

qu'est-ce qu'un gène ?

Une petite histoire du concept

par Hervé Le Guyader

Université Pierre-et-Marie-Curie (Paris 6), UMR 7622, Biologie du développement, équipe Développement et Évolution
Bâtiment B, 7^e étage, 9, quai Saint-Bernard, case 241, 75252 Paris cedex 05 France
Herve.Le-Guyader@snv.jussieu.fr

Pour tout un chacun, l'avènement de la biologie moléculaire a permis l'élaboration d'une définition claire du gène ; cela va du simple « fragment d'ADN » à une description plus fine, comme « une séquence transcrite, flanquée de séquence(s) régulatrice(s) ». Qu'il y ait une relation évidente entre gène et ADN, nul n'ira le discuter ; mais cette relation est maintenant si complexe que, à l'inverse des idées communément admises, le concept de gène semble avoir perdu de sa limpidité. En fait, il a tellement évolué que, pour comprendre les difficultés actuelles d'appréciation et leurs conséquences, il est indispensable de suivre le double cheminement théorique et expérimental qui a rythmé l'évolution de la conception du gène depuis la création de cette discipline biologique majeure qu'est la génétique.

Le facteur mendélien

À tout seigneur, tout honneur ; c'est bien Gregor Mendel (1822-1884) qui est le premier à publier un article de génétique, son fameux mémoire de 1865, *Recherches sur des hybrides végétaux*, même si le terme de *gène* a été forgé presque 20 ans après sa mort. Sans entrer dans une exégèse du texte¹, quelques remarques doivent être faites. Celui-ci est construit en deux parties : la description des croisements des variétés de pois tout d'abord, la conceptualisation ensuite. Pour comprendre le résultat des croisements, Mendel réfléchit sur le rôle des cellules sexuelles et postule que ces dernières transmettent des « facteurs » qui sont responsables des caractères observables chez les organismes issus de la fécondation, comme la couleur (jaune ou vert) ou l'aspect (lisse ou ridé) de la graine. Ces facteurs sont amenés pour parts égales par le gamète mâle et le gamète femelle, font partie intégrante de la cellule et sont distincts des caractères qu'ils gouvernent. Bien évidemment, Mendel n'a jamais vu ces « facteurs » ; en revanche, il a observé les caractères. C'est un raisonnement conceptuel qui l'a amené à postuler l'existence de tels facteurs dans la structure intime de la cellule. Le facteur mendélien, objet conceptuel, est tout d'abord virtuel.

On peut se demander pourquoi la communauté scientifique n'a pu alors recevoir ce travail novateur. Pourtant, Mendel avait expédié son mémoire aux grands noms de la biologie européenne. Des explications ont été proposées, de différents types suivant les spécialités des biologistes consultés. Par

¹ Voir le chapitre 1, « *Un peu plus d'un siècle de génétique...* », rédigé par J. Deutsch dans : Rossignol J.L., 2000. *Génétique, gènes et génomes : cours et questions de révision*. Dunod, Paris, 231 p.

exemple, il est vraisemblable que nombreux sont ceux qui ont cru que Mendel avait décrit une originalité du pois, sans en voir toute l'extraordinaire portée. Mais il y a le mystère Charles Darwin (1809-1882). En effet, *L'origine des espèces*, publié peu de temps auparavant (en 1859), aurait été un ouvrage encore plus puissant s'il s'était appuyé sur des fondements corrects de l'hérédité. On sait que Darwin a reçu le mémoire de Mendel, mais qu'il ne l'a pas lu² ; mais d'autres biologistes de son entourage l'ont fait, et auraient pu attirer son attention³. Visiblement, il n'y a pas eu rencontre entre les deux hommes. Pourquoi ? Très vraisemblablement, l'origine de la non-rencontre est à chercher du côté de la théorie : le concept-clé de Darwin est celui de *descendance avec modification*, qui postule que les caractères héréditaires sont variables. Or, par le brassage de facteurs *a priori* immuables, Mendel montrait comment, à travers le désordre apparent des caractères, il y avait un ordre sous-jacent, gouverné par le ballet des « facteurs » au moment de la formation des gamètes puis de la fécondation. En d'autres termes, si Darwin était transformiste, Mendel était fondamentalement fixiste, étant donné que, privé du concept de mutation, il ne pouvait pas imaginer, à son niveau d'expérimentation, une variabilité génétique.

La génétique et le darwinisme se rencontrent enfin

Au tout début du XX^e siècle, en 1900, le hollandais Hugo De Vries (1848-1935), l'autrichien Erich von Tschermak (1871-1962) et l'allemand Cari Correns (1864-1933) refont indépendamment des expériences semblables à celles de Mendel, en faisant d'ailleurs référence au mémoire de 1865⁴. Mais aux « lois de Mendel », ils (particulièrement De Vries) ajoutent la mutation.

