

# CHAPITRE 1

## PRINCIPAUX CONCEPTS ET MÉTHODES RELATIFS A LA DESCRIPTION DE LA DIVERSITÉ GÉNÉTIQUE D'UNE ESPÈCE

R. GUYOMARD

INRA-CRJ, Laboratoire de Génétique des Poissons, 78350 JOUY-EN-JOSAS

### I - STRUCTURE DE L'ESPÈCE

#### 1. Concept biologique de l'espèce

L'appartenance de deux individus à une même espèce a longtemps été déterminée d'après leur degré de ressemblance morphologique. Il est peu à peu apparu que l'application de ce critère (qualifié de typologique, car il définit une espèce par un type morphologique) conduisait parfois à des conclusions erronées\*. Un exemple fameux est constitué par le mâle et la femelle du Colvert qui avaient été classés à l'origine dans deux espèces différentes, respectivement *Anas boschas* et *Anas platyrhynchos*. Les difficultés rencontrées dans l'application du concept typologique et l'émergence de la théorie de la sélection naturelle, accordant une place importante à la diversité interindividuelle au sein de l'espèce, ont peu à peu conduit les taxonomistes à adopter l'interfécondité, et non plus la similitude, comme principal critère d'appartenance à une même espèce. Celle-ci est donc désormais définie comme un ensemble d'individus contemporains qui, dans les conditions naturelles, ont une probabilité non nulle d'engendrer une descendance commune fertile (Mayr, 1974). Ce concept est qualifié de biologique, car sa définition est de nature biologique.

L'apparition du concept biologique de l'espèce a constitué un progrès remarquable pour la compréhension des phénomènes évolutifs et la résolution de nombreux problèmes taxonomiques. S'il est désormais unanimement reconnu par les biologistes, son application n'est pas toujours facile et la définition de nombreuses espèces, notamment de poissons d'eau douce, repose encore sur des critères morphologiques.

#### 2. Notion de population

L'aire de distribution d'une espèce peut être très étendue et morcelée; les individus qui la composent ont alors tendance à s'accoupler en priorité localement. Il faut donc admettre qu'une espèce peut être constituée de sous-unités reproductives privilégiées, appelées populations, qui peuvent se maintenir sur de nombreuses générations. Idéalement, une population se définit comme une communauté dans laquelle les individus de sexes opposés et sexuellement mûrs s'associent au hasard pour se reproduire. On nomme panmixie ce mode de reproduction. Par contre, entre deux populations distinctes, mais appartenant à la même espèce, il existera un isolement reproducteur partiel. L'une des principales difficultés auxquelles se heurtent les taxonomistes est de déterminer si l'isolement reproductif entre deux populations est partiel ou total et si celles-ci appartiennent ou non à la même espèce. D'un point de vue génétique, la population peut être décrite comme un ensemble de gènes (le terme de "pool génique" est fréquemment utilisé) partiellement isolé des autres populations. A chaque génération, la reproduction sexuée assure le brassage des gènes dans la population et l'apparition de nouvelles combinaisons géniques (fig 1). La sélection naturelle favorise les plus efficaces d'entre elles et les gènes présents dans la population atteignent ainsi un degré de coadaptation qu'ils n'ont pas nécessairement avec ceux des autres populations. La population constitue donc une unité majeure de l'évolution de l'espèce.

#### 3. Facteurs d'évolution de la variabilité génétique des populations

Il est extrêmement rare de ne pas observer de variabilité interindividuelle pour n'importe quel caractère lorsqu'on parcourt le domaine d'une espèce. Méconnaître la nature de cette variabilité ainsi que les facteurs qui en sont à l'origine peut conduire à des interprétations totalement erronées quant à sa signification.

Les caractères observés sont des phénotypes, c'est-à-dire le résultat de l'interaction d'une information génétique et de l'environnement. Il faut donc se garder d'affirmer, *a priori*, que toute variabilité phénotypique reflète une variabilité génétique car elle peut être, dans certains cas, purement environnementale. Il est nécessaire d'analyser la diversité observée selon des méthodes appropriées (cf la partie II de ce chapitre), afin de déterminer s'il existe une variabilité génétique sous-jacente.

La nature chimique, l'organisation, ainsi que le mode de répllication et de transmission de l'information génétique sont bien connus (éventuellement, se reporter à des ouvrages de base

\* en classant les individus interféconds dans des espèces différentes ou, au contraire, en regroupant des entités morphologiquement semblables, mais évolutivement distinctes (espèces jumelles).

comme, par exemple, Prévost 1976 ; cf également la figure 1 et le glossaire situé en fin de fascicule). En revanche, il est peut-être utile de rappeler la nature et le rôle, chez les organismes diploïdes sexués, des différents facteurs d'évolution de cette information génétique. A côté de la *sélection* qui constitue la notion maîtresse de la théorie de l'évolution des espèces, il existe quatre autres facteurs d'évolution qui sont la *mutation*, la *recombinaison*, la *migration* et la *dérive génétique*.

