

IV. La ressemblance entre apparentés

La ressemblance entre apparentés est un phénomène de constatation courante chez l'homme. Cependant, les proverbes « tel père, tel fils » et « à père avare, fils prodigue » montrent bien la difficulté d'appréhender correctement l'hérédité de caractères complexes. C'est par l'analyse statistique de la ressemblance entre apparentés que, dès la fin du XIX^{ème} siècle, les biométriciens ont tenté de mettre à jour les bases de l'hérédité des caractères quantitatifs. On rappelle que deux individus sont apparentés s'ils possèdent au moins un ancêtre commun. Quand les généalogies sont connues, il est possible de créer dans une population des groupes au sein desquels l'apparentement est homogène : par exemple, des familles de pleins frères-sœurs, des paires parent-enfant, etc. Cette structuration permet d'estimer le niveau de ressemblance phénotypique entre individus apparentés et de juger s'il y a un lien entre ce niveau de ressemblance et le degré d'apparentement. Dans ce chapitre, nous allons présenter quelques résultats, obtenus sur des populations de différentes espèces, permettant d'établir ce lien. Ensuite, nous montrerons comment le modèle que nous avons développé au chapitre précédent permet de rendre compte de ces résultats et de proposer des lois générales de transmission des caractères quantitatifs. Nous redéfinirons le sens du concept d'héritabilité dans ce contexte et montrerons son utilité, en particulier pour prédire la réponse à la sélection.

A. Observations statistiques au sein de populations

Pour essayer de comprendre comment se transmettent les caractères quantitatifs, la solution la plus immédiate consiste à observer les liens qui existent entre parents et descendants ou ceux qui existent entre des individus d'une même fratrie, c'est-à-dire qui ont un parent en commun (demi-frères-soeurs) ou les deux (pleins frères-soeurs). De fait, ces cas se rencontrent très fréquemment tant chez les plantes que chez les animaux et ce sont des types d'apparentés très souvent analysés. Chez l'homme, on utilise également fréquemment des données mesurées sur des jumeaux. En termes statistiques, le phénomène qui nous intéresse peut être illustré par la distribution des différences de valeur phénotypique entre individus apparentés ; le plus souvent, il est appréhendé par la corrélation et la régression linéaire.

1. La liaison entre membres d'une même fratrie

La figure 14 est une illustration de la ressemblance entre frères chez l'homme en ce qui concerne la taille. On constate tout d'abord qu'il subsiste des écarts phénotypiques pour des individus génétiquement identiques, des vrais jumeaux, élevés ensemble : ces écarts sont imputables, comme nous l'avons vu au chapitre II, à des conditions de micro-milieu différentes au cours du développement des deux jumeaux. Ces écarts entre vrais jumeaux sont faibles néanmoins, avec une moyenne de 1,8 cm et 70% des écarts inférieurs à 2 cm. La moyenne des écarts est de 4,7 cm pour des faux jumeaux et de 4,5 cm pour les pleins frères ; ces deux valeurs, non significativement différentes, sont significativement supérieures à celle relative aux vrais jumeaux. Les écarts de taille sont donc moindres pour des paires de vrais jumeaux que pour des paires génétiquement moins proches. Les écarts observés au sein des paires de pleins frères (faux jumeaux ou non, ce qui est équivalent sur le plan génétique) sont imputables, d'une part, au fait que deux individus ayant même père et même mère reçoivent en partie des gènes différents de leurs parents et, d'autre part, à des facteurs de micro-milieu.

Figure 14. 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 : Stern, 1960.

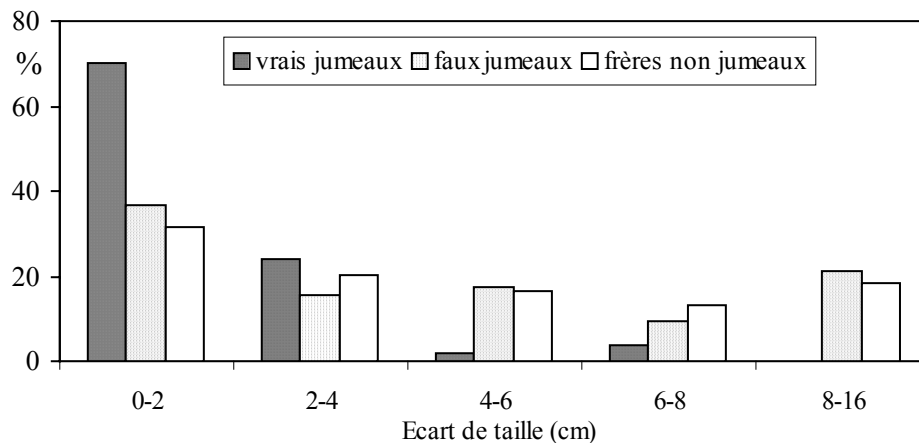
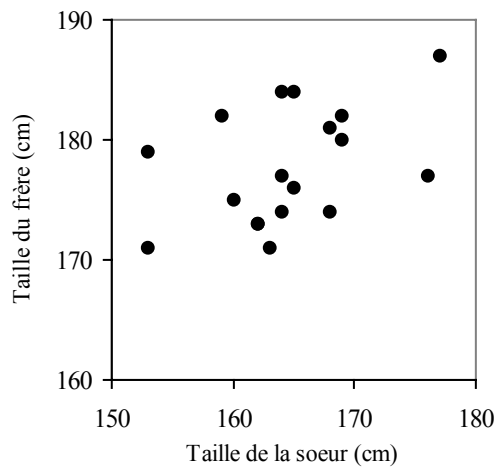


Figure 15. Liaison entre la taille du frère et celle de la sœur chez l'homme (18 familles).

Source : Pearson et Lee, 1902.



La liaison entre membres d'une même fratrie s'analyse plus généralement au travers de la corrélation. Pour mesurer cette corrélation, on peut effectuer un tirage au sort sans remise de paires d'individus, comme suit. Le premier individu de la paire est tiré au hasard dans la population ; le second individu de la paire est tiré au hasard parmi les autres membres de la famille du premier. Quand on a constitué un grand nombre de paires, il est alors possible de calculer la corrélation entre les valeurs observées pour chacun des deux membres d'une paire. La représentation graphique des données issues d'une telle analyse est visible à la figure 15. On note une tendance d'association positive entre les valeurs des frères et celles des sœurs, mais la taille extrêmement réduite de l'échantillon ne permet pas de conclure à ce stade.

Le tableau 10 présente les résultats de corrélation entre valeurs phénotypiques pour différents caractères mesurés chez le porc, en fonction du lien de parenté existant entre les membres de paires constituées comme décrit plus haut. Effectuée sur des effectifs conséquents, cette analyse permet de dégager plusieurs enseignements.

Tableau 10. 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. sign. = signification statistique par rapport à zéro : ns : corrélation non significativement différente de zéro ; * : différence significative au seuil de 5 % ; *** : différence significative au seuil de 1%. Source : INRA, 1993

Caractère mesuré		non apparentés	demi frères-soeurs 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	*
	n	3 836	2 770	2 732
Age à 100 kg	r	0,024	0,068	0,168
	sign.	ns	***	***
	n	3 835	2 770	2 728
Épaisseur moyenne de lard dorsal à 100 kg	r	0,009	0,107	0,238
	sign.	ns	***	***

- Quel que soit le caractère, la corrélation n'est pas significativement différente de zéro pour des individus non apparentés entre eux, comme attendu (ils ont été choisis au hasard et n'ont aucune raison *a priori* de se ressembler). A l'opposé, on observe des corrélations significatives pour les membres d'une même fratrie.
- Pour un lien de parenté donné, la corrélation dépend du caractère mesuré : elle croît du pH musculaire (qui est un facteur de la qualité technologique de la viande) à l'âge à 100 kg (qui traduit la vitesse de croissance) et à l'épaisseur de lard dorsal (caractère qui permet de prédire l'adiposité de la carcasse).
- Pour un caractère donné, la corrélation entre frères-sœurs de même portée est plus élevée qu'entre demi-frères-sœurs (plus que le double). Cette différence s'interprète comme la conséquence d'un apparentement plus fort (deux parents communs au lieu d'un seul) et d'un environnement commun (même mère et même portée).



