



L'informatique génétique : du code génétique au message héréditaire (1)

M. Gillois

► To cite this version:

M. Gillois. L'informatique génétique : du code génétique au message héréditaire (1). Génétique sélection évolution, 1983, 15 (2), pp.299-322. <hal-00893572>

HAL Id: hal-00893572

<https://hal.science/hal-00893572v1>

Submitted on 11 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Tribune libre

L'informatique génétique : du code génétique au message héréditaire (1)

M. GILLOIS

*I.N.R.A. Laboratoire de Génétique cellulaire, Chemin de Borde-Rouge,
Centre de Recherches de Toulouse - F 31320 Castanet-Tolosan*

Résumé

Cet essai scientifique est une tentative pour intégrer les résultats de la génétique mendélienne, de la génétique moléculaire et de l'hérédité quantitative.

La permanence héréditaire de la morphogénèse au sens de R. THOM découle des dégénérescences successives du code génétique, des mécanismes de la synthèse protéique et de la stabilité structurale des systèmes dynamiques à contrôle polygénique.

L'analyse des propriétés du système dynamique traduisant l'induction de l'opéron lactose de *E. coli* montre le bien fondé et les limites de la théorie mendélienne de l'hérédité quantitative élaborée par R. FISHER. Ces propriétés éclairent les points communs des sélections naturelle et artificielle.

Enfin l'invention génétique est considérée comme l'essence de l'Évolution. Les traits principaux des mécanismes moléculaires, génétiques et épigénétiques susceptibles de concourir à l'invention génétique sont discutés.

Mots clés : système dynamique, hérédité polygénique, sélection, évolution moléculaire, invention génétique.

Summary

The genetic information : from the genetic code to the hereditary message

This scientific essay is an attempt to integrate the results of Mendelian genetics, molecular genetics and quantitative inheritance.

The hereditary permanence of morphogenesis, in the sense used by R. THOM, proceeds from successive degenerations of both the genetic code and the mechanisms of protein synthesis, and from the structural stability of dynamic systems under polygenic control.

Analysis of the properties of one dynamic system, that which brings about the induction of the Lactose operon of *E. coli*, shows how well founded is the Mendelian Theory of quantitative inheritance, as elaborated by R. FISHER, as well as showing its limitations. These properties illustrate the points in common between natural and artificial selection.

(1) Conférence du 5 juin 1982 des Séminaires Interdisciplinaires du Collège de France. Professeurs A. LICHNEROWICZ, F. PERROUX, C. GADOFFRE. A paraître in LICHNEROWICZ A., PERROUX F., GADOFFRE C. Information et Communication. (Séminaires interdisciplinaires du Collège de France), collection « Recherches interdisciplinaires », Maloine, Paris 1983. Publié également dans Génét. Sél. Evol. avec l'autorisation de M. P. DELATTRE, Directeur de la collection « Recherches interdisciplinaires ».

Finally « genetic invention » is considered as the essence of evolution. The principal features of the molecular mechanisms, genetic and epigenetic, which are likely to work towards « genetic invention » are discussed.

Key words : *dynamical systems, polygenic inheritance, selection, molecular evolution, genetic invention.*

I. Introduction

Actuellement la génétique est un domaine scientifique éclaté en îlots de recherche séparés par leurs méthodologies d'analyse. La génétique formelle, la génétique de population, la génétique quantitative, la génétique biochimique, la génétique moléculaire, et le génie génétique accumulent des résultats spécifiques mais parcellaires. En particulier un fossé existe entre la génétique quantitative et la génétique moléculaire. La génétique quantitative est la modélisation des propriétés de l'hérédité des caractères quantitatifs et leur exploitation dans les programmes de sélection artificielle. La génétique moléculaire est l'analyse poussée à l'extrême de la nature biochimique du gène et de ses propriétés de codage et de décodage. Ainsi l'approche globale de la génétique quantitative, qui ignore les propriétés moléculaires des gènes et leurs relations fonctionnelles, s'oppose à l'approche analytique de la génétique moléculaire qui, elle, ne rend pas compte des propriétés de l'hérédité des caractères quantitatifs. De ce fait, une présentation générale et rationnelle de l'information génétique n'est pas encore possible.

Le thème de l'information génétique pourrait être traité avec les seuls résultats acquis en génétique moléculaire. Récemment M. MORANGE (1982) a très bien évoqué l'introduction de la vision informative des problèmes de la génétique sous l'impulsion de E. SCHRÖDINGER (1944) et des physiiciens convertis à la Biologie après la Seconde Guerre mondiale. Dès lors les buts essentiels de la recherche devenaient la nature, la structure, la lecture du support du message héréditaire. Par la suite le développement théorique et technique des ordinateurs a donné, sinon un cadre, du moins une vision informatique de ces problèmes. L'idée de message se traduit actuellement en génétique par un vocabulaire spécifique dans lequel apparaissent les mots : information, code, codage, codon, transcription, traduction, etc. Ce vocabulaire décrit les objets moléculaires (ADN, ARN messenger, ARN de transfert, ARN ribosomal, ribosomes, peptides) et les événements auxquels ils participent (duplication, transcription, traduction, maturation). Cette description constitue le code génétique : un gène est lié à l'état discret et fonctionnel de la macromolécule peptidique qu'il code.

Le code génétique n'est qu'un élément de l'information génétique, le message héréditaire en est un autre.

La seule interprétation génique de l'hérédité des caractères « tout ou rien » a été si convaincante que le mot génétique s'est substitué à celui d'hérédité. Cette interprétation génique de l'hérédité, même complétée des dernières découvertes du code génétique apportées par la Biologie moléculaire, n'est pas en mesure de rendre compte de l'hérédité d'une forme, de la morphogenèse, de l'embryogenèse, de l'ontogenèse, de la convergence fonctionnelle, d'une fonction physiologique, de la permanence structurale et de la flexibilité dans et entre les populations au cours du temps.

A la notion de code génétique propre à la génétique moléculaire correspond celle, complémentaire, de message héréditaire de la génétique quantitative. Ce message héréditaire est l'ensemble des propriétés liées au fonctionnement de nombreux gènes associés pour la réalisation d'un caractère. Il tient compte essentiellement des propriétés dynamiques des macromolécules participant à l'expression d'un phénotype sous contrôle polygénique. Car pour la vie, il s'agit non seulement de pouvoir faire, ce qu'explique le code génétique, mais surtout d'avoir le temps d'assembler les informations prélevées sur de nombreux gènes.

Les résultats bien établis relatifs au message héréditaire sont issus essentiellement d'une étude unique, celle du fonctionnement de l'opéron lactose de *E. coli*. C'est peu, mais ils sont exemplaires, aussi seront-ils considérés comme la pierre angulaire de cet exposé sur l'information génétique. Les relations du code génétique et du message héréditaire seront traitées sous trois angles.

Le premier, le contrôle polygénique du phénotype d'un caractère, sera une introduction au contrôle génétique de la morphogenèse au sens de R. THOM (1972).

Le second, les variations du message héréditaire dans les populations, éclairera certains points communs à la sélection artificielle et à la sélection naturelle, mécanisme de la microévolution.

Le troisième, l'apparition, l'établissement, l'annulation des messages héréditaires au cours des générations débouchera sur la propriété essentielle de l'évolution, sensu stricto : les capacités et les incapacités d'innovation.

II. Le contrôle polygénique du phénotype d'un caractère

A. L'expression phénotypique d'un ensemble de gènes se réalise en trois étapes séquentielles. La première fait correspondre à la séquence nucléotidique de chaque gène la séquence des acides aminés d'une chaîne polypeptidique. La seconde développe à partir de chaque séquence d'acides aminés une forme, elle seule douée d'une fonction, qui est le sens premier du message codé dans le génome. La troisième est le déploiement dans l'espace et au cours du temps des interactions fonctionnelles entre les éléments moléculaires synthétisés.

B. La première étape, qui réunit les événements conduisant des lectures des gènes à l'apparition des produits primaires, est considérée comme le système de « communication génétique ». Bien que les lectures successives par excision, épissage, maturation, des Eucaryotes soient beaucoup plus complexes que celles des Procaryotes, le modèle de communication génétique présente une analogie avec le schéma de la communication d'un message à codage linéaire sur une ligne de transmission. Dans le cadre de la théorie de l'information appliquée à la transcription et à la traduction génétiques, le problème est de rendre compte de la stabilité et de la sécurité de la traduction de l'ARN messager en une séquence d'acides aminés. Les mécanismes de la traduction associent trois partenaires majeurs : l'ARN messager mature ou message codé, les ribosomes, et les ARN de transfert chargés spécifiquement de leur acide aminé. Un seul enzyme d'activation, spécifique d'un acide aminé, le charge sur son ou ses ARN de transfert. Un ARN de transfert lit un ou plusieurs codons spécifiques de l'acide aminé qu'il transfère.

On a d'abord pensé que la reconnaissance stérique et statique entre un codon de l'ARN messager et l'anticodon de l'ARN de transfert suffisait à assurer la stabilité et la sécurité de la traduction. La réalité est différente ; la stabilité et la sécurité sont dues à une dynamique. L'écoulement d'un laps de temps est nécessaire à la réalisation d'une liaison peptidique. Tout ARN de transfert placera son acide aminé dans la séquence peptidique en croissance s'il reste en place suffisamment longtemps. Le mécanisme naturel assurant la sécurité de lecture vis-à-vis du code consiste à abaisser la durée de vie de toutes les associations codon-anticodon possibles, de telle sorte que seules les associations conformes au code génétique aient une durée de vie suffisante pour permettre la réalisation de la liaison peptidique. Le rôle de l'Inosine et des dérivés de l'Uracile dans les anti-codons serait d'affaiblir les capacités d'association de l'anticodon pour éviter les erreurs de traduction. Cette notion a été proposée par J. NINIO (1971) et J. HOPFIELD (1974). A. SPIRIN et ses collaborateurs (1976) lui ont donné une confirmation expérimentale.