C'est le réel début de la génétique ; le danois Wilhelm Johannsen (1857-1927) apparaît l'homme-clé, en créant le terme de « gène », qui remplacera celui de « facteur », et en proposant la distinction fondamentale entre le génotype (l'ensemble des gènes) et le phénotype (l'ensemble des caractères). L'anglais William Bateson (1861-1926), auteur du mot « génétique » (*genetics*), et le français Lucien Cuénot (1866-1951) étendent les principes mendéliens aux espèces animales. Le gène est toujours une unité abstraite et le problème de son support biologique va encore être laissé de côté, pour bien peu de temps. Mais, si le gène est une unité de transmission, c'est maintenant une unité de mutation.

Il faudra quelques années pour fonder la génétique des populations. En 1908, la loi dite de Hardy et Weinberg, permettant de calculer les fréquences génotypiques dans une population à l'équilibre, est proposée indépendamment par le mathématicien anglais Godfrey H. Hardy (1877-1947) et le médecin allemand Wilhelm Weinberg (1862-1937). Étant donné qu'un gène peut présenter différentes formes (ou allèles), on peut imaginer que certaines d'entre elles favorisent l'organisme qui les porte. La sélection intervient donc et c'est John B. S. Haldane (1892-1964) qui, en 1924, a su poser un modèle mathématique simple de sélection. C'est de cette manière que mendélisme et darwinisme se sont retrouvés. En effet, les gènes sont variables par le jeu de la mutation et les phénotypes associés subissent la sélection. En complément, les travaux de Ronald A. Fisher (1890-1962) et surtout de Sewall Wright (1889-1988) vont donner corps à une génétique des populations qui sera l'un des piliers de la théorie synthétique de l'évolution des années 1940 .

² P. Corsi, communication personnelle.

³ Certains pensent que le botaniste allemand Heinrich Hoffmann aurait mis Darwin au courant. Voir Blanc M., 1984. Gregor Mendel : la légende du génie méconnu. *La Recherche*, 15, 51, p. 48.

⁴ Voir le numéro spécial « 1900 Redécouverte des lois de Mendel » des *Comptes Rendus de l'Académie des Sciences*, Série III, tome 323, n°12, 1033-1196, décembre 2000.

⁵ Voir Montalenti G., 1996. Génétique. In P. Tort : *Dictionnaire du darwinisme et de l'évolution*. PUF, Paris, pp. 1825-1835.



... DIEU MERCI!
DE TELLES
INFIRMITÉS SONT
DEVENUES
RARISSIMES!

Rouso

La théorie chromosomique de l'hérédité

La question-clé qui a été posée immédiatement a trait au support matériel du gène : quelle est sa localisation dans la cellule ? Est-il lié à une structure particulière ? Dès 1902, W. Sutton remarque que les chromosomes à la méiose présentent un comportement identique à celui que l'on attend de la part des gènes. Il propose la théorie chromosomique de l'hérédité, à savoir que les chromosomes seraient le support des gènes. Cette théorie sera brillamment démontrée par les travaux sur la drosophile de l'équipe de Thomas H. Morgan (1866-1945), puis acceptée par une importante fraction de la communauté biologique en 1914. Les facteurs mendéliens acquièrent non seulement une réalité physique mais aussi une localisation précise, suivant les cartes chromosomiques réalisées pour la première fois par Alfred H. Sturtevant (1891-1970), élève de Morgan. L'ensemble de ces études permet d'expliquer le comportement de gènes indépendants (gènes portés par des chromosomes différents) et de gènes liés (gènes portés par un même chromosome). De plus, la recombinaison génétique intrachromosomique est mise en évidence : le crossing-over, dont la réalité est étayée au niveau cytologique, autorise un échange entre chromosomes homologues à la méiose. Le gène apparaît donc comme une nouvelle unité, celle de recombinaison.

On peut remarquer qu'avec la théorie chromosomique de l'hérédité, le concept de gène présente une grande pureté conceptuelle : support d'un caractère héréditaire, il est à la fois unité de transmission, unité de mutation et unité de recombinaison.

Le gène : quelle fonction ?

On arrive maintenant à la troisième question-clé : quelle est la fonction du gène ? C'est l'américain Georges W. Beadle (1903-1989) qui sera le moteur de cette recherche. Tout d'abord, dans les années 1930, avec Boris Ephrussi (1901-1979), il réalise une série d'expériences sur la synthèse des pigments des yeux de la drosophile. Puis il change de matériel et utilise un champignon ascomycète, *Neurospora crassa*, simple à cultiver en laboratoire. Avec Edward L. Tatum (1909-1975), il isole des mutants auxotrophes, c'est-à-dire incapables de synthétiser certaines molécules indispensables, comme des vitamines ou des acides aminés. Déterminant alors que les gènes ont une fonction-clé dans le métabolisme de ces molécules essentielles, Beadle et Tatum proposent en 1941 l'adage « un gène - un enzyme ». Or cet enzyme, c'est une protéine. Cette étude importante est un début d'explication du passage du génotype au phénotype, par l'intermédiaire de la chimie cellulaire. La biochimie, qui s'est développée de manière indépendante, commence à rejoindre la génétique par l'intermédiaire de la fonction du gène.

