisation théorique

- ▶ On considère **un seul caractère** mesuré dans une population composée de I génotypes distincts.
- ▶ À chaque génotype $i \in \{1, \dots, I\}$ est associée une **valeur génotypique** g_i .
- ▶ Dans la théorie classique de la génétique quantitative :

$$g_i = a_i + d_i + \zeta_i$$

- ▶ où :
 - ▶ a_i : effet **additif** (valeur transmissible),
 - ▶ d_i : effet de **dominance**,
 - ▶ ζ_i : effet d'**épistasie**.
- ▶ Ces composantes sont supposées **indépendantes**.
- ▶ La composante additive a_i est la seule transmise en moyenne à la descendance : elle est appelée *breeding value*.

Hypothèse infinitésimale et variance additive

- ▶ Le caractère est supposé contrôlé par un grand nombre de loci ayant chacun un effet faible : **hypothèse infinitésimale**.
- ▶ Sous cette hypothèse :

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

- ▶ En particulier, pour la composante additive :

$$a_i \sim \mathcal{N}(0, \sigma_a^2)$$

où σ_a^2 est la **variance génétique additive**.

Hypothèse infinitésimale et variance additive

- ▶ En écriture multivariée :

$$\mathbf{a} \sim \mathcal{N}_I(0, \sigma_a^2 A)$$

- ▶ **A est la matrice de relations génétiques additives, dérivable :**
 - ▶ du pédigrée (espérance mendélienne),
 - ▶ ou directement des données de génotypage.
- ▶ Tant que $\sigma_a^2 > 0$, la sélection des individus ayant les plus grandes valeurs additives $a_i^{(s)}$ permet d'augmenter la moyenne génotypique au fil des générations.

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Modélisation statistique

- ▶ Les valeurs génotypiques $\{g_i\}$ ne sont **pas observables directement**.
- ▶ On dispose uniquement d'observations phénotypiques :

$$\{y_n\}_{1 \leq n \leq N}, \quad \mathbb{E}[y_n] = \mu_0, \quad \text{Var}[y_n] = \sigma_p^2$$

- ▶ Hypothèses statistiques de base :
 - ▶ données indépendantes,
 - ▶ distribution Normale,
 - ▶ absence de covariance entre erreurs.
- ▶ La vraisemblance marginale s'écrit alors :

$$y_n \mid \mu_0, \sigma_p^2 \stackrel{\text{i.i.d}}{\sim} \mathcal{N}(\mu_0, \sigma_p^2)$$

- ▶ Objectif : utiliser les phénotypes pour **estimer ou prédire** les effets génétiques.

Modélisation statistique

- ▶ En pratique, on observe plusieurs répétitions par génotype :

$$N = I \times J$$

- ▶ Le modèle phénotypique s'écrit :

$$y_{ij} = \mu + g_i + \epsilon_{ij}$$

avec :

$$\epsilon_{ij} \mid \sigma^2 \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$$

- ▶ Conditionnellement aux effets génétiques :

$$y_{ij} \mid \mu, g_i, \sigma^2 \sim \mathcal{N}(\mu + g_i, \sigma^2)$$

- ▶ g_i est traité comme un **effet aléatoire** : on est en présence d'un **modèle linéaire mixte**.

Variances et héritabilité

- ▶ Après intégration des effets génétiques :

$$y_{ij} \mid \mu_0, \sigma_g^2, \sigma^2 \sim \mathcal{N}(\mu_0, \sigma_g^2 + \sigma^2)$$

- ▶ Décomposition de la variance phénotypique :

$$\sigma_p^2 = \sigma_g^2 + \sigma^2$$

- ▶ Les composantes de variance peuvent être estimées par *ReML*.
- ▶ Les prédictions des effets génétiques g_i reposent sur les équations de Henderson (BLUP).

Variances et héritabilité

- ▶ Corrélation entre génotype et phénotype :

$$\rho_{g,y} = \frac{\text{Cov}(g, y)}{\sigma_g \sigma_p}$$

- ▶ En l'absence de covariance génotype–environnement :

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2}$$

- ▶ H^2 mesure la capacité du dispositif expérimental à révéler les différences génétiques.

