

II. HEREDITE ET MILIEU

A. Définitions

1. Le phénotype et le génotype

Le phénotype est la 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. Ainsi, on appelle *valeur phénotypique* le résultat de la mesure effectuée sur un individu. Dans le milieu professionnel de la sélection, animale ou végétale, on parle également de performance. Notons qu'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 peut se définir à l'échelle de la parcelle : elle représente en fait une moyenne de valeurs phénotypiques individuelles.

Le génotype d'un individu en un locus est l'ensemble des gènes qu'il possède à ce locus (deux gènes homologues pour un individu diploïde). Le génotype en plusieurs locus¹ est l'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 (encore ces exceptions ne concernent-elle que le génotype). En pratique, nous emploierons les termes "génotype" et "phénotype" 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.

¹ Locus étant un mot d'origine latine, le pluriel latin « loci » est encore souvent utilisé en français, comme c'est le cas en anglais (prononcer « loçai »). Toutefois, nous utiliserons le pluriel « locus », conformément à l'usage qui veut que lorsqu'un mot étranger est d'un usage suffisamment courant, son pluriel soit francisé.

2. Le milieu

a. Le milieu : ensemble des causes non génétiques de variation

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

La première acception est certainement la plus intuitive. En production végétale, on range dans cette catégorie des facteurs tels que 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, on considère des facteurs tels que 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.

Parmi les caractéristiques physiologiques propres à l'individu que l'on considère comme facteurs de milieu, l'âge concerne les cultures permanentes et les animaux : le rendement en fruits d'un cep de vigne ou d'un arbre fruitier varie en fonction de son âge ; il en est de même chez tous les animaux en ce qui concerne la taille et le poids, chez les femelles de mammifères en ce qui concerne l'intensité de la production laitière, etc. Le stade physiologique à un âge donné joue également sur les valeurs phénotypiques : cycle et stade de végétation des plantes fourragères, stade de lactation des femelles laitières, etc.

L'observateur, enfin, a une influence au travers du protocole de mesure qu'il applique, de la précision de ses instruments de mesure, et des erreurs de mesure qu'il peut commettre.

b. Facteurs de milieu contrôlés et non contrôlés

Dès maintenant, il faut introduire une autre classification des facteurs de milieu, qui est déterminante dès lors que l'on est au stade de l'expérimentation ou de l'analyse des données recueillies sur le terrain.

Les facteurs de milieu contrôlés sont des facteurs de milieu que l'on identifie comme tels, dont on sait ou dont on pense qu'ils ont un effet important sur les caractères étudiés : on parle également de *macro-milieu*. Ici, non seulement on peut enregistrer les conditions propres à plusieurs individus, mais on peut agir dessus : par exemple, on peut apporter à une culture une fumure azotée plus ou moins forte, mais contrôlée. Les exemples donnés plus haut relèvent en fait tous de cette catégorie. Ainsi, en même temps que l'on mesure le rendement en fruits d'un arbre fruitier, on enregistre l'année correspondante, la parcelle, l'âge de l'arbre, etc. Lorsque l'on mesure la production laitière d'une vache, on enregistre également le numéro du troupeau, le rang de lactation de l'animal, sa date de mise-bas (donc de démarrage de la lactation), etc. Au stade de l'analyse, on considère qu'un effet de macro-milieu s'applique en commun à tous les individus se trouvant dans une catégorie de milieu donnée.

Les facteurs de milieu non contrôlés sont des facteurs, soit que l'on ne maîtrise pas car ils échappent à l'observateur, soit que l'on n'enregistre pas car le recueil de l'information correspondante est trop compliqué ou trop coûteux. On pense que ces facteurs non contrôlés induisent chacun des faibles variations, car résultant de microphénomènes locaux, s'appliquant à de manière différente à chaque individu : on parle aussi de *micro-milieu*.

 *Cette distinction entre macro-milieu et micro-milieu comporte une part d'arbitraire liée à l'échelle d'observation à laquelle on se situe. Ne pas enregistrer un facteur de milieu qui, en réalité, induit des variations importantes, conduit à des erreurs d'analyse ; dans la pratique de la sélection, où l'on doit corriger les données pour les effets des facteurs de milieu contrôlés, une telle omission a des répercussions relativement graves.*

B. Mise en évidence des facteurs de variation génétique

Les caractères quantitatifs présentant une variation continue, la mise en évidence des facteurs génétiques à l'origine de cette variation est en général plus compliquée que pour les caractères pour lesquels les individus observés se rangent dans un petit nombre de catégories. Quelques résultats expérimentaux vont nous permettre de donner des éclaircissements et de poser certaines hypothèses quant à l'hérédité de ces caractères.

1. L'hypothèse multifactorielle

Plusieurs expériences ont été réalisées au début du XX^{ème} siècle afin de confronter la variation des caractères quantitatifs à la théorie mendélienne (notons que Mendel lui-même, au travers de ses expériences sur la couleur de la fleur chez *Phaseolus multifloris*, avait déjà posé les bases de l'hérédité multifactorielle). Pour se ramener dans les mêmes conditions, les parents utilisés pour les croisements étaient des « lignées pures », c'est-à-dire homozygotes à tous les locus. Cet état est facilement obtenu chez des plantes se reproduisant naturellement par autofécondation (plantes autogames) : après plusieurs générations d'autofécondation, le coefficient de consanguinité au sein d'une lignée tend vers 1, c'est-à-dire que tous les individus de la lignée sont homozygotes, à chacun des locus, pour le même allèle (voir GP).

La figure 4 présente les résultats d'une expérience de croisement réalisée par East (1916) à partir de deux lignées consanguines d'une espèce ornementale de tabac différant pour la longueur des fleurs. Compte tenu du régime d'autofécondation, à l'intérieur de chacune des deux lignées parentales (P1 et P2), tous les individus sont génétiquement identiques. De même, tous les individus de la première génération de croisement (F1) sont génétiquement identiques : notamment, ils sont tous hétérozygotes aux locus où les deux lignées parentales diffèrent. Ainsi, la variance observée chez les parents et en F1 est d'origine strictement environnementale. Par contre, l'accroissement substantiel de la variance observée en F2 (F1 x F1) évoque une ségrégation de gènes. Si le polymorphisme en un seul locus était responsable des variations observées de la longueur de fleur, la distribution de la F2 devrait être un mélange des distributions respectives de P1, F1 et P2 dans des proportions mendéliennes (1/4, 1/2 et 1/4), ce qui n'est manifestement pas le cas.