L'ADN support de l'information génétique

En 1944, à la fondation Rockefeller de New-York, Oswald Th. Avery (1877-1955), avec l'aide de Me Leod et Me Carthy, montre que l'ADN est le principe actif de la transformation du pneumocoque, phénomène découvert par Griffith en 1928. Il constate que, chez cette bactérie, le changement d'un caractère héréditaire - présence de paroi - ne peut être réalisé que par transfert de cette molécule. De plus, en 1952, Alfred D. Hershey (1908-1997) et Martha Chase prouvent que le seul matériel infectant d'un bactériophage est la molécule d'ADN. C'est donc elle qui porte l'« information » génétique. Enfin, une année plus tard, James D. Watson et Francis Crick proposent le célèbre modèle en double hélice qui explique comment, par le jeu de la succession des bases azotées, l'information peut être portée sur cette molécule et transmise inchangée de génération en génération.

C'est seulement maintenant que le gène semble changer de statut : d'une entité conceptuelle, il devient une entité matérielle. Mais, si la conception « gène = segment d'ADN » est alors largement affichée, les biologistes de l'époque étaient nourris de génétique morganienne, et savaient tout de l'histoire du concept de gène. Dorénavant, la question essentielle est : quelle(s) est (sont) la (les) caractéristique(s) de la molécule d'ADN qui rende(nt) compte des propriétés du gène ? En d'autres termes, il convient de savoir si le gène *est* une molécule d'ADN, ou bien si l'ADN est *le support matériel* du gène.

La régulation génique

Paradoxalement, c'est par le biais de la compréhension de la régulation de l'expression du gène que l'on va pouvoir approcher la manière dont l'« information » portée par l'ADN passe au niveau des protéines, et donc à l'activité enzymatique. C'est le beau travail des années 1960 de François Jacob et Jacques Monod sur l'opéron qui conduira au décodage de la transcription (synthèse de l'ARN messenger) et de la traduction (synthèse des protéines). Le principe du « code génétique » est admis, puis dévoilé. Enfin, ces découvertes permettent de redéfinir de manière fine l'unité de mutation et l'unité de recombinaison, car, maintenant, on est passé à l'exploration intragénique. L'atome de la génétique n'est plus le gène, mais une partie du gène : l'unité de mutation et l'unité de recombinaison sont une seule et même chose, le nucléotide.

De plus, avec l'opéron lactose, Jacob et Monod montrent que l'expression de l'information génétique est finement régulée et répond aux besoins immédiats de la bactérie dans son environnement. On démontrera par la suite que le schéma de cette régulation peut, dans une certaine mesure, être généralisé à l'ensemble des organismes, en particulier les eucaryotes. La régulation de la transcription (synthèse de l'ARN messenger) fait appel à des protéines, les « régulateurs de transcription », qui peuvent jouer le rôle d'activateur ou de répresseur. Les boucles de régulation protéine/ADN/protéine commencent à être détectées et on comprend très vite que de telles régulations permettent de contrôler l'activité séquentielle des gènes impliqués lors d'un processus cellulaire (différenciation) ou embryologique (organogenèse). À cette époque, le travail a d'abord été appliqué à la construction de bactériophages. Du point de vue du développement, la notion de « programme génétique » dépendant de l'ADN prend corps.

Le concept d'information arrive ainsi en génétique par deux canaux différents et de manière presque concomitante. D'une part, la découverte du code génétique permet de postuler ce que l'on a tout de suite appelé « le dogme de la biologie moléculaire », à savoir que l'information coule de l'ADN vers les protéines, via l'ARN messenger. D'autre part, la régulation génique permet de mettre en évidence l'existence de séquences d'ADN non traduites, mais qui jouent un rôle-clé dans l'expression des gènes. Ces séquences régulatrices sont des éléments importants du programme, et qui dit programme dit information. Ce point explique l'émergence de l'ambiguïté du concept de gène, car une séquence codante n'a de rôle biologique que par le jeu des séquences régulatrices qui reçoivent des signaux de l'environnement immédiat. Le gène ne peut plus être une séquence codante seule, mais une séquence codante munie de ses éléments de régulation.

Trois découvertes fondamentales vont permettre un accès aisé au génome. Ce sont : 1°) les enzymes de restriction qui, coupant l'ADN en des sites précis, autorisent le génie génétique et le clonage de gènes ; 2°) les méthodes de séquençage de l'ADN, facilitant l'accès à la structure intime des gènes ; 3°) la réaction de polymérisation en chaîne (PCR) permettant d'amplifier une séquence d'ADN *in vitro*. Ces découvertes ouvriront l'ère de la génomique et de la transgénèse.