Exclusion de l'anticodon GCG (Arginine) au profit de GCI
 Replacement of GCG by GCI anticodons

Logique du code génétique que la nature refuse

<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: right;">Codon</td> <td style="text-align: center;">1 2 3</td> <td></td> </tr> <tr> <td></td> <td style="text-align: center;">U G C</td> <td>Cysteine</td> </tr> <tr> <td style="text-align: right;">Anticodon</td> <td style="text-align: center;">A C G</td> <td>Cysteine</td> </tr> <tr> <td></td> <td style="text-align: center;">3 2 1</td> <td></td> </tr> </table>	Codon	1 2 3			U G C	Cysteine	Anticodon	A C G	Cysteine		3 2 1		<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: right;">Codon</td> <td style="text-align: center;">1 2 3</td> <td></td> </tr> <tr> <td></td> <td style="text-align: center;">U G C</td> <td>Cysteine</td> </tr> <tr> <td style="text-align: right;">Anticodon</td> <td style="text-align: center;">G C G</td> <td>Arginine</td> </tr> <tr> <td></td> <td style="text-align: center;">3 2 1</td> <td></td> </tr> </table>	Codon	1 2 3			U G C	Cysteine	Anticodon	G C G	Arginine		3 2 1		<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: right;">Codon</td> <td style="text-align: center;">1 2 3</td> <td></td> </tr> <tr> <td></td> <td style="text-align: center;">C G U</td> <td>Arginine</td> </tr> <tr> <td style="text-align: right;">Anticodon</td> <td style="text-align: center;">G C G</td> <td>Arginine</td> </tr> <tr> <td></td> <td style="text-align: center;">3 2 1</td> <td></td> </tr> </table>	Codon	1 2 3			C G U	Arginine	Anticodon	G C G	Arginine		3 2 1	
Codon	1 2 3																																					
	U G C	Cysteine																																				
Anticodon	A C G	Cysteine																																				
	3 2 1																																					
Codon	1 2 3																																					
	U G C	Cysteine																																				
Anticodon	G C G	Arginine																																				
	3 2 1																																					
Codon	1 2 3																																					
	C G U	Arginine																																				
Anticodon	G C G	Arginine																																				
	3 2 1																																					
Association	de longue durée	d'assez longue durée	de longue durée																																			
Elongation peptidique	possible	possible	possible																																			

Pratique d'un code génétique trafiqué par la nature

<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: right;">Codon</td> <td style="text-align: center;">1 2 3</td> <td></td> </tr> <tr> <td></td> <td style="text-align: center;">U G C</td> <td>Cysteine</td> </tr> <tr> <td style="text-align: right;">Anticodon</td> <td style="text-align: center;">A C G</td> <td>Cysteine</td> </tr> <tr> <td></td> <td style="text-align: center;">3 2 1</td> <td></td> </tr> </table>	Codon	1 2 3			U G C	Cysteine	Anticodon	A C G	Cysteine		3 2 1		<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: right;">Codon</td> <td style="text-align: center;">1 2 3</td> <td></td> </tr> <tr> <td></td> <td style="text-align: center;">U G C</td> <td>Cysteine</td> </tr> <tr> <td style="text-align: right;">Anticodon</td> <td style="text-align: center;">G C I</td> <td>Arginine</td> </tr> <tr> <td></td> <td style="text-align: center;">3 2 1</td> <td></td> </tr> </table>	Codon	1 2 3			U G C	Cysteine	Anticodon	G C I	Arginine		3 2 1		<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: right;">Codon</td> <td style="text-align: center;">1 2 3</td> <td></td> </tr> <tr> <td></td> <td style="text-align: center;">C G U</td> <td>Arginine</td> </tr> <tr> <td style="text-align: right;">Anticodon</td> <td style="text-align: center;">G C I</td> <td>Arginine</td> </tr> <tr> <td></td> <td style="text-align: center;">3 2 1</td> <td></td> </tr> </table>	Codon	1 2 3			C G U	Arginine	Anticodon	G C I	Arginine		3 2 1	
Codon	1 2 3																																					
	U G C	Cysteine																																				
Anticodon	A C G	Cysteine																																				
	3 2 1																																					
Codon	1 2 3																																					
	U G C	Cysteine																																				
Anticodon	G C I	Arginine																																				
	3 2 1																																					
Codon	1 2 3																																					
	C G U	Arginine																																				
Anticodon	G C I	Arginine																																				
	3 2 1																																					
Association	de longue durée	de courte durée	de longue durée																																			
Elongation peptidique	possible	impossible	possible																																			

FIG. 1

A = Adenine ; C = Cytosine ; G = Guanine ; I = Inosine ; U = Uracile.

A = Adenine ; C = Cytosine ; G = Guanine ; I = Inosine ; U = Uracil.

C. La traduction de l'ARN messager donne naissance à une séquence d'acides aminés. Cette chaîne peptidique nouvellement synthétisée s'isole ; déployée dans l'espace elle se replie et acquiert une conformation native, propre et biologiquement active : c'est la seconde étape.

Les séquences d'acides-amino et la conformation tridimensionnelle de la protéine sont deux messages liés par un deuxième mécanisme de traduction qui est le repliement de la chaîne peptidique native, B. ROBSON et R.H. PAIN (1971), Y. CHOU et G.D. FASMAN (1974), E. SUZUKI et B. ROBSON (1976) B. ROBSON et E. SUZUKI (1976), J. GARNIER *et al.* (1978). L'épellation du message génétique s'achève, les mots apparaissent.

Les protéines ont un squelette linéaire dont les replis et les détours sont des combinaisons de quelques motifs structuraux de base bien répertoriés. L. PAULING et R. COREY (1951) ont ouvert ce lexique en découvrant l'hélice α et les feuillets plissés. Les chaînes d'acides aminés ont une tendance assez forte à adopter localement soit une conformation en hélice α , soit une conformation en zigzag. Deux zigzags côte à côte forment un feuillet plissé β parallèle ou antiparallèle suivant que les chaînes progressent dans le même sens ou non. Ces structures et les structures qui les rejoignent, comme les boucles, les tournants β ou γ , sont déterminées par la nature et la séquence des acides aminés qui les composent. Le code qui lie la séquence d'acides aminés d'une protéine à sa conformation tridimensionnelle n'est pas encore complètement élucidé. La première étape dans la solution de ce problème consiste à prédire la structure locale, ou conformation secondaire, à laquelle appartient un acide aminé en connaissant la séquence peptidique qui l'entoure. Les chercheurs dans ce domaine ont introduit la quantité d'information I apportée sur la structure locale entourant un acide aminé par la connaissance de la séquence des acides aminés. Des considérations mathématiques et expérimentales permettent d'exprimer I comme une somme d'informations indépendantes apportées par chaque acide aminé de la séquence.

De plus, l'expérience montre que pour prédire une structure locale au voisinage de la position J, l'information apportée par la connaissance de la nature des acides aminés dans la séquence s'annule au-delà de J-8 et de J+8, J. GARNIER *et al.* (1978). Des tables de I sont publiées et corrigées au fur et à mesure qu'augmente le nombre des protéines séquencées et dont la conformation est connue.

Décodage des structures secondaires locales des séquences peptidiques

HELICE - α									
	-4	-3	-2	-1	0	+1	+2	+3	+4
Alanine	-20	-30	-40	-50	-50	-40	-30	-20	-10
Leucine	-20	-30	-40	-50	-56	-20	-10	0	0
Glutamine	-20	-30	-40	-45	-47	-20	-0	10	5
Glycine	10	30	55	55	57	40	0	0	0
Tyrosine	15	15	20	25	29	25	20	15	15
Proline	50	70	10	-90	36	90	10	0	0

TOURNANT									
	-4	-3	-2	-1	0	+1	+2	+3	+4
Alanine	30	40	50	60	65	60	50	40	30
Leucine	20	25	28	30	32	30	28	25	20
Glutamine	10	20	60	70	78	78	78	78	78
Glycine	-30	-40	-50	-60	-86	-60	-50	-40	-30
Tyrosine	-25	-30	-35	-40	-45	-40	-35	-30	-25
Proline	-80	-100	-120	-140	-77	-60	-30	-20	-10

FIG. 2

Les nombres, -4, -3, -2, -1, 0, +1, +2, +3, +4, indiquent le rang de l'acide aminé. Les nombres du tableau sont des quantités d'information exprimée en nats $\times 1\ 000$. D'après GARNIER, 1978, Op. Cit.

The numbers, -4, -3, -2, -1, 0, +1, +2, +3, +4, designate the position of the amino acid in respect to the conformation observed at position 0. The values recorded are a measure of the conformational information expressed as nats $\times 1\ 000$. See GARNIER, 1978, Op. Cit.

Ces motifs de structure locale forment des motifs d'ordre supérieur en respectant des règles de composition. Ces règles servent dès maintenant à rédiger des programmes informatiques de prévision et de visualisation des conformations.

D. Les conformations des protéines ne sont significatives que par les fonctions qu'elles remplissent. Ces fonctions sont, soit exclusivement des assemblages macromoléculaires mettant en jeu des forces de faible niveau énergétique (ex. : Van der Waals, London,...), soit des assemblages du type précédent et des transformations chimiques utilisant des forces de liaison de haut niveau énergétique (ex. : liaison covalente). Les premières fonctions donnent naissance aux structures macromoléculaires, les secondes aux activités enzymatiques. Ces fonctions sont caractérisées par leur spécificité, leur cinétique,

leur durée de vie. C'est la troisième étape, finale, de la transmission du message héréditaire dans une cellule. L'étude des cinétiques conjointes nécessite l'emploi de l'outil mathématique. La traduction mathématique doit satisfaire à plusieurs conditions :

1 - Elle doit être un résumé fidèle du fonctionnement du système dans les différentes conditions de milieu, ce qui exige un accord qualitatif avec l'ensemble des résultats expérimentaux.

2 - Elle doit être capable de traduire une variabilité génétique par un choix de paramètres assurant une relation claire entre un gène, ses allèles et leurs effets.

3 - Elle doit être formulée en termes dynamiques, car les effets d'un ensemble de gènes sont décrits par des associations moléculaires, des cinétiques enzymatiques, des taux d'incorporation ou de dégradation.

4 - Elle doit avoir une structure mathématique assez simple pour permettre d'exprimer en termes généraux les propriétés qualitatives du système.

Quel est le bilan des modèles mathématiques proposés pour répondre à ces exigences ? Ils se rattachent à deux types de méthodes, l'algébrique et l'analytique.

R. ROSEN (1972) décrit les liaisons structurelles entre divers éléments d'un système métabolique soumis à des régulations, avec le langage de la théorie des catégories. R. THOMAS (1973) utilise un formalisme booléen qui distingue deux états pour chacun des éléments, présent ou non, en fonction ou non, en quantité suffisante ou non. Ces modèles algébriques sont essentiellement statiques et qualitatifs, ils décrivent seulement les états stationnaires et ils ne tiennent pas compte de la variabilité génétique.

La représentation analytique à l'aide de systèmes dynamiques traduit, par des équations différentielles liées, les variations de concentration des ARN messagers, des enzymes et des substrats. Certains de ces essais, comme celui de B. GOODWIN (1963) d'une mécanique statistique des processus moléculaires et de leurs régulations, et comme celui de H. KACSER et A. BURNS (1981) de la dynamique des flux dans un réseau métabolique, sont très généraux et ne possèdent qu'une valeur de métaphore. D'autres travaux traitent au contraire de systèmes particuliers et s'appuient sur les données structurelles et les mesures acquises par l'expérimentation comme les études des premières étapes de la glycolyse contrôlées par la phospho-fructo-kinase menées par J. HIGGINS (1967), E. SEL'KOV (1968) et J. DEMONGEOT (1981). De même l'opéron lactose de *E. coli* a été étudié par W. KNORRE (1973), B. GOODWIN (1969), A. BABLOYANTZ et M. SANGIER (1972), M. SANGIER et G. NICOLIS (1976). Tous ces travaux ont été entrepris avec pour objectif la recherche de structures biologiques susceptibles d'engendrer les oscillations entretenues d'une « horloge biologique ». Ces modèles de l'opéron lactose comportent des simplifications injustifiées, concernant les vitesses relatives des différentes étapes, traduisent de façon erronée les mécanismes de l'induction et de la répression catabolique, enfin ils ne tiennent pas compte des études expérimentales des souches mutantes de *E. coli* dont l'analyse a été déterminante dans la compréhension du système. P. VON HIPPEL (1974), W. MANDECKI (1979) et T. MANABE (1981) limitent leurs modèles aux mécanismes moléculaires du seul domaine des gènes de régulation.