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.

2. La liaison parent-descendant

Depuis des temps très anciens, l'homme a essayé d'élucider les lois gouvernant la ressemblance des enfants avec leurs parents. Plusieurs auteurs grecs suggéraient que, simplement, les garçons ressemblent à leur père et les filles à leur mère. De façon plus convaincante, on trouve dans certains textes des observations empiriques correspondant à des faits aujourd'hui bien établis. Ainsi, dans l'Odyssée d'Homère, Télémaque, le fils d'Ulysse, se lamente à propos de son impuissance à chasser les prétendants qui occupent la maison de son père ; l'exhortant au courage, Athéna lui dit alors : « *peu d'enfants sont pareils à leur père (quand celui-ci est hors du commun, comme pouvait l'être Ulysse) : la plupart sont pires ; il en est peu qui aient plus de mérite* »(chant II, vers 275-277).

C'est Francis Galton (1822-1911), cousin de Charles Darwin, qui, le premier, chercha à élucider les lois de l'hérédité des caractères quantitatifs en étudiant la liaison entre parent et descendant. Il introduisit le concept de régression pour décrire le fait que les individus exceptionnels donnent des descendants qui, en moyenne, ont des valeurs phénotypiques plus proches de la moyenne que les leurs (fait évoqué dans la citation ci-dessus). Galton a réalisé de nombreuses études et collecté des masses impressionnantes de données en vue de trouver une loi de transmission des caractères au cours des générations. Ses principaux résultats dans ce domaine ont été publiés dans la période 1875-1890. Dans une série d'articles publiés entre 1898 et 1903, un autre biométricien, Karl Pearson, a proposé une formulation mathématique des lois de l'hérédité des caractères quantitatifs mathématiquement plus satisfaisante que celle proposée par Galton lui-même. Toutefois, l'approche biométrique et l'approche mendélienne ont mis du temps à s'unifier et à cohabiter dans l'esprit des biologistes⁵.

Une des études les plus marquantes de Galton a porté sur la taille à l'âge adulte chez l'homme. Dans cette analyse, les données ont été corrigées pour tenir compte du fait qu'en moyenne les hommes sont plus grands que les femmes et chaque couple de parents a été caractérisé par la moyenne des deux individus le composant (notion de « parent moyen »). Sur la base de la valeur moyenne des deux parents, Galton a ensuite constitué des catégories régulièrement espacées, le pas de classe étant de 1 inch (= 2,54 cm).

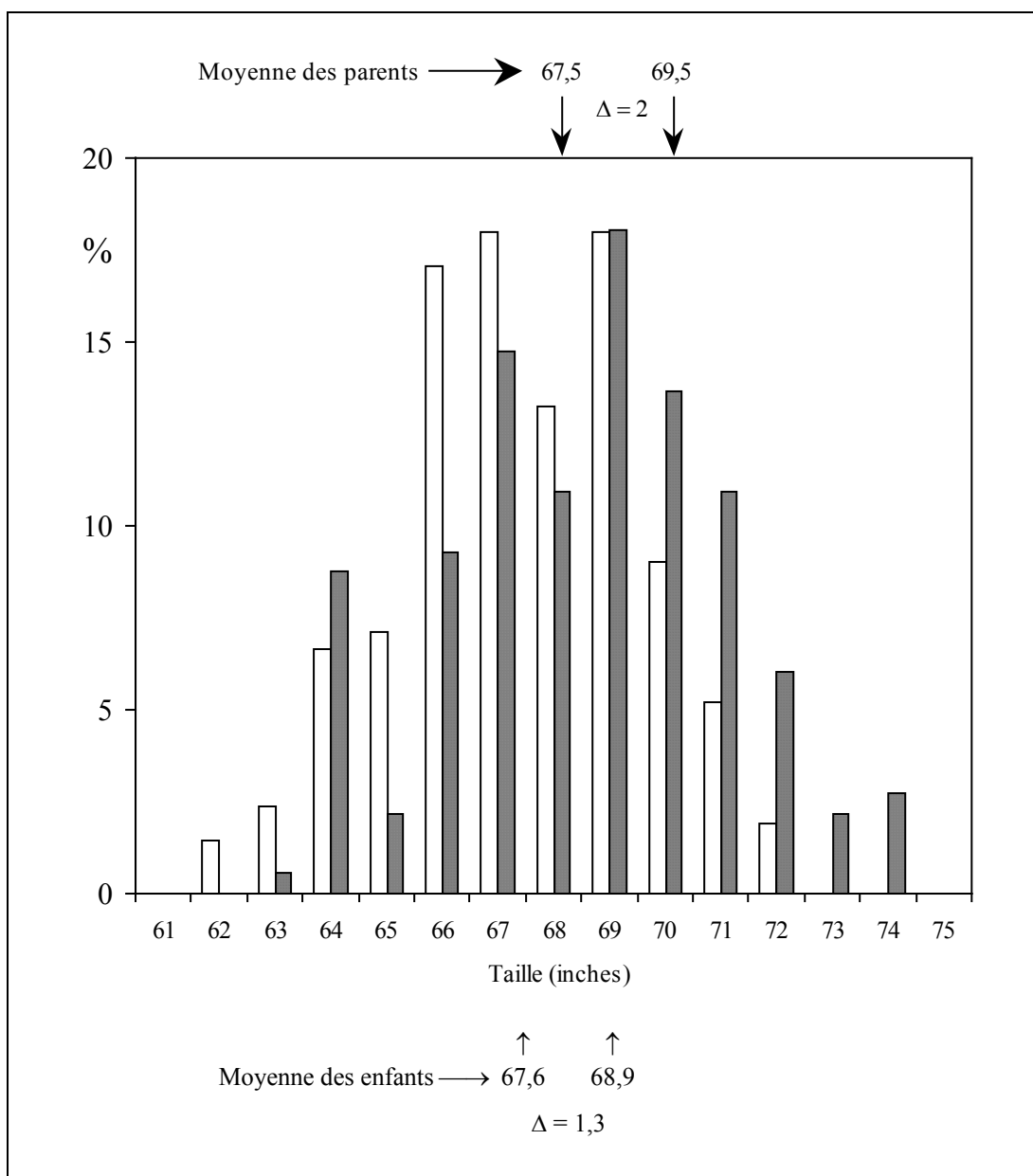
⁵ Pour plus de détails sur ces aspects historiques, voir l'ouvrage de B. Provine, *The origin of theoretical population genetics*, The University of Chicago Press.

Figure 16. 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 66 et 67 inches. Foncé = 183 enfants issus de 41 couples de parents dont la taille moyenne est comprise entre 68 et 69 inches.

La moyenne de chacune des deux catégories parentales est indiquée en haut du graphe, celle des enfants correspondants en bas.

d'après Galton, 1886



La figure 16 présente un extrait des résultats de cette étude, où l'on compare la descendance de deux catégories de parents d'effectifs similaires. Plusieurs points sont à souligner :

- La descendance des deux groupes montre une grande variabilité phénotypique, variabilité qui est du même ordre de grandeur d'une descendance à l'autre.
- Les deux distributions se chevauchent largement : les enfants issus des parents les plus grands ne sont pas toujours plus grands que ceux issus des parents les plus petits.
- Si l'on s'intéresse aux moyennes, on constate alors que la moyenne des descendants des parents les plus grands est plus élevée que celle des parents les plus petits.
- L'écart entre les moyennes de descendants est inférieur à l'écart qui existait entre les moyennes des deux catégories de parents : ici, a été réduit d'environ un tiers.