Un premier retour sur le concept de gène

Avant 1950, le gène était principalement une unité de « calcul », permettant de réaliser des prédictions sur la descendance de croisements. Après 1950, le gène paraît devenir une molécule- ou une partie de molécule - accessible à l'expérimentation. Le gène aurait-il donc cessé d'être une entité théorique pour seulement devenir un élément physique accessible à l'observation ?

La définition classique du gène des années 1960 peut être, par exemple, celle que Jacob donne dans *La logique du vivant*⁶ : « Quel que soit le type d'analyse employé, le terme ultime est le même : c'est le gène qui représente à la fois l'unité de fonction, de mutation et de recombinaison. Ainsi le matériel de l'hérédité se résout en unités élémentaires qu'on ne peut fractionner. Les gènes deviennent comme les atomes de l'hérédité. »

On en était alors à une cohérence entre les données conceptuelles et les résultats de laboratoire, cohérence telle que tout un chacun garde en mémoire cette définition comme un paradis perdu, en oubliant les difficultés qui, à cette date, étaient à venir. Or cette vision simple et cohérente des choses s'estompée devant les progrès récents des biologies moléculaire, cellulaire et du développement. Ainsi, l'idée qu'un segment particulier d'ADN ne code qu'une seule protéine ne tient plus. Par exemple, une même séquence d'ADN peut être traduite par deux phases différentes de lecture, donnant naissance à deux protéines différentes⁷ ; plus fondamentalement, la structure en mosaïque des gènes (introns codants, exons non codants) remet en cause le concept de gène, étant donné que, par épissage différentiel (c'est-à-dire choix de certains introns parmi la panoplie présente), un même ARN messager peut donner naissance à une famille de protéines⁸. Il n'y a donc pas de relation biunivoque entre protéine et séquence codante, c'est pourquoi un nouvel acronyme a été forgé : l'ORF (*Open Reading Frame* : cadre ouvert de lecture) désigne une séquence codante, quelles que soient les possibilités de choix des introns⁹.

Il devient donc plus difficile de dire quand et où le gène prend réalité ; dans le cas d'un épissage différentiel, le gène est-il la séquence d'ADN (c'est-à-dire l'ORF) ou la séquence de TARN messager mature ? Dans le premier cas, le gène doit être considéré comme une structure transitoire ; dans le deuxième cas, on doit admettre qu'il y a plusieurs gènes pour une séquence codante d'ADN¹⁰.

On arrive à un paradoxe : alors que, quittant le domaine de l'abstrait, sa description moléculaire devient d'une grande précision, le concept de gène ne paraît pas se simplifier. Comme test, à cet égard, la double définition que l'on trouve dans un rapport de l'Académie des Sciences est exemplaire¹¹ : « Gène : en génétique classique : toute région du génome qui confère un caractère phénotypique à un organisme. En génétique moléculaire : segment d'ADN comprenant les régions transcrites en ARN (messager, de transfert, ribosomique...) et les régions régulatrices adjacentes. »

Et l'évolution ?

Toute une série de travaux théoriques et expérimentaux en biologie évolutive porte sur la sélection. Suivant une logique darwinienne stricte, seul le phénotype est la cible de la sélection ; on raisonne

⁶ Jacob F., 1970. *La logique du vivant*. Gallimard, Paris, 357 p.

⁷ Comme, par exemple, chez le phage ϕ 80 ; voir le chapitre 1, « *Un peu plus d'un siècle de génétique...* », rédigé par J. Deutsch dans : Rossignol J.L., 2000. *Génétique, gènes et génomes*. Dunod, Paris, p. 19.

⁸ Sur l'épissage, voir par exemple Rossignol J.L., 1996. *Génétique*. Masson, Paris, p. 197.

⁹ Par exemple, un même ORF permet la synthèse de la maturase et du cytochrome b ; voir le chapitre 1, « *Un peu plus d'un siècle de génétique...* », rédigé par J. Deutsch dans : Rossignol J.L., 2000. *Génétique, gènes et génomes*. Dunod, Paris, p. 20.

¹⁰ Gayon J., 1999. Génétique. In D. Lecourt : *Dictionnaire d'Histoire et Philosophie des Sciences*. PUF, pp. 453-454.

¹¹ Académie des Sciences - CADAS. *Les techniques de transgénèse en agriculture*, rapport commun n° 2, octobre 1993. Voir le glossaire.

alors au niveau de l'organisme et, dans le concept de *struggle for life*, c'est bien ce dernier qui paraît être sélectionné. Dans le cadre de la sélection artificielle, ce sont bien des chiens, des moutons, des chevaux, des rosiers ou des plants de blé... qui sont sélectionnés par l'éleveur. Dans l'étude paradigmatique de H. B. D. Kettlewell sur le mélanisme industriel de la Phalène du bouleau (*Biston betularia*, Lépidoptère Géométridé), ce sont les papillons qui sont mangés par les oiseaux, donc cibles de la sélection¹².