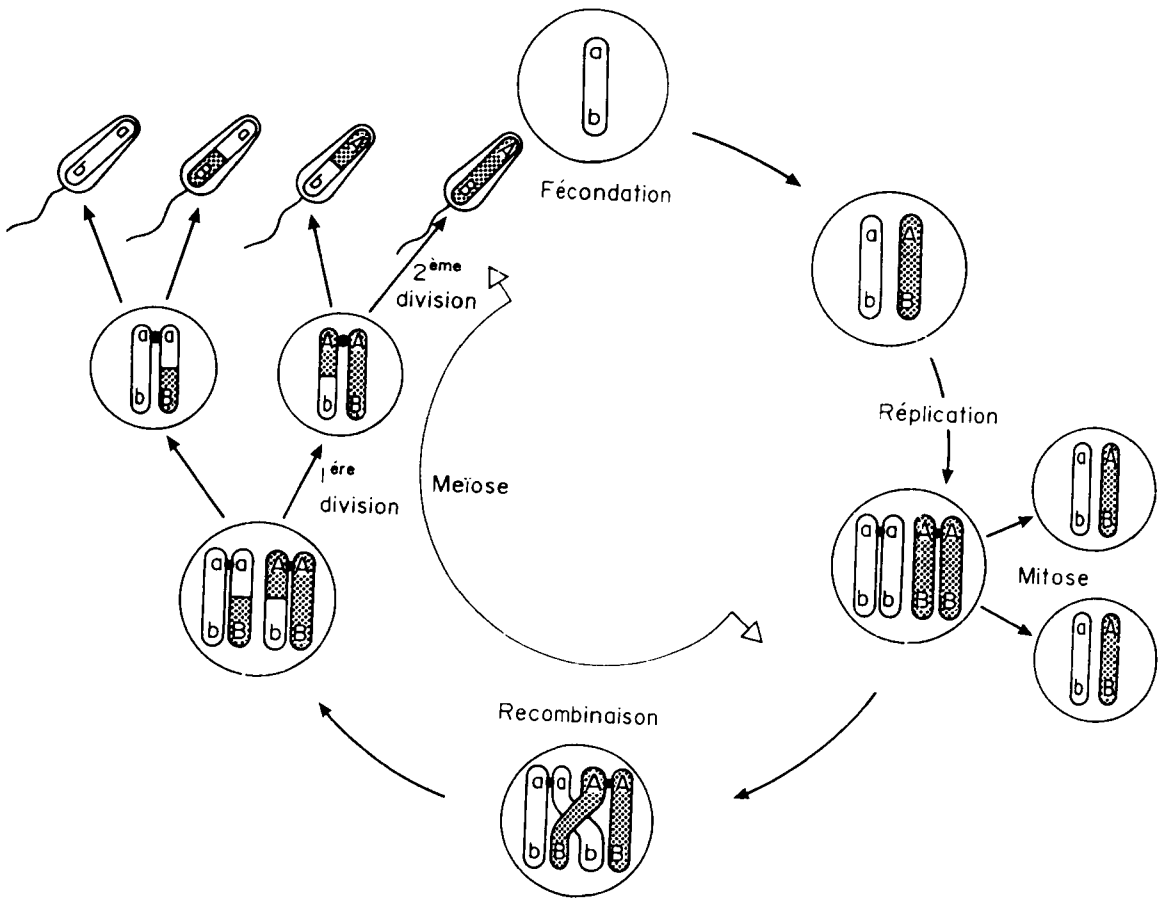
La sélection favorise les gènes, ou associations de gènes, qui confèrent aux individus les meilleures chances de survie et de reproduction (en fait, la sélection n'agit qu'indirectement sur les gènes ou les génotypes par l'intermédiaire des phénotypes). Elle constitue le seul facteur qui imprime, à la population, une évolution dans un sens bien précis. Cette évolution sera cependant très différente selon les modalités d'action de cette sélection. Ceci peut être montré simplement dans le cas d'un locus présentant un polymorphisme dû à un couple d'allèles A et a. Trois génotypes sont possibles : AA, aa (génotypes homozygotes) et Aa (génotype hétérozygote). Chacun d'eux peut être affecté d'une valeur sélective (fitness chez les anglo-saxons). Celle-ci est approximativement le produit de deux facteurs : d'une part, la probabilité de survie du génotype de la fécondation au stade reproducteur ; d'autre part, le succès reproducteur, exprimé par le nombre de gamètes efficaces (ou de descendants) produits, de ce génotype. Le devenir du polymorphisme va dépendre de ces valeurs sélectives. Si l'un des génotypes homozygotes possède toujours, au cours des générations, la plus forte valeur sélective, la population ne conservera finalement que l'allèle correspondant au génotype avantageux (en l'absence de mutation ou de migration ; cf plus loin) et le locus se fixera pour cet allèle. Mais la sélection agit parfois, au contraire, comme un facteur de maintien du polymorphisme. C'est le cas lorsque le génotype hétérozygote a toujours une plus forte valeur sélective que les deux homozygotes. Le polymorphisme peut aussi être stable si les valeurs sélectives ne sont pas constantes et que les deux génotypes homozygotes sont alternativement avantageux. La cause des variations des valeurs sélectives peut être les fluctuations des facteurs du milieu, mais aussi la composition génétique de la population : il a ainsi été démontré que certains génotypes sont avantageux lorsqu'ils sont rares, puis désavantageux lorsqu'ils deviennent fréquents.

Quel que soit le mécanisme impliqué, la sélection opère nécessairement à travers des mortalités et des fécondités différentielles. Il est admis que le taux moyen de substitution d'un gène par sélection naturelle est forcément faible au sein d'une population pour rester compatible avec les taux de mortalités et de fécondités observés, même dans le cas d'espèces très fécondes (Nei, 1975). Par contre, à l'échelle de l'espèce, des modifications génétiques rapides peuvent se produire si celle-ci est structurée en populations génétiquement très différentes et si l'une ou plusieurs d'entre elles sont éliminées par sélection.

Contrairement à la sélection, les quatre autres facteurs modifient au hasard la composition génétique de la population. Trois d'entre eux, la mutation, la migration et la recombinaison sont générateurs de variabilité. La mutation constitue la source première de nouveauté génétique. Il semble toutefois que les taux des mutations soient, en général, très faibles (de  $10^{-5}$  à  $10^{-7}$ ) et que la grande majorité d'entre elles n'aient qu'un effet très léger sur le phénotype. Les conséquences immédiates de la mutation sur la diversité génétique des organismes sexués sont vraisemblablement minimales lorsqu'on les compare à celles de la recombinaison génique. Celle-ci est le résultat de deux événements qui caractérisent la reproduction sexuée : la méiose et la fécondation. Les échanges entre chromosomes paternels et maternels homologues intervenant au cours de la première division de méiose, puis la ségrégation des chromosomes en fin de méiose aboutissent à la production d'un nombre considérable de nouveaux types gamétiques (fig 1). Le nombre possible de types zygotiques issus de l'association, au hasard, de ces gamètes est encore plus impressionnant. La recombinaison (il faudrait dire la reproduction sexuée) est, de loin, le principal facteur de renouvellement de la variabilité génétique dans la population.

L'apport de variabilité dû à la migration est évidemment bien moindre que celui de la recombinaison génique, mais reste certainement supérieur à celui de la mutation dans la plupart des cas (Mayr, 1974). Le rôle majeur du courant de gènes est de maintenir la cohésion génétique entre populations qui composent l'espèce. De ce point de vue, son efficacité est remarquable ; en effet, des taux de migration extrêmement faibles (de l'ordre d'un immigrant dans une population de  $10^4$  à  $10^5$  individus) peuvent suffire à prévenir toute différenciation, même en présence de forces sélectives relativement fortes (Li, 1976 ; Nei, 1987). Par contre, dès que le courant de gène est interrompu entre deux populations, celles-ci peuvent diverger très rapidement (Nei, 1975). L'interruption d'un courant de gène peut être, *a priori*, due à des causes intrinsèques — ce sont les mécanismes d'isolement reproducteur acquis par une population vis-à-vis d'une autre (Mayr, 1974) — ou extrinsèques — c'est l'isolement géographique. Les généticiens admettent pratiquement tous qu'il ne peut y avoir apparition de mécanisme d'isolement sans mise en place préalable d'un isolement géographique. Plus concrètement, la différenciation sur un même territoire d'une population en deux populations génétiquement différenciées semble impossible (à l'exception toutefois de la spéciation par polyploïdisation ; cf Mayr, 1974).