Relation entre parents et descendants

- ▶ On cherche à relier le phénotype des enfants au phénotype de leurs parents.
- ▶ On considère :
 - ▶ $y_{\text{mère}}$: phénotype de la mère,
 - ▶ $y_{\text{père}}$: phénotype du père,
 - ▶ $y = \frac{y_{\text{mère}} + y_{\text{père}}}{2}$: **parent moyen**,
 - ▶ y_e : phénotype de l'enfant.

Relation entre parents et descendants

- ▶ Décomposition des phénotypes parentaux :

$$y_{\text{mère}} = a_{\text{mère}} + d_{\text{mère}} + \zeta_{\text{mère}} + \epsilon_{\text{mère}}$$

$$y_{\text{père}} = a_{\text{père}} + d_{\text{père}} + \zeta_{\text{père}} + \epsilon_{\text{père}}$$

- ▶ Le phénotype de l'enfant s'écrit :

$$y_e = a_{\text{mère}} + a_{\text{père}} + \epsilon_e$$

- ▶ Seule la composante **additive** est transmise à la descendance.

Régression enfants–parents

- ▶ Pour caractériser la relation parents–enfants, on considère la régression linéaire :

$$y_e = \alpha + \beta_{\text{enfants,parents}} y + \eta$$

- ▶ La pente de la droite est donnée par :

$$\beta_{\text{enfants,parents}} = \frac{\text{Cov}[y, y_e]}{\text{Var}[y]}$$

- ▶ Hypothèses clés :
 - ▶ panmixie (accouplements aléatoires),
 - ▶ absence de sélection,
 - ▶ absence de covariance génotype–environnement,
 - ▶ absence de transmission d'effets environnementaux.
- ▶ Sous ces hypothèses, de nombreux termes de covariance s'annulent.

Covariance parents–enfants

- ▶ La covariance s'écrit :

$$\text{Cov}[y, y_e] = \text{Cov} \left[\frac{a_{\text{mère}} + a_{\text{père}}}{2}, a_{\text{mère}} + a_{\text{père}} \right]$$

- ▶ En utilisant l'indépendance entre parents :

$$\text{Cov}[y, y_e] = \frac{\text{Var}[a_{\text{mère}}] + \text{Var}[a_{\text{père}}]}{2}$$

- ▶ En panmixie, la variance génétique additive totale est :

$$\sigma_a^2 = \text{Var}[a_{\text{mère}}] + \text{Var}[a_{\text{père}}]$$

- ▶ Ainsi :

$$\text{Cov}[y, y_e] = \frac{\sigma_a^2}{2}$$

Héritabilité au sens strict

- ▶ La variance du parent moyen est :

$$\text{Var}[y] = \frac{\text{Var}[y_{\text{mère}}] + \text{Var}[y_{\text{père}}]}{4} = \frac{\sigma_p^2}{2}$$

- ▶ La pente de la régression devient :

$$\beta_{\text{enfants,parents}} = \frac{\sigma_a^2/2}{\sigma_p^2/2} = \frac{\sigma_a^2}{\sigma_p^2}$$

- ▶ On définit :

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2} \quad (\text{héritabilité au sens strict})$$

- ▶ h^2 est aussi la corrélation entre phénotype et valeur génétique additive :

$$\rho_{a,y} = h$$

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du sélectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

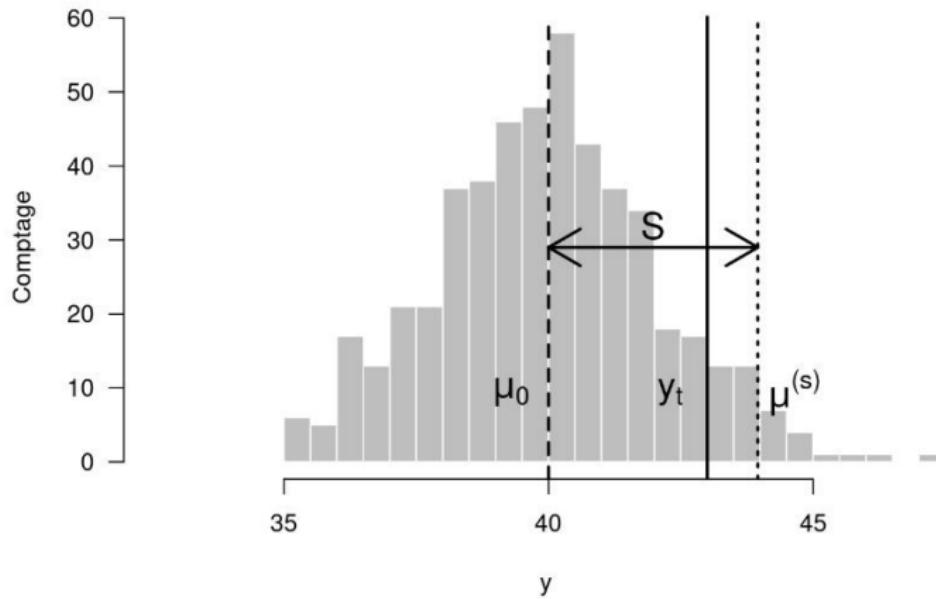
Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Differentiel de selection