Pourtant, tout a changé quand on s'est rendu compte qu'il n'y avait peut être pas identité entre l'entité qui est soumise à la sélection et celle qui est réellement sélectionnée. Le suédois G. Ostergren a été pionnier en ce sens, en postulant dès 1945 que « les unités de sélection ne sont en aucun cas les individus biologiques, mais leurs gènes et leurs chromosomes. » De nombreuses études récentes, pour beaucoup seulement possibles grâce aux techniques de la biologie moléculaire, plaident en faveur de cette hypothèse. Deux types de découvertes paraissent alors capitales. Tout d'abord, on découvre l'existence des transposons, courtes séquences d'ADN de quelques centaines à quelques milliers de bases, codant une ou plusieurs protéines lui permettant de se multiplier et, par là-même, d'envahir le génome où il est intégré. L'étonnement des généticiens a été grand quand ils ont compris que ces transposons, comme l'élément P chez la drosophile, ne paraissent pas d'un intérêt immédiat pour son hôte¹³. Ce sont des éléments dits « égoïstes », qui jouent pour eux-mêmes et non « pour le bien » du génome auquel ils participent. Concomitamment, on mettait en évidence des conflits au niveau du (des) génome(s)¹⁴ : entre génome nucléaire et génome(s) cytoplasmique(s) (mitochondrial et/ou chloroplastique), entre chromosomes sexuels, entre chromosomes non sexuels, entre gènes (suivant leur origine, paternelle ou maternelle)...

De telles études amènent à revoir complètement les impressions premières ; disparaît alors la belle harmonie intuitivement postulée, d'une part, à l'intérieur d'un même génome et, d'autre part, dans la relation génotype-phénotype : 1°) s'il y a conflit, il y a sélection ; c'est donc bien un gène, ou un ensemble de gènes, qui est sélectionné ; 2°) s'il y a conflits à l'intérieur d'un même génome ou entre génomes d'une même cellule, la sélection ne s'applique pas seulement sur l'organisme mais sur toute entité plus ou moins autonome comportant un ou plusieurs gènes. Il est toujours difficile pour certains d'admettre que des éléments qui participent à un même programme puissent être en conflit...

On est donc obligé de distinguer différents niveaux de sélection, suivant le niveau d'organisation de l'entité impliquée¹⁵ : niveau moléculaire (comme celui des transposons) ; niveau cellulaire (une cellule cancéreuse « gagne » momentanément) ; niveau de l'organisme (le niveau classique des sélectionneurs) ; peut-être un niveau de groupe. Le point-clé concerne la règle simple qui gouverne les sélections participant à ces différents niveaux : toute forme de sélection à un niveau donné prend le pas sur une forme de sélection agissant à un niveau supérieur. Ceci veut dire que si un gène est sélectionné au niveau de l'organisme, c'est qu'il l'a été avec succès aux niveaux moléculaire et cellulaire.

Un nouveau retour sur le concept de gène

Nous sommes maintenant en mesure de répondre à notre question : quelle(s) est (sont) la (les) caractéristique(s) de la molécule d'ADN qui rende(nt) compte des propriétés du gène ? Il faut tout d'abord se rappeler que l'entité qui *est sélectionnée* est celle qui est transmise de génération en

¹² Voir Ridley M., 1997. *Évolution biologique*. De Boeck Université ; sur *Biston betularia* : pp. 103-109.

¹³ Voir Bazin C. *et al.*, 1998. Séquences d'ADN mobiles et évolution du génome. In Le Guyader H. : *L'évolution*. Bibliothèque Pour la Science, Paris, pp. 138-143.

¹⁴ Voir Hurst L. D., Atlan A., Bengtsson B. O., 1996. Genetic conflicts. *The Quarterly Review of Biology*, 71, 3, 317-364, et Atlan A., 1998. La guerre froide des chromosomes sexuels. *La Recherche*, 306, 42-44.

¹⁵ *Ni Dieu ni gène*, livre récent de J.J. Kupiec et P. Sonigo (2000, au Seuil) généralise cette idée aux niveaux moléculaire et cellulaire.

génération et que l'entité qui *subit la sélection* est celle qui transmet l'entité sélectionnée. Ainsi le gène ne doit pas être appréhendé sous son aspect matériel mais plutôt sous celui de l'information génétique qui détermine le phénotype particulier corrélé. La molécule d'ADN n'est en fait que le support, le véhicule, de cette information. Si, au début des années 1970, on pensait qu'il existait une relation biunivoque entre le comportement de l'information génétique et celui de la molécule support, ce qui donnait une représentation cohérente, on sait maintenant que les choses sont plus compliquées, mais également prodigieusement intéressantes d'un point de vue biologique. En effet, on découvre une subtilité d'utilisation du codage et de la combinatoire.