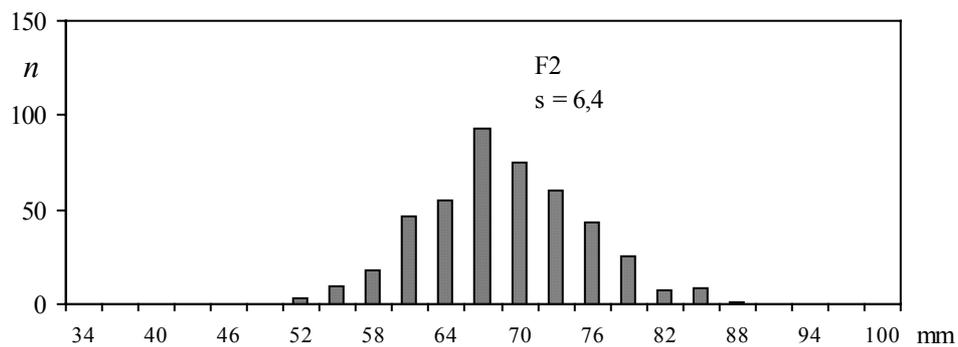
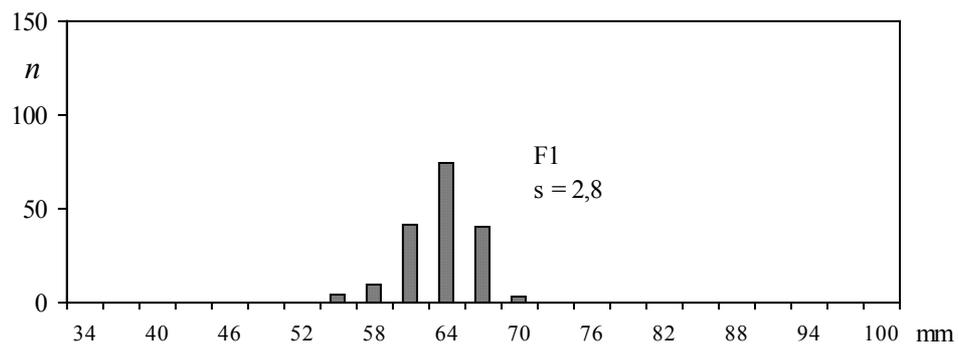
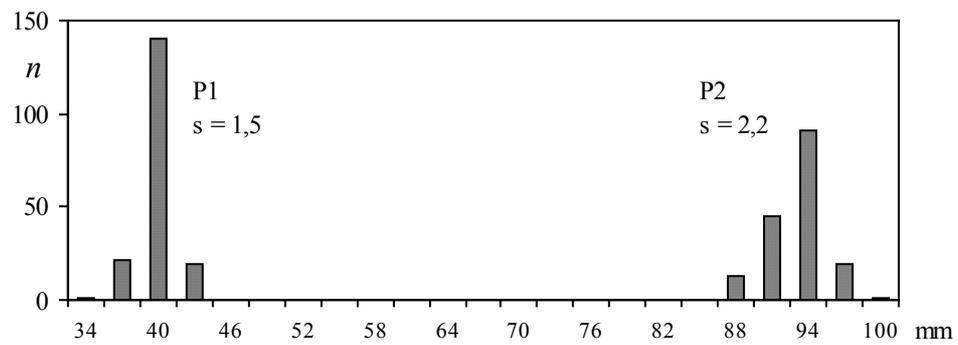
Cette expérience fut poursuivie par l'analyse de la réponse à la sélection en F3. L'ensemble des résultats a permis de suggérer que les lignées parentales différaient non pas à un locus, mais pour plusieurs ayant une influence sur le caractère mesuré. En effet, les distributions globales observées s'accordent bien avec l'hypothèse d'une distribution binomiale 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.

Figure 4. Distribution de la longueur de la fleur (mm) dans un croisement entre deux variétés consanguines (P1 et P2) d'une espèce ornementale de tabac. s = Ecart-type.

Source : East, 1916.



2. Mise en évidence de locus ayant un effet sur un caractère quantitatif

a. Description détaillée d'une expérience sur drosophile

La mise en évidence de gènes induisant des variations pour un caractère quantitatif suppose une variation allélique aux locus concernés. Pour cela, il est utile de prendre comme matériel d'étude des populations différant nettement pour le caractère observé. Cela peut se trouver dans des populations naturelles, ou bien être obtenu par sélection divergente (sélection de deux lignées dans des sens opposés pour un même caractère) : si en plusieurs locus, il existe des allèles qui sont « favorables » à l'obtention de fortes valeurs du caractère et des allèles « défavorables », on peut raisonnablement penser (voir GP) que la fréquence des allèles « favorables » sera plus élevée dans la lignée sélectionnée pour une augmentation de la moyenne du caractère que dans la lignée sélectionnée vers une diminution.

La figure 5 présente les résultats d'une expérience de ce type réalisée sur drosophile. Le caractère étudié est le taux de survie à une certaine dose de DDT. Les individus observés sont ceux d'une F2 obtenue après le croisement de parents issus de deux lignées différentes, une lignée sélectionnée pendant plusieurs générations pour la résistance au DDT et une lignée de contrôle. Dans ces lignées, les chromosomes comportent des supprimeurs de recombinaison et sont porteurs de marqueurs génétiques qui permettent d'identifier la provenance de chaque chromosome chez les individus de la F2.

La figure 5 montre clairement que les chromosomes contribuent aux variations de résistance au DDT : très schématiquement, plus la proportion de chromosomes provenant de la lignée résistante est importante dans le génome des individus, plus ceux-ci sont résistants. Ces résultats sont cohérents avec l'hypothèse multifactorielle évoquée plus haut, puisqu'ils mettent en évidence l'action d'au moins trois locus (au minimum un par chromosome, en réalité plus) sur les variations de résistance au DDT chez la drosophile.

Une analyse plus détaillée permet de préciser l'action des chromosomes (et donc des locus qui y sont portés). Ainsi, il apparaît que la substitution d'un chromosome 3 de la lignée de contrôle par un chromosome 3 de la lignée résistante confère une résistance plus élevée quelle que soit la combinaison aux autres chromosomes (comparer les valeurs respectives des types 1 et 4, 5 et 8, 12 et 15). Ceci n'est pas vrai pour le chromosome X : les individus de type 2

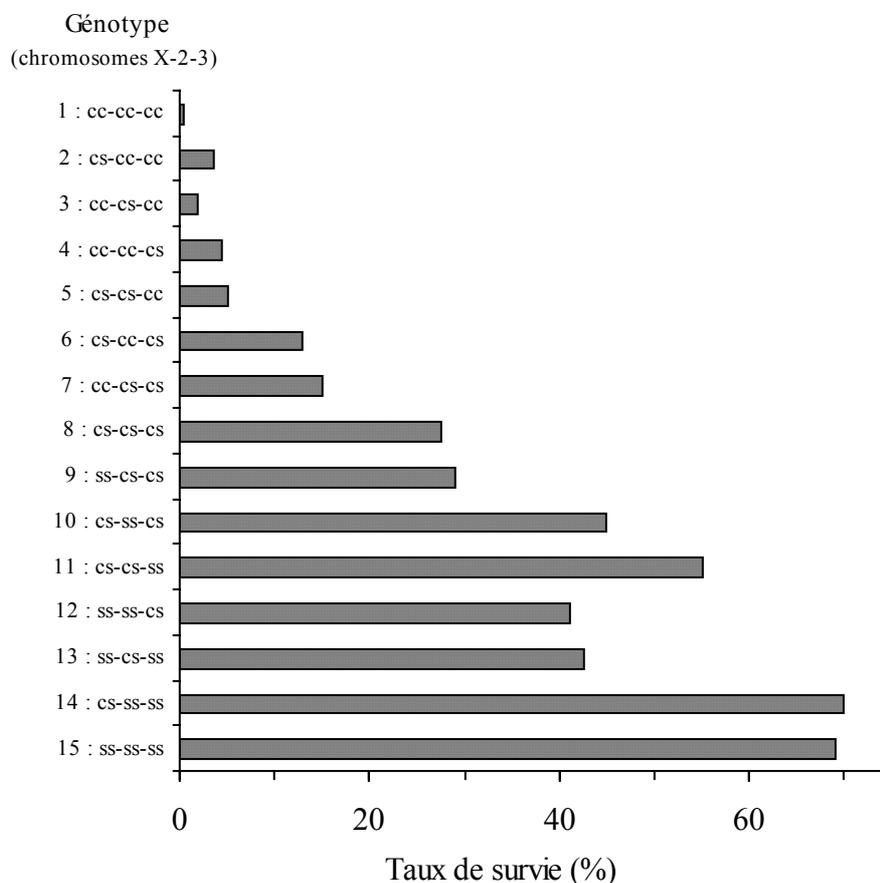
sont plus résistants que ceux de type 1, les individus de type 8 le sont plus que ceux de type 7, mais il n'y a pas de différence entre les individus de type 14 et 15. Par ailleurs, la comparaison des individus de type 1 et 6, d'une part, et de ceux de type 8 et 13, d'autre part, montre que les effets "bénéfiques" des chromosomes X et 3 de la lignée résistante ont tendance à se cumuler (ce qui ne signifie pas qu'ils s'ajoutent strictement).