Nous avons tenté, C. CHEVALET, F. CORPET, A. MICALI et M. GILLOIS, une synthèse mathématique des connaissances analytiques accumulées à propos de l'opéron lactose. (1982). L'intérêt de cette démarche est l'étude du fonctionnement conjoint d'un ensemble de gènes et l'espoir d'une meilleure compréhension des mécanismes de l'expression de l'hérédité polygénique.

Le modèle mathématique de l'induction de l'opéron lactose est un système de onze équations différentielles non-linéaires avec arguments retardés — six équations décrivent les probabilités des états de la région de contrôle de l'ADN, deux équations traduisent les cinétiques de la synthèse des enzymes β -galactosidase et perméase, trois équations donnent les cinétiques de l'apparition dans la cellule du

lactose, du glucose et de l'allolactose. L'établissement de ce modèle tient compte des hiérarchies des nombres d'éléments et des hiérarchies des vitesses. Les petits nombres d'effecteurs autour de l'opérateur, les nombres plus grands d'enzymes et encore plus grands de substrats justifient la représentation probabiliste de la région de contrôle et la représentation cinétique des activités enzymatiques. La plus grande vélocité de l'interaction entre inducteur et répresseur par rapport à l'interaction entre répresseur et opérateur d'une part, et la dégradation de l'ARN messager plus rapide que les variations du taux de transcription d'autre part, permettent d'exprimer le taux moyen d'apparition des enzymes en fonction du taux de transcription à un moment antérieur. Cette propriété biologique impose l'introduction des arguments retardés dans les équations.

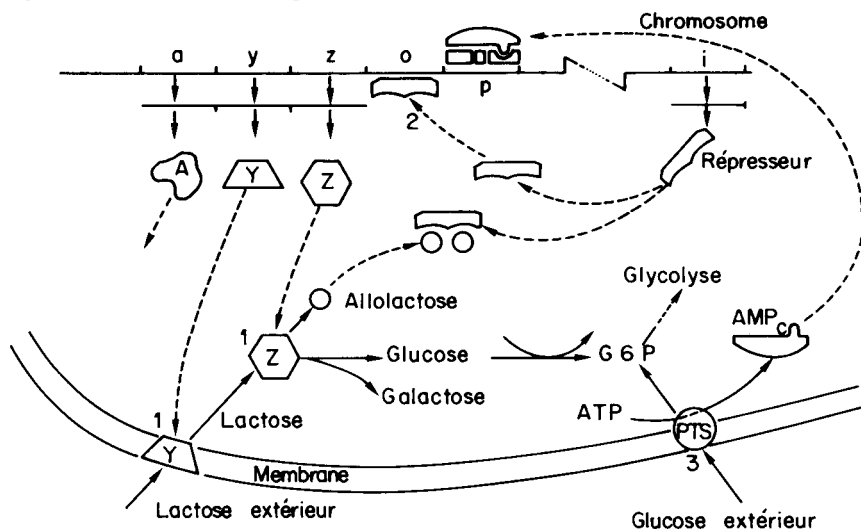


FIG. 3

Schéma de l'opéron lactose de *E. COLI*

1. La β -galactosidase (Z) possède deux fonctions : l'hydrolyse du lactose (lactose \rightarrow glucose + galactose) et l'isomérisation du lactose en allolactose, qui est l'inducteur naturel I.
2. Répression et induction de la transcription des gènes de structure (z, y, a) de l'opéron : en absence d'inducteur I le répresseur R, produit du gène i , s'associe au chromosome bactérien au niveau de l'opérateur O ($R + O \rightleftharpoons RO$) et bloque ainsi la transcription ; l'inducteur I est susceptible de s'associer au répresseur R ($I + RI_k \rightleftharpoons RI_k$; $k = 1, 2, 3, 4$) et au complexe RO dont il facilite la dissociation ($RO + I \rightleftharpoons ORI \rightleftharpoons RI + O$).
3. Le système de perméation (PTS) assure soit la phosphorylation du glucose, soit la conversion de l'ATP en AMP cyclique : ces deux réactions sont compétitives car les enzymes correspondants sont liées à la même structure membranaire. L'AMP cyclique, en s'associant à un complexe protéique au niveau du promoteur P, joue un rôle d'amplificateur de la transcription de l'ADN.

Diagram of the lactose operon of *E. COLI*

1. β -galactosidase (Z) has two functions : hydrolysis of lactose (lactose \rightarrow glucose + galactose) and isomerisation of lactose to allolactose, which is the natural endogenous inductor I.
2. Repression and induction of transcription of the structural genes (z, y, a) of the operon : in the absence of inductor I the repressor R, product of gene i , binds to the operator O in the bacterial chromosome ($R + O \rightleftharpoons RO$), blocking transcription ; the inductor I may bind both to the repressor R ($I + RI_{k-1} \rightleftharpoons RI_k$; $k = 1, 2, 3, 4$) and to the RO complex, facilitating its dissociation ($RO + I \rightleftharpoons ORI \rightleftharpoons RI + O$).
3. The permease system (PTS) leads to either the phosphorylation of glucose or the conversion of ATP to cAMP : these two reactions are competitive, as a consequence of the corresponding enzymes being bound to the same membrane component. The cAMP, when bound to a protein complex at the promoter P, leads to the amplification of transcription of the lactose operon.

Les propriétés structurelles du système déterminent les traits communs aux différents comportements, induits par les variations des grandeurs numériques attachées aux paramètres génétiques et aux paramètres du milieu. L'analyse qualitative des solutions du système, et la recherche des domaines de variation des paramètres à l'intérieur desquels les solutions présentent des similarités renseignent sur les propriétés structurelles.

Ainsi les études numériques montrent que la région de contrôle sur l'ADN, qui peut être dans l'un de six états fonctionnels exclusifs, possède des réponses quasi-instantanées aux variations de concentration des molécules régulatrices. Sur un plan théorique, cela signifierait que l'opéron lactose fonctionne comme un automate sans mémoire ; sur un plan technique, cette première propriété permet une réduction de la dimension du système d'équations différentielles autorisant un traitement analytique.

La seconde propriété est l'unicité de l'état d'équilibre que le génotype soit sauvage ou soit mutant (I-, I^q, I^s, I^s, PI, PII, PIII). Les conditions suffisantes d'unicité ont une interprétation biologique :

1 - La somme des activités des perméases doit être inférieure à celle des β -galactosidases. L'activité de la perméase est le facteur limitant de l'expression de l'opéron. C'est une observation expérimentale ancienne.

2 - La durée moyenne de vie des perméases doit être plus longue que celle de la β -galactosidase. C'est une conjecture, car il n'existe pas de mesure directe de la durée de vie pour un composant membranaire.

3 - Le répresseur doit être suffisamment sensible à l'inducteur.

Ces trois propriétés sont possédées par les souches de *E. coli* de génotypes sauvages ou mutants connus.

La troisième propriété est la stabilité pratique de l'équilibre qui est une condition de son observation. Mathématiquement elle se traduit soit par un point d'équilibre stable, soit par une trajectoire présentant des fluctuations de très faible amplitude. Les conditions de stabilité s'expriment d'une façon analogue à celles de l'unicité par des inégalités entre les paramètres du modèle. Le génotype sauvage et tous les mutants connus vérifient les conditions de stabilité. Ce n'est pas une surprise, sinon ils n'auraient pas pu être isolés et caractérisés par les expérimentateurs. C'est aussi une confirmation de la validité du modèle. Les conditions de stabilité peuvent s'exprimer en fonction des valeurs des paramètres caractéristiques du génotype sauvage. Ainsi est-il possible de définir deux conditions, de signification génétique, conduisant à l'instabilité du point d'équilibre pour certaines concentrations du lactose extérieur à la cellule :

1 - Les perméases doivent devenir plus efficaces (constante catalytique accrue) ou plus nombreuses (synthèse plus forte) mais moins stables (taux de dégradation plus grand) que chez le sauvage.

2 - L'affinité de l'inducteur pour le répresseur doit être diminuée par un facteur de l'ordre de 100 à 1000. Cette dernière propriété existe dans les mutants I^s. Mais il n'existe pas de mutants décrits possédant les deux mutations. Peut-être des mutants possédant simultanément ces deux propriétés ont-ils été isolés par des généticiens, qui les ont déclarés « leaky » ou bon à rien.

Si l'on admet, comme le suggèrent nos résultats numériques (1982), que les concentrations intracellulaires des substrats lactose et allolactose s'ajustent rapidement à celles des enzymes perméase et β -galactosidase, alors il existe une condition suffisante de stabilité dont la traduction biologique est la permanence de l'ordre de lecture des gènes sur l'ADN de l'opéron lactose.

La quatrième propriété est l'influence des retards dus à la synthèse protéique. Les retards τ_y et τ_z de la synthèse des enzymes perméases et β -galactosidases n'ont aucun effet sur la stabilité des équilibres tant pour le génotype sauvage que pour ceux des mutants isolés et décrits.

Nous avons conçu deux mutants raisonnables qui transgressent les conditions de stabilité lorsque les retards de transcription et de traduction sont nuls.

Le mutant « Florence » porte deux mutations : le coefficient d'association du répresseur à l'inducteur est cinq cents fois plus petit ; la perméase devient beaucoup plus instable que la β -galactosidase et acquiert une constante catalytique accrue, de 300 molécules de lactose transférées par seconde et par perméase. Cette valeur est la borne supérieure des estimations expérimentales faites pour le génotype sauvage. Le mutant « Claude » porte aussi deux mutations : le coefficient d'association du répresseur à l'inducteur est divisé par cinquante ; la perméase devient beaucoup plus instable que la β -galactosidase mais est produite en excès par la duplication de son gène.

L'étude en fonction de τ_y et τ_z du nombre de valeurs propres à parties réelles positives du système linéarisé autour du point d'équilibre donne les couples de τ_y , τ_z pour lesquels ce nombre est nul, ce qui entraîne la stabilité de l'équilibre.

Pour des retards τ_z de 4 minutes et τ_y de 5 minutes, considérés comme normaux, le mutant « Florence » est stable, des oscillations amorties le conduisent à son équilibre ; le mutant « Claude » est instable. Dans le domaine où τ_y et τ_z sont compris entre 1 et 16 minutes, le système admet des situations de stabilité et d'autres d'instabilité dont la détermination dépend des gènes contrôlant la machinerie de transcription et de traduction, et de la concentration du substrat inducteur dans le milieu de culture.

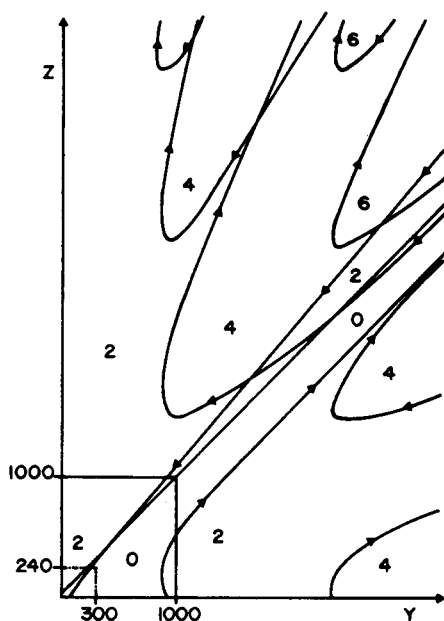


FIG. 4

Stabilité fonctionnelle de l'opéron lactose du mutant « Florence » de E. COLI.