Quel que soit le nom donné à ce qui est transmis de génération en génération, et par là-même tout au long du processus évolutif, ce n'est pas d'une entité matérielle *sensu stricto*, mais plutôt d'une information, au sens le plus large du terme, qu'il s'agit. Un évolutionniste¹⁶ donnera comme définition du gène : « information portée par une séquence définie de nucléotides sur l'ADN, accompagnée des signaux nécessaires à son expression, et qui permet la réalisation d'une fonction élémentaire (production d'une chaîne polypeptidique, d'un ARNt ou d'un ARNr). » Les points importants de cette définition concernent *l'information, l'expression et la fonction élémentaire*. Mais elle ne tient pas compte des ambiguïtés citées plus haut, relatives à l'absence de liaison simple entre « une séquence de nucléotides » et « une chaîne polypeptidique ». Entre les deux, il y a une série de boucles de régulation qui s'appliquent non pas seulement à la transcription des ARN, mais à l'ensemble des niveaux d'expression : maturation de l'ARN messenger, traduction, modifications post-traductionnelles des protéines, localisation cellulaire des ARN ou des protéines ; tout cela participe du programme génétique.

Donc un gène n'a de réalité que vu comme partie prenante d'un programme, programme qui s'est construit pas à pas au cours de l'évolution, suivant le couperet implacable de la sélection appliquée à tous les niveaux, suivant la règle énoncée ci-dessus. C'est en fait une généralisation complète du concept darwinien¹⁷. On peut d'ailleurs rattacher à cette idée le concept de terrain (ou fond) génétique, c'est-à-dire le fait que l'expression d'un gène est souvent dépendante de sa localisation dans le génome et est, de manière générale, sous l'influence de l'ensemble du génome¹⁸. Le gène ne peut donc être considéré comme un élément isolé de son contexte.

L'ambiguïté du concept de programme

La métaphore de l'ADN-programme a une double provenance : 1°) la découverte du code génétique amène à dire que le programme nécessaire à la construction d'une protéine est porté par la séquence d'ADN ; 2°) comme les molécules-clés de la matière vivante sont les protéines, l'ensemble des programmes codant ces protéines constituerait le programme complet de l'organisme. Selon Mayr¹⁹, l'ADN contient un programme, une détermination, « une information codée ou pré-ordonnée qui contrôle un processus ou un comportement en le menant vers une fin donnée. » C'est cette manière de voir qui fait dire à Dawkins²⁰, à propos des chatons des saules : ils contiennent une portion d'ADN « dont les caractères codés énoncent des instructions spécifiques pour l'élaboration de saules pleureurs qui répandront une nouvelle génération de graines duveteuses. [...] Dehors, il pleut des instructions ; il pleut des programmes ; il pleut des algorithmes créateurs d'arbres et semeurs de duvet. Ce n'est pas

¹⁶ Gouyon P.H. *et al.* 1997. *Les avatars du gène*. Belin, Paris, p. 197.

¹⁷ Voir Kupiec J.J., Sonigo P., 2000. *M Dieu ni gène*. Le Seuil, Paris, 230 p.

¹⁸ Voir le chapitre rédigé par C. Lamour-Isnard (chap. 3, « Classification fonctionnelle des mutations ») sur le fond génétique, et celui rédigé par M. Fellous (chap. 4 : « Les interactions génétiques ») sur l'épistasie, dans : Rossignol J.L., 2000. *Génétique, gènes et génomes*. Dunod, Paris.

¹⁹ Cité dans Atlan H., 1999. *La fin du « tout génétique » ?* INRA Éditions, Paris, 96 p.

²⁰ Dawkins R., 1989. *L'horloger aveugle*. Robert Laffont, Paris, p. 136.

une métaphore ; c'est la simple vérité. Aussi simple que s'il pleuvait des disquettes. » Simple, peut-être, simpliste, assurément.

Dès les années 1970, des théoriciens se sont très vite rendu compte des difficultés liées à une telle simplification. Ils ont essayé de les résoudre par l'exploitation, par exemple, du concept d'auto-organisation ou de l'analogie avec les automates autoreproducteurs de von Neumann ou avec la machine de Turing universelle. Pourtant, dès la compréhension du fonctionnement de l'opéron lactose, il était clair que les informations ne sont pas codées dans leur ensemble par le génome, étant donné que, par l'intermédiaire du lactose, l'environnement est déterminant.