Les locus responsables des variations des caractères quantitatifs ont des effets qui se cumulent.

Figure 5. Taux de survie à une dose de DDT chez la drosophile. Les individus observés comportent diverses combinaisons de chromosomes, décrites pour les chromosomes X, 2 et 3, respectivement : c = chromosome originaire de la lignée de contrôle, s = chromosome originaire de la lignée sélectionnée pour la résistance au DDT.

Source : Crow, 1957.



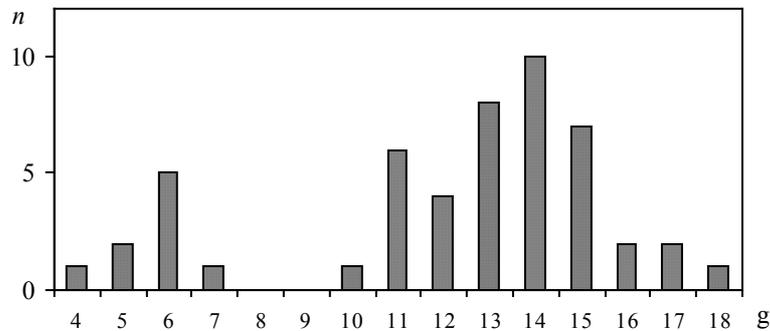
b. Gènes marqueurs et gènes majeurs

Une analyse plus fine peut être réalisée si, au lieu de repérer l'origine du génome des individus chromosome par chromosome, on peut le faire sur des fragments chromosomiques plus petits. Ceci est maintenant possible grâce au marquage moléculaire pratiqué sur des organismes recombinant normalement. Plusieurs types de marqueurs moléculaires sont aujourd'hui disponibles : RFLP, AFLP, microsatellites, SNP, etc. A côté d'espèces modèles (l'homme, la souris, la drosophile, *Arabidopsis*, le riz, certaines levures ou bactéries..., où l'on en est au stade du séquençage du génome entier), les études chez les espèces majeures de plantes cultivées ou d'animaux d'élevage ont fourni un grand nombre de marqueurs, capables de baliser l'ensemble du génome, et qui, une fois ordonnés sur une carte génétique, autorisent l'étude fine des variations des caractères quantitatifs. L'analyse des descendances obtenues avec recombinaison permet de révéler des co-ségrégations entre les allèles aux marqueurs et les valeurs observées pour des caractères quantitatifs. Couplées à la carte génétique des marqueurs, ces analyses conduisent à la localisation chromosomique de locus dont le polymorphisme induit des variations sur les caractères étudiés, locus nommés QTL (de l'anglais *Quantitative Trait Locus*). Le principe de la détection de QTL et les applications de cette méthodologie sont présentés au chapitre VI.

Il n'a pas été besoin d'attendre les progrès de la génétique moléculaire pour mettre en évidence des locus polymorphes responsables de variations importantes pour un caractère quantitatif. De simples expériences de sélection divergente et/ou de croisement ont permis dans divers cas de détecter certains locus de ce type, par simple observation des distributions. A titre d'exemple, la figure 6 présente la distribution du poids à 60 jours dans une lignée de souris sélectionnée pour de faibles poids, et dans laquelle une forme de nanisme est apparue. Les individus observés ayant été élevés rigoureusement dans les mêmes conditions de milieu, l'hypothèse de la ségrégation d'un gène induisant de grandes différences de poids a été avancée afin d'expliquer l'aspect bimodal de la distribution. Les résultats d'accouplements d'individus issus de chacun des deux sous-ensembles de la lignée ont permis de valider cette hypothèse. Il a été établi que le locus concerné (dénommé « pygmée ») possède deux allèles, l'allèle muté, responsable du nanisme, étant quasi-récessif par rapport à l'allèle normal. Ainsi, tous les individus du sous-ensemble inférieur de la distribution sont des homozygotes pour l'allèle muté.

Figure 6. Poids à 60 jours (en g) de 50 souris mâles issues d'une lignée sélectionnée pour de faibles poids.

Source : King, 1950.



Bien sûr, pour que l'action d'un gène soit décelable au vu des seules distributions, il faut que les différences induites par ce gène soient grandes en regard de la variation totale observée. Dans l'exemple de la figure 6, l'écart-type total est de 3,5 g pour l'échantillon mesuré. La différence entre la valeur moyenne des individus nains (homozygotes mutés) et celle des individus normaux (autres génotypes) est de l'ordre de 8 g, soit environ 2,25 unités d'écart-type, ce qui est considérable : à titre de comparaison, cela équivaut à un gène qui induirait chez l'homme un écart d'une quinzaine de cm pour la taille. Aussi, de tels gènes sont-ils appelés **gènes majeurs**, du fait de l'ampleur des différences phénotypiques qu'ils occasionnent.

Dans de nombreuses espèces domestiques, l'existence de gènes majeurs pour certains caractères quantitatifs a pu être décelée par analyse des distributions. Cependant, des analyses fines de génétique quantitative mettent en évidence une composante héréditaire autre que celle relative aux locus concernés (c'est le cas pour les données de la figure 6). On ne peut toutefois pas individualiser les autres locus impliqués ; aussi, suppose-t-on qu'ils sont relativement nombreux et que le polymorphisme en chacun de ces locus induit de faibles différences : on parle de **polygènes**. Enfin, il existe beaucoup de caractères pour lesquels aucun gène majeur n'a pu encore être mis en évidence.