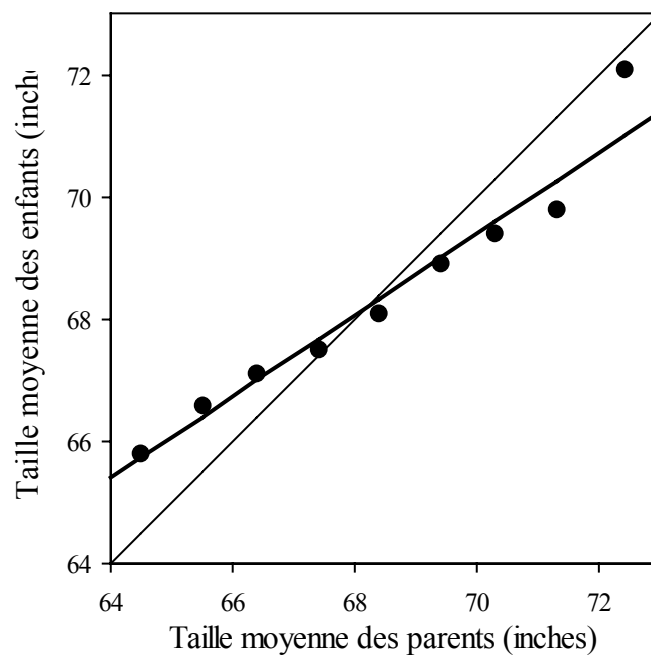
L'utilisation de la régression linéaire pour analyser la liaison parent-descendant permet de généraliser ces résultats et de les exposer de façon plus synthétique. La figure 17 montre ainsi clairement la liaison positive qui existe entre la taille du parent moyen et celle de ses descendants : les parents les plus grands ont en moyenne des descendants supérieurs à la moyenne et vice versa. Toutefois, quand les parents sont supérieurs à la moyenne, leur descendance tend à être inférieure à eux et quand les parents sont inférieurs à la moyenne, leur descendance tend à être supérieure à eux. Ainsi, selon l'expression de Galton, « *il y a une régression des valeurs phénotypiques* » vers la moyenne. Dans le cas précis de la taille chez l'homme, le coefficient de régression de la taille du descendant sur celle de son parent moyen est de $2/3$. D'autres études montrent que la valeur du coefficient de régression dépend largement du caractère mesuré. Ce coefficient de régression peut s'interpréter comme la fraction des différences phénotypiques observées chez les parents qui sont transmises à la descendance. Evidemment, cette tendance générale admet des exceptions : certains individus supérieurs à la moyenne ont aussi des descendants meilleurs qu'eux. Si tel n'était pas le cas, la population tendrait rapidement vers une perte de variabilité phénotypique, tous les individus ayant la même valeur, ce que l'on ne constate pas dans la réalité.



Les écarts phénotypiques qui existent entre parents se retrouvent en partie, mais en partie seulement, entre leurs descendances. L'importance de la fraction transmise à la descendance dépend du caractère étudié.

Figure 17. 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). Chaque cercle représente la moyenne de plusieurs familles. La droite en trait fin est la bissectrice (pente = 1). Celle en trait épais s'ajuste aux points observés (pente $\approx 2/3$), le poids des points extrêmes étant réduit car correspondant à une faible proportion d'individus. La différence entre la droite ajustée et la bissectrice représente, selon l'expression de l'auteur, « *the regression toward mediocrity* », ce qui est traduit de manière à atténuer le jugement de valeur par « la régression vers la moyenne » (le terme « régression » ayant toutefois été maintenu).

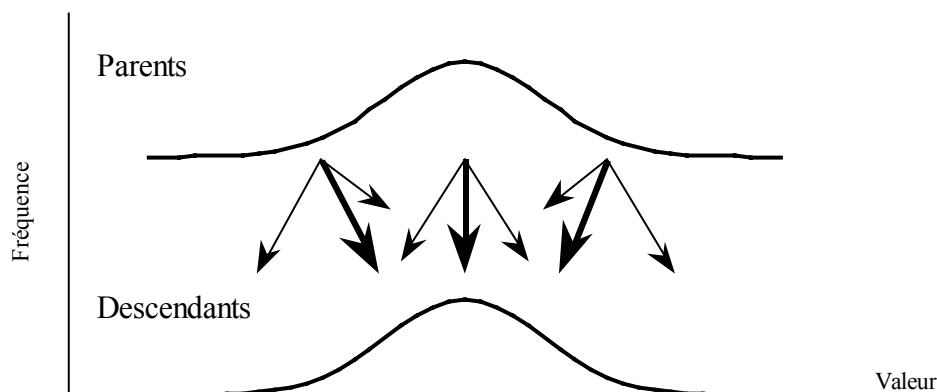
d'après Galton, 1886



L'ensemble de ces résultats peuvent se résumer comme indiqué à la figure 18. Les parents dont la valeur se situe autour de la moyenne ont en général des descendants moyens, les parents supérieurs à la moyenne ont en général des descendants supérieurs à la moyenne mais moins bons qu'eux et les parents inférieurs à la moyenne ont en général des descendants inférieurs à la moyenne mais meilleurs qu'eux. Dans chaque cas, il existe des descendants, en proportion moindre, qui s'écartent notablement de la tendance générale. Ceci permet le maintien de la variabilité d'une génération à l'autre. D'une génération à l'autre, les sources de variation d'origine génétique sont constituées de la variabilité génétique existant au sein des parents, de la manière dont les couples se forment et des hasards de la méiose et des recombinaisons chromosomiques.

Dans une optique de sélection au sein d'espèces domestiques, on peut déduire la conclusion pratique suivante : lorsque l'on choisit bien des parents on n'est jamais totalement sûr d'avoir un bon descendant mais, de cette façon, on a plus de chances d'avoir un bon descendant plutôt qu'en choisissant les parents au hasard.

Figure 18. Représentation schématique du passage d'une génération à l'autre. Les deux courbes se rapportent à la distribution d'un caractère quelconque au sein d'une génération parentale, d'une part, et de la génération de descendants qui en est issue, d'autre part.



B. Prédiction de la covariance entre apparentés

Les observations que nous venons de décrire ont permis de dégager certaines tendances générales concernant l'hérédité des caractères quantitatifs : les descendants ressemblent à leurs parents, sans leur être identiques ; deux individus qui partagent un ou deux parents en commun se ressemblent, sans être identiques entre eux. Nous avons souligné que le degré de ressemblance dépend du lien de parenté et du caractère étudié. Afin d'être en mesure de faire des prédictions, il est nécessaire de quantifier ces phénomènes et de les interpréter selon les termes du modèle développé aux chapitres précédents. La ressemblance entre apparentés est concrètement mise en évidence par la corrélation ou par la régression linéaire. Les coefficients correspondants faisant intervenir un terme de covariance, nous allons voir comment le modèle nous permet de calculer la covariance entre les valeurs phénotypiques d'apparentés. Pour préciser les liens génétiques entre apparentés, nous aurons besoin du coefficient de parenté (voir GP). Nous rappelons, dans l'encadré 5, la définition et le mode de calcul du coefficient de parenté entre deux individus, ainsi que la définition de la consanguinité individuelle, qui est une notion proche mais qu'il est nécessaire de bien distinguer de la notion de parenté.