La biologie moléculaire du développement a, au cours de ces 20 dernières années, éclairé de manière précise la notion de programme de développement. Elle a mis en particulier l'accent sur le fait que pour qu'il y ait programme, il est indispensable que l'expression des gènes se fasse dans un environnement physico-chimique d'une grande précision. Ainsi, au tout début de l'embryogenèse de l'œuf de drosophile, les gènes-clés qui, comme *bicoid*, déterminent les polarités de l'embryon, ne peuvent jouer leur rôle, dans le programme en cours, que dans un environnement extrêmement contrôlé (situation particulière des ARNm, structuration particulière du cytosquelette microtubulaire, taille précise de l'ovocyte...). La dissection moléculaire des promoteurs de certains gènes de segmentation (*hairy*; *fushi tarazu*) montre que leur expression ne se fait que dans un environnement protéique déterminé directement par les étapes antérieures du développement. Si l'on ne s'intéresse qu'à la succession causale d'expressions des gènes, on peut être tenté de dire que le développement de l'embryon est complètement déterminé par l'ensemble des gènes, qui constituent le programme génétique. Mais c'est oublier que, dans cette succession, les gènes qui s'expriment à un instant t réalisent, par leurs produits, l'environnement des gènes qui s'expriment à $(t + 1)$. Il y a apparence de programme parce que les environnements ainsi créés sont stables et répétables ; mais une altération fortuite de l'environnement peut amener un développement tératologique, ce qui est à la base du concept d'épigénèse. Il convient donc d'admettre que le gène n'a pas réellement l'autonomie et l'existence physique que certains lui prêtent, mais que ce qui est biologiquement actif, c'est une dynamique du génome en interaction avec son environnement cellulaire. C'est pourquoi séquencer un génome ne donne certainement pas tout le programme... Lors du clonage par transfert de noyau, la reprogrammation consiste en réalité en un « mouvement important des protéines du cytoplasme receveur vers le noyau [...]. Plus de 75% des protéines préexistantes dans le noyau transféré sont alors éliminées »²¹. C'est donc l'environnement protéique du génome qui est crucial pour le développement harmonieux du zygote.

En fait, le gène est un concept, une idéalisation dont, à la limite, on pourrait actuellement se passer dans une description moléculaire fine des phénomènes cellulaires. On peut mesurer ceci à l'aune des sigles créés par les molécularistes pour désigner des séquences nucléiques présentant des caractéristiques précises (ORF, EST...). La publication récente du génome humain a donné lieu à un battage médiatique sur le faible nombre de gènes de l'homme (30 000, au lieu des 100 000 attendus)²². En réalité, le décompte porte à la fois sur des ADN et sur des ORF... Subtilité ? Non, une séquence codante et un gène, ce n'est pas la même chose. En fait, le relativement faible nombre obtenu prouve que les interactions géniques et les régulations de l'expression sont encore plus complexes qu'on ne l'imaginait.

Malheureusement, il faut se rendre compte que le terme de gène a acquis son succès médiatique et populaire au moment où, au laboratoire, il a perdu de sa clarté théorique.

²¹ Kikyo N., Wolffe A.P., 2000. La clef du clonage : le noyau reprogrammé. *Biofutur*, 203, 36-39.

²² Weissenbach J., Esnault Y., 2000. Ce que révèle le génome humain. *Biofutur*, 209, 49-55.

Et les OGM ?

Suivant le *Journal officiel*, un organisme génétiquement modifié est un organisme dont le matériel génétique a été modifié autrement que par multiplication ou recombinaison naturelle. Il est donc essentiel d'analyser non seulement le « plus » amené en terme de performance, mais également la manière dont le « gène » s'intègre dans le programme général de développement d'un organisme, ce qui nécessite des précisions principalement sur : 1°) l'endroit précis du génome où est intégré le transgène ; 2°) le moment de l'expression du transgène ; 3°) l'interaction du transgène et du reste du génome. En d'autres termes, il convient de distinguer si le gène étranger est intégré comme une entité autonome qui « ajoute » une performance supplémentaire à un organisme, ou comme partie prenante d'un programme existant.

Une lecture des difficultés techniques permet de donner quelques éléments de réponse. Commentons, par exemple, le résumé fourni pour les étudiants par Louis-Marie Houdebine²³ : « Des observations qui se sont accumulées ont fait émerger des règles générales. Un transgène est d'autant plus mal exprimé 1) qu'il contient un ADNc plutôt qu'une région génomique avec ses introns ; 2) qu'il contient des séquences d'ADN très étrangères aux cellules animales et, en particulier, des séquences bactériennes riches en GC ; 3) qu'il est intégré dans le génome en plus grand nombre de copies organisées en tandem ou dans la configuration tête-queue ; 4) que le promoteur utilisé est court. »

Houdebine propose diverses explications : « Tous ces faits concordent pour considérer que le mauvais fonctionnement des transgènes doit résulter de plusieurs faits indépendants : 1°) ils sont probablement souvent mal construits car ils ne contiennent pas tous les éléments régulateurs présents dans la chromatine ; 2°) leur expression est favorisée s'ils sont intégrés dans une région ouverte et transcriptionnellement active de la chromatine et, au contraire, inhibée s'ils sont intégrés dans une région fermée de la chromatine (hétérochromatine) comme les centromères et les télomères ; 3°) les transgènes mal construits ou en trop grand nombre de copies induisent dans la cellule une réaction de défense qui a pour effet de les inactiver ; ces mécanismes semblent être ceux qui inactivent les transposons et les séquences rétrovirales. Ils induisent localement la formation d'hétérochromatine accompagnée ou provoquée par la désacétylation des histones, l'accumulation de nucléosomes et la méthylation de l'ADN. »

