Thèse de Doctorat

de l'Institut national agronomique Paris Grignon

Présentée par Thomas Heams pour obtenir le titre de

Docteur de l'Institut national agronomique Paris Grignon

APPROCHE ENDODARWINIENNE DE LA VARIABILITE INTERCELLULAIRE DE L'EXPRESSION GENETIQUE

Thèse soutenue le 19 Mai 2004 devant le Jury composé de

Monsieur le Professeur Jean-Claude Meunier, Président du Jury

Madame le Docteur Nadine Peyrieras, Rapporteur

Monsieur le Professeur Jean-Claude Ameisen, Rapporteur

Monsieur le Professeur Jean Gayon, Examinateur

Monsieur le Professeur Andras Paldi, Examinateur

Monsieur le Professeur Carlos Sonnenschein, Examinateur

Monsieur le Docteur Jean-Jacques Kupiec, Examinateur et Directeur de Thèse.

Cette thèse est écrite à la mémoire de Dominique Heams.

Elle est dédiée à mes parents, Elisabeth et Houpa,

et à ma compagne, Célia.

TABLE DES MATIÈRES

TABLE DES MATIÈRES 3

TABLE DES FIGURES 7

REMERCIEMENTS 9

RESUME 12

INTRODUCTION 14

I OBJECTIF GENERAL ***15***
I. A UNE PROPOSITION THEORIQUE. 15

I. B QU'EST CE QUE LE DETERMINISME EN BIOLOGIE ? 22

I. C EXISTE-T-IL UNE CRISE DU DETERMINISME? 23
I. C 1 Un concept robuste 24
I. C 2 Des bases fragiles 27
I. C 2 a La spécificité 27
I. C 2 b Le génocentrisme. 34

II. ÉNONCE DE L'HYPOTHESE ENDODARWINIENNE. ***42***
II. A UN PARALLELE HISTORIQUE 42

II. B L'HYPOTHESE ENDODARWINIENNE 44

II. C NECESSITE POUR L'HYPOTHESE ENDODARWINIENNE D'EXPLIQUER LA REPRODUCTIBILITE DES PHENOMENES. 45
II. C 1 Un paradoxe apparent 45
II. C 2 Les arguments pour résoudre ce paradoxe. 46
II. C 2 a La stabilisation limite les choix 46
II. C 2 b Un phénomène aléatoire est statistiquement reproductible 47
II. C 2 c Transmissibilité d'un profil d'expression aléatoire 48

II. D RELATION ENTRE PARADIGME DETERMINISTE ET PARADIGME PROBABILISTE. 50
II. D 1 Un déterminisme rationalisé tolère une dimension stochastique. 50
II. D 2 Le paradigme probabiliste englobe le paradigme déterministe 53

APPROCHE BIBLIOGRAPHIQUE ***56***

I LES OBSERVATIONS EXPERIMENTALES DE VARIABILITE NON GENETIQUE ET DE STOCHASTICITE DE L'EXPRESSION GENETIQUE ***58***
I. A VARIETE DES PHENOMENES BIOCHIMIQUES STOCHASTIQUES 58
I. A 1 Large présence des phénomènes stochastiques 58
I. A. 2 De nombreux systèmes où une variabilité non génétique est observée 60
I. A. 2 a Variabilité chez les procaryotes 60
I. A. 2 b Variabilité intercellulaire dans les voies de différenciation 61
I. A. 2 c Hétérogénéité intercellulaire d'expression génétique dans un contexte homogène 62

I A. 3 Variabilité et stabilisation 63
I. A 3 a Des précurseurs 63
I. A 3 b Deux théories majeures 64
-La théorie de la sélection clonale. 64
-La théorie de la stabilisation sélective des neurones 66
I. A 3 c D'autres systèmes reposant sur une dynamique de hasard/sélection 67

I. B L'EXPRESSION STOCHASTIQUE DES GENES 69
I. B 1 Bases moléculaires de l'expression stochastique des gènes. 71
I. B 1 a La haute dynamicité des échanges moléculaires 71
I. B 1 b La dimension stochastique apparaît avec l'effet d'échantillonnage. 74
I. B 1 c La structure des boucles de régulation peut conduire à une expression stochastique. 77
I. B 2 Le contrôle du bruit 80
I. B 4 Génération de variété phénotypique. 82

I. C. LES LIMITES DES DYNAMIQUES ENDODARWINIENNES. 86
I. C 1 Des limites déjà atteintes ? 86
I. C 2 Des limites prévisibles 88