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

3. Conséquences de l'hypothèse multifactorielle en matière de distribution

Il est aisé d'établir que, 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 :

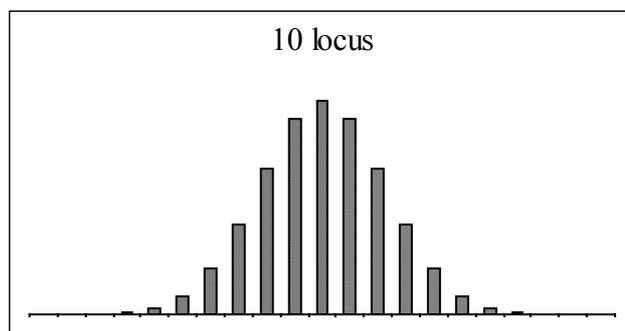
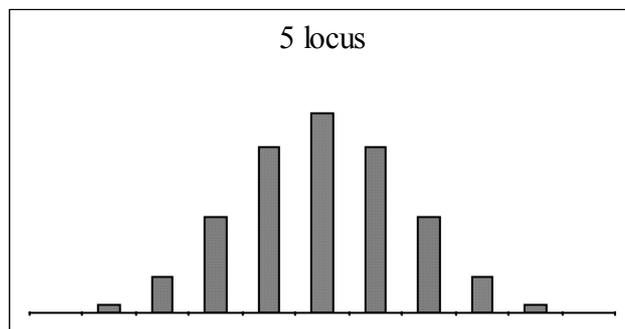
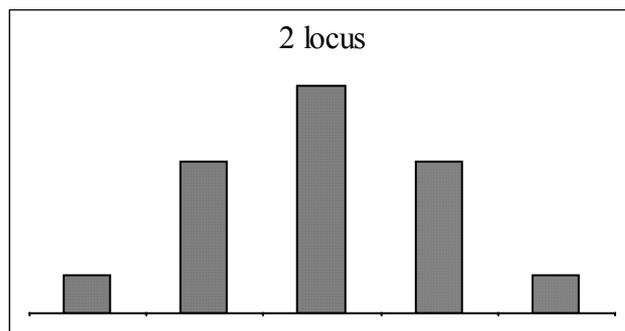
$$N = [k(k + 1)/2]^L$$

Ce nombre devient très vite extrêmement important, ainsi que le montre le tableau 2. Si l'on fait abstraction des effets du milieu, dès que le nombre de locus impliqués dépasse quelques unités, on ne s'étonnera pas de constater que la distribution des valeurs prises par un caractère devient continue. De plus, si le polymorphisme en chaque locus induit effectivement de faibles variations, la distribution des valeurs tend vers la distribution caractéristique de la loi normale, ainsi que l'indique la figure 7 (nous trouvons ici une illustration du théorème central limite, voir Stat). Ainsi, l'hypothèse multifactorielle permet de concilier la nature discontinue de l'hérédité mendélienne, d'une part, et, d'autre part, la continuité des caractères quantitatifs et la normalité souvent rencontrée de leur distribution. Cette conclusion demeure valable si l'on tient compte des variations de plus ou moins grande ampleur induites par le milieu. Il est cependant nécessaire de préciser les effets respectifs du génotype et du milieu et d'analyser la manière dont ces derniers se combinent.

Tableau 2. Nombre de génotypes 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

Figure 7. Distribution des valeurs selon le nombre de locus. La valeur d'un génotype est la somme des valeurs apportées par chaque allèle. A chaque locus, 2 allèles de même fréquence, l'un apportant une valeur nulle et l'autre une valeur de 1 : la valeur des génotypes va donc de 0 (homozygotes « 0 » à tous les locus) à 2 fois le nombre de locus (homozygotes « 1 » à tous les locus). On ne considère pas d'effet de milieu.



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

1. Approche générale

L'analyse génétique des caractères quantitatifs nécessite de passer par une représentation mathématique de la réalité : c'est la modélisation. Nous avons vu au paragraphe précédent que les variations phénotypiques observées pour un caractère quantitatif étaient imputables (1) à des différences de génotype entre individus et (2) à des différences de conditions de milieu dans lesquelles sont placés les individus mesurés. Rappelons que ces deux composantes ne peuvent être mises en évidence, et donc prises en considération, que si elles présentent elles mêmes des variations. Lorsque nous parlons de l'influence du génotype sur un caractère, implicitement nous ne considérons que les locus qui présentent un polymorphisme dans la population étudiée. Par exemple, chez les animaux, le locus codant pour l'hormone de croissance a un rôle physiologique indéniable sur la croissance ; cependant, dans une population où un allèle au locus correspondant serait fixé, tous les individus auraient le même génotype et le gène de l'hormone de croissance ne serait pas responsable des variations phénotypiques observées.

Une difficulté dans l'analyse des caractères quantitatifs réside dans le fait que l'on est incapable de déterminer le nombre de locus responsables des variations phénotypiques. Même lorsque l'on est capable de détecter un gène majeur, ou de mettre en évidence grâce à des marqueurs quelques QTL, on ignore le nombre et les effets des autres gènes en ségrégation. Aussi est-on amené à raisonner globalement à partir des variations de valeur phénotypique, qui sont les seules faciles d'accès pour l'observateur.

Dans cette section, nous ne considérons pas les effets des facteurs de milieu contrôlé : toutes les variations de nature environnementale sont supposées résulter de facteurs de micro-milieu (cf. § A.2.b). Sur le plan statistique, les effets des facteurs contrôlés (macro-milieu) sont considérés comme fixes, c'est-à-dire qu'on n'émet aucune hypothèse *a priori* quant à leur distribution. A l'opposé, les effets des facteurs non contrôlés (micro-milieu) sont considérés comme des variables aléatoires, avec une distribution *a priori* (généralement supposée normale).

2. Notion de valeur génétique (ou valeur génotypique²)

Chez certains organismes, essentiellement végétaux, on est capable de répliquer le même génotype en de nombreux exemplaires : la multiplication végétative, le développement de lignées totalement consanguines, l'haplo-diploïdisation, sont des techniques que l'on peut utiliser pour cela. L'observation de nombreux individus de même génotype dans diverses conditions de micro-milieu montre une variation des valeurs phénotypiques (cf. la figure 4 du § B.1 du présent chapitre). ***Pour caractériser les individus de même génotype, il est naturel (cf. chapitre I) de considérer leur valeur phénotypique moyenne.*** Par exemple, les lignées parentales de tabac de la figure 4 ont des moyennes de longueur de fleur de 40 mm et 93 mm respectivement (ce qui est significativement différent compte tenu des écarts-types associés).

Par définition, ***la moyenne phénotypique de nombreux individus de même génotype, exprimée en écart à la moyenne de la population, représente la valeur génétique attachée au génotype considéré.*** Certes, pour toutes les espèces pour lesquelles on ne peut pas répliquer en grand nombre le même génotype, cette notion de valeur génétique est une abstraction. Elle n'en garde pas moins son sens : elle représente l'écart à la moyenne de la population qui est imputable à l'effet du génotype.



La valeur génétique représente l'effet moyen du génotype sur un caractère donné. Elle s'exprime en unités du caractère mesuré.

Appelons μ la moyenne des valeurs phénotypiques dans une population (P) :

$$\mu = E(P)$$

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

$$G = E(P|\text{génotype}) - \mu$$

² En génétique végétale, « valeur génétique » a toujours la signification qui est donnée ici. En génétique animale, ce terme est souvent employé dans un sens plus restreint (voir Ch.III) et l'on emploie alors parfois « valeur génotypique » pour désigner la valeur d'un génotype dans son ensemble, terme que nous ne retiendrons cependant pas ici.

L'écart entre la valeur phénotypique centrée d'un individu et sa valeur génétique est imputable aux conditions de micro-milieu ; par définition cet écart (E) est désigné sous le terme d'écart environnemental. En définitive, la valeur phénotypique d'un individu se décompose ainsi :

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

3. L'héritabilité au sens large

Afin de quantifier l'importance des deux sources de variation des caractères quantitatifs, il faut raisonner à l'échelle d'une population dans son ensemble et en termes de variances. En effet, il est pertinent, et souvent utile, de pouvoir déterminer quelle est la part de la variabilité phénotypique observée dans une population qui est imputable à la variabilité génétique. A ce stade, comme précisé plus haut, on ne considère dans la variation environnementale que les facteurs de micro-milieu : les concepts que nous allons définir concernent les valeurs phénotypiques corrigées pour tous les effets systématiques des facteurs de milieu contrôlés.