Commentons pas à pas ces explications :

Ils sont probablement mal construits... : La séquence n'est pas pourvue de toutes les caractéristiques biologiques ; c'est une molécule artificielle, fabriquée de la main de l'homme. Affirmer que faire un OGM revient à mimer la nature, comme on l'entend parfois, relève d'un double mensonge : 1°) le transfert horizontal de gène (c'est-à-dire le passage d'un organisme d'une espèce à un organisme d'une autre espèce), s'il est fréquent chez les archées et les eubactéries, n'est que très exceptionnel chez les eucaryotes et, en particulier, chez les animaux et les plantes vertes terrestres ; 2°) dans le cas d'un OGM, le transfert n'intéresse pas une séquence recombinante naturelle, c'est-à-dire une séquence résultant de processus biologiques mis en œuvre au cours de l'histoire évolutive d'un organisme, mais une séquence artificielle, chimère, résultant d'un assemblage en laboratoire, et qui possède les caractéristiques biologiques que, pour le moment, on attribue au gène.

Leur expression est favorisée ... : C'est une preuve que le programme ne se résout pas exclusivement à la séquence, ni aux régions régulatrices proches ; l'endroit du génome où le gène est intégré revêt toute son importance.

²³ Houdebine L.M., 2001. *Transgénèse animale et clonage*. Dunod, Paris, p. 101.

Les transgènes mal construits [...] induisent [...] une réaction de défense... : c'est une belle démonstration de l'importance de l'épigénèse, c'est-à-dire de la mise en œuvre de l'ensemble des boucles de régulation de la cellule.

S'il fallait enfoncer le clou, l'utilisation des isolateurs²⁴ est là pour nous apprendre la manière dont un transgène est, pour le moment, intégré dans le génome. Son expression est d'autant plus active qu'il est isolé du reste du génome par des séquences non codantes²⁵, « *appelées isolateurs dans la mesure où elles semblent préserver le gène contre son environnement chromatinien.* » Maintenant, la chose est claire : en fait, actuellement, la transgénèse ne fonctionne bien que si le transgène est isolé du reste du génome, c'est-à-dire *n'est pas intégré fonctionnellement au génome de l'organisme transformé.* En d'autres termes, on ajoute un petit élément qui, pour être opératoire, doit avoir le moins possible de relations avec les gènes de l'hôte.

À la suite de Catherine et Raphaël Larrère, on peut évidemment dire²⁶ : «... cette technique, pour spectaculaire qu'elle soit, relève plus du bricolage que d'un procédé maîtrisé et spectaculaire ». Ils en déduisent que c'est un bluff technologique. Réellement ? Le tout est de savoir de quel côté on se place : comme travail de laboratoire, c'est quand même pas mal... Comme application, on ne peut qu'être d'accord avec ces auteurs. Mais n'oublions pas que dans ce domaine, il peut arriver que l'on utilise quelque chose dont on ne comprend pas l'action (jusqu'à peu, l'aspirine, par exemple) ; *a contrario* on peut bien comprendre quelque chose et ne pouvoir rien en faire (une anomalie chromosomique, par exemple)... De toute manière, ceci nous donne des pistes pour intervenir dans le débat « pour ou contre les OGM ».

Comme tout n'est pas connu sur le gène, il convient d'en continuer l'étude en laboratoire, et les OGM en sont un bel outil de recherche fondamentale ; on voit par là que le concept de gène ne doit pas être abandonné, car c'est grâce à lui que l'on peut questionner le vivant ; ce n'est pas avec le concept d'ORF que les généticiens seront inventifs...

Les OGM représentent également un bel outil en biologie intégrative, pour comprendre le rôle d'un gène dans un réseau de régulation physiologique ou en biologie du développement, chez l'animal comme chez le végétal ; là encore, la recherche fondamentale doit être continuée. C'est là que les OGM paraissent pouvoir fournir le matériel de laboratoire idéal pour réifier le concept de « programme ».

Quant aux applications agronomiques, attention au «bluff technologique»... La réponse est certainement à chercher au cas par cas, en utilisant le principe de précaution dans son acception la plus forte.

²⁴ Voir Frey S., 2001. Les isolateurs du génome. *Biofutur*, 213, 18.

²⁵ Houdebine L.M., 2001. *Transgénèse animale et clonage*. Dunod, p. 103.

²⁶ Voir Larrère C, Larrère R., 2001. Les OGM entre hostilité de principe et principe de précaution. *Le Courrier de l'environnement de l'INRA*, 43, 15-23.