Les effets dus aux fluctuations d'échantillonnage sur la variabilité génétique d'une population peuvent être illustrés à l'aide d'un cas extrême, celui d'une population constituée d'un seul mâle et d'une seule femelle à chaque génération (figure 2). A une certaine génération, ces deux individus



**figure 1** : Devenir d'une paire de chromosomes homologues et ségrégation des allèles présents à deux loci situés sur ces chromosomes au cours d'un cycle biologique chez un organisme diploïde à reproduction sexuée. Grâce à la recombinaison, à la ségrégation des chromosomes (1<sup>re</sup> et 2<sup>e</sup> divisions de méiose) et à la fécondation, la reproduction sexuée produit un nombre considérable de nouvelles combinaisons de gènes à chaque génération.

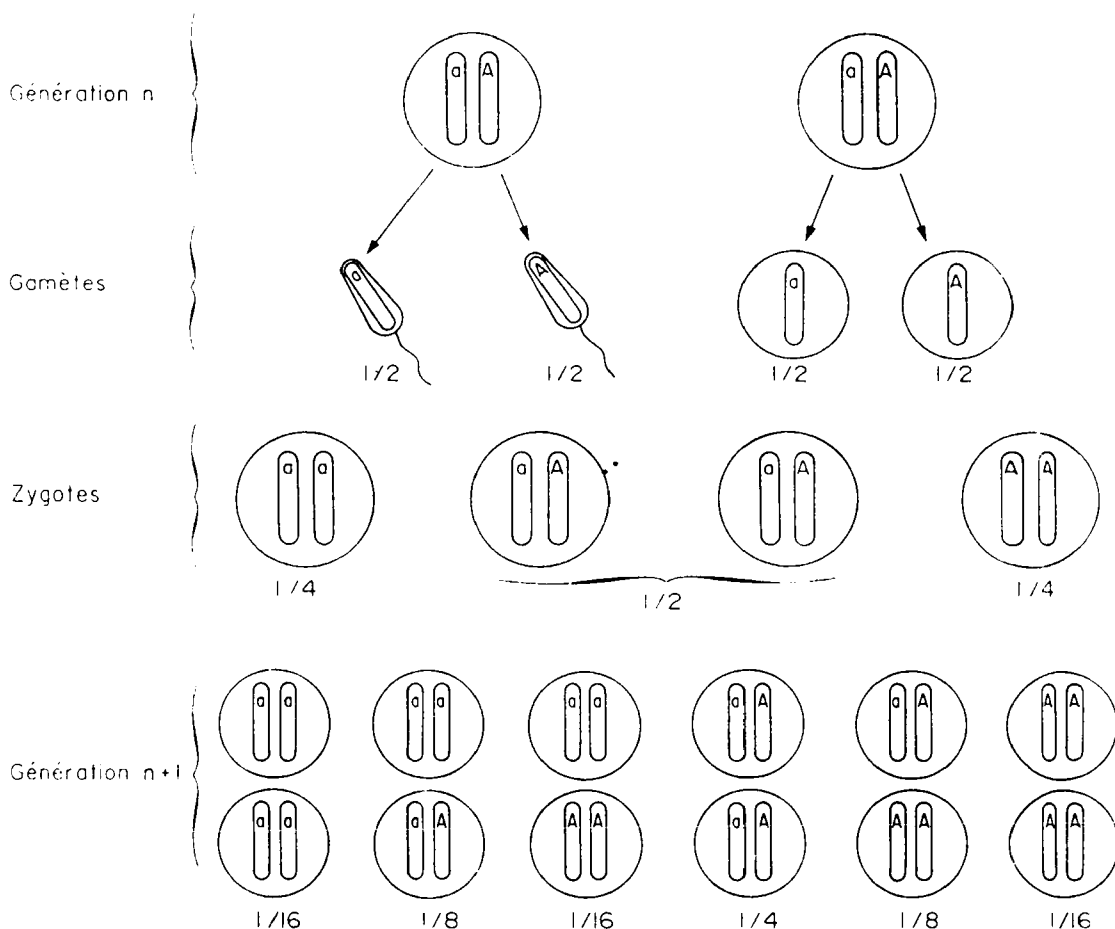
**figure 1** : Gene segregation at loci localized on the same pair of chromosomes in a diploid, sexually reproducing, organism. Recombination, chromosome segregation and fertilization produce a considerable number of new gene combinations at each generation.

peuvent être hétérozygotes et produisent alors des descendants hétérozygotes et homozygotes. Il n'y a qu'une chance sur quatre pour que deux individus tirés au hasard parmi l'ensemble des descendants possibles soient l'un et l'autre hétérozygotes. La nouvelle génération a donc une probabilité plus forte d'être différente de la précédente. Ce sont ces fluctuations d'échantillonnage des fréquences géniques que l'on appelle dérive génétique. Elle se produit dans toutes les populations, mais se fait particulièrement sentir quand l'effectif de la population est ou devient petit. Elle peut alors conduire à une perte importante de variabilité. Il existe une école de pensée, dite neutraliste, qui, sans nier l'existence de la sélection, considère que la plus grande partie de la variabilité génétique observée, au sein d'une espèce, est sélectivement neutre, sans signification adaptative et que l'évolution des espèces comporte essentiellement des événements aléatoires. A l'opposé, on trouve les sélectionnistes qui estiment que les changements sont principalement dus à la sélection.

## II. LES MÉTHODES DE DESCRIPTION DE LA DIVERSITÉ GÉNÉTIQUE DE L'ESPÈCE

En général, la diversité génétique d'une espèce ne peut être mesurée à partir d'observations directement faites dans la nature ; l'existence probable d'effets de milieu ne permettant pas d'interpréter les variations observées. Le premier problème que doit résoudre le généticien est celui de la recherche et de l'acquisition de méthodes adéquates permettant d'estimer la variabilité génétique.