1. Décomposition de la covariance phénotypique entre deux individus

Selon le modèle que nous avons développé aux chapitres précédents, la valeur phénotypique (P) est la somme d'une valeur génétique (G) et d'une valeur environnementale (E). ***Dans tout ce qui suit, nous supposons que E traduit les seuls effets du micro-milieu***, c'est-à-dire que l'on a pu correctement corriger les données pour tout effet systématique de macro-milieu. Soit deux individus i et j appartenant à la même population, et dont nous analysons les valeurs phénotypiques ; nous avons, en centrant sur la moyenne de population (μ) :

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

Du fait de la répartition aléatoire des différents génotypes dans les différentes conditions de micro-milieu, les variables G et E sont non corrélées ; la covariance entre les deux valeurs phénotypiques s'exprime donc ainsi :

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

Pour deux individus tirés au hasard et indépendamment l'un de l'autre, les trois covariances ci-dessus sont nulles. Il n'apparaîtra de covariances non nulles que si les deux individus i et j ne sont pas tirés indépendamment l'un de l'autre. Ce sera le cas si le degré d'apparentement est utilisé lors de l'échantillonnage, pour constituer des paires d'individus apparentés entre eux : on s'attend alors à observer une covariance génétique non nulle entre les individus des différentes paires ainsi constituées. Il est également possible de faire apparaître une covariance environnementale si les individus d'une même paire partagent un environnement commun de micro-milieu, non corrigé au titre des effets de macro-milieu (voisins immédiats sur une parcelle, dans un bâtiment, etc.). Si, pour des raisons biologiques, des individus apparentés partagent aussi un environnement commun (jeunes mammifères issus d'une même portée, par exemple), la covariance phénotypique aura ses deux composantes non nulles.



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.***

Il est fondamental de pouvoir identifier les cas où la composante d'environnement commun est non négligeable. On est dans cette situation quand un facteur de milieu non contrôlé, ou dont on ne peut pas corriger l'effet, affecte non pas un seul individu mais plusieurs. Ainsi, chez les espèces d'animaux polytoques (où chaque femelle a plusieurs descendants en une seule fois), plusieurs caractères sont soumis à ce que l'on appelle des effets maternels : les caractéristiques de la mère jouent comme un effet de milieu sur les valeurs phénotypiques des jeunes. Chez les mammifères, c'est le cas du poids à la naissance (caractéristiques morphologiques du tractus génital de la mère, métabolisme de la mère), ou de la croissance avant sevrage (comportement maternel et lactation). Pour ces caractères, les individus d'une même portée sont soumis aux mêmes effets maternels, ce qui donc constitue un environnement commun (cf. tableau 10). De la même façon, chez les plantes, la qualité des semences est très dépendante de la plante-mère et influe fortement sur la croissance des plantes-filles au stade jeune. Chez l'homme, enfin, les individus d'une même fratrie partagent dans une large mesure et jusqu'à un âge avancé des conditions de vie communes. L'effet d'environnement commun à des pleins frères-soeurs est en général impossible à corriger du fait de la confusion qu'il y a entre cet effet et les effets des gènes parentaux, qui sont eux aussi en partie partagés par les membres de la fratrie.

Encadré 5

Notions de parenté et de consanguinité - Calcul du coefficient de parenté

Deux individus sont apparentés s'ils ont un ancêtre commun.

Sur le plan génétique, l'apparement de deux individus a pour conséquence la possibilité que ces individus aient reçu chacun, en un locus quelconque, une copie du même gène présent chez l'ancêtre commun. Deux gènes qui sont la copie mendélienne d'un même gène ancêtre sont dits *identiques*.

Un individu est consanguin si ses deux parents sont apparentés.

Sur le plan génétique, la consanguinité a pour conséquence la possibilité que l'individu concerné ait reçu, en un locus quelconque, deux copies du même gène présent chez l'ancêtre commun à ses deux parents.

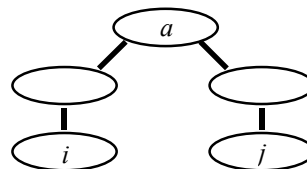
Les coefficients mesurant l'intensité de ces deux phénomènes sont définis pour un locus autosomal non soumis à la sélection ni aux mutations. La définition de ces coefficients s'appuie sur l'approche probabiliste introduite par Malécot (1948). **Le coefficient de parenté** entre deux individus i et j (Φ_{ij}) est la probabilité pour que deux gènes tirés au hasard au même locus, l'un chez i et l'autre chez j , soient identiques. **Le coefficient de consanguinité** d'un individu x (F_x) est égal à la probabilité pour qu'en un locus les deux gènes qui s'y trouvent soient identiques. Les deux gènes de x étant issus d'un tirage au hasard d'un gène parmi les deux de son père et d'un gène parmi les deux de sa mère, le coefficient de consanguinité d'un individu est égal au coefficient de parenté entre ses deux parents ; soit x de père p et de mère m , nous avons :

$$F_x = \Phi_{pm}$$

Ne pas confondre ces deux notions. La parenté concerne des paires d'individus, la consanguinité concerne des individus seuls. La confusion est fréquente car dans le langage courant, consanguin désigne le fait de descendre d'une même « souche ». Par exemple, le mariage de deux cousins germains est très souvent qualifié de mariage consanguin. Les généticiens parlent de mariage entre apparentés et réservent le terme consanguin pour les enfants nés d'un tel mariage.

Les coefficients de parenté et de consanguinité étant définis pour un locus neutre, leur calcul repose sur la seule connaissance des généalogies des individus. Les résultats que l'on peut obtenir sur des populations réelles, et en particulier domestiques, dépendent toutefois de l'ancienneté et de la fiabilité du recueil des généalogies (état civil chez les animaux et chez l'homme). La méthode que nous présentons n'est qu'une des méthodes disponibles pour le calcul de ces coefficients.

Deux individus apparentés sont reliés l'un à l'autre par une suite d'individus, comprenant l'ancêtre commun et constituant une **chaîne de parenté** (cf schéma ci-dessous). Une chaîne de parenté ne peut pas passer deux fois par un même individu. Une chaîne de parenté n'admet qu'un seul changement de direction dans l'échelle de temps, au niveau de l'ancêtre commun.



Pour une chaîne de parenté donnée, la probabilité d'identité $\Pr(\equiv)$ d'un gène de i et d'un gène de j dépend du nombre (n) d'individus sur la chaîne (i et j compris) et du coefficient de consanguinité de l'ancêtre (F_a) :

$$\Pr(\equiv) = \left(\frac{1}{2}\right)^n (1 + F_a)$$

Encadré 5

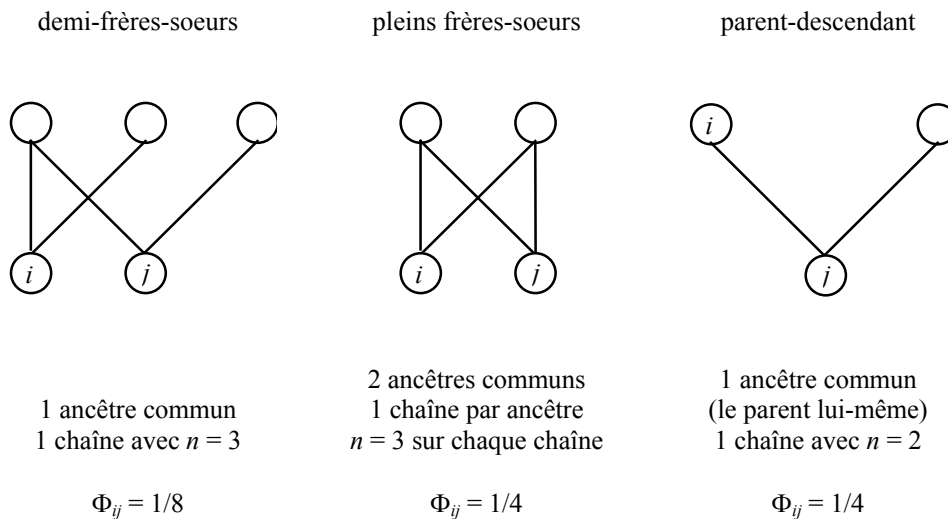
(suite)