Dans une population constituée d'individus de génotypes différents, la valeur génétique (G) est considérée comme une variable aléatoire, dont la distribution dépend des fréquences des génotypes (voir chapitre III). La valeur environnementale (E) est alors un résidu aléatoire, d'espérance nulle par construction, quel que soit le génotype. En expérimentation, on s'assure de l'indépendance entre G et E par la "randomisation", qui consiste à répartir les individus mesurés au hasard dans les micro-milieus : semis au hasard des graines sur une parcelle, etc. En général, on fait l'hypothèse que la distribution de E , et notamment sa variance, est la même pour tous les génotypes. Dans ce cas, G et E ont des distributions indépendantes et la variance phénotypique est la somme de la variance génétique et de la variance environnementale³ :

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

$$V_P = V_G + V_E$$

$$\sigma_P^2 = \sigma_G^2 + \sigma_E^2$$

³ On peut montrer aussi que, par construction, E étant un résidu d'espérance nulle pour tout individu, la covariance entre G et E est nulle même si la variance de E n'est pas identique d'un génotype à l'autre.

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

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

L'hérabilité au sens large est donc une proportion : c'est la part de variance phénotypique qui est d'origine génétique. Elle peut s'exprimer en pourcentage, ce qui est parfois plus parlant. Si l'on pouvait entretenir tous les individus d'une population strictement dans les mêmes conditions de micro-milieu (ce qui est une utopie !), 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 absolument pas du génotype des individus aurait une hérabilité au sens large nulle.

En génétique végétale, comme nous l'avons déjà signalé plus haut, l'entité de mesure est parfois un groupe d'individus. Ainsi, on est parfois amené à définir une hérabilité au sens large au niveau parcellaire. Sur un ensemble (une parcelle) d'individus de même génotype, les valeurs de micro-milieu ont tendance à se compenser. La variance entre parcelles des effets de micro-milieu devient alors très faible : la variance d'une moyenne de valeurs indépendantes (les valeurs de E) est inversement proportionnelle au nombre d'individus mesurés (voir Stat). Ainsi, l'hérabilité au sens large d'un caractère tend rapidement vers 1 dès lors que ce paramètre est défini au niveau d'un groupe (voir exercice n° 2).

Notons enfin qu'un caractère peut s'exprimer plusieurs fois dans la vie d'un même individu : productions annuelles d'un arbre fruitier ou d'un cep de vigne, lactations successives d'une femelle laitière, etc. Dans ce cas, la valeur environnementale (effet du micro-milieu) se décompose en un effet permanent, c'est-à-dire qui se répète d'une performance à l'autre, et un effet transitoire, variant d'une performance à l'autre (voir exercice n° 3). Comme exemple de facteur de micro-milieu permanent, on peut citer certaines caractéristiques du sol spécifiques à la micro-zone autour d'un plan ou la place dans une stabulation entravée : il s'agit de facteurs de micro-milieu identifiés mais non enregistrés (cf. les définitions de micro- et macro-milieu au § A.2.b du présent chapitre).

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

1. L'hypothèse d'additivité des effets du macro-milieu et du génotype, et ses implications

Des conditions de macro-milieu peuvent induire d'importantes variations des caractères quantitatifs. Pour caractériser une modalité particulière d'un facteur de milieu contrôlé, là encore, il est naturel de calculer la valeur moyenne des individus qui sont soumis à cette modalité. Exprimée en écart à la moyenne de la population, cette valeur moyenne est, par définition, l'effet moyen de la modalité considérée. Cette analyse suppose cependant que les différents génotypes soient répartis de façon équilibrée entre les différentes modalités de milieu (l'encadré 1 illustre la notion de confusion entre l'effet du génotype et l'effet d'un facteur contrôlé de milieu).

Encadré 1

Possibilité de confusion entre effet du génotype et effet d'un facteur de milieu contrôlé

Imaginons que nous souhaitions comparer le rendement en grains de deux variétés de blé. Les deux variétés ont été cultivées lors du même essai, au même endroit, mais pas avec le même niveau d'intensification :

niveau 1 : 250 graines / m² au semis, 60 unités d'azote / ha
niveau 2 : 300 graines / m² au semis, 120 unités d'azote / ha

Soit les résultats moyens suivants :

Variété A, niveau 2 : 85 qtx / ha
Variété B, niveau 1 : 80 qtx / ha

Avec de tels résultats, on ne peut pas conclure, car il y a une confusion totale entre le génotype (ici, la variété) et un facteur de milieu. On ne sait pas si l'on doit attribuer la supériorité du premier résultat à l'effet de la variété ou aux conditions plus intensives de culture. Pour faire une comparaison valable, il faut évidemment répartir les différents génotypes dans les différentes modalités de milieu. Soit le tableau de résultats suivant :

	Variété A	Variété B
niveau 1	75	80
niveau 2	85	90

On voit que la variété B est supérieure à la variété A dans les deux conditions de culture. Si l'on veut comparer les deux variétés sur l'ensemble des conditions de milieu, il faut corriger les données pour l'effet du niveau d'intensification. Si le dispositif est équilibré, c'est-à-dire qu'il y a les mêmes effectifs dans chacun des 4 cas considérés, c'est très simple : on constate que le niveau 2 se traduit, pour chacune des variétés, par un gain de 10 qtx / ha par rapport au niveau 1, et il suffit, par exemple, de retrancher 10 aux valeurs obtenues dans le niveau 2. Si le dispositif est déséquilibré (effectifs inégaux dans les différentes « cellules »), les choses se compliquent et il faut utiliser des méthodes statistiques plus élaborées; mais on peut toujours faire des comparaisons valables du moment qu'il n'y a pas une confusion entre un génotype et un facteur de milieu.

Dans le cas où il n'y a qu'un seul facteur de macro-milieu, ce dernier est représenté par plusieurs modalités (par exemple, différentes doses de fumure azotée, différentes températures d'élevage pour les volailles, etc.). Le modèle d'analyse le plus simple suppose que *l'effet moyen (m) d'une modalité* (tel que défini au paragraphe précédent) *s'ajoute* aux différents effets définis plus hauts, à savoir, l'effet moyen du génotype (G) et l'effet environnemental résiduel (E) :

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

Si plusieurs facteurs de milieu sont contrôlés, nous supposons que les effets moyens des différentes modalités des différents facteurs s'ajoutent les uns aux autres. Par exemple, si l'expérimentation décrite à l'encadré 1 était réalisée trois années de suite, nous pourrions proposer le modèle suivant pour en analyser les résultats :