Les méthodes dont on dispose aujourd'hui peuvent être classées en trois catégories, selon qu'elles se rapportent à l'étude de la variabilité des caractères quantitatifs, des variations chromosomiques ou du polymorphisme moléculaire. Quel que soit le caractère analysé, sa diversité génétique au sein de l'espèce est généralement décrite en terme de variabilité intra et interpopulations. Cette démarche traduit l'importance que l'on accorde à la notion de population (cf. 1,2).



**figure 2** : Dérive génétique à un locus auquel ségrègent deux allèles (A et a) dans une population de taille constante égale à 2 et constituée d'un mâle et d'une femelle hétérozygotes à la génération n. Les fractions portées sur la figure correspondent aux probabilités de transmission de A et a aux gamètes mâles et femelles, d'apparition de zygotes de génotype AA, Aa ou aa et des différentes combinaisons de deux génotypes à la génération n+1. La probabilité que la génération n+1 soit, comme la génération n, composée de deux individus hétérozygotes n'est que de 1/4.

**figure 2** : Genetic drift at a locus possessing 2 alleles (A et a) in a population of constant size of 2 individuals. The figures correspond to the probabilities of (1) allele sampling in gametes, (2) occurrence of homozygous and heterozygous zygotes, (3) occurrence of the different possible generations n+1 when both parents are heterozygous. The probability that generation n+1 is composed of two heterozygous individuals at generation n is 1/4 only.

### 1. L'analyse des caractères quantitatifs

Les caractères quantitatifs sont des caractères à variations interindividuelles continues (par exemple, la taille) ou susceptibles de prendre un nombre relativement élevé de valeurs (par exemple, le nombre de coeca pyloriques). Leur déterminisme génétique (nombre de gènes contrôlant le caractère) n'est pas connu et est supposé polygénique. Ces caractères peuvent être sensibles aux variations de milieu et ils doivent être nécessairement étudiés sur des individus élevés dans des conditions identiques de milieu et selon des protocoles expérimentaux qui, chez les poissons, s'avèrent lourds à réaliser (nécessité d'élever indépendamment un nombre de lots qui augmente

rapidement avec le nombre de populations comparées). De plus, les résultats obtenus risquent de n'être valables que dans les conditions de milieu dans lesquelles ont été faites les comparaisons, s'il existe d'importantes interactions génotype-milieu pour les caractères étudiés. Ces méthodes sont surtout utilisées pour l'analyse de la variabilité génétique de traits zootechniques importants, dans le cadre de programmes d'amélioration génétique. Elles sont présentées, ainsi que les concepts qui s'y rattachent (variabilité génétique additive, de dominance, d'épistasie, hérabilité), dans le chapitre suivant.

## 2. Les méthodes caryologiques

L'ADN, support de l'information génétique, est structuré en chromosomes qui peuvent être observés en microscopie. La méthode de base consiste à bloquer les cellules en division à un stade particulier (métaphase) lorsque les chromosomes sont compacts et à visualiser ceux-ci à l'aide de colorants ayant une affinité pour l'ADN (fig. 3). Le nombre et la forme de ces chromosomes sont presque toujours les mêmes pour une espèce donnée. Toutefois, dans certaines espèces, des variations interindividuelles des caryotypes sont parfois observées. La figure 4 présente un exemple de polymorphisme chromosomique dû à la fusion de chromosomes par les centromères chez la truite arc-en-ciel. Les conséquences de mélanges de populations conspécifiques ayant des formules chromosomiques différentes peuvent être graves (problèmes d'appariement des chromosomes à la méiose et formation de gamètes déséquilibrés) et justifient ces études caryologiques. L'expérience montre cependant que les méthodes de base n'apportent, en général, que peu d'informations sur la variabilité intraspécifique. Des techniques plus performantes, dites de banding (car elles font apparaître des bandes sur les chromosomes) ont été mises au point chez les vertébrés supérieurs, mais leur transposition aux poissons s'est avérée décevante. L'efficacité de la caryologie pour décrire la diversité génétique intraspécifique est donc, pour l'instant, très réduite chez les poissons.

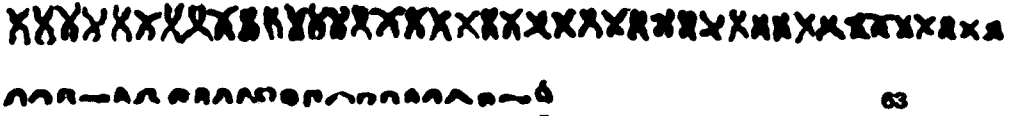


**figure 3** : Chromosomes présents dans le noyau d'une cellule somatique de truite arc-en-ciel (le nombre de chromosomes est ici de 63). Chaque chromosome est dédoublé en deux chromatides-sœurs encore reliés par le centromère, car la réplication de l'ADN qui précède la formation des deux cellules-filles issues d'une mitose a déjà eu lieu. Les chromosomes ont donc des formes en X (chromosomes métacentriques) ou en V (chromosomes acrocentriques). Outre le nombre de chromosomes, on utilise fréquemment un autre paramètre pour décrire le caryotype d'un individu : le nombre fondamental (NF) qui est le nombre total de bras chromosomiques, chaque chromosome métacentrique comptant pour deux bras et chaque chromosome acrocentrique pour un. NF est généralement beaucoup plus constant que le nombre de chromosomes dans une espèce et traduit mieux la stabilité de la quantité d'ADN au sein de l'espèce. Chez la truite arc-en-ciel, NF est toujours égal à 104 (cf. figure 4). Cliché de D. CHOURROUT, INRA.

**figure 3** : Chromosome preparation from rainbow trout embryonic cells. The chromosome number is 63 in this individual. All the chromosomes are divided in 2 sister chromatids linked by the centromere because the replication which precedes the cell division has already occurred. The X - and V - shaped chromosomes are termed metacentric and acrocentric respectively. The metacentric chromosomes have 2 arms and the acrocentric ones have only one. The total arm number found in a metaphase is termed "fundamental number" and is more constant than the chromosome number within a species. In rainbow trout,  $NF = 104$  (cf. figure 4). Photography by D. CHOURROUT, INRA.