Z = Retard de synthèse de la β -galactosidase en secondes.

Y = Retard de synthèse de la perméase en secondes.

0, 2, 4, ..., $2p$ sont les nombres de valeurs propres à partie réelle positive du système linéarisé autour du point d'équilibre. L'équilibre est stable quand ce nombre est nul.

Functional stability of the lactose operon in the mutant "Florence" of E. COLI.

Z = Delay, in seconds, of β -galactosidase synthesis.

Y = Delay, in seconds, of permease synthesis.

0, 2, 4, ..., $2p$ are positive and real eigen values of the system linearised in respect to the equilibrium point. The equilibrium is stable when that number is 0.

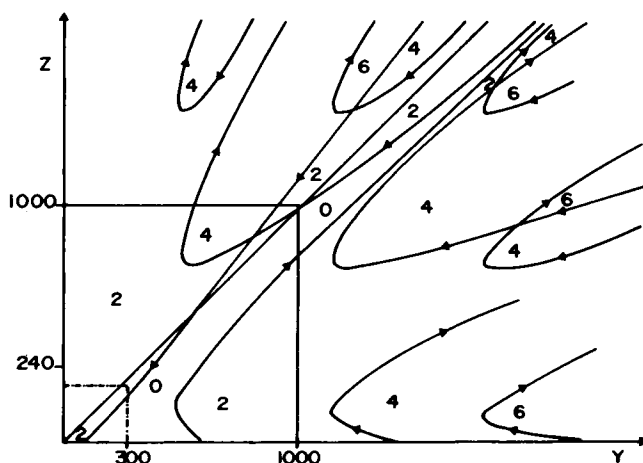


FIG. 5

Stabilité fonctionnelle de l'opéron lactose du mutant « Claude » de E. COLI.

Z = Retard de synthèse de la β -galactosidase en secondes.

Y = Retard de synthèse de la perméase en secondes.

0, 2, 4, ..., $2p$ sont les nombres de valeurs propres à partie réelle positive du système linéarisé autour du point d'équilibre. L'équilibre est stable quand ce nombre est nul.

Functional stability of the lactose operon in the mutant "Claude" of E. COLI.

Z = Delay, in seconds, of β -galactosidase synthesis.

Y = Delay, in seconds, of permease synthesis.

0, 2, 4, ..., $2p$ are positive and real eigen values of the system linearised in respect to the equilibrium point. The equilibrium is stable when that number is 0.

En conclusion, ce travail sur le système polygénique de l'opéron lactose d'*E. coli* montre que les propriétés héréditaires du phénotype ne dépendent pas exclusivement du code génétique qui spécifie la fonction des protéines. Ces propriétés héréditaires dépendent des relations fonctionnelles entre les gènes, des positions relatives des gènes du système sur le chromosome, enfin des phénotypes intracellulaires des autres systèmes génétiques. Ainsi l'inversion de position entre le gène *z* de la β -galactosidase et le gène *y* de la perméase pourrait abolir la propriété de stabilité de l'équilibre unique. Certains doubles mutants de l'opéron lactose auraient des propriétés de stabilité dépendantes de l'efficacité de la synthèse protéique.

L'expression de l'équilibre du système en fonction des paramètres du milieu est la traduction des notions bien connues des biologistes que sont la plasticité ou la marge d'adaptation physiologique d'un génotype dans des milieux différents. L'expression de l'équilibre du système en fonction des paramètres géniques permet une étude renouvelée des variabilités géniques et phénotypiques d'une population et de ses réponses à la sélection.

Les mécanismes de l'expression phénotypique de l'information génétique pendant l'embryogénèse ou la morphogénèse sont encore inconnus. Les nombreuses études logiques et mathématiques suscitées par ce domaine de recherche sont trop métaphoriques, R. ROSEN (1970), R. ROSEN (1973), G. NICOLIS et I. PRIGOGINE (1977), H. HAKEN (1978). Souhaitons que les outils mathématiques qu'ils ont forgés, seront adaptés à la synthèse des résultats expérimentaux de l'analyse génétique et moléculaire lorsque ces résultats seront disponibles. Alors, le moment sera venu de donner une base génétique à la morphologie rationnelle au sens de B. GOODWIN (1981).

III. Les variations du message héréditaire et la sélection

Au chapitre précédent l'exemple de l'opéron lactose a montré d'une part que le génotype sauvage et tous les génotypes mutants connus possèdent les mêmes propriétés d'unicité et de stabilité de leurs équilibres phénotypiques, et d'autre part que les coordonnées de ces équilibres sont différentes et dépendent des paramètres génétiques, c'est-à-dire des allèles réunis dans chaque génotype. Les mécanismes de reproduction dans une population assurent la réunion, la ségrégation, l'apparition, la disparition des allèles.

L'information génétique s'exprime dans une cellule ou un individu, y est dupliquée puis transmise à des descendants. La duplication avec mutation, amplification, transposition des ADN, et la ségrégation des ADN et de leurs supports induisent des variations de l'information génétique. Les lois de la transmission de l'information génétique sont définies par la structure du chromosome et l'existence ou non de phases haploïdes, diploïdes. Ainsi les lois de Mendel ne s'appliquent qu'aux organismes eucaryotes, qui possèdent à la fois la succession haplophase, fusion, diplophase, méiose, et une structure chromosomique à corps central protéique support de l'ADN, J.R. PAULSON et U.K. LAEMMLI (1977). C'est à ces organismes que la suite de ce chapitre se réfère.

La ressemblance entre les individus d'une population est d'autant plus grande qu'ils sont apparentés, c'est-à-dire partagent les mêmes ancêtres. Deux individus apparentés peuvent avoir reçu ou non des gènes d'un ancêtre commun. Si ces gènes sont les copies non modifiées des gènes présents dans l'ancêtre, leurs expressions phénotypiques dans la descendance seront équivalentes. Les mécanismes moléculaires à contrôle polygénique des caractères quantitatifs héréditaires sont encore inconnus. Aussi les gènes en cause ne peuvent être décrits que par trois attributs : l'homologie, l'isoaction, l'identité, M. GILLOIS (1964). Deux gènes sont homologues s'ils sont au même lieu sur le chromosome ; deux gènes sont isoactifs s'ils ont la même action dans la réalisation d'un même caractère. Deux gènes sont identiques s'ils dérivent par copies biochimiques successives sans mutation d'un même gène ancêtre.

Les traces géniques du message héréditaire dans les populations sont les filiations ou pedigrees liant les individus et les appréciations statistiques de la ressemblance de leurs caractères héréditaires.

La relation d'identité est une relation binaire d'équivalence. Elle induit sur tout ensemble de gènes homologues d'une population une partition en sous-ensembles disjoints de gènes identiques, qui est une situation d'identité. La réalisation d'une situation d'identité dépend d'une part des lois d'appariement des reproducteurs, traduites par les filiations ou pedigrees, et d'autre part des ségrégations des gènes traduites par les lois de l'hérédité mendélienne.

Si deux gènes sont identiques, la connaissance de la classe d'isoaction ou état fonctionnel de l'un suffit pour définir l'appartenance de l'autre à cette même classe. La connaissance de l'identité réalisée entre deux ou plusieurs gènes suffit à définir la loi conjointe conditionnelle de leurs états fonctionnels. Par contre, lorsque deux gènes sont non identiques, il n'est plus possible de définir les lois conjointes des états géniques fonctionnels ou d'isoaction. Cette indécision est levée si la loi d'appariement des reproducteurs est la panmixie. La loi conjointe conditionnelle des états d'isoaction est alors le produit des lois marginales non-conditionnelles (M. GILLOIS, 1964, 1967). La connaissance des pedigrees permet de calculer les probabilités de ces conditions, ce sont les coefficients d'identité, en utilisant les propriétés algébriques des ensembles ordonnés ou inf-1/2-faisceaux (M. GILLOIS, 1964). La connaissance du nombre des reproducteurs mâles et femelles d'une population panmixique permet le calcul de l'évolution des coefficients d'identité à l'aide de systèmes d'équations aux différences finies (M. GILLOIS, 1964). Donc dans une population panmixique quelconque, les lois conjointes des états fonctionnels ou d'isoaction des gènes sont conditionnelles aux situations d'identité. Lorsque la loi d'appariement des reproducteurs n'est pas la panmixie mais l'homogamie, la définition des lois conjointes conditionnelles des états fonctionnels ou d'isoaction des gènes est plus riche. Les conditions des lois conjointes sont

alors au nombre de trois : l'isocontrainte, l'hétérocontrainte, la non-contrainte (M. GILLOIS *et al.* ; 1967, 1966). Les probabilités de ces trois conditions sont elles aussi calculables (M. GILLOIS *et al.* ; 1969). Le très grand intérêt de ces lois conjointes des états fonctionnels des gènes est de pouvoir exprimer a priori la variance ou la covariance de caractères héréditaires à contrôle polygénique pour des populations d'effectifs quelconques (M. GILLOIS, 1964 ; M. GILLOIS *et al.*, 1971 ; C. CHEVALET, 1969). Mais ce calcul nécessite de définir l'effet quantitatif de la substitution d'un gène à un autre gène d'une autre classe d'isoaction. La définition la plus efficace et la plus robuste est celle de R. FISHER (1918). Ainsi l'une quelconque des expressions obtenues (M. GILLOIS, 1964) présente la forme générale suivante :

$$\text{COV} (Z_1 Z_2) = \sum_i \delta_i^1 \gamma_i^{12}$$

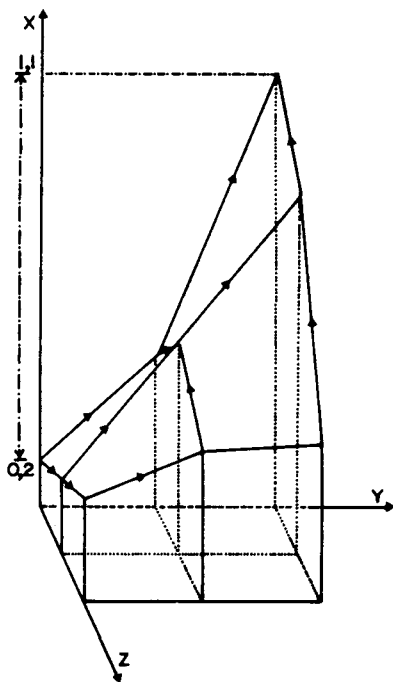


FIG. 6

Surface des valeurs du phénotype de l'opéron lactose de *E. COLI* en fonction de deux variables génétiques.

axe X : phénotype, production de glucose,

unité : 1.10^3 molécules par seconde et par bactérie ;

axe Y : génotype, durée de vie, ou stabilité de la perméase ;

axe Z : génotype, durée de vie ou stabilité de la β -galactosidase.

Sur l'axe X l'intervalle 0,2 à 1,1 borne les possibilités de sélection.

Le lecteur peut simuler une sélection sur l'accroissement de la production de Glucose. La trajectoire de la réponse phénotypique à la sélection sur la surface des valeurs du phénotype dépend de la stratégie génétique de substitution des allèles dans les génotypes de la population. Par exemple, substitution d'allèles du gène de la perméase, puis de celui de la β -galactosidase. Une autre stratégie correspond à la situation inverse. Une troisième stratégie joue avec les allèles des deux gènes simultanément. Le lecteur doit retrouver des courbes de réponse du type de celles publiées par D.S. FALCONER. (1973, *Genet. Res., Camb.*) 22, 291-321.