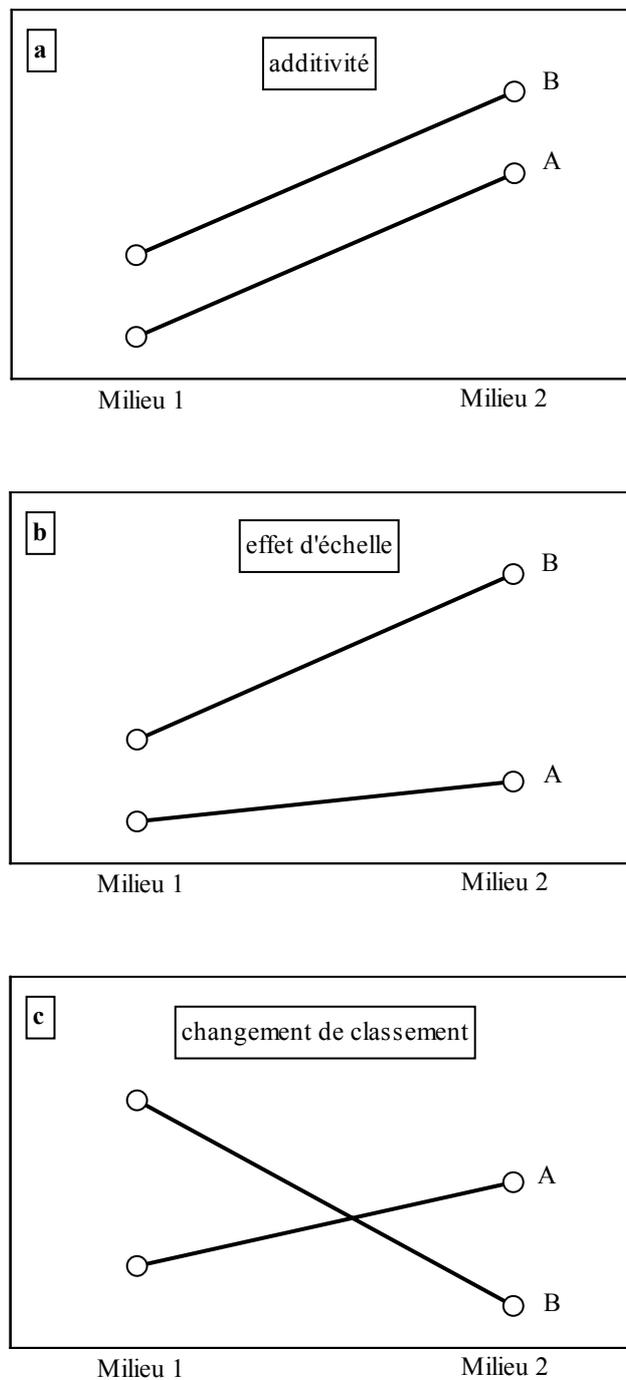
$$\begin{aligned} P = & \mu + \text{effet moyen de l'année (3 modalités} \rightarrow 3 \text{ valeurs)} \\ & + \text{effet moyen du niveau d'intensification (2 modalités} \rightarrow 2 \text{ valeurs)} \\ & + G \\ & + E \end{aligned}$$

Les implications de l'additivité supposée des effets du génotype et du milieu sont illustrées à la figure 8.a. et développées dans l'encadré 2. Si le passage d'une condition de milieu à une autre se traduit, pour un génotype donné, par un certain écart de moyenne phénotypique, alors cet écart doit se retrouver pour un autre génotype à l'occasion du même changement de conditions de milieu : l'écart entre les moyennes relatives aux différentes modalités de milieu est le même pour tous les génotypes. Pareillement, le classement des génotypes et l'écart entre leurs valeurs moyennes est le même sous différentes conditions de milieu.



La L'additivité entre les effets du génotype et du milieu implique qu'une modalité de milieu donnée affecte tous les génotypes de la même façon, et qu'un génotype donné confère le même (dés)avantage quel que soit le milieu dans lequel se développe l'individu correspondant.

Figure 8. Représentation de différentes possibilités de combinaison des effets du génotype et du milieu (voir texte). A et B représentent deux génotypes différents. En ordonnée, moyenne phénotypique pour un génotype donné dans une modalité de milieu donnée.



Encadré 2

Implications de l'hypothèse d'additivité entre les effets du génotype et ceux du macro-milieu

Soit deux génotypes (A et B) comparés dans deux conditions de milieu (1 et 2). Raisonnons sur les moyennes : soit μ_{1A} la moyenne dans le milieu 1 du génotype A, etc. Le modèle posé au § D.1 nous dit :

$$\mu_{1A} = \mu + m_1 + G_A, \quad \mu_{2A} = \mu + m_2 + G_A, \quad \mu_{1B} = \mu + m_1 + G_B, \quad \mu_{2B} = \mu + m_2 + G_B$$

On peut alors évaluer les conséquences d'un changement de milieu en calculant, pour un génotype donné, l'écart entre les moyennes obtenues dans les 2 modalités du facteur de milieu considéré :

$$\text{Génotype A : } \mu_{2A} - \mu_{1A} = \mu + m_2 + G_A - (\mu + m_1 + G_A) = m_2 - m_1$$

$$\text{Génotype B : } \mu_{2B} - \mu_{1B} = \mu + m_2 + G_B - (\mu + m_1 + G_B) = m_2 - m_1$$

Ainsi, le passage du milieu 1 au milieu 2 se traduit par la même différence de moyenne pour les deux génotypes. De la même façon, on peut comparer les deux génotypes dans chacune des deux modalités de milieu :

$$\text{Milieu 1 : } \mu_{1B} - \mu_{1A} = \mu + m_1 + G_B - (\mu + m_1 + G_A) = G_B - G_A$$

$$\text{Milieu 2 : } \mu_{2B} - \mu_{2A} = \mu + m_2 + G_B - (\mu + m_2 + G_A) = G_B - G_A$$

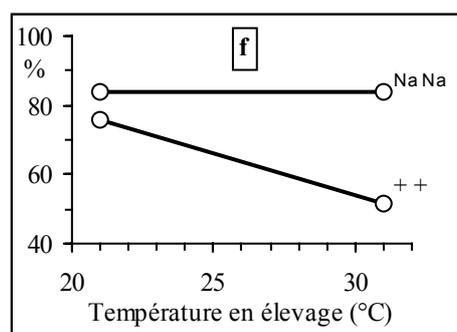
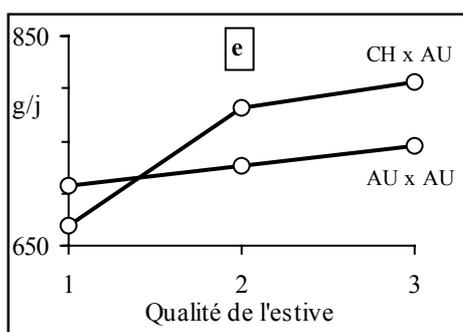
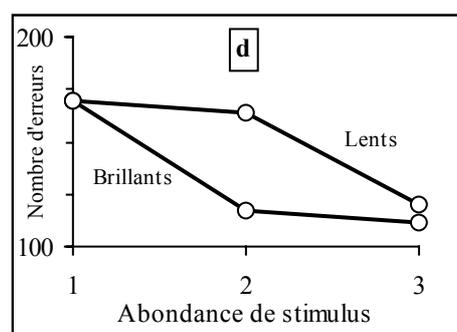
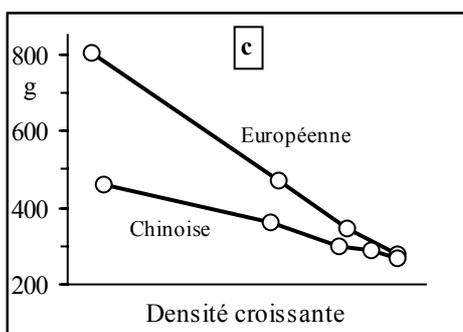
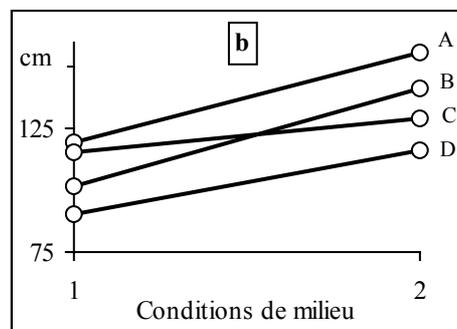
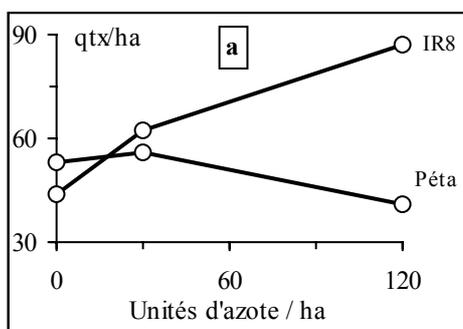
Ainsi, la différence de moyenne entre les deux génotypes est la même sous les deux conditions de milieu.