La généralisation nécessite une description précise des pedigrees : on recense tous les ancêtres communs à i et à j et toutes les chaînes de parenté reliant i et j . On calcule le coefficient de consanguinité de chaque ancêtre commun (F_a) et les probabilités d'identité des gènes relatives à chacune des chaînes ; la somme de ces probabilités élémentaires est égale au coefficient de parenté recherché (Φ_{ij}) :

$$\Phi_{ij} = \sum_{\text{ancêtres, chaînes}} (1/2)^n (1 + F_a)$$

Applications numériques

Dans ces pedigrees simplifiés, les individus sont représentés par des cercles ; ceux pour lesquels on calcule un coefficient de parenté sont désignés i et j respectivement. Les parents occupent la ligne du haut, les descendants celle du bas. Dans tous les cas, on suppose que les parents ne sont pas consanguins et qu'ils ne sont pas apparentés entre eux. Un individu peut être considéré comme son propre ancêtre.



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

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



La covariance entre valeurs phénotypiques s'interprète en fonction de la seule composante génétique uniquement si le caractère étudié n'est pas sujet à des effets d'environnement commun. La non prise en compte dans l'analyse d'un effet réel d'environnement commun aboutit à des conclusions erronées.

2. Décomposition de la covariance entre valeurs génétiques

Nous négligeons ici les effets d'épistasie : la valeur génétique (G) d'un individu est la somme sur tous les locus des effets moyens du génotype aux différents locus (cf. § B.1.a, ch. III). Pour établir la covariance entre deux apparentés quelconques, i et j , on exprime leur valeur génétique (G) selon les deux composantes, génétique additive (A) et de dominance (D) :

$$G_i = A_i + D_i \quad \text{et} \quad G_j = A_j + D_j$$

Par construction, les variables A et D sont non corrélées (cf. § C.2.a, ch. III). La covariance existant entre les valeurs génétiques des deux individus s'écrit donc :

$$\text{Cov}(G_i, G_j) = \text{Cov}(A_i, A_j) + \text{Cov}(D_i, D_j)$$

Pour que la composante additive de la covariance soit non nulle, les deux individus doivent posséder des gènes en commun, provenant d'un ancêtre commun. Pour que la composante de dominance soit non nulle, les deux individus doivent avoir reçu, via leurs deux parents respectifs, le même couple de gènes. La covariance est donc non nulle si les deux individus ont reçu chacun au moins une copie d'un même gène présent chez un ancêtre commun. Deux gènes qui sont la copie mendélienne d'un même gène ancêtre sont dits *identiques*. Le calcul de la covariance entre apparentés fait intervenir les probabilités d'identité des gènes⁶. En un locus donné, deux situations d'identité des gènes sont distinguées :

- L'un des deux gènes de i est identique à l'un des deux gènes de j . La composante génétique additive de la covariance est alors non nulle. La probabilité de cet événement est égale au coefficient de parenté entre i et j (Φ_{ij}).
- Les deux gènes de i sont identiques aux deux gènes de j (identité des gènes deux à deux). La composante de dominance de la covariance est alors non nulle. La probabilité de cet événement, ici notée Ψ_{ij} , est parfois désignée par le joli terme de « coefficient de fraternité ». Dans quelques cas simples, cette probabilité est égale au produit des coefficients de parenté liant respectivement les deux pères (Φ_{pp}) et les deux mères (Φ_{mm}) des individus i et j . Il ne s'agit cependant pas là d'une généralité.

⁶ La formule de la covariance entre apparentés a tout d'abord été établie par Fisher en 1918, sous l'hypothèse d'absence d'épistasie. Sous cette même hypothèse, elle a été formulée selon une approche probabiliste par Malécot en 1948. Le concept d'identité des gènes et la description exhaustive des situations d'identité entre deux individus en un locus ont été introduits par Gillois en 1964.

En définitive, sous l'hypothèse d'absence d'effet d'épistasie, l'expression de la covariance génétique, dont les éléments ci-dessus esquissent la démonstration, est :

$$\text{Cov}(G_i, G_j) = 2\Phi_{ij}V_A + \Psi_{ij}V_D$$

Le tableau 11 donne les valeurs des probabilités utilisées et l'expression de la covariance génétique entre différents types d'apparentés. Dans les cas de liaison entre parent et descendant et demi-frères-soeurs, la covariance génétique ne fait pas intervenir de termes de dominance : en effet, il n'y a pas de raison pour que deux individus dont l'apparentement ne provient que d'une seule voie parentale (leur père ou leur mère) aient des couples de gènes en commun. A l'opposé, la liaison entre pleins frères-soeurs fait intervenir un terme de dominance : deux individus qui ont à la fois même père et même mère ont une probabilité égale à $\frac{1}{4}$ d'avoir, en un locus neutre, des gènes qui sont identiques deux à deux. On voit bien que le simple coefficient de parenté ne suffit pas à caractériser le degré de ressemblance génétique entre deux individus : si l'on peut s'exprimer ainsi, la ressemblance entre deux pleins frères-soeurs est plus du double de la ressemblance entre deux demi-frères-soeurs.

Tableau 11. Probabilités relatives à certaines situations d'identité des gènes (Φ_{ij} et Ψ_{ij} ⁷, voir texte) et covariance génétique entre différents types d'apparentés exprimée en termes de variances génétique additive (V_A) et de dominance (V_D).

Lien de parenté	Φ_{ij}	Ψ_{ij}	Expression de la covariance génétique
Non apparentés	0	0	0
Demi-frères-sœurs	1/8	0	$\frac{1}{4}V_A$
Pleins frères-sœurs	1/4	1/4	$\frac{1}{2}V_A + \frac{1}{4}V_D$
Vrais jumeaux	1/2	1	$V_A + V_D = V_G$
Parent-descendant	1/4	0	$\frac{1}{2}V_A$

⁷ Les probabilités affectant V_A et V_D sont parfois notées Φ_A et Φ_D respectivement.

C. Interprétation de la notion d'héritabilité et estimation

1. Confrontation des modèles : cas des liens frères-soeurs

Nous avons vu (§ A.1) qu'il est possible de caractériser la ressemblance entre individus d'une même fratrie par le coefficient de corrélation liant leurs valeurs phénotypiques. A partir d'un modèle simple, on peut caractériser *a priori* la covariance entre apparentés (section B). Il est donc possible de confronter les prédictions du modèle génétique (*a priori*) et les résultats des analyses statistiques (*a posteriori*). Reportons nous au protocole permettant de calculer une corrélation entre membres d'une même fratrie (cf. § A.1). La valeur phénotypique du premier individu d'une paire (P_i) est une variable aléatoire d'espérance μ et de variance V_P ; il en est de même pour la valeur phénotypique du second individu de la paire (P_j). Nous pouvons donc définir le coefficient de corrélation (r) entre ces deux variables. D'après la décomposition de la covariance phénotypique vue au § B.1., ce coefficient de corrélation s'écrit :

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

Si l'on peut admettre l'absence d'effet d'environnement commun, le coefficient de corrélation s'exprime alors en fonction de la seule covariance génétique, soit en reprenant les expressions données au tableau 11 :

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

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

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

Ainsi, la valeur attendue du coefficient de corrélation entre valeurs phénotypiques de membres d'une même fratrie s'exprime en fonction de paramètres génétiques du caractère étudié : l'héritabilité au sens étroit (h^2) ou au sens large (h_{sl}^2) et le rapport de la variance de dominance sur la variance phénotypique. Ceci amène deux remarques.