II. UNE LECTURE NON DETERMINISTE DU DEVELOPPEMENT. ***90***
II. A INTRODUCTION 90

II. B LE DETERMINISME EN EMBRYOLOGIE : UN CHOIX D'OBSERVATION CONTESTABLE 92
II. B 1 Le déterminisme apparent des expériences fondatrices pose des difficultés d'interprétation 92
II. B 2 Un programme dégénéré et non spécifique est-il un cadre explicatif satisfaisant ? 95
II. B 3 Les gènes du développement, nouvel avatar du déterminisme ? 98
II. B 4 La plasticité des devenirs cellulaires pose un problème à l'interprétation programmatique 101
-Dédifférenciation-Transdifférenciation 101
- Le cas des cellules de la crête neurale 105

II. C UNE RELATION TROPHIQUE AU CŒUR DU DEVELOPPEMENT 107
II. C 1 Les sous entendus de la métaphore informationnelle 107
II. C 2 La relation trophique au cœur des échanges intercellulaires 108
- Portée et limite d'une métaphore de substitution 108
- Relation trophique et vascularisation 112
- Relation trophique et apoptose 113

II.D DES SURPRODUCTIONS TRANSITOIRES. 114

II E DES GREFFES QUI DEVOILENT DES COOPERATIONS CELLULAIRES SUR UNE BASE PLUS FONCTIONNELLE QUE GENETIQUE 117

II F CONCLUSION 119

III L'APOPTOSE CONTREDIT-ELLE LES DYNAMIQUES ENDODARWINIENNES ? ***121***
III. A INTRODUCTION 121

III B LA TYPOLOGIE DES MORTS CELLULAIRES A DES CONTOURS FLOUS QUI METTENT PARTIELLEMENT EN CAUSE LA SPECIFICITE DES PROCESSUS APOPTOTIQUES 122
III. B 1 La nécrose 122
III. B 2 L'apoptose 123

III. B 3 Des formes intermédiaires présentent des caractéristiques mixtes 125

III. C LA NON-SPECIFICITE DES MOLECULES IMPLIQUEES DANS L'APOPTOSE 128
III. C 1 Certains acteurs de l'apoptose ne sont pas toujours indispensables. 128
III. C 2 Des molécules antagonistes ont des structures très proches. 132
III. C 3 L'apoptose comme une rupture d'équilibre 133
III. C 4 Les déclencheurs de l'apoptose sont souvent des " signaux par défaut " 134

III. D UNE APPROCHE ECOSYSTEMIQUE PEUT FOURNIR DES CLEFS DE REINTERPRETATION DE LA MORT CELLULAIRE. 136

III. E L'APOPTOSE CHEZ LES UNICELLULAIRES. 139
III. E 1 Le paradoxe de l'apoptose chez les organismes unicellulaires 139
III. E 2 Deux modèles rendent intelligible une apoptose chez les unicellulaires. 141
-Le premier modèle est celui dit de l'addiction. 141
-Le second modèle est celui du "péché originel" 143

III. F CONCLUSION 144

APPROCHE EXPERIMENTALE 146

I INTRODUCTION : UNE PREDICTION DU MODELE ENDODARWINIEN NECESSITE D'ETRE TESTEE PAR UNE APPROCHE SUR CELLULES ISOLEES. 147

II. LES TECHNIQUES D'ETUDE DE L'EXPRESSION GENETIQUE SUR CELLULES ISOLEES 149
II. A CONTEXTE DE L'APPARITION DE CES TECHNIQUES 149

II. B DEUX GRANDS TYPES DE METHODES POUR ANALYSER LES VARIATIONS INTERCELLULAIRES 150
II. B 1 Les méthodes permettant de visualiser l'expression génétique par des marqueurs 150
II. B 2 Les méthodes reposant sur l'amplification du matériel génétique sur cellules isolées comportent plusieurs variantes. 153
Les méthodes reposant sur l'amplification linéaire de l'ARN (aARN). 153
-Les méthodes reposant sur des adaptations de la RT-PCR. 154

II. C UNE MODIFICATION DU PROTOCOLE DE TPEA **158**

II. D CONCLUSION SUR LES ETUDES D'EXPRESSION GENETIQUE SUR CELLULES ISOLEES. 165

III L'ETUDE DE LA VARIABILITE EPIGENETIQUE AU COURS DE LA DIFFERENCIATION. 167
III A LA PROBLEMATIQUE DE L'ETUDE DE LA METHYLATION 167
III B METHODOLOGIE 169
III C RESULTATS 172
III D DISCUSSION 175