2. L'interaction génotype x milieu : définition, exemples et interprétations

Par définition, on dit qu'il y a interaction génotype x milieu lorsque l'on s'écarte de la situation d'additivité, c'est-à-dire lorsqu'une condition de milieu affecte différemment les génotypes comparés et que l'écart entre les génotypes n'est pas le même selon les conditions de milieu. Ce phénomène peut avoir plus ou moins d'ampleur et des conséquences pratiques différentes. Notamment, on distingue le cas où l'écart entre les génotypes dépend des conditions de milieu mais leur classement est inchangé (figure 8.b), et le cas où le classement des génotypes est modifié (figure 8.c). Fondamentalement, l'interaction entre les effets du génotype et du milieu est un concept statistique, dont la mise en évidence passe par les techniques de l'analyse de variance (voir Stat). Quelques exemples empruntés aux plantes cultivées, aux animaux de laboratoire et aux animaux domestiques vont nous permettre de souligner la signification biologique du phénomène (figure 9). Notez qu'ici, hormis le cas des variétés consanguines de riz et de tabac, les moyennes relevées correspondent à des groupes d'individus génétiquement proches (appartenant à la même variété, à la même lignée, à la même race) mais pas forcément strictement identiques : ces valeurs traduisent l'effet d'un ensemble de génotypes, ayant un certain degré de ressemblance, et non pas strictement l'effet d'un seul génotype.

Figure 9. Quelques exemples d'interaction génotype x milieu.

- a : Rendement en grains (qtx / ha) de deux variétés de riz en fonction de la dose d'azote apportée. IR8 = variété naine de 1968 ; Péta = variété ancienne de taille normale. (Source : Jennings, 1976)
- b : Hauteur (cm) de la plante entière de 4 lignées de tabac cultivées dans deux conditions de milieu différentes. (Source : Jinks, cité par Falconer, 1981)
- c : Gain de poids (grammes) en 5 mois de deux souches de carpe, l'une d'origine européenne, l'autre d'origine chinoise, selon la densité d'animaux par bassin d'élevage. (Source : Moav et al., 1975)
- d : Nombre d'erreurs au cours d'un test de labyrinthe effectué à l'âge de 65 jours pour deux lignées de rats, « lents » et « brillants », selon l'abondance de stimulus dans le milieu d'élevage (notée de 1 « pauvre » à 3 « riche »). (Source : Cooper et Zubek, 1958)
- e : Vitesse de croissance (g/jour) réalisée en estive par des génisses issues d'une mère de race Aubrac (AU), selon la qualité de l'estive (notée de 1 « médiocre » à 3 « bonne ») et la race de leur père. CH x AU = père de race Charolaise ; AU x AU = père de race Aubrac. (Source : Vissac, 1970)
- f : Taux d'éclosion (%) en fonction de la température ambiante d'élevage des reproductrices, dans deux lignées de poules différant pour leur génotype à un gène responsable d'une réduction de l'emplumement (gène « cou nu »). Na Na = génotype à emplumement réduit ; + + = génotype à emplumement normal. (Source : Mérat et al., 1989)



Ces graphiques illustrent les différences entre génotypes dans leur réponse aux changements de milieu et le fait que le milieu peut agir comme un facteur limitant. L'ancienne variété de riz Péta (figure 9.a) ne répond pas à une intensification du milieu alors que la variété IR8 y répond très bien. Le rendement plus faible de la variété Péta lorsque le niveau de fertilisation azotée est très élevé s'explique par la hauteur de paille de cette variété, qui ne lui permet pas de supporter des poids de grain importants. A l'opposé, un apport azoté nul constitue manifestement un facteur limitant à l'expression des aptitudes de productivité de la variété IR8. On observe exactement les mêmes phénomènes quand on compare les variétés anciennes et actuelles de blé. Les figures 9.c et 9.e illustrent également cette notion de facteur de milieu limitant. Dans le premier cas (9.c), la croissance des animaux est limitée par leur densité excessive dans les bassins (compétition accrue pour l'accès à la ressource alimentaire) : pour les fortes densités, on n'observe plus de différence entre les deux souches de carpes. Dans le second cas (9.e), c'est l'aptitude de croissance de la race bovine Charolaise (spécialisée dans la production de viande) qui ne s'exprime pas quand les conditions du pâturage en estive sont mauvaises.

La figure 9.d montre, de façon exemplaire, comment un facteur de milieu peut gommer d'éventuelles différences entre génotypes. Dans un milieu « standard », deux lignées différentes de rats présente une grande différence d'aptitude à résoudre le test du labyrinthe, d'où les dénominations respectives de « lents » et de « brillants ». Si le milieu d'élevage est très riche en stimulus (sous forme de divers objets dans les cages), les valeurs respectives des deux lignées sont similaires, les individus « lents » en milieu « standard » ayant fait de nets progrès dans ce milieu « amélioré ». Placées dans un milieu dépourvu de stimulus, les deux lignées commettent également le même nombre d'erreurs, mais il est très élevé.

Les caractères d'adaptation aux contraintes du milieu jouent un rôle important dans l'existence d'interaction génotype x milieu. Quand les conditions de milieu sont difficiles, les veaux de race pure Aubrac ont une vitesse de croissance plus élevée que les veaux croisés Charolais (figure 9.e), sans doute grâce à leur comportement qui les incite à effectuer des déplacements importants à la recherche de nourriture ainsi qu'à leur aptitude à diversifier leur alimentation et à valoriser des fourrages grossiers. De même, chez la poule, la lignée homozygote « cou nu » est indifférente à l'élévation de la température en ce qui concerne le taux d'éclosion, alors que la lignée normale y est très sensible ; ceci est en liaison avec l'emplumement réduit des animaux qui induit une thermotolérance accrue.