a. On constate tout d'abord que **le modèle développé permet de rendre compte d'observations statistiques faites sur des populations réelles**. En effet, les expressions ci-dessus sont en accord (cf. tableau 10), (i) avec l'accroissement de la corrélation observée entre valeurs phénotypiques lorsque le lien de parenté est plus étroit, (ii) avec le fait que la corrélation entre pleins frères-soeurs vaut plus du double de la corrélation entre demi-frères-soeurs et (iii), avec le fait que, 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 (cf. § C.2, ch. III).

b. Par ailleurs, on constate que **l'analyse de données réelles fournit un moyen d'estimer certains paramètres génétiques du caractère mesuré**. Par exemple, la corrélation entre demi-frères-soeurs s'exprimant *a priori* seulement en fonction de l'héritabilité au sens étroit, nous déduisons facilement un estimateur (désigné par un chapeau) de ce paramètre :

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

De la même façon, on donne un estimateur de l'héritabilité au sens large :

$$\hat{h}_{sl}^2 = \text{corrélation entre vrais jumeaux ou entre individus d'un clone}$$

 **Les estimations des paramètres génétiques à partir de l'analyse statistique des données n'ont de sens qu'en absence d'effets d'environnement commun**. En effet, si de tels phénomènes existent, les expressions ci-dessus ne sont plus valables : l'expression du coefficient de corrélation doit comporter également une composante environnementale. Le fait de négliger un effet d'environnement commun alors qu'il a une importance réelle revient à incorporer une composante environnementale dans les causes d'origine génétique de ressemblance. En conséquence, en cas d'effet d'environnement commun, les estimateurs ci-dessus surestiment les paramètres génétiques. Ce problème est particulièrement critique pour l'analyse de tous les caractères soumis à effet maternel, ainsi que pour l'analyse des caractères mesurés chez l'homme où les membres d'une même fratrie partagent, durant un bon nombre d'années, un même milieu familial.

2. La liaison parent-descendant

Effectuons dans ce cas la même démarche consistant à confronter le modèle génétique au modèle statistique. Appelons P_i la valeur phénotypique d'un descendant, P_p celle de son père et P_m celle de sa mère. Admettons que les pères et les mères sont représentatifs de l'ensemble de la population et qu'ils s'unissent au hasard. Supposons enfin qu'il n'y a pas d'effet d'environnement commun entre parent et descendant.

a. Cas où les deux parents sont connus et mesurés

Nous raisonnons sur la valeur du "parent moyen" [$\bar{P} = (P_p + P_m) / 2$]. Compte tenu de nos hypothèses et selon les expressions de covariance vues plus haut, nous avons :

$$\text{Cov}(P_i, \bar{P}) = \frac{1}{2} [\text{Cov}(P_i, P_p) + \text{Cov}(P_i, P_m)] = \frac{1}{2} \left[\frac{1}{2} V_A + \frac{1}{2} V_A \right] = \frac{1}{2} V_A$$

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

L'absence de covariance entre P_p et P_m est assurée s'il y a panmixie, ce qui implique que les deux parents sont non apparentés entre eux. Par ailleurs, on assimile la variance des valeurs parentales à celle de l'ensemble de la population. On déduit la valeur du coefficient de régression du descendant sur son "parent moyen" :

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

b. Cas d'un seul parent connu ou mesuré

Considérons que le parent connu est le père (si c'est la mère, le résultat est le même). En procédant comme dans le cas de deux parents connus, nous avons :

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

c. Conclusions : interprétation de la notion d'héritabilité au sens étroit

La régression de la valeur phénotypique du descendant sur celle d'un parent, ou sur celles des deux parents, ne s'exprime qu'en fonction de l'héritabilité au sens étroit.



L'héritabilité au sens étroit apparaît comme un paramètre fondamental pour décrire la transmission héréditaire des caractères quantitatifs.

En revenant à la notion de valeur génétique additive (cf. Ch.III), on peut comprendre ces résultats. Par définition, l'héritabilité au sens étroit est la part de la variance phénotypique qui est d'origine génétique additive. La valeur additive est la fraction de la valeur génétique dont on peut facilement prédire la transmission. Nous admettons par ailleurs que, sauf en cas d'effet d'environnement commun, le milieu ne se transmet pas. L'héritabilité au sens étroit représente donc la part de variance imputable aux seuls effets transmissibles de parent à descendant. Quant au coefficient de régression du descendant sur son ou ses parents, il peut s'interpréter comme la part des différences phénotypiques, observées au sein de l'ensemble des parents, qui sont susceptibles d'être transmises à la descendance. Il en découle que le coefficient de régression caractérisant la liaison parent-descendant s'exprime uniquement en fonction de l'héritabilité au sens étroit. Un reproducteur ne transmettant, en espérance, que la moitié de sa valeur génétique additive, ce coefficient de régression est égal à $\frac{1}{2} h^2$ quand on ne connaît la valeur d'un seul parent et à h^2 quand on connaît la valeur des deux parents.



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

Cette interprétation permet de rendre compte des résultats statistiques obtenus sur des populations. La régression constatée par Galton (cf. figure 17) vient de ce que les différences phénotypiques observées à l'échelon parental ne correspondent qu'à une fraction h^2 transmissible à la descendance. Signalons que, là encore, l'analyse statistique permet de fournir des estimations de l'héritabilité du caractère observé : par exemple, on interpréterait aujourd'hui les résultats de Galton en estimant à $\frac{2}{3}$ l'héritabilité au sens étroit de la taille chez l'homme. Toutefois, compte tenu de la structure des populations humaines (choix des partenaires, milieu familial commun), cette valeur représente sans doute une surestimation.

D. Introduction à la notion de sélection

1. Définitions

Il y a sélection au sein d'une population quand l'espérance de la taille de descendance dépend du génotype des parents (voir GP). Pour un individu donné, l'espérance de la taille de sa descendance dépend de sa probabilité de devenir reproducteur et, ensuite, de sa fécondité. La sélection est un phénomène auquel sont soumises les populations naturelles : on parle de sélection naturelle. La sélection affecte également les populations domestiques, la sélection exercée par l'homme, ou sélection artificielle, se surimposant à la sélection naturelle. La sélection artificielle consiste essentiellement à exclure de la reproduction certains individus, le principal moyen étant l'élimination physique des individus exclus. Sans doute dès le début de l'agriculture et/ou de l'élevage, au néolithique, l'homme a exercé une sélection artificielle au sein des espèces qu'il a domestiquées. On peut d'ailleurs raisonnablement penser que les premiers caractères sur lesquels ont porté les débuts de la sélection artificielle étaient des caractères facilitant la domestication elle-même : pour les plantes, l'aptitude au bouturage, l'absence d'égrainage spontané chez les céréales ; pour les animaux, la docilité et la grégarité.

Le choix des individus reproducteurs s'effectue, consciemment ou inconsciemment, sur la base de l'idée que l'on se fait de leur génotype. Dans le cas des caractères quantitatifs, les informations dont on dispose sont les valeurs phénotypiques. La procédure de sélection la plus simple consiste alors à choisir les reproducteurs au vu de leur valeur phénotypique pour le caractère qui nous intéresse : si l'on souhaite aller vers une augmentation de la moyenne du caractère, on retient pour la reproduction les individus qui ont les meilleurs valeurs phénotypiques. Ce mode de sélection est désigné sous le vocable sélection individuelle ou, plus souvent, sélection massale, de l'anglais « *mass selection* » qui signifie « sélection dans la masse » (des candidats à la sélection). Si le caractère que l'on souhaite améliorer n'est mesurable que sur les individus d'un seul sexe, cette sélection ne s'applique évidemment qu'à la moitié de la population : c'est le cas, par exemple, des caractères de production laitière ou de production d'œufs chez les animaux, des caractéristiques des fruits ou des grains chez les végétaux, et. Dans le cadre des enseignements d'amélioration des plantes et d'amélioration des animaux, des méthodes de sélection plus élaborées que la sélection massale seront présentées.