DISCUSSION 179

BIBLIOGRAPHIE 190

ANNEXES 221

TABLE DES FIGURES

Figure 1 : Le déterminisme suppose une spécificité de réponse à un signal.....Face **p.22**
Figure 2 : Modèle de l'Opéron Lactose chez *E. coli*. .....Face **p. 25**
Figure 3 : Modèle de Britten et Davidson (simplifié). .....Face **p. 26**
Figure 4 : Gradients morphogènes et information de position. .....Face **p. 39**
Figure 5 : L'hypothèse endodarwinienne. .....Face **p. 44**
Figure 6 : Stabilisation d'un profil aléatoire..... Face **p. 47**
Figure 7 : Variabilité du positionnement des chromosomes. .....Face **p. 49**
Figure 8 : Le déterminisme, cas particulier du probabilisme. .....Face **p. 54**
Figure 9 : Technique de FRAP. .....Face **p. 71**
Figure 10 : Équilibre bistable. .....Face **p. 77**
Figure 11 : Limite des dynamiques probabilistes. .....Face **p. 88**
Figure 12 : Expérience de Spemann démontrant l'existence d'un organisateur. Face **p.92**
Figure 13 : Exemple de greffe interespèces. ....Face **p. 106**
Figure 14 : Schéma simplifié des deux grandes voies d'apoptose..... Face **p. 124**
Figure 15 : La famille des protéines Bcl-2. .....Face **p. 132**
Figure 16 : Le module d'addiction. .... Face **p. 141**
Figure 17 : Une prédiction de l'hypothèse endodarwinienne. ..... Face **p. 147**
Figure 18 : l'amplification ARN. ....Face **p. 153**
Figure 19 : La RT-PCR. .... Face **p. 154**
Figure 20 : Le principe du TPEA. .... Face **p. 158**
Figure 21 : Les simplifications du protocole de TPEA. .... Face **p. 159**
Figure 22 : Résultats d'amplification pour trois gènes par TPEA modifié..... Face p. **163**
Figure 23 : Contrôle de TPEA. ....Face p. **164**
Figure 24 : Comparaison entre les formules de la cytosine non méthylée et méthylée des îlots CpG. .... Face p. **167**
Figure 25 : Différence d'action du Bisulfite de Sodium sur la cytosine et la méthyl-cytosine. .... Face p. **169**
Figure 26 : Principe du traitement des cellules au bisulfite. .... Face p. **170**
Figure 27 : Structure de la zone de méthylation différentielle amplifiée.....Face p. **171**
Figure 28 : Evolution du taux de méthylation au cours d'une cinétique de différenciation - Cellules non différenciées (cinétique témoin). .... Face p. **172**
Figure 29 : Evolution du taux de méthylation au cours d'une cinétique de différenciation - Cellules différenciées. ....Face p. **173**
Figure 30 : Evolution du taux de méthylation au cours d'une cinétique de différenciation - Cellules mortes. .... Face p.**173**

Remerciements

Mes remerciements vont tout d'abord conjointement à Jean-Jacques Kupiec et Pierre Sonigo. Ma gratitude envers eux est immense. Il m'ont proposé ce sujet périlleux mais passionnant. Ils m'ont fait constamment confiance, tout en me laissant autonome. J'ai eu plaisir à voir s'effondrer nombre de mes certitudes en biologie à leur contact, non par effet de séduction, mais parce qu'ils ont toujours su avec moi accepter le débat et la contradiction, qui donnent de la force aux audaces. Jean-Jacques a dirigé ma thèse dans cet esprit, avec cœur, engagement et conviction, mais m'a aussi laissé me l'approprier. Sa relecture minutieuse de mon manuscrit l'a considérablement amélioré. Pierre a toujours été attentif à ma place dans son équipe, je l'ai toujours senti à mon écoute, et lui aussi m'a énormément nourri.
Leur empreinte scientifique sur ce travail est, bien sûr, prépondérante.

Je remercie les membres du Jury d'avoir accepté d'évaluer ce travail. Parmi eux, j'ai une reconnaissance toute particulière pour Jean-Claude Meunier, qui a accepté que ce projet fasse l'objet d'un travail de thèse et qui m'a soutenu pendant ce parcours.