Ces résultats, ainsi que de nombreux autres, permettent de dégager des tendances générales quant à l'existence possible d'interaction génotype x milieu. Tout d'abord, il faut bien sûr que les conditions de milieu soient enregistrées : comme déjà souligné, seuls des facteurs de macro-milieu sont, de ce fait, mis en cause dans ce phénomène. De surcroît, on observe d'autant plus facilement une interaction que les modalités du milieu sont tranchées. De même, ce phénomène s'observe souvent quand les différences entre génotypes sont importantes : comparaison de deux variétés ou de deux races. Il a été cependant mis en évidence des interactions impliquant les génotypes individuels au sein d'une même variété ou d'une même race. Notons enfin que, l'interaction peut dépendre des génotypes comparés : dans le cas de la hauteur de plante chez le tabac (figure 9.b), il aurait été impossible de conclure à l'existence d'une interaction si seules les lignées A et D avaient été observées.

3. Conséquences pour la pratique de l'amélioration des espèces domestiques

Dans la pratique de l'amélioration des plantes ou des animaux, l'existence d'interaction pose deux problèmes majeurs : (1) dans un milieu donné, quel est le meilleur génotype (quel est le génotype qui répond le mieux aux objectifs de l'amélioration) et (2) dans quel milieu pratiquer la sélection ? Ces deux questions sont en fait les deux facettes du même problème, qui a d'autant plus d'acuité que l'interaction est forte. En pratique, on s'en préoccupe plus particulièrement quand il y a un changement de classement entre génotypes, puisque dès lors aucun génotype n'est meilleur que tous les autres dans l'ensemble des conditions de milieu. La question du choix du meilleur génotype se pose notamment quand on souhaite utiliser dans un milieu à fortes contraintes des génotypes qui ont été améliorés dans des conditions plus favorables. Tant en production végétale qu'en production animale, l'importation dans les pays en voie de développement de génotypes améliorés pour une forte productivité dans les conditions des pays industrialisés en est un exemple criant. De la même façon, lorsque l'on poursuit ou démarre une sélection, il faut éviter de trop artificialiser le milieu dans lequel on entretient les candidats à la sélection : certes, une telle pratique peut permettre de mieux révéler les différences entre individus, le milieu n'étant pas limitant, mais s'il y a un trop grand décalage entre le milieu de sélection et le milieu de production (où se développeront les descendants des individus sélectionnés), tout le progrès réalisé par la sélection en milieu favorable peut être annihilé du fait d'un phénomène d'interaction. Les exemples de l'encadré 3 permettent d'illustrer ces problèmes, dont l'importance est sans doute trop souvent négligée dans la pratique de l'amélioration des espèces domestiques.

Encadré 3

Interaction génotype x milieu et amélioration des espèces domestiques

Choix des types génétiques pour le repeuplement porcin en Haïti

En Haïti, le porc est un animal qui tient une très grande place dans l'économie de subsistance des familles paysannes. L'élevage (un animal par famille) se fait dans des conditions difficiles. Notamment, les ressources alimentaires sont très irrégulières dans l'année et comportent une part non négligeable de produits celluloseux. En 1982, suite à une épidémie de peste porcine déclarée en 1978, le gouvernement haïtien décréta l'abattage de tout le cheptel, soit plus d'un million de têtes, dont 95 % de race Créole. En 1983, débuta le repeuplement avec des animaux appartenant à 7 races d'Amérique du Nord, plus productives dans leur pays d'origine que ne l'était la race Créole en Haïti. Très vite, il est apparu que ces animaux étaient incapables, dans les conditions de l'élevage familial haïtien, d'exprimer leurs aptitudes de productivité. Souvent même, les animaux dépérissaient au bout de quelques semaines seulement. Ainsi, ces races n'ont pu convenir qu'à une minorité d'éleveurs possédant la trésorerie nécessaire à l'entretien d'installations appropriées, à l'achat d'aliments concentrés et à la rémunération d'un minimum de main d'œuvre. Face à cet échec, dû à un manque de bon sens, un programme de coopération franco-haïtien s'est mis en place dès 1985, en vue d'un repeuplement fondé sur des génotypes mieux adaptés, obtenus par croisement des races Créole des Antilles françaises, chinoise Meishan et Gasconne.

Décalage entre milieu de sélection et milieu de production dans le cas de la riziculture

La variété de riz IR8, déjà évoquée plus haut, est une des variétés sélectionnées à partir du début des années 60 par un organisme international (IRRI, sis aux Philippines) sur lesquelles devait se fonder la "révolution verte". Cette variété possède un gène induisant une forte réduction de la hauteur de tige. Elle a été sélectionnée dans des conditions de bonne maîtrise de l'eau, de la fumure et des traitements phytosanitaires. L'utilisation de cette variété dans des conditions beaucoup moins maîtrisées a conduit à de nombreux échecs (à Madagascar notamment). En effet, les variétés à paille courte subissent d'importants dégâts au cours d'inondations qui peuvent se produire quand on ne possède pas la maîtrise de l'eau. Par ailleurs, le surcroît de productivité de la variété IR8 suppose l'utilisation de fumure azotée (voir figure 9.a). Ainsi, selon les pays et les choix des producteurs, soit cette variété n'a pu convenir qu'à une minorité, soit son utilisation a contribué à faire émerger une situation d'endettement des paysans.

EXERCICES

Exercice n°1

On mesure le poids de grain par plante chez le maïs (g). L'écart-type phénotypique dans une grande population est de 15. L'écart-type phénotypique dans une lignée pure (consanguine) est de 12. Calculer l'héritabilité au sens large du poids de grain par plante, (1) dans la lignée consanguine et (2) dans la grande population. Expliciter votre raisonnement et souligner les réserves que ce dernier peut susciter.

Exercice n°2

Dans une espèce de plante cultivée, on mesure le rendement moyen sur des ensembles de nombreuses répétitions du même génotype. L'emplacement de chaque plante sur la parcelle expérimentale est choisi au hasard. Soit P_{ij} la valeur mesurée sur la j ème répétition du génotype i . Le modèle d'analyse à l'échelon individuel est le suivant :

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

avec μ la moyenne générale, G la valeur génétique et E la résiduelle environnementale. Pour chaque génotype, on dispose de n répétitions dont la moyenne ($P_{i\bullet}$) s'écrit :

$$P_{i\bullet} = \frac{1}{n} \sum_{j=1}^n P_{ij}$$

En écrivant le modèle d'analyse pour les valeurs phénotypiques moyennes ($P_{i\bullet}$), montrer que l'héritabilité au sens large calculée à l'échelle de moyennes de génotypes tend vers 1 quand n devient très grand.

Exercice n°3

Un échantillon de 10 brebis a fourni les valeurs suivantes de production laitière (kg) cours de leurs première et deuxième lactations :

Individu	1	2	3	4	5	6	7	8	9	10
1ère lactation	150	121	183	144	178	116	154	163	125	196
2ème lactation	217	143	262	170	215	144	191	180	181	172

- 1) Comparer les moyennes des valeurs pour les deux rangs de lactation considérés. Qu'en concluez-vous quant à l'effet du rang de lactation sur la production laitière et quelles observations complémentaires devrait-on faire afin de mieux affermir les conclusions ?
- 2) Par définition, **la répétabilité** d'un caractère (r) est égale à la corrélation entre deux mesures de ce caractère chez le même individu. Calculer, sur cet échantillon, la répétabilité de la production laitière chez la brebis. Quelles sont les limites de validité du résultat obtenu ?
- 3) Une brebis (non présente dans l'échantillon) a produit 130 kg de lait en première lactation. Quelle prédiction peut-on faire quant à sa production en deuxième lactation ?