Phénotypes des parents et différentiel de sélection



$$S = \mu^{(s)} - \mu_0$$

Intensité de sélection

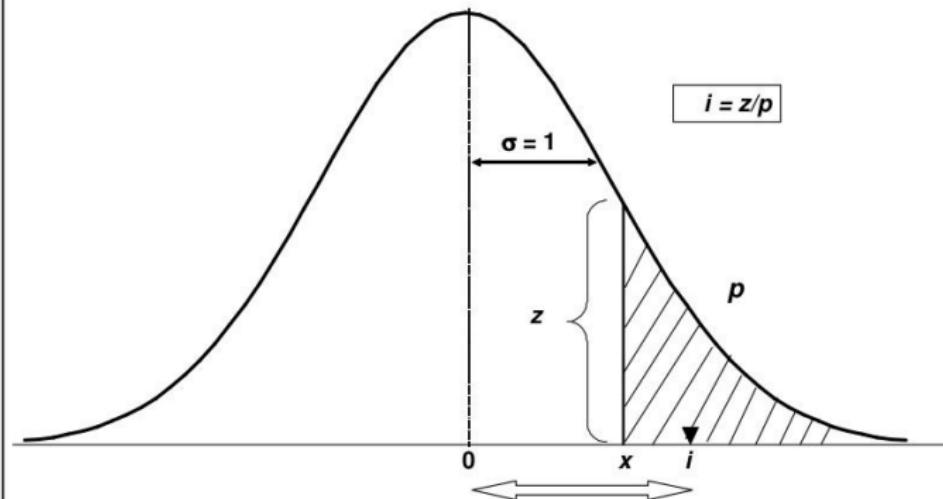
- ▶ S est dépend de l'unité de mesure du phénotype.

Pour comparer la sélection sur différents caractères, il est donc recommandé de travailler avec une valeur standardisée, l'*intensité de sélection*, notée i (à ne pas confondre avec l'indice i du modèle statistique):

$$i = \frac{S}{\sigma_p}$$

Intensité de sélection

Derivation de l'intensité de la sélection dans une loi normale



Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Réponse à la sélection : idée centrale

- ▶ L'héritabilité au sens strict vérifie :

$$0 \leq h^2 \leq 1$$

- ▶ La droite de régression des enfants sur les parents est donc **moins pentue que la droite identité**.
- ▶ Cette observation est à l'origine du terme de *régression vers la moyenne* (Galton).
- ▶ Notons :
 - ▶ μ_0 : moyenne phénotypique de la population parentale,
 - ▶ $\mu^{(s)}$: moyenne des parents sélectionnés,
 - ▶ μ_1 : moyenne phénotypique de la descendance.
- ▶ On définit :

$$S = \mu^{(s)} - \mu_0 \quad (\text{différentiel de sélection})$$

$$R = \mu_1 - \mu_0 \quad (\text{réponse à la sélection})$$

- ▶ La sélection des parents induit donc un changement de moyenne chez les descendants.

Équation du sélectionneur

- ▶ La pente de la régression enfants-parents s'écrit :

$$\beta_{\text{enfants,parents}} = \frac{\mu_1 - \mu_0}{\mu^{(s)} - \mu_0} = \frac{R}{S}$$

- ▶ Or, on a montré précédemment que :

$$\beta_{\text{enfants,parents}} = h^2$$

- ▶ On obtient alors la relation fondamentale :

$$R = h^2 S$$

- ▶ Cette équation combine :
 - ▶ une information d'hérédité (h^2),
 - ▶ une information de sélection (S),
 - ▶ pour prédire le changement intergénérationnel (R).
- ▶ Elle est connue sous le nom d'**équation du sélectionneur**.