64



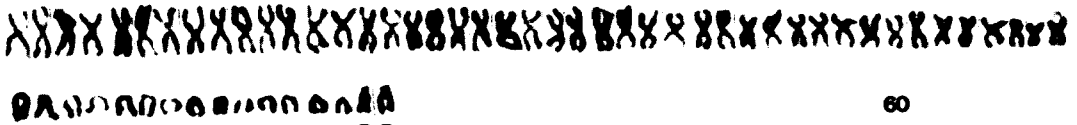
63



62



61



60



59

figure 4 : Chromosomes de six truites arc-en-ciel différant par leurs nombres chromosomiques (59 à 64), mais ayant le même nombre fondamental (NF = 104). Ces variations du nombre de chromosomes peuvent s'expliquer à partir des fissions de chromosomes métacentriques en deux acrocentriques ou, au contraire, par des fusions de deux acrocentriques donnant un métacentrique. Ainsi, un individu à 60 chromosomes possède deux acrocentriques de plus, mais un métacentrique de moins, qu'un individu à 59 chromosomes (d'après CHOURROUT et HAPPE, 1986).

figure 4 : Chromosomes of six individuals differing in their chromosome numbers (59 to 64), but having the same fundamental number (NF = 104). These variations can be explained by fissions of metacentric chromosomes in two acrocentric ones or, on the contrary, fusions of two acrocentric chromosomes in one metacentric (from CHOURROUT and HAPPE, 1986).

### 3. Le polymorphisme moléculaire

#### a. Principe de la technique électrophorétique

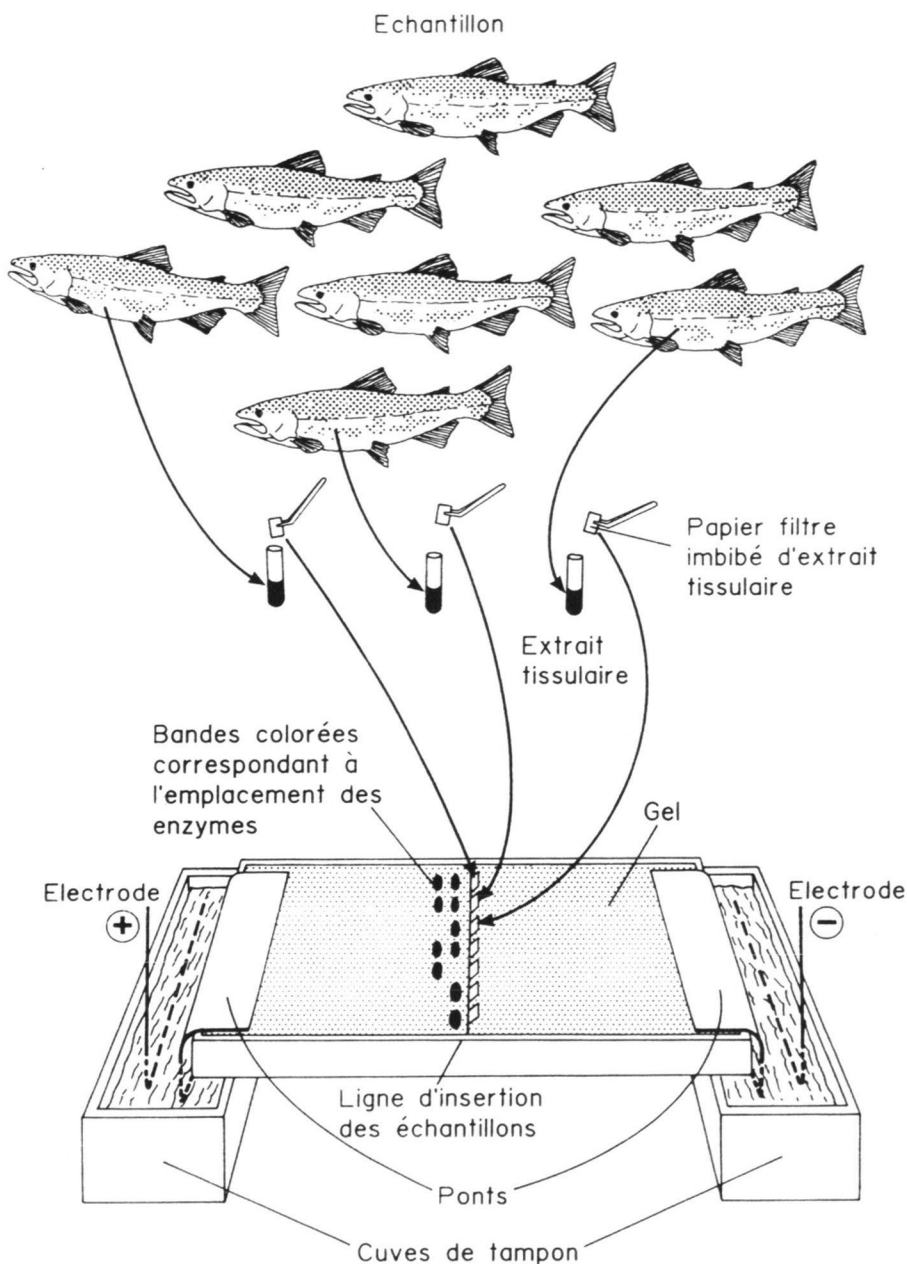
La génétique dispose d'un outil puissant pour mettre en évidence le polymorphisme moléculaire, l'électrophorèse. Il s'agit d'une technique permettant, entre autres choses, de séparer les molécules biologiques solubles. En effet, ces molécules acquièrent une charge électrique lorsqu'elles sont en solution et se déplacent donc en présence d'un champ électrique. Leur vitesse de déplacement dépend de leur charge et de leur taille. L'électrophorèse a été largement utilisée pour étudier la variabilité génétique des enzymes. Considérons un locus codant pour une enzyme donnée et présentant deux allèles A et B. Ces deux allèles peuvent très bien produire deux types de protéines, PA et PB, présentant des différences de taille ou de charge. Les deux formes seront donc séparées sur un support d'électrophorèse (notre laboratoire utilise le gel d'amidon). Après arrêt de l'électrophorèse, la présence des enzymes PA et PB peut être visualisée sur le gel à l'aide d'une réaction biochimique, utilisant celle que l'enzyme effectue dans la cellule; une tache ou bande apparaîtra sur le gel à l'endroit où se trouve chacune des formes. Si l'individu analysé a le génotype AA (respectivement BB) au locus codant pour cette enzyme, seule la forme PA (respectivement PB) apparaîtra; s'il est hétérozygote, les deux formes PA et PB seront révélées simultanément (fig. 5). La simple lecture des bandes obtenues permet de déterminer le génotype d'un individu au locus codant cette enzyme. Nous venons de décrire un cas très simple où un seul locus avec deux allèles code pour l'enzyme étudiée. Le nombre de locus et allèles impliqués peut être plus élevé. De plus, dans l'exemple précédent, nous avons supposé que seulement deux isoenzymes (enzymes ayant la même fonction), correspondant aux deux protéines PA et PB, existaient. L'enzyme est dite monomère. Cependant pour être fonctionnelles, certaines enzymes doivent être constituées de l'association de deux (enzymes dimères), trois (enzymes trimères) ou quatre (enzymes tétramères) chaînes polypeptidiques. Dans le cas d'une enzyme toujours contrôlée par un locus avec deux allèles, mais cette fois dimère, trois isoenzymes seront formées: PAPA, PAPB et PBPB; les deux homozygotes ne présenteront toujours qu'une seule bande, mais l'hétérozygote en donnera trois car il pourra produire les trois types d'isoenzymes. La figure 6 montre la complexité des électrophorégrammes de certains systèmes enzymatiques chez les salmonidés. Pour interpréter génétiquement les variations électrophorétiques d'une enzyme, il est nécessaire de faire des hypothèses sur la structure de cette enzyme; le nombre de locus et d'allèles codant pour celle-ci (cf. Pasteur *et al.* pour une présentation détaillée des aspects techniques de l'électrophorèse). La validité des hypothèses doit être vérifiée à l'aide de croisements de contrôle (fig. 7). Ces croisements permettent également de savoir si les variations alléliques observées aux différents locus sont liées (ce qui correspond généralement au fait que les locus sont situés sur un même chromosome); information nécessaire pour interpréter d'éventuels déséquilibres de liaison (cf annexe) mis en évidence dans les populations naturelles (fig. 7).