2. Prédiction de la réponse à une génération de sélection massale

Nous considérons ici une population de grande taille, à l'équilibre, et où les valeurs phénotypiques suivent une loi normale (ce qui correspond à la majeure partie des caractères d'intérêt agronomique ou zootechnique ; cf. Chapitre .I) Nous supposons aussi que l'espèce concernée est à générations séparées : les reproducteurs d'une génération donnée ne participent pas à la reproduction de la génération suivante ; ils sont remplacés par de nouveaux reproducteurs choisis parmi leurs descendants. Cette situation recouvre le cas des plantes annuelles et de quelques espèces d'élevage (volailles, salmonidés).

a. Présentation générale

Si les reproducteurs sont pris au hasard, il n'y a pas de raison pour que la moyenne de la population évolue dans un sens ou dans un autre. *A contrario*, il y a sélection quand on retient comme parents de la génération suivante les meilleurs individus ou, plus généralement, un groupe d'individus supérieurs à la moyenne. Dans un sexe donné, on appelle **différentielle de sélection (S)** l'écart entre la moyenne phénotypique des individus retenus pour la reproduction (μ^*) et la moyenne de la population (μ) :

$$S = \mu^* - \mu$$

La différentielle de sélection moyenne (\bar{S}) est égale à la moyenne des différentielles de sélection réalisées chez les mâles (indice m) et chez les femelles (indice f) :

$$\bar{S} = \frac{S_m + S_f}{2} = \frac{\mu_m^* + \mu_f^*}{2} - \mu$$

Lorsque l'on pratique la sélection massale chez des espèces monoïques ou hermaphrodites, la différentielle de sélection est la même dans les deux sexes. Quand le caractère ne s'exprime que dans un seul sexe, on considère que les parents du sexe où l'on ne peut pas faire de mesure ne sont pas sélectionnés et que la différentielle de sélection correspondante est nulle⁸.

⁸ Ceci est vrai si l'on ne peut rien faire d'autre que de la sélection massale. Dans la pratique, dès lors que les généalogies sont connues, les individus sur lesquels un caractère ne peut pas être mesuré sont sélectionnés sur la base des performances de leurs apparentés connus : ascendants, collatéraux ou descendants.

On appelle *progrès génétique* (ΔG), ou *réponse à la sélection*, l'écart attendu entre la valeur des descendants issus des parents sélectionnés et leur valeur s'ils étaient issus de parents pris au hasard. Du fait de la normalité des distributions, le meilleur prédictor de la valeur d'un descendant sachant ses parents est obtenu par régression linéaire (voir Stat). Raisonons sur la notion de parent moyen : la différentielle de sélection sur le parent moyen est \bar{S} , définie plus haut. Nous avons vu à la section précédente que, selon l'expression de la covariance entre apparentés, le coefficient de régression de la valeur du descendant sur celle de son parent moyen est égal à l'héritabilité au sens étroit. Ainsi, l'espérance du progrès génétique s'écrit :

$$E(\Delta G) = h^2 \bar{S}$$

En supposant que les conditions de milieu restent stables, la valeur phénotypique moyenne (μ') attendue pour la génération issue des parents sélectionnés est :

$$E(\mu') = \mu + h^2 \bar{S}$$



L'héritabilité au sens étroit est un paramètre clef de l'efficacité de la sélection. Ceci découle de la définition de la valeur génétique additive (voir encadré 6).

Il sera vu dans les enseignements d'amélioration des plantes et d'amélioration des animaux que la valeur de l'héritabilité au sens étroit conditionne l'efficacité de la sélection, d'une part, et oriente le choix d'une méthode de sélection, d'autre part.

Encadré 6

Lien entre l'espérance du progrès génétique et la notion de valeur génétique additive

On peut montrer que l'héritabilité au sens étroit (h^2) peut servir à prédire, par régression linéaire, la valeur génétique additive (A_i) d'un individu sachant sa valeur phénotypique (P_i) :

$$E(A_i | P_i) = h^2 (P_i - \mu)$$

L'espérance de l'écart génétique additif moyen d'un groupe de parents sélectionnés, par rapport à la moyenne de la population, se calcule donc comme le produit $h^2 \times S$, S étant la différentielle de sélection réalisée dans le sexe considéré. Un reproducteur transmet en espérance la moitié de sa valeur génétique additive (cf. Ch.III). On en déduit l'espérance du progrès génétique, en désignant par S_m et S_f les différentielles de sélection chez les mâles et les femelles, respectivement :

$$E(\Delta G) = \frac{1}{2} h^2 S_m + \frac{1}{2} h^2 S_f = h^2 \bar{S}$$

b. Cas d'une sélection par troncature

En sélection artificielle, il est fréquent de retenir pour la reproduction tous les individus dont la valeur se situe au-delà d'un certain seuil (figure 19) : on parle de sélection par troncature. La définition du seuil dépend en général du nombre de reproducteurs que l'on souhaite retenir. Dans ce cas, on peut raisonner l'efficacité attendue de la sélection uniquement à partir de paramètres démographiques. Afin de disposer d'un paramètre indépendant de la variabilité du caractère sélectionné, on définit l'intensité de sélection (i) comme la différentielle de sélection exprimée en unités d'écart-type :

$$i = S/\sigma_P$$

On montre que, dans le cas d'une sélection par troncature, l'intensité de sélection se déduit directement de la proportion d'individus sélectionnés (p) : soit z l'ordonnée de la loi normale au seuil de sélection, l'intensité de sélection vaut :

$$i = z/p$$

Le tableau 12 donne les valeurs du seuil de sélection, en unités d'écart-type, et de l'intensité de sélection en fonction des valeurs de p (Attention ! Dans ce tableau, p est donnée en pourcentage pour des raisons de lisibilité ; ci-dessus, p est une proportion, variant de 0 à 1). En désignant par les indices m et f ce qui concerne les mâles et les femelles respectivement, nous définissons l'intensité de sélection moyenne (\bar{i}) et exprimons en fonction de ce paramètre la différentielle de sélection moyenne (\bar{S}) :

$$\bar{i} = \frac{i_m + i_f}{2} \quad \text{et} \quad \bar{S} = \frac{i_m + i_f}{2} \sigma_P = \bar{i} \sigma_P$$

Le progrès génétique attendu s'exprime alors comme suit :

$$E(\Delta G) = \bar{i} h^2 \sigma_P = \frac{1}{2} (i_m + i_f) h^2 \sigma_P$$


 *Quand on fait la moyenne sur les deux sexes, c'est en termes de différentielle (S) ou d'intensité (i) qu'il faut raisonner, et non pas en termes de taux (p) de sélection. En effet, l'intensité de sélection moyenne (\bar{i}) n'est pas égale à l'intensité de sélection correspondant à la moyenne des taux de sélection.*

Figure 19. Sélection par troncature dans une distribution normale centrée réduite.