The envelope values of the phenotype of the lactose operon of E. COLI as a function of two genetic variables.

X axis : phenotype, glucose production ; units : 1.10^3 molecules per second per bacterium ;

Y axis : genotype, life time or stability of permease ;

Z axis : genotype, life time of stability of β -galactosidase.

On the X axis selection may be exercised within the interval 0,2 to 1,1.

The reader can simulate a selection for the increase in the production of glucose. The course of the phenotypic response upon the envelope of phenotypic values depends upon the sequence of allelic substitution in the population. There may be an initial substitution of a permease allele followed by a β -galactosidase allele, or vice versa or simultaneous substitution of alleles of both the permease and β -galactosidase genes. The reader should discover phenotypic response curves of the type published by D.S. Falconer. (1973. Genet. Res., Camb.), 22, 291-321.

δ_i étant un coefficient d'identité et γ_i un moment d'une variable aléatoire 12 associée aux effets différentiels des gènes et définis par rapport à une population d'origine.

Ces résultats théoriques servent à construire des fonctions d'estimation, au sens statistique, de paramètres significatifs d'un caractère héritable dans une population. Les estimées de ces paramètres sont employées dans des fonctions prédictrices permettant de classer les reproducteurs contemporains par rapport à l'espérance des performances de leurs descendants. Ces fonctions de prédiction, ou de décision génétique, sont les clefs de voûte de tous les programmes de sélection animale ou végétale réalisés par les agronomes. Toutes les entreprises de sélection n'ont pas réussi, mais certaines ont réussi de façon exemplaire. Ces succès ne permettent guère de douter pour ces dernières de la pertinence de l'hypothèse de R. FISHER. Alors quels rapports a-t-elle avec les modèles dynamiques des phénotypes sous contrôle polygénique.

Revenons au modèle de l'opéron lactose de E. COLI. L'observation de la surface des points d'équilibres en fonction des paramètres génétiques nous donne des éléments de réponse.

Par exemple, si l'on considère la surface des équilibres repérée par trois axes, l'axe de la production de glucose, c'est-à-dire le phénotype, l'axe génique de la stabilité de la perméase, l'axe génique de la stabilité de la β -galactosidase, alors deux gènes n'ayant pas les mêmes repères sur les axes géniques n'assurent pas les mêmes valeurs de production de glucose. La génétique moléculaire et les modèles dynamiques associés apportent une justification à la notion d'effet quantitatif dû à la substitution d'un allèle par un autre introduite par R. FISHER. Cet effet peut être grand ou nul. Les possibilités de la sélection artificielle, mais aussi naturelle, d'un caractère héritable à déterminisme polygénique, dépendent de la grandeur de la projection sur l'axe phénotypique de la surface des équilibres phénotypiques. Si cette projection se réduit à un point ou à un intervalle étroit, le caractère héritable n'est pas sélectionnable. Les réponses à la sélection caractérisées par leurs vitesses et leurs limites ou plateaux de sélection traduisent les trajectoires de la valeur moyenne du caractère dans la population sur la surface des équilibres phénotypiques. Les formes de cette surface, dômes, cuvettes, cols, repliements, définissent les trajectoires possibles de la réponse à une sélection.

De nouvelles études de modèles dynamiques associés à des systèmes de contrôle polygénique, dont les descriptions biochimique et génétique seraient soigneusement réalisées, permettront de confirmer et d'affiner ces premiers résultats puis d'interpréter des notions comme la dominance ou l'épistasie.

IV. De la stabilité à l'invention héréditaire source d'évolution

A. La stabilité, ou la quasi-stabilité, des formes héréditaires au cours des générations successives est due à des mécanismes jouant aux différents niveaux de l'expression phénotypique. Les recombinaisons chromosomiques, les mutations, les duplications, les amplifications géniques, les déplacements des

transposons sont les causes autonomes des perturbations de la transmission de l'information génétique d'un ascendant à son descendant, d'une génération à la suivante. Cette notion de perturbation et celle de variation, envisagée au chapitre précédent méritent d'être comparées. Les variations sont des substitutions d'allèles ayant des niveaux différents de fonction, voire nuls. Les perturbations sont des variations, au sens précédent, et des modifications des nombres de gènes, de la nature de leurs fonctions et de leurs relations.

Quels sont les mécanismes de résistance ou d'inertie qui tempèrent l'effet de ces perturbations sur l'expression phénotypique et en assurent la régularité au cours du temps ? Pourquoi l'information génétique reste-t-elle stable ?

Quels sont les mécanismes d'invention qui biaisent avec l'inertie ou l'homéostasie des précédents et ouvrent aux génotypes l'exploration de nouvelles possibilités ? Comment l'information génétique s'enrichit-elle et se transforme-t-elle ?

B. Abordons les mécanismes de la résistance aux mutations qui assurent la stabilité. Le code génétique est dégénéré, de sorte que chaque acide aminé, ou chaque signe de ponctuation, est déterminé par une classe d'un ou plusieurs codons synonymes. L'existence de ces classes de synonymie traduit la redondance du codage considérée comme une protection contre les mutations qui échangent les codons entre eux.

A la suite des travaux de T. SONNEBORN (1965), l'idée d'une résistance du code génétique aux effets des mutations a conduit à la recherche de codes optimaux ; c'est-à-dire de trouver un code dont la définition des classes de synonymie minimise la probabilité pour qu'une mutation fasse sortir un codon de sa classe. Par exemple, pour un code doublet de n codons ayant un sens et de $(16-n)$ codons n'ayant pas de sens, sur 3 003 codes seulement 108 ont une résistance maximum (J.M. LABOUYGUES, 1981). Cette étude mériterait d'être poursuivie, d'une part dans le cas de codes triplets et d'autre part en tenant compte du rôle des acides aminés dans la structure secondaire des peptides.

Une autre approche est due à R. GRANTHAM (1975) qui définit le potentiel mutagène d'une classe de codons synonymes par le rapport c/n , où c désigne la fréquence d'utilisation d'un codon ayant seulement deux nucléotides communs avec la classe considérée (codons mutationnellement connexes) et n désigne la fréquence des codons non-connexes. Si f désigne la fréquence d'utilisation d'un codon de la classe considérée, \bar{f} la moyenne de f sur toutes les classes, alors l'équation $f = \bar{f} (c/n)$ est vérifiée en première approximation. Donc, une classe de codons est d'autant plus utilisée qu'elle est le lieu d'aboutissement d'un plus grand nombre de mutations.

D'autres chercheurs, comme R. JORRE (1974) et R. CURNOW (1975), T. WU, W. FITCH et E. MAR-GOLIASH (1974) et C. FUCHS (1980), ont abordé le problème de la résistance aux mutations sous un autre angle. Un modèle mathématique, une chaîne de Markov homogène, décrit l'évolution d'une séquence peptidique sous l'influence des mutations. Ce modèle permet le calcul, à partir de la matrice de transition, des temps de passage d'un acide aminé à un autre. Ce calcul a été appliqué à des protéines de synthèse qui diffèrent par un seul acide aminé soit de l'ocytocine, soit de la vasopressine. La classification de ces protéines par ordre croissant de leur activité biologique est négativement corrélée avec la classification par ordre croissant des temps des passages de l'acide aminé initial à l'acide aminé substitué. Les substitutions les plus longues à se réaliser entraînent la plus forte diminution d'activité biologique. C'est un nouvel argument en faveur de l'optimalité du code génétique et de sa résistance aux mutations.

Le décodage de la séquence des acides aminés en structures secondaires locales d'hélices α , de zig-zags, de boucles, de tournants β et γ assure aussi une résistance à l'effet des mutations. Chacun de ces motifs secondaires peut être réalisé par plusieurs séquences équivalentes d'acides aminés. Ainsi l'alanine, la leucine, l'acide glutamique sont fréquents dans les hélices α ; la méthionine, la valine, l'isoleucine prédominent dans les zig-zags ; la tyrosine, la proline, la glycine établissent les tournants. La substitution d'un acide aminé par un autre n'altère la séquence des motifs structuraux secondaires que dans des cas très rares. Généralement la nature et la fréquence de ces motifs structuraux secondaires

res sont difficilement altérées par les mutations. En conséquence, dans une espèce, ou un phylum d'espèces, s'accumulent, à partir d'une même séquence peptidique, des séquences variables réalisant toujours la même fonction avec des propriétés biophysiques identiques ou distinctes : vitesse de réaction, stéréospécificité, thermosensibilité, etc. L'essence de l'information génétique tient dans une séquence de motifs structuraux susceptibles de s'assembler en une structure d'ordre supérieur. Les changements que les mutations additionnent dans les séquences d'acides aminés au cours de l'écoulement des générations persistent dans la mesure où l'essence du message héréditaire, qu'est la fonction, n'est pas fortement altérée. Cette situation est décrite par le polymorphisme génétique des protéines.

Ces mécanismes de dégénérescences successives assurent la résistance aux perturbations d'une information génétique pour permettre son expression phénotypique conforme de génération en génération. Plusieurs séquences linéaires de trois bases nucléiques (ou codons) codent pour un même acide aminé, objets en trois dimensions ; plusieurs séquences linéaires d'acides aminés codent pour une même structure locale secondaire en trois dimensions ; plusieurs séquences de structures locales secondaires codent pour une même structure tertiaire en trois dimensions. L'assemblage des structures tertiaires engendre un système dynamique qui possède des propriétés de stabilité dans le temps.

C. Abordons maintenant les mécanismes qui exploitent les perturbations d'une information héréditaire pour inventer de nouvelles fonctions précédemment inconnues dans l'expression du génome. L'existence de plusieurs mécanismes, qui assurent conjointement la résistance aux mutations du message héréditaire, suggère que les processus évolutifs se développent en exploitant les possibilités offertes par une expression phénotypique normale. On peut invoquer par exemple le rôle des ambiguïtés de la stéréospécificité des macromolécules régulatrices et des enzymes.

Les enzymes ont une spécificité très précise pour leur substrat, mais ont aussi des capacités d'association non négligeables avec d'autres substrats soit toujours absents dans la cellule soit en concentration faible et à catabolisme innocuitaire. De très belles expériences d'évolution génétique, c'est-à-dire d'invention de nouvelles fonctions non codées dans l'ADN, ont été réalisées par R. MORTLOCK et WOOD (1977), E. LIN (1976), B. HARTLEY (1974), P. CLARKE (1974), B. HALL (1977) et J. KEMPER (1974). Présentons les grandes lignes de l'une d'entre elles. La bactérie *Klebsiella aerogenes* est équipée pour utiliser une grande variété de substrats de la famille des alcools polyhydriques : glycérol, ribitol, arabitol, mannitol, sorbitol. Naturellement cette bactérie est incapable d'utiliser le xylitol, un alcool polyhydrique qui n'existe pas dans la nature. Un traitement mutagène des bactéries suivi d'une sélection directe sur xylitol est strictement inopérant. Aucun mutant capable d'utiliser le xylitol n'apparaît. Par contre, sans aucun traitement mutagène préalable une sélection différentielle douce à nombreuses étapes successives permet d'isoler des bactéries capables d'utiliser le xylitol comme source exclusive de carbone. Les mécanismes mis en œuvre par la bactérie pour acquérir une nouvelle fonction héréditaire méritent l'attention.