De façon générale, les systèmes enzymatiques restent codés par un faible nombre de locus et leurs variations électrophorétiques sont presque toujours interprétables en termes de fréquences alléliques. En outre, ces variations, contrairement à celles des caractères quantitatifs, sont peu perturbées par les variations de milieu; ce qui permet l'étude de populations provenant directement de rivières ou piscicultures différentes. La stratégie d'échantillonnage des populations dans le milieu naturel dépendra de la distribution géographique de l'espèce et, éventuellement, de certaines caractéristiques de sa biologie (par exemple, la présence de phénotypes différents sur un même territoire). La description de la diversité génétique sera d'autant plus précise que l'on examinera un nombre élevé d'individus et de locus.

#### b. Méthodes d'analyse des données électrophorétiques

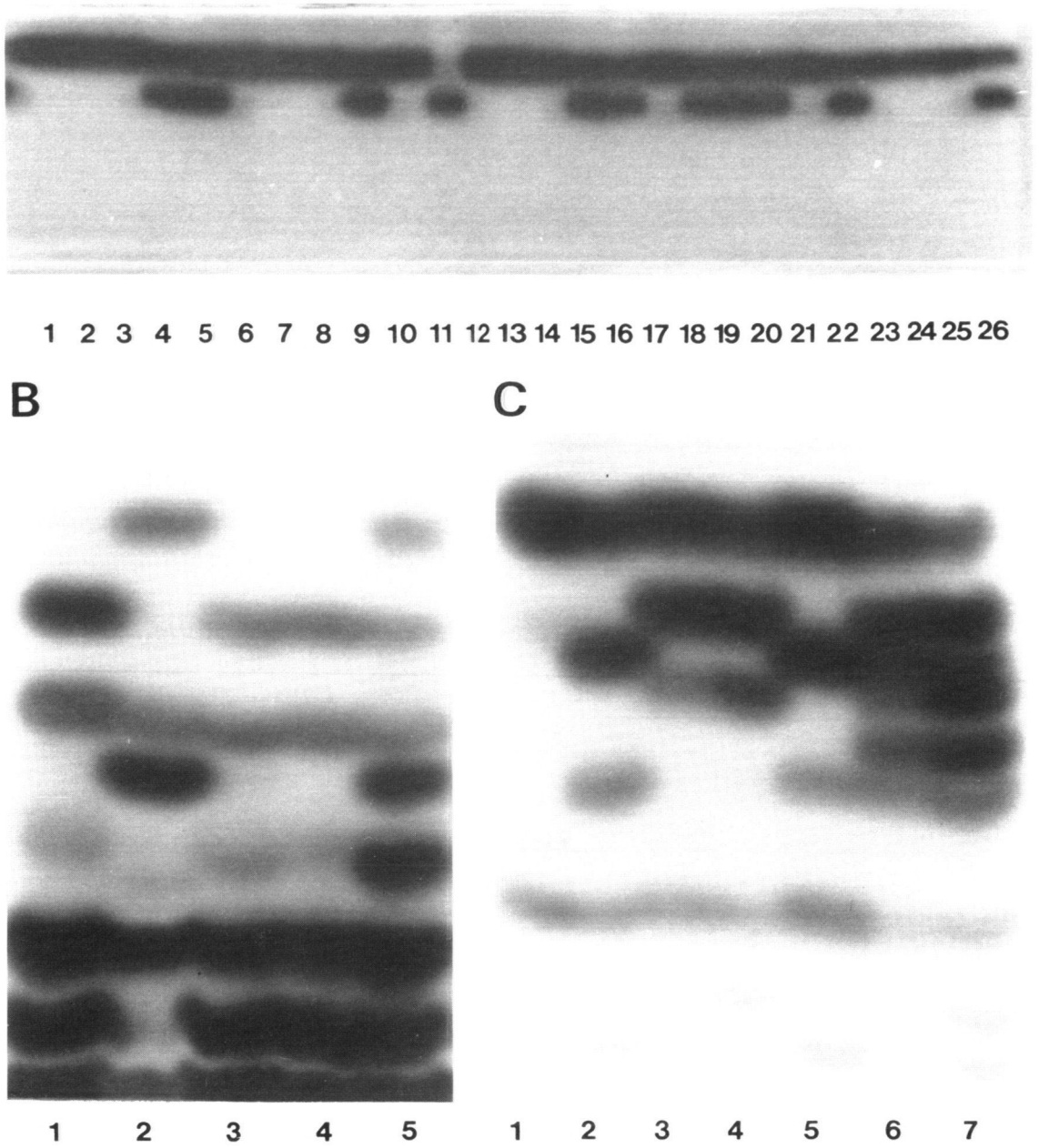
On dispose finalement pour chaque individu d'une "carte d'identité génétique" regroupant les génotypes observés sur l'ensemble des locus. Cette information doit être présentée de façon plus synthétique pour pouvoir être interprétée. Cela peut être fait en traitant directement le tableau des données brutes (individus x génotypes) par des méthodes d'analyses multifactorielles. Ces méthodes permettent de représenter l'information contenue dans le tableau initial sous forme de plans factoriels sur lesquels chaque individu est localisé en un point dépendant de sa constitution génétique. Plus deux individus se ressembleront, plus ils seront proches l'un de l'autre sur les plans factoriels et la dispersion des individus sur le graphe permet d'apprécier les variabilités intra et interpopulations.