Interprétation et optimisation de la sélection

- ▶ Le différentiel de sélection peut s'écrire :

$$S = i \sigma_p$$

où i est l'intensité de sélection.

- ▶ La réponse devient :

$$R = i h^2 \sigma_p = i h \sigma_a$$

- ▶ Le terme $h = \frac{\sigma_a}{\sigma_p}$ est aussi :

$$h = \rho_{a,y}$$

la corrélation entre phénotype observé et valeur génétique additive, appelée **précision** ou *accuracy* (r).

- ▶ On peut donc écrire :

$$R = i r \sigma_a$$

Interprétation et optimisation de la sélection

- ▶ En résumé :

Réponse = intensité × précision × variance additive

- ▶ Une sélection trop intense augmente R à court terme, mais réduit la variance génétique disponible à long terme.

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Effet de la sélection sur la variance génétique

- ▶ **Question clé :** La sélection modifie les fréquences alléliques aux locus impliqués dans la variation d'un trait. Peut-on s'attendre à une modification rapide de la variance génétique d'une population ?

Effet de la sélection sur la variance génétique

- ▶ Question clé : La sélection modifie les fréquences alléliques aux locus impliqués dans la variation d'un trait. Peut-on s'attendre à une modification rapide de la variance génétique d'une population ?
- ▶ Pour un locus à effets additifs : Si l'effet de chaque allèle est faible, le changement de variance au locus est très limité sur quelques générations.

Effet de la sélection sur la variance génétique

- ▶ Question clé : La sélection modifie les fréquences alléliques aux locus impliqués dans la variation d'un trait. Peut-on s'attendre à une modification rapide de la variance génétique d'une population ?
- ▶ Pour un locus à effets additifs : Si l'effet de chaque allèle est faible, le changement de variance au locus est très limité sur quelques générations.
- ▶ Pour un caractère polygénique (somme de plusieurs locus) : La sélection crée une covariance négative entre locus non liés.

Effet de la sélection sur la variance génétique

- ▶ Question clé : La sélection modifie les fréquences alléliques aux locus impliqués dans la variation d'un trait. Peut-on s'attendre à une modification rapide de la variance génétique d'une population ?
- ▶ Pour un locus à effets additifs : Si l'effet de chaque allèle est faible, le changement de variance au locus est très limité sur quelques générations.
- ▶ Pour un caractère polygénique (somme de plusieurs locus) : La sélection crée une covariance négative entre locus non liés.
- ▶ Cette covariance négative réduit la variance additive totale observée dans la population. C'est l'**effet Bulmer**.

Effet de la sélection sur la variance génétique

- ▶ Question clé : La sélection modifie les fréquences alléliques aux locus impliqués dans la variation d'un trait. Peut-on s'attendre à une modification rapide de la variance génétique d'une population ?
- ▶ Pour un locus à effets additifs : Si l'effet de chaque allèle est faible, le changement de variance au locus est très limité sur quelques générations.
- ▶ Pour un caractère polygénique (somme de plusieurs locus) : La sélection crée une covariance négative entre locus non liés.
- ▶ Cette covariance négative réduit la variance additive totale observée dans la population. C'est l'**effet Bulmer**.
- ▶ Cette diminution de variance n'est pas définitive : elle est stockée dans le **déséquilibre de liaison** et peut se relâcher si la pression de sélection diminue.

Effet de la sélection sur la variance génétique

- ▶ En résumé :
 - ▶ La sélection modifie les fréquences alléliques très progressivement à chaque locus.
 - ▶ La variance diminue surtout à cause de covariances négatives entre loci.
 - ▶ La variance totale peut se récupérer avec le temps et le recombinage.

Outline

Motivation

Réponse à la sélection

Modélisation théorique

Modélisation statistique

Equation du selectionneur

Differentiel de selection

Réponse à la sélection

Effet de la selection sur la variance

Activité - 45-60 min

Activité

- Code R