Ma sympathie la plus grande va à mon équipe, le Laboratoire de Génétique des Virus de l'Institut Cochin, qui a accueilli en son sein mes recherches, avec une curiosité et un intérêt permanents. Je n'aurai bien sûr pas pu transformer ce projet en travail quotidien sans l'alchimie particulière de ce groupe, composé d'individus entiers, généreux et passionnants. Je voudrais ici saluer tout particulièrement mes deux compagnons de thèse : mon désormais vieil ami Olivier Delelis et ma si chère complice Audrey Brussel, ainsi que Thierry Leste-Lasserre, l'indispensable.
Je n'oublie bien sûr pas Caroline Petit, et aussi Isabelle Tardieux et France Pietri-Rouxel, dont l'expérience humaine et scientifique m'ont permis de persévérer.
Mais plus généralement, c'est à tout le laboratoire que je suis infiniment reconnaissant.

Je remercie l'ensemble de ceux qui, à l'institut Cochin, ont rendu mes recherches possibles.

J'ai une dette évidente envers les membres du "groupe du Vendredi". Derrière cette appellation mystérieuse se cache une réunion informelle que j'ai eu l'honneur d'organiser pendant plusieurs mois, autour des hypothèses probabilistes en biologie,

et qui a suscité l'intérêt et l'enthousiasme de chercheurs à l'esprit ouvert, qui m'ont ainsi nourri de manière décisive. Je les remercie tous, et leur promets de reprendre bientôt ces réunions !
Plus généralement, mes discussions nombreuses avec des scientifiques que mon approche séduisait ou rebutait ont toutes contribué à façonner ma réflexion, et je leur en suis reconnaissant.

Mais une thèse ne se fait pas qu'entre les quatre murs d'un laboratoire. Elle est un bloc de vie intense et parfois solitaire où le soutien de ses proches est indispensable. Les sautes d'humeur imprévisibles du thésard font, de plus, que ce soutien n'est pas toujours simple à témoigner.
Je veux dire ici qu'il ne m'a jamais manqué. Mes amis ont toujours été là. Mais plus encore, sans la patience et les encouragements permanents de ma compagne Célia, et le soutien total et tendre de mes parents, Elisabeth et Houpa, rien de ce qui suit n'aurait pu voir le jour. Ma reconnaissance est indicible. J'ai tenté de transformer tout cela en un travail digne. J'espère que, bien que modeste, il pourra témoigner, là où les mots me manquent, que les concessions faites n'ont pas été vaines, et que derrière la rigueur à laquelle j'ai tenté de ne pas déroger, c'est bien une passion dense et secrète qui m'a permis d'aboutir. A tous, j'espère pouvoir la transmettre en retour. Ce serait ma manière à moi de leur rendre ce trésor que fut leur présence à mes côtés.

RESUME

La biologie moléculaire repose sur des bases déterministes qui ont longtemps été fécondes mais sont désormais un handicap pour la compréhension des phénomènes biologiques. Le déterminisme en génétique est un concept robuste mais dont les fondations apparaissent de plus en plus fragiles, notamment le concept de spécificité et l'importance centrale accordée au génome.
À partir de ces constats, nous faisons une proposition théorique. Nous proposons une nouvelle grille de lecture des relations intercellulaires, que nous appelons endodarwinienne. Il s'agit de postuler que les cellules ont un comportement aléatoire *a priori*, notamment dans leur expression génétique et que seules celles qui adoptent un profil d'expression adapté à leur environnement sont dans un second temps sélectionnées. Dans ce modèle, les cellules ne répondent pas à des "signaux" de manière prévisible. Dans une large mesure, les cellules sont mues par leur intérêt immédiat. Ce modèle est une réponse à la crise du déterminisme car il rend inutile de postuler une spécificité des interactions moléculaires, et permet de sortir de la notion rigide de "programme génétique", sans pour autant nier le rôle évident des gènes. Il y a un paradoxe à proposer que des phénomènes reproductibles soient fondés sur des dynamiques stochastiques, c'est pourquoi nous cherchons à montrer que ce paradoxe n'est qu'apparent, que l'aléa de réactions individuelles est précisément une force structurante, et que l'approche probabiliste englobe l'approche déterministe.
Une synthèse de données expérimentales met en évidence l'importance sous-évaluée de la variabilité intercellulaire de l'expression génétique, dans des situations très diverses. Dans celles-ci, de toute évidence, des cellules clonales se comportent de manière aléatoire et ne répondent pas de manière homogène à des signaux. Cette variabilité peut être la base d'une sélection. Par ailleurs, des données récentes confirment que la