Une autre stratégie consiste tout d'abord à calculer, pour chaque locus, les fréquences géniques dans chaque population, puis, à partir du tableau de fréquences obtenu, les variabilités intrapopulation (taux d'hétérozygotie) et interpopulations (distance génétique). Cette façon de procéder offre l'avantage d'être beaucoup moins lourde que la précédente, mais a l'inconvénient de ne pas utiliser toute l'information disponible. Il est, en effet, évident que les données initiales, c'est-à-dire les génotypes de chaque individu, ne peuvent pas être reconstituées. De plus, cette méthode impose de regrouper les individus, en général selon leur échantillon d'origine, avant le calcul des fréquences; ce qui n'est pas nécessairement justifié d'un point de vue génétique. Il est donc recommandé de s'assurer au préalable que les échantillons analysés ne sont pas issus de



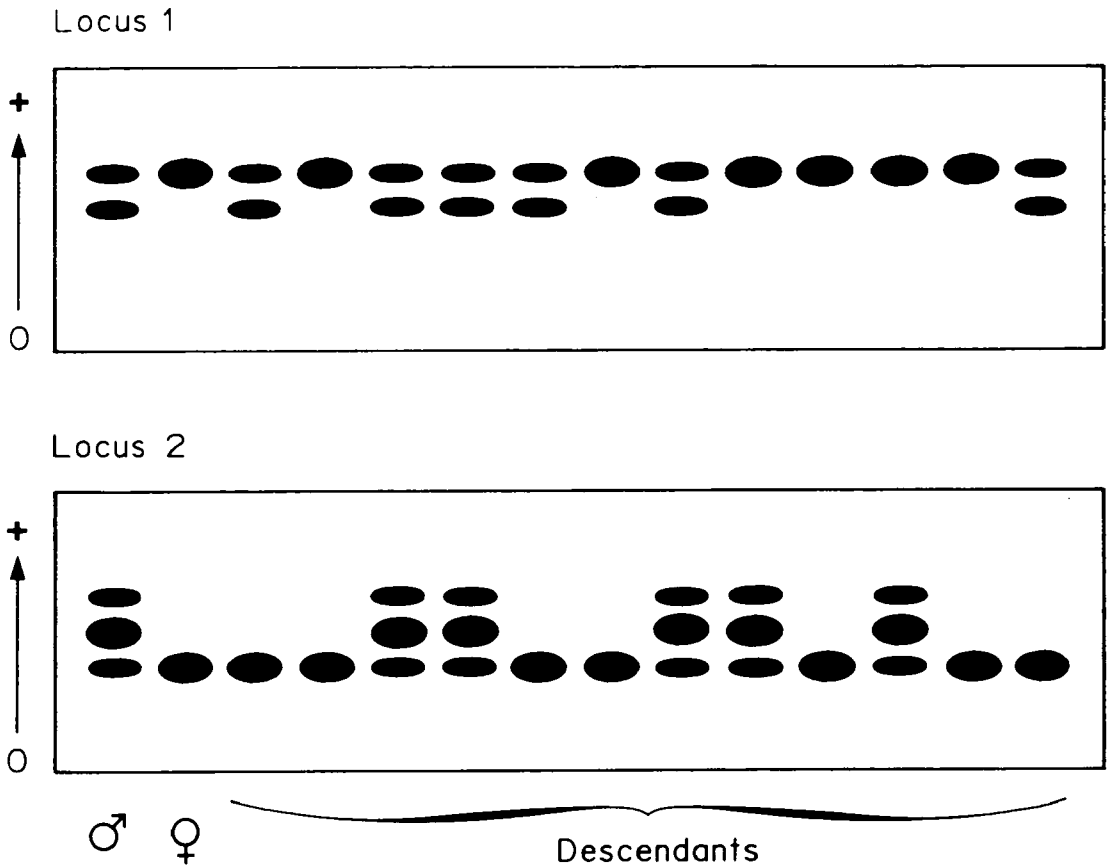
**figure 5** : L'électrophorèse de protéines comporte trois étapes : (1) la préparation d'extraits protéiques à partir de tissus de différents individus, (2) la migration et la séparation des protéines contenues dans ces extraits sous l'action d'un courant électrique continu, les différents échantillons préparés étant insérés côte à côte dans le gel, (3) à l'issue de la migration qui dure deux à trois heures, la révélation d'une enzyme qui est obtenue en recouvrant le gel d'une solution de révélation spécifique de cette enzyme. Sur le schéma, nous avons figuré une enzyme monomère codée par un locus auquel ségrègent deux allèles.

**figure 5** : Electrophoresis includes three steps : (1) preparation of crude protein extracts from several individuals, (2) migration and separation of the proteins of these extracts under an electric field, (3) after a migration of two to three hours, visualization of a particular enzyme with a specific staining solution poured on the gel. The figure shows a monomeric enzyme encoded by one locus possessing two alleles.



**figure 6** : Variations électrophorétiques observées chez la truite commune et la truite arc-en-ciel. A : PMI (phosphomannose isomérase), enzyme dimère codée par un locus dans le muscle de la truite commune ; les variations observées, ici chez 26 individus, sont dues à deux allèles. B et C : MDH (malate déshydrogénase), enzyme dimère codée par deux locus, Mdh-1 et 2, dans le foie des salmonidés (en B, variations observées dans le foie de cinq truites communes et dues à deux allèles ségrégeant au locus Mdh-2) et deux autres locus, Mdh-3 et 4, dans le muscle (en C, variations dues à quatre allèles ségrégeant aux locus Mdh-3 et 4 chez sept truites arc-en-ciel).

**figure 6** : Electrophoretic variations in brown trout and rainbow trout. A : phenotypes of a two-allele polymorphism for one locus encoding a monomeric enzyme, phosphomannose isomerase, from skeletal muscle of twenty-six brown trouts. B: phenotypes of a two-allele polymorphism for one of the two loci encoding a dimeric enzyme, malate dehydrogenase (MDH), from livers of five brown trouts. C : phenotypes of a four-allele polymorphism for two loci encoding MDH from skeletal muscle of seven rainbow trouts.