La bactérie est mise sur un milieu contenant les alcools naturels ribitol et arabitol et un alcool inconnu le xylitol. L'expression phénotypique du génotype de la cellule est la synthèse de deux triades d'enzymes, la perméase, la déhydrogénase, la kinase, l'une spécifique du ribitol, l'autre de l'arabitol.

La première étape est l'exploration par la bactérie de la nouvelle frontière biochimique grâce aux ambiguïtés stéréospécifiques des enzymes synthétisées. Le xylitol pénètre dans la cellule en fraude en trompant la D-arabitol perméase. Les quelques molécules intracellulaires de xylitol sont oxydées par la D-ribitol-déhydrogénase qui les transforme en xylulose. Le xylulose est le produit normal de l'activité d'une autre enzyme la D-arabitol-déhydrogénase. Une fonction parasite s'est donc installée qui aboutit à un produit, le xylulose, normalement répertorié dans la mémoire héréditaire de la cellule. Cette fonction parasite acquiert alors une signification biologique sanctionnée dans cet exemple par un apport énergétique supplémentaire.

La seconde étape est la colonisation de ce nouvel espace biochimique assurée par les innovations génétiques. D'abord la régulation de l'enzyme D-ribitol-déhydrogénase est abolie. La synthèse de

l'enzyme devient constitutive à la suite d'une mutation de son gène de régulation. Cette synthèse ne peut plus être réprimée par le ribitol. Ensuite, l'enzyme est produite en plus grande quantité par amplification génique sans modification de sa faible spécificité pour le nouveau substrat, le xylitol. Cette amplification génique augmente le nombre de copies du même gène de structure de l'enzyme et de son gène de régulation dans le génome. Les mutations s'accumulent sur ces copies du même gène. Si certains gènes restent équivalents au gène d'origine, d'autres acquièrent une bien plus grande spécificité pour le nouveau substrat, le xylitol, enfin peut-être quelques-uns perdent-ils toute signification.

La troisième étape est la différenciation de plusieurs systèmes génétiques acquérant chacun une régulation propre. Le génome d'origine possédait deux systèmes génétiques régulés, l'un pour le ribitol, l'autre pour l'arabitol. Le génome nouveau possède trois systèmes génétiques régulés pour le ribitol, l'arabitol, le xylitol. Ce sont des mutations aux niveaux des gènes de régulation qui isolent ces trois systèmes génétiques contrôlant chacun une voie métabolique. Un nouveau système génétique est apparu à côté des deux systèmes initiaux restaurés.

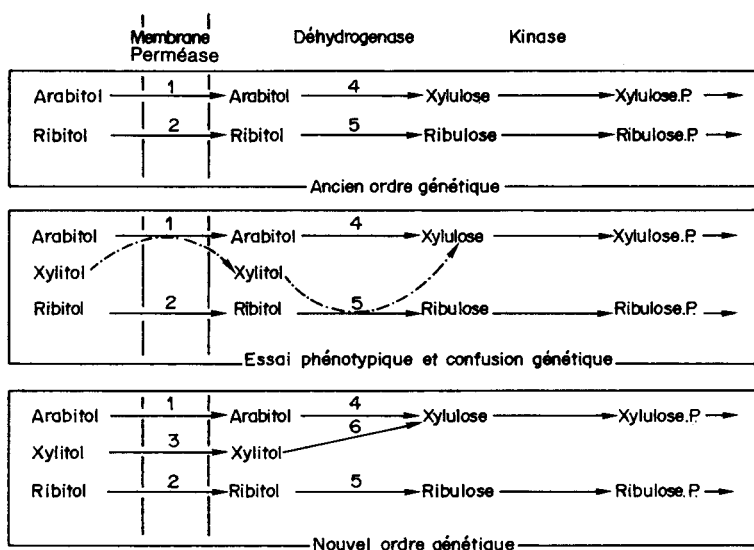


FIG. 7

Séquence schématique d'une invention héréditaire.

- 1 = Arabitol - perméase
- 2 = Ribitol - perméase
- 3 = Xylitol - perméase
- 4 = Arabitol - déshydrogénase
- 5 = Ribitol - déshydrogénase
- 6 = Xylitol - déshydrogénase

Sequence of steps involved in a genetic innovation.

- 1 = Arabitol - permease
- 2 = Ribitol - permease
- 3 = Xylitol - permease
- 4 = Arabitol - dehydrogenase
- 5 = Ribitol - dehydrogenase
- 6 = Xylitol - dehydrogenase

La stratégie qu'utilise le vivant pour augmenter sa mémoire héréditaire met en jeu trois mécanismes. Le premier est l'ambiguïté stéréospécifique de faible niveau de toutes les macromolécules codées génétiquement. Ces molécules assurent leur tâche essentielle avec une très forte probabilité et participent à des associations marginales avec une très faible probabilité. Ces événements marginaux aléatoires sont une sorte de bruit de fond phénotypique. Le second est l'apparition dans le bruit de fond phénotypique de séquences d'événements significatifs. Elles se greffent sur les réseaux des événements phénotypiques majeurs si elles portent une signification biologique, quelle que soit sa valeur sélective. L'évolution commence par une exploration phénotypique des possibilités de modifications ultérieures du message héréditaire. L'apparition du xylitol à côté du ribitol et de l'arabitol, puis le remplacement progressif du ribitol par le xylitol ont été les critères de choix parmi les possibilités qu'offrait la préadaptation génétique des ambiguïtés stéréospécifiques. Ce second mécanisme est celui de l'incertitude génétique où une fonction se construit en utilisant des macromolécules codées pour d'autres fonctions établies. Le troisième est l'acquisition par cette nouvelle séquence d'événements rares, mais qui représente déjà une nouvelle fonction, d'un contrôle génétique autonome et d'un statut normalisé dans le phénotype. C'est la phase d'évolution génétique vers un nouvel équilibre phénotypique stable.

L'exemple de *Klebsiella aerogenes* est confirmé par d'autres expériences. L'importance de ces exemples est d'insister sur le rôle de l'amplification génique sans laquelle les mutations, et la variabilité génique resteraient impuissantes. La stabilité des génomes, en dehors des périodes d'innovation génétique, laisse supposer que leur propension à l'extension est combattue. Dans le premier chapitre, il a été montré que l'amplification du gène de la perméase de l'opéron lactose pourrait contribuer à la perte des propriétés de stabilité de l'équilibre du système dynamique de fourniture de glucose. Cette perte jouerait contre le maintien de l'amplification puisque la cellule incorporerait le glucose moins régulièrement ce qui la défavoriserait vis-à-vis du génotype d'origine. Les propriétés structurales des systèmes dynamiques sous contrôle polygénique maintiendraient la stabilité de la nature des gènes et de leurs relations fonctionnelles.

L'exemple de l'évolution artificielle de *Klebsiella aerogenes* permet de cerner ce que nous croyons être le vrai problème de l'Évolution : à un stade donné quels sont les facteurs qui déterminent les possibilités et les limites d'un saut évolutif ? Le facteur principal est la possibilité de greffer des événements phénotypiques nouveaux sur un réseau d'événements phénotypiques existants. Aussi l'étude des propriétés des systèmes dynamiques à contrôle polygénique est-elle essentielle. On peut se demander si de tels mécanismes pourraient rendre compte de l'évolution physiologique et morphologique des Méta-zoaires. Avec prudence nous avancerons une réponse positive, car ces caractères sont sous un contrôle polygénique, qu'il reste à décrire il est vrai. Les différences que nous voyons sont d'une part que le rôle du xylitol est tenu par une macromolécule codée génétiquement, et d'autre part que les caractères nouveaux n'ont pas forcément une valeur sélective. Ce sont parfois des caractères potentiels en attente dont le coût n'est pas trop élevé dans l'expression phénotypique du génome.

V. Conclusion

Dans la conclusion les thèmes précédemment exposés seront utilisés d'abord pour justifier le maintien d'une distinction entre génétique et hérédité, ensuite pour discuter les thèses actuelles de l'Évolution.

A. La Génétique est communément définie comme la Science de l'Hérédité. Cette acception peut paraître insuffisante.

L'hérédité a un support matériel dont l'étude est la génétique et un logiciel dont l'étude est l'épigénétique. Deux arguments illustreront cette affirmation de la nature duale de l'hérédité. Le premier est l'histoire du déplacement de l'intérêt porté à l'hérédité vers la génétique. Le second est un survol des efforts accomplis pour établir une interprétation rationnelle et générale des interactions épistatiques.

A la fin du 19^e siècle l'un des plus importants problèmes biologiques était celui de la découverte des mécanismes de l'hérédité. Deux approches méthodologiques concurrentes ont réussi. L'une est celle de GALTON qui décrit les lois statistiques de la transmission des phénotypes, par exemple prévision de la valeur moyenne d'un caractère ou d'une performance d'un descendant connaissant la valeur d'un parent ou celles des deux parents. L'autre est celle de MENDEL qui décrit les lois probabilistes de la transmission des génotypes sous des conditions expérimentales très strictes dont l'une est le choix de la plus simple des relations possibles entre le génotype et son phénotype, l'autre l'existence de différences héréditaires très marquées. Dans le cadre de la méthodologie mendélienne, l'analyse probabiliste des faits de l'hérédité dévoile le portrait-robot d'objets inaccessibles à la description biologique, physiologique ou chimique de l'époque. Un nom est donné à ces objets invisibles, doués de mémoire héréditaire, ce sont les facteurs, d'où la naissance de la théorie de l'hérédité factorielle. Ce nom de facteur sera supplanté par celui de gène introduit par le danois JOHANNSEN.

La génétique s'est développée avec l'étude de la topologie, de la nature biochimique, de la fonction, de la réplication, de la variation et de la transmission du gène.

Au début du XX^e siècle le problème crucial était la propriété de généralité de cette théorie. Le test était sa capacité à rendre compte de l'hérédité des caractères quantitatifs (taille, poids, performance, forme). Ces caractères ne pouvaient pas être analysés à l'aide de la méthodologie mendélienne. Ce moment a été celui de la confrontation entre la théorie de l'hérédité « mélangée » ou « continue » et celle de l'hérédité factorielle.

R. FISHER clôt le débat en 1918, en postulant que les caractères quantitatifs héréditaires seraient gouvernés par un grand nombre de gènes, à ségrégation mendélienne, aux états indépendants et ayant des effets petits et de même ordre de grandeur. La théorie polygénique et mendélienne de l'hérédité quantitative était née. Son développement a donné la génétique mathématique des populations, les bases rationnelles de la sélection artificielle et l'interprétation synthétique de l'Évolution. La moisson des résultats acquis en génétique mathématique et génétique évolutive des populations est impressionnante, pourtant un malaise subsiste alimenté par les nombreux problèmes irrésolus. Par exemple, le mode de pensée mendélien impose l'énoncé des fréquences génotypiques comme description génétique d'une population, alors que la substance des changements évolutifs se réalise au niveau des phénotypes.

Cette situation a pour principale origine le maintien dans les raisonnements d'une relation génotype-phénotype trop restrictive, ne prenant en compte que la structure logique du gène et jamais la structure logique des relations fonctionnelles entre les gènes.