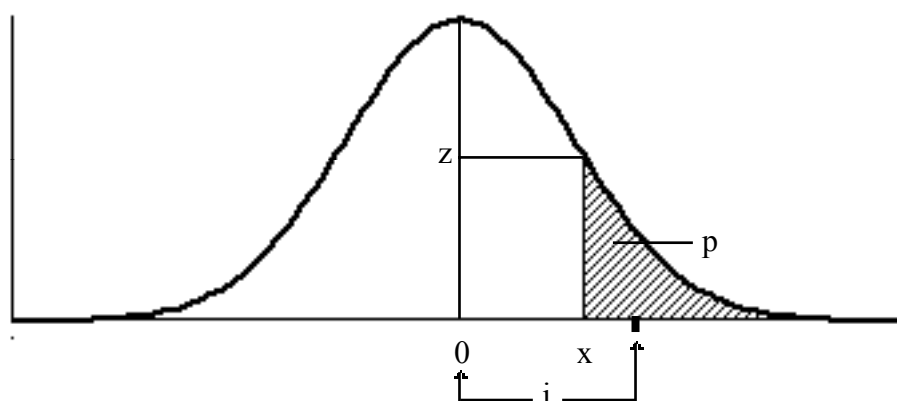


Tableau 12. Intensité de sélection (i) et seuil de sélection (x) en fonction de la proportion d'individus retenus pour la reproduction (p , exprimée en pourcentage). Les valeurs du seuil (x) pour $p > 50\%$ s'obtiennent par réciprocity (cf. figure 19).

p (%)	i	x	p (%)	i	x	p (%)	i	x
0,01	3,958	3,719	11	1,710	1,226	41	0,948	0,227
0,02	3,789	3,540	12	1,667	1,175	42	0,931	0,202
0,03	3,687	3,432	13	1,628	1,126	43	0,914	0,176
0,04	3,613	3,353	14	1,590	1,080	44	0,896	0,151
0,05	3,555	3,290	15	1,555	1,036	45	0,880	0,125
0,06	3,506	3,239	16	1,521	0,994	46	0,863	0,100
0,07	3,465	3,195	17	1,489	0,954	47	0,846	0,075
0,08	3,429	3,156	18	1,458	0,915	48	0,830	0,050
0,09	3,397	3,121	19	1,428	0,878	49	0,814	0,025
			20	1,400	0,841	50	0,798	0,000
0,1	3,368	3,090	21	1,373	0,806			
0,2	3,171	2,878	22	1,346	0,772	51	0,782	
0,3	3,050	2,748	23	1,320	0,739	52	0,766	
0,4	2,962	2,652	24	1,295	0,706	53	0,751	
0,5	2,893	2,576	25	1,271	0,674	54	0,735	
0,6	2,834	2,512	26	1,248	0,643	55	0,720	
0,7	2,784	2,457	27	1,225	0,613	56	0,704	
0,8	2,741	2,409	28	1,202	0,583	57	0,689	
0,9	2,701	2,366	29	1,181	0,553	58	0,674	
			30	1,159	0,524	59	0,659	
1	2,666	2,326	31	1,138	0,496	60	0,644	
2	2,421	2,054	32	1,118	0,467			
3	2,269	1,881	33	1,098	0,440	65	0,570	
4	2,155	1,751	34	1,078	0,412	70	0,497	
5	2,063	1,645	35	1,058	0,385	75	0,424	
6	1,986	1,555	36	1,039	0,358	80	0,350	
7	1,918	1,476	37	1,021	0,332	85	0,274	
8	1,859	1,405	38	1,002	0,305	90	0,195	
9	1,805	1,341	39	0,984	0,279	95	0,109	
10	1,755	1,281	40	0,966	0,253	100	0,000	

3. Prise en compte de plusieurs générations de sélection

Lorsque l'on raisonne sur plusieurs générations, les progrès génétiques se cumulent. Soit ΔG_t le progrès génétique entre la génération $t-1$ et la génération t , calculé comme ci-dessus, l'espérance du progrès génétique cumulé sur T générations ($\Delta G_{0 \rightarrow T}$) s'écrit :

$$E(\Delta G_{0 \rightarrow T}) = \sum_{t=1}^T E(\Delta G_t) = \sum_{t=1}^T \bar{i}_t h_t^2 \sigma_{P_t}$$

Il ne faut pas en déduire que le progrès génétique sur plusieurs générations est linéaire. En effet, les paramètres du progrès génétique sur une génération (i , h^2 , σ_P) peuvent varier d'une génération à l'autre, soit pour des raisons démographiques en ce qui concerne le premier, soit pour des raisons de modification de structure génétique de la population en ce qui concerne les deux derniers paramètres. Les expériences de sélection menées sur le long terme illustrent bien ces phénomènes : d'une part, la réponse à la sélection peut se maintenir pendant de nombreuses générations (de quelques dizaines à une centaine pour les expériences les plus spectaculaires menées sur le maïs ou sur la drosophile) mais, d'autre part, le progrès génétique dégagé sur une génération n'est pas constant dans le temps et peut connaître des phases de ralentissement, suivies parfois d'accélération, liées à l'évolution de la variabilité génétique au sein de la population sélectionnée.

EXERCICES

Exercice n°1

En décomposant la valeur phénotypique d'un individu (cf. Ch.II-III) et en exprimant la valeur génétique additive d'un individu en fonction de celles de ses parents (cf. Ch.III), démontrer sans utiliser la formule de covariance entre apparentés que dans une population répondant aux hypothèses de Hardy-Weinberg, le coefficient de régression de la valeur phénotypique d'un descendant sur la moyenne des valeurs phénotypiques de ses parents est égal à l'héritabilité au sens strict du caractère considéré.

Exercice n°2

Le tableau ci-dessous donne, chez l'homme, la taille (en centimètres) de 11 paires de pleins frères-soeurs mesurés à l'âge adulte (il s'agit des données de la figure 15). Calculer, à partir de ces données, le coefficient de corrélation reliant la taille du frère et celle de la sœur. En déduire une estimation de l'héritabilité de la taille chez l'homme. Quelles réserves pouvez-vous émettre quant à ce résultat ?

N° de famille	1	2	3	4	5	6	7	8	9	10	11
Taille du frère	180	173	168	170	178	180	178	185	183	165	168
Taille de la sœur	175	163	165	160	165	157	165	163	168	150	157

Exercice n°3

En sélection aquacole, afin d'établir les généalogies et du fait de l'impossibilité de marquer les animaux à leur naissance (voir la taille d'un alevin !), on élève généralement les membres d'une même famille de pleins frères-sœurs dans un même bassin séparé des autres bassins où sont élevées d'autres familles. Dans un élevage de sélection de truites arc-en-ciel, on a analysé les corrélations entre les valeurs phénotypiques pour le poids à 1 an d'individus apparentés et élevés de façon contemporaine. Après correction des données pour l'effet du sexe, on a obtenu les résultats suivants (valeur estimée \pm écart-type d'erreur) :

- Corrélation entre pleins frères-sœurs (individus ayant même père et même mère) :
 $r_1 = 0,29 \pm 0,01$
- Corrélation entre demi frères-sœurs de père (individus ayant même père et des mères différentes) :
 $r_2 = 0,07 \pm 0,02$
- Corrélation entre demi frères-sœurs de mère (individus ayant même mère et des pères différents) :
 $r_3 = 0,08 \pm 0,02$

- 1) En vous fondant sur l'expression de la covariance entre apparentés, donnez 3 estimations de l'héritabilité au sens étroit (h^2) du poids à 1 an chez la truite arc-en-ciel (à partir de chacune des 3 valeurs ci-dessus) et donnez les écarts-types d'erreur correspondants.
- 2) Ces valeurs sont-elles cohérentes entre elles ? En particulier, quelles hypothèses pouvez-vous avancer pour expliquer la différence constatée entre la valeur de h^2 estimée à partir de r_1 , d'une part, et celles obtenues à partir de r_2 et r_3 , d'autre part ? Que pensez-vous en définitive, dans le cas présent, de l'estimation de l'héritabilité à partir de la corrélation entre pleins frères-soeurs ?