**figure 7** : Principe des analyses de ségrégations simples et conjointes permettant de vérifier la nature génétique des variations électrophorétiques observées et l'indépendance des locus étudiés. Une enzyme monomère et une enzyme dimère ont été examinées chez un mâle et une femelle ainsi que dans leur descendance. Faisons l'hypothèse que la variabilité observée pour la première enzyme est due à un polymorphisme génétique à un locus, que le mâle est hétérozygote et la femelle homozygote. Dans ce cas (et si aucune sélection ne se produit), les fréquences d'individus hétérozygotes et homozygotes ne doivent pas être significativement différentes chez les descendants. La même hypothèse peut être testée pour la deuxième enzyme. Si toutes ces hypothèses sont vérifiées, l'indépendance des deux locus polymorphes peut être testée. Pour cela il faut dénombrer les quatre différents types d'individus reconnus dans la descendance à partir des variations des deux enzymes : les deux phénotypes identiques à l'un ou l'autre des parents sont appelés types parentaux (par exemple, le deuxième descendant examiné est homozygote aux deux locus et identique à la mère) et les deux autres, types recombinés. Si les deux locus sont indépendants, les nombres de types parentaux et recombinés ne doivent être significativement différents.

**figure 7** : Single and joint segregation analysis at enzyme loci in full-sib families. Two enzymes, one monomeric and the other dimeric, have been examined in one male, one female and their progeny. We assume that (1) the polymorphism observed for the first enzyme results from genetic variation at one locus (2) the male is heterozygous and the female homozygous. Under this hypothesis and if no selection occurs, the number of heterozygous and homozygous offspring should not differ significantly. The variation of the second enzyme can be tested for the same hypothesis. Four different phenotypes can be recognized if we consider the two systems simultaneously. Two of them are identical to the parents (parental types) ; the two others are termed "recombinant types". If the two enzymes are encoded by two independent polymorphic loci, the number of parental and recombinant individuals should not significantly differ in the progeny.

mélanges de populations distinctes. On considère que cette condition est remplie si les échantillons respectent la loi de Hardy-Weinberg et ne présentent pas de déséquilibre de liaison (cf annexe).

Des études théoriques et expérimentales ont montré que la précision des estimations de distances génétiques (variabilité interpopulations) et des taux d'hétérozygotie (variabilité intrapopulation) ne dépend pratiquement plus que du nombre de locus analysés dès lors que le nombre d'individus étudiés par population est supérieur à 10 (Nei, 1978 ; Gorman et Renzi, 1979).

L'électrophorèse de protéine présente deux limites majeures. D'une part, elle ne met en évidence que les différences génétiques se traduisant par des différences électrophorétiques, c'est-à-dire environ un tiers de la variabilité génétique totale des gènes étudiés. D'autre part, une part importante du génome (les gènes de régulation et certains gènes de structure) est exclue de son champ d'analyse. Ces problèmes sont aujourd'hui surmontés grâce au développement de techniques d'analyse du polymorphisme de l'ADN permettant d'accéder à l'ensemble du génome.

Malgré certaines faiblesses, l'électrophorèse de protéine reste une méthode privilégiée pour obtenir rapidement des informations sur la phylogénie des espèces et leur diversité génétique. Dans le cas des espèces de poissons, elle a contribué à résoudre certains problèmes taxonomiques délicats, au niveau de l'espèce (cf. par exemple le cas des formes sympatriques naines et normales d'Omble chevalier ; Hindar *et al.*, 1986) ou du genre, et fournit la plupart des données relatives à la variabilité génétique intraspécifique. Nous avons illustré l'apport de l'électrophorèse de protéine à partir de l'exemple de la truite commune (cf. 3<sup>e</sup> chapitre).

#### Annexe

##### Loi de Hardy-Weinberg :

On admet que, lorsque cette loi est vérifiée, la population dont est issu l'échantillon examiné est panmictique. Considérons un locus auquel ségrègent deux allèles, A et a. A partir de fréquences génotypiques observées des génotypes AA, Aa et aa, on calcule les fréquences géniques de A et a, soient p et q respectivement. Si les géniteurs s'accouplent au hasard (et s'il n'y a ni sélection, ni mutation), les fréquences génotypiques observées de AA, Aa et aa ne doivent pas différer significativement des proportions de Hardy-Weinberg :  $p^2$ ,  $2pq$ ,  $q^2$  respectivement (cf. Génormont, 1979, pour plus de détails).

##### Déséquilibre de liaison :

Si, dans un échantillon, certains allèles d'un locus polymorphe sont plus fréquemment associés à certains allèles d'un autre locus polymorphe que ne le laissent prévoir les fréquences alléliques observées à ces locus dans l'échantillon, on dit que la population d'origine de celui-ci est en déséquilibre de liaison. Lorsqu'on mélange deux populations ayant des fréquences alléliques différentes à deux locus, un déséquilibre de liaison apparaît automatiquement à ces deux locus. Celui-ci persistera pendant plusieurs générations même si le mélange constitué par les deux populations se reproduit comme une seule unité panmictique (cf. Nei, 1975, pour plus de détails sur ce phénomène).

**REMERCIEMENTS :** Les figures 1, 2, 5 et 7 ont été réalisées par J. GALLÉ (INRA).

#### BIBLIOGRAPHIE

- CHOURROUT D., HAPPE A., 1986. Improved methods of direct chromosome preparation in rainbow trout, *Salmo gairdneri*. *Aquaculture*, 52, 255-261.
- GÉNORMONT J., 1979. Les mécanismes de l'évolution, Dunod, Paris, 235 p.
- GORMAN G.C., RENZI J., 1979. Genetic distance and heterozygosity estimates in electrophoretic studies : effects of sample size. *Copeia*, 242-249.
- HINDAR K., RYMAN N., STAHL G., 1986. Genetic differentiation among local populations and morphotypes of arctic char, *Salvelinus alpinus*. *Biol. J. Linn. Soc.*, 27, 269-285.
- LI W.H., 1976. Effect of migration on genetic distance. *Amer. Natur.*, 110, 841-847.
- MAYR E., 1974. Populations, espèces et évolution, Hermann, Paris, 496 p.
- NEI M., 1975. Population genetics and molecular evolution, North holland, Amsterdam and New York, 228 p.
- NEI M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583-590.
- NEI M., 1987. Genetic distance and molecular phylogeny. In N. RYMAN and F. UTTER eds., Population genetics and fishery management, 193-223, University of Washington, Seattle and London.
- PASTEUR N., PASTEUR G., BONHOMME F., CATALAN J., BRITTON-DAVIDIAN J., 1987. Manuel technique de génétique par électrophorèse des protéines. Technique et documentation, Lavoisier, Paris, 217 p.
- PRÉVOST G., 1976. Génétique, Hermann, Paris, 299 p.