La totalité des interactions entre les gènes qui contrôlent le phénotype est l'Epigénotype. Ce mot a été créé pour rassembler un grand nombre de concepts des plus simples comme la dominance, l'épistasie à de plus riches comme la coadaptation génique de Dobzhanski, l'homéostasie génétique de Lerner, la canalisation de Rendel, les gènes régulateurs de Britten et Davidson, les gènes répresseurs de Glass et le paysage épigénétique de Waddington.

L'étude de l'epigénotype est bien loin d'être accomplie. Je pense que l'association de la génétique somatique, de la biologie moléculaire et des biomathématiques donnera un cadre expérimental à la réalisation de travaux équivalents à celui du modèle dynamique de l'opéron lactose que j'ai exposé.

Comme l'informatique, l'hérédité a deux composantes, l'une matérielle, l'autre logicielle. La composante matérielle, ce sont les éléments autoreproductibles, variables, signifiants que sont les gènes. La composante logicielle, ce sont les logiques de leurs relations fonctionnelles. Derrière la notion d'interactions épistatiques existent des systèmes logiques dont les types sont probablement en nombre limité,

les algèbres de leurs compositions à découvrir, la description mathématique de leurs propriétés dynamiques à entreprendre.

Le message héréditaire, transmis des ascendants à leurs descendants, est l'union du génotype et de l'épigénotype. Un gène, seul, n'a pas de sens, il n'en acquiert un que par rapport au système logique dans lequel sa fonction s'insère.

B. L'Évolution est un concept dont la signification englobe toutes les situations de changements décrites par la Paléontologie, l'Anatomie, l'Embryologie et la Physiologie comparées.

Ces disciplines ont analysé l'ampleur, les vitesses, les rythmes et les directions de ces changements. La situation la plus simple est celle des croissances des tailles dans la Microévolution. Ces croissances sont soit isométriques si les proportions respectives des organes sont conservés, soit allométriques si ces proportions sont modifiées. G.C. SIMPSON (1945) réunit sous le terme de Mégaévolution les situations les plus complexes caractérisées par l'invention des morphologies et des fonctions nouvelles.

Au cours du temps, et dans les différents segments des branches d'un phylum les vitesses d'évolution diffèrent. Une vitesse d'évolution extrêmement lente, pour ne pas dire parfois nulle, est la Bradytélie qui favorise le maintien des premiers représentants. Une vitesse moyenne d'évolution, caractéristique de la plupart des segments des branches d'un phylum est l'Orotélie qui accompagne la diversification d'un groupe. Enfin, en quelques rares de ces segments, des vitesses très élevées trahissent la précipitation temporaire de l'Évolution. O. ABEL (1911) et G.C. SIMPSON (1945) ont systématisé les directions de l'évolution en introduisant le principe de « l'inertie évolutive » ou tendance à évoluer dans une seule direction et le principe du « mouvement évolutif » ou tendance à poursuivre l'évolution dans une seule direction jusqu'à un extrême, souvent fatal. Le premier principe est l'orthogénèse, le second l'hypertélie. Enfin, la description des modalités de l'orthogénèse insiste sur les situations d'évolution divergente ou évolution éruptive de P.E. CLOUD (1948) et sur les sauts évolutifs associés à la tachytélie.

Les thèses actuelles de l'Évolution utilisent quatre notions essentielles : celle de changements graduels due à C. LYELL ; celle d'une relation univoque simple entre le gène (ou facteur) et le phénotype due à G. MENDEL, notion étendue aux caractères héréditaires quantitatifs par R. FISHER ; celle de mutations aléatoires à faibles fréquences des gènes due à TH. MORGAN ; et celle de sélection naturelle due à C. DARWIN.

La théorie sélective, sous sa forme renouvelée de « théorie » synthétique » de J. HUXLEY (1942) attribue au phénotype vis-à-vis du milieu et du génome une passivité et une rigidité qu'il n'a pas. L'essence de cette théorie est due à des modèles mathématiques concernant les fluctuations des fréquences des gènes dans les populations en fonction du nombre d'individus et du temps. Malheureusement ces modèles mathématiques ignorent les interactions fonctionnelles des gènes dans l'expression du phénotype et ne rendent pas compte de l'invention. En général, la sélection et l'évolution ne portent pas sur les états alléliques d'un ou de plusieurs gènes mais sur les propriétés des systèmes dynamiques à contrôle polygénique.

De ce même point de vue la thèse neutraliste, dont M. KIMURA (1969) est le champion, est insuffisante. Les dégénérescences des codes successifs nécessaires à l'expression des gènes et les propriétés de stabilité des systèmes autorisent l'accumulation des mutations. Cette accumulation quand elle est observable est une mesure du temps phylogénétique, et la preuve de l'existence des mécanismes de résistance aux mutations.

Principalement il manque à ces théories l'imagination des mécanismes possibles qui expliquent à la fois la stabilité héréditaire des phénotypes et l'invention génétique de fonctions et de morphologies inconnues précédemment, qui rendent compte des variations de vitesse de l'Évolution, et tout particulièrement de la simultanéité des accélérations de vitesse avec les éruptions d'inventions génétiques et phénotypiques.

Considérons l'invention de fonctions nouvelles. L'installation d'une nouvelle fonction utilise les ambiguïtés d'association à faible fréquence des macromolécules génétiquement codées : n'est-ce pas là au niveau moléculaire une notion proche de celle de la réadaptation des biologistes ? A tout moment dans un organisme l'ensemble de ces associations stéréospécifiques et de ces activités enzymatiques irrégulières forme le bruit de fond phénotypique. Le coût énergétique de ce bruit de fond phénotypique est d'autant plus grand que les probabilités des associations et des activités irrégulières sont elles-mêmes élevées. On comprend aisément que la croissance de ces probabilités vers l'unité abolit la continuité héréditaire. On serait tenté de penser que la décroissance de ces probabilités jusqu'à une valeur nulle représente la perfection de la continuité héréditaire au moindre coût énergétique et que les mutations favorisent l'évolution en affectant les gènes existants par échanges de leurs caractéristiques alléliques ou par augmentation de leur nombre. Cette conception des mécanismes géniques de l'Evolution est entre autres, celle de R. FISHER dans son hypothèse de la variation de la dominance ou bien celle de S. WRIGHT, de G. MALECOT et de M. KIMURA dans leurs études des lois de distributions des équilibres géniques en fonction du temps. Toutes les théories classiques de l'Evolution ignorent ou négligent l'existence et le rôle du bruit de fond phénotypique. Or c'est dans ce bruit de fond phénotypique qu'apparaissent des séquences d'événements porteurs d'une nouvelle signification biologique. L'installation d'une nouvelle signification biologique est entreprise, sans aucun changement du message génétique existant, par des essais phénotypiques qui créent un nouveau système dynamique épigénétique parasitant le génome et ne possédant pas encore son support polygénique propre et autonome.

La signification biologique est l'acquisition par un système dynamique épigénétique de la stabilité structurelle définie par les mathématiciens. La stabilité structurelle est la propriété de posséder une morphologie caractéristique et stable malgré les perturbations. D'un point de vue concret et biologique c'est une forme comme celles répertoriées par les anatomistes et les embryologistes, c'est une fonction comme le batttement cardiaque des physiologistes, ce sont les cinétiques d'une séquence d'activités enzymatiques, etc... Cette acquisition de la stabilité structurelle confère aux gènes codant et à leurs interrelations une identité de système polygénique. Ainsi la signification biologique ne se définit pas par rapport à une valeur adaptative mais par rapport à un ordre épigénétique.

A ce niveau de la discussion deux questions apparaissent :

La première : quels sont les facteurs qui feront que ces systèmes dynamiques épigénétiques, illégitimes, seront éliminés, seront tolérés, ou bien seront légitimés avec l'acquisition d'un contrôle polygénique autonome ?

La seconde : quel est le rôle des mutations dans l'essence et les propriétés du bruit de fond phénotypique ?

Tous les systèmes dynamiques épigénétiques nouveaux naissent sur un ensemble de gènes dispersés dans des systèmes polygéniques déjà existants. C'est pour ces gènes un pléiotropisme. Considérons un système dynamique ancestral avec son contrôle polygénique bien établi, et un système dynamique nouveau parasitant une partie des gènes du système dynamique ancestral. Ces deux systèmes se perturbent mutuellement, principalement au niveau des régulations. Le système dynamique ancestral est soumis à des contraintes découlant de son intégration dans l'individu par coadaptation avec les autres systèmes dynamiques ancestraux. La sélection naturelle, au sens darwinien, joue sur les génotypes des individus pour rétablir une régulation et un contrôle géniques autonomes, non perturbés, du système dynamique ancestral. Cette pression sélective a pour effet d'isoler à nouveau l'ensemble de contrôle polygénique du système dynamique ancestral du parasitisme qu'exerce le système dynamique nouveau. Ce sont les mécanismes moléculaires autonomes et aléatoires, comme l'amplification génique suivie par l'effet des mutations, qui permettent l'action de la sélection naturelle. La sélection naturelle et l'homéostasie favorisent les génotypes qui après amplification génique et mutations ont un contrôle polygénique restauré pour le système dynamique ancestral.

Quant au système dynamique nouveau son devenir est variable ; il comporte trois solutions. Il peut disparaître dans le bruit de fond phénotypique d'où il a émergé. Il peut se maintenir avec un

contrôle polygénique autonome plus ou moins élaboré. Ce maintien n'est possible que si le coût énergétique pour l'ensemble du génotype ou de l'individu n'est pas trop lourd. Beaucoup des caractères morphologiques recensés par les anatomistes ou les paléontologues comme n'ayant aucune valeur adaptative, mais répondant strictement aux contraintes de la stabilité structurelle, seraient des exemples de ce maintien. Enfin reste à discuter la dernière issue, celle pour laquelle le système dynamique nouveau, bien qu'imparfait, confère aux êtres qui le possèdent la possibilité d'entreprendre des expériences impossibles ou de supporter des situations fatales à d'autres génotypes. Au milieu trop sélectionneur des successeurs de C. DARWIN, s'oppose le système héréditaire affairé, entreprenant, puis conquérant ou failli. Un système héréditaire est affairé et entreprenant quand il dispose d'un système dynamique nouveau à expression phénotypique et contrôle génique imparfaits, quand il vient de quitter un ordre génétique ancien et qu'il tend vers un nouvel ordre génétique réalisable parmi plusieurs solutions possibles. N'est-ce pas le moment de l'accélération de l'Evolution, des phases de sauts évolutifs rapides de la Macroévolution et de la Mégaévolution ? La sélection naturelle n'est donc pas le *Deus ex Machina* de l'Evolution au sens si strict de A.R. WALLACE. Elle a un rôle plus modeste, elle ne crée rien, elle récapitule inlassablement les contraintes physiques qui orientent vers de meilleures adaptations un élément héréditaire nouveau qu'elle n'a pas sollicité, et qui apparaît dans le bruit de fond phénotypique. En fait, il faut considérer l'essence du bruit de fond phénotypique d'une part, son spectre et son ampleur d'autre part. L'existence du bruit de fond phénotypique découle des contraintes physiques, seuls le spectre et l'ampleur dépendent vraiment des mutations.

Il est important de souligner la notion de pléiotropisme des gènes. Ce n'est pas une curiosité de l'analyse génique à propos de l'ambiguïté de fonction des gènes, ou une illustration d'un principe d'économie. Le pléiotropisme des gènes est la marque et la conséquence de la diversification des systèmes polygéniques au cours de l'Evolution.

L'étude expérimentale des systèmes polygéniques de différenciation ou d'embryogenèse devient possible par le biais de la biologie moléculaire couplée à la Génétique Somaticque. Dans les années à venir les éléments de la thèse que nous avons développés pourront être testés.

Enfin nous réinsistons sur l'importance des systèmes dynamiques à contrôle polygénique car leurs propriétés structurelles imposent la continuité héréditaire, permettent et limitent les possibilités de l'Evolution. Les conditions de milieu sont des défis. Pour relever ces défis les systèmes polygéniques ont, ou n'ont pas, le temps et la possibilité d'élaborer des stratégies distinctes plus ou moins efficaces pour s'adapter ou conquérir.

Reçu le 29 septembre 1982.

Accepté le 23 février 1983.

Références

- O. ABEL, 1911. *Grundzüge der Palaeobiologie der Wirbeltiere*. Stuttgart.
- A. BABLOYANTZ et M. SANGIER, 1972. Chemical instabilities of « all or none » type in β -galactosidase induction and active transport, *FEBS letters*, **23**, 364-366.
- J.L. BETZ, P.R. BROWN, M.J. SMYTH et P.H. CLARKE, 1974. Evolution in Action. *Nature*, (247), 261-264.
- C. CHEVALET, 1971. Calcul a priori intra et inter - populations des variances et covariances génotypiques entre apparentés quelconques. *Ann. Génét. Sé. anim.*, **3**, 463-477.
- C. CHEVALET, M. GILLOIS, F. CORPET et A. MICALI, 1982. On a dynamical model of the lactose operon of *Escherichia Coli*. 2nd world conference on Mathematics at the service of man. (sous presse).
- Y. CHOU et G.D. FASMAN, 1974. Prediction of Protein Conformation. *Biochemistry*. **13** (2), 222-245.
- P.E. CLOUD, 1948. Some Problems and Patterns of Evolution exemplified by fossil invertebrates. *Evolution II*, **4**, 322-350.
- F. CORPET, C. CHEVALET, M. GILLOIS et A. MICALI, 1982, Dynamics and stability of induction of the lactose operon of *Escherichia coli*. *Lecture notes in Biomathematics*, J. Demongeot, ed., Springer-Verlag.(sous presse).

- J. DEMONGEOT, 1981. Existence de solutions périodiques pour une classe de systèmes différentiels gouvernant la cinétique de chaînes enzymatiques oscillantes : Application à la glycolyse, in : *Modèles Mathématiques en Biologie*, C. Chevalet et A. Micali ed., Lecture notes in Biomathematics, Springer Verlag, Berlin, **41**, 40-62.
- R. FISHER, 1918. The correlations between relatives on the supposition of mendelian inheritance. *Trans. roy. Soc. Edinburgh*, **52**, 399-433.
- C. FUCHS, 1980. On the distribution of the nucleotides in seven completely sequenced DNA's. *Gene*, **10**, 371-373.
- J. GARNIER *et al.*, 1978. Analysis of the Accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. mol. Biol.*, **120**, 97-120.
- M. GILLOIS, 1964. *La relation d'identité en génétique*. Thèse de doctorat d'État. Fac. Sci. Paris. 294 p.
- M. GILLOIS, 1966. La notion de génotype. *Ann. Génét.*, **10** (4), 201-202.
- M. GILLOIS, 1967. Les lois conjointes des variables aléatoires génétiques. *Ann. Génét.*, **10** (4), 203-206.
- M. GILLOIS, J. BOUFFETTE et A.R. BOUFFETTE, 1969. Etude des populations d'effectif limité homogames phénotypiques et panmictiques. *Ann. Inst. Henri Poincaré*, **B**, **5** (1), 69-86.
- M. GILLOIS, J. BOUFFETTE et A.R. BOUFFETTE, 1969. Covariance génotypique a priori dans les populations homogames. *Ann. Inst. Henri Poincaré*, **B**, **5** (1), 87-89.
- B.C. GOODWIN, 1963. *Temporal organization in cells*. Academic Press, London.
- B.C. GOODWIN, 1969. Control dynamics of β -galactosidase in relation to the bacterial cell cycle. *Eur. J. Biochem.*, **10**, 515-522.
- B. GOODWIN, 1978. A fied approach to developmental physiology. *Lecture notes in biomathematics*, C. Chevalet et A. Micali ed., **41**, 14-27.
- R. GRANTHAM, 1975. *Le code génétique : contribution à l'étude de son origine, sa structure et son fonctionnement. Vers un modèle moléculaire de l'évolution*. Thèse Doct. Sci. Université Claude-Bernard, Lyon I, 1-79 p.
- B.G. HALL et N.D. CLARKE, 1977. Regulation of newly evolved enzymes : IV, Evolution of the *ebg* Repressor during selection for enhanced lactase activity. *Genetics*, **85**, 193-201.
- H. HAKEN, 1978. *Synergetics, an introduction*. 2nd Edition - Springer-Verlag - Berlin.
- J. HIGGINS, 1967. Oscillating reactions. *And. Eng. Chem.*, **5**, 18-62.
- J.J. HOPFIELD, 1974. Kinetic Proofreading : A New Mechanism for Reducing Errors in Biosynthesis Processes Requiring High Specificity. **71**, (10), 4135-4139.
- J. HUXLEY, 1942. *Evolution, the modern synthesis*. Allen et Unwin, London.
- C.B. INDERLIED et R.P. MORTLOCK, 1977. Growth of *Klebsiella aerogenes* on xylitol : implications for bacterial enzyme evolution. *J. mol. Evol.*, **9**, 181-190.
- R.P. JORRE et R.N. CURNOW, 1975. The evolution of the genetic code. *Biochimie*, **57**, 1141-1146 ; 1147-1154.
- H. KACSER et J.A. BURNS, 1981. The control of flux. *Symp. Soc. exp. Biol.*, **27**, 65-104.
- J. KEMPER, 1974. Evolution of a new gene substituting for the leud gene of salmonella typhimurium origin and nature of supq and new mutations. *J. Bact.*, **120**, (3), 1176-1185.
- M. KIMURA, 1969. The rate of molecular evolution considered from the standpoint of population genetics. *Proc. natl. Acad. Sci.*, **63**, 1181-1188.
- W. KNORRE, 1973. Oscillations in the epigenetic system : biophysical model of the β -galactosidase control system, in : *Biological and Biochemical oscillators*. B. Chance, E. Pye, A. Ghosh, B. Hess Ed., Academic press, London, 449-457.
- J.M. LABOYNGUES et J.M. GOURGAND, 1981. Optimisation contre la survenue de terminateurs pour codes théoriques doublets. *C.R. Acad. Sci.*, **3**, **292**, (19), 1105-1108.
- E.C.C. LIN, A.J. HACKING et J. AGUILAR, 1976. Experimental models of acquisitive evolution. *Bioscience*, **26**, (9), 548-555.
- T. MANABE, 1981. Mathematical expression for the rate of initiation and certain prediction on the non-competitive Binding operon : The *E. Coli* lactose operon, *J. theor. Biol.*, **89**, 271-302.
- W. MANDECKI, 1979. A kinetic model for interaction of regulatory protein and RNA polymerase with the control region of the *lac* operon of *E. Coli*, *J. theor. Biol.*, **81**, 105-122.
- M. MORANGE, 1982. La révolution silencieuse de la biologie moléculaire : d'Avery à Hershey. *Le Débat*, **18**, 62-75.
- G. NICOLIS et I. PRIGOGINE, 1977. Self-organisation in nonequilibrium systems - *From dissipative structures to order through fluctuations*. Wiley-Interscience, New-York.
- J. NINIO, 1971. Codon-anticodon Recognition : The Missing Triplet Hypothesis. *J. mol. Biol.*, **56**, (1), 63-74.
- L. PAULING, R.B. COREY et H.R. BRANSON, 1951. The structure of proteins : two hydrogen-bonded Helical Configurations of the Polypeptide chain, *Proc. natl. Acad. Sci.*, **37**, (4), 205-211.
- L. PAULING et R.B. COREY, 1951. The pleated sheet, a new layer Configuration of polypeptide chains. *Proc. natl. Acad. Sci.*, **37**, (5), 251-256.
- L. PAULING et R.B. COREY, 1951. The structure of fibrous proteins of the collagen-gelatin group. **37**, (5), 272-281.
- J.R. PAULSON et U.K. LAEMMLI, 1977. The structure of histone-depleted metaphase chromosomes. *Cell*, **12** (3), 817-828.
- P.W.J. RIGBY, B.D. BURLEY et B.S. HARTLEY, 1974. Gene duplication in experimental enzyme evolution. *Nature*, (251), 200-204.

- B. ROBSON et R.H. PAIN, 1971. Analysis of the code Relating Sequence to conformation in proteins : possible implications for the mechanism of formation of helical regions. *J. mol. Biol.*, **58**, (1), 237-259.
- B. ROBSON et E. SUZUKI, 1976. Conformational properties of Amino Acid Residues in Globular Proteins. *J. mol. Biol.*, **107**, 327-356.
- R. ROSEN, 1970. *Dynamical system theory in biology*, Wiley-Interscience, New-York.
- R. ROSEN, 1972. Some relational cell models : the metabolism-repair systems, in : *Foundations of mathematical biology*, R. Rosen, ed., Academic Press, Vol II, 217-253.
- R. ROSEN ed., 1973. *Foundations of mathematical biology, subcellular systems, cellular systems, supercellular systems*. Academic Press, New-York.
- M. SANGLIER et G. NICOLIS, 1976. Sustained oscillations and threshold phenomena in an operon control circuit, *Biophysical Chemistry*, **4**, 113-121.
- E. SCHRÖDINGER, 1944, *What is life*. Cambridge University Press.
- E. SEL'KOV, 1968. Self oscillations in glycolysis I. A simple kinetic model, *Eur. J. Biochem*, **4**, 79-86.
- G.G. SIMPSON, 1945. *Tempo and mode of evolution*. N.Y., Colombia University Press. 1950, *Rythme et Modalités de l'Évolution* (trad. Saint Seine). Albin Michel.
- T. SONNEBORN, 1965. *Evolving genes and proteins*. V. Bryson, K. Vogel ed., Academic Press, 337-397.
- A.S. SPIRIN et L.S. ASATRYAN, 1976. The effect of ribosomal peptidyl transferase inhibitors is antagonized by elongation factor G with GTP. *FEBS letter*, **70**, 101-104.
- E. SUZUKI et B. ROBSON, 1976. Relationship between helix-coil transition parameters for synthetic polypeptides and helix conformation parameters for globular proteins. A simple model. *J. mol. Biol.*, **107**, 357-367.
- R. THOMAS, 1973. Boolean formalization of genetic control circuits, *J. theor. Biol.*, **42**, 563-585.
- R. THOM, 1972. Stabilité structurelle et morphogénèse : Essai d'une théorie générale des modèles. W.A. Benjamin, INC.
- P. VON HIPPEL, A. REVZIN, C.A. GROSS, A.P. WANG, 1974. Non-specific DNA binding of genome regulating proteins as a biological control mechanism : I. The *lac* operon : Equilibrium aspects, *Proc. natl Acad. Sci., USA*, **71**, 4808-4812.
- T.T. WU, W.H. FITCH et E. MARGOLIASH, 1974. The information content of protein aminoacid sequences. *A. Rev. Biochem.*, **43**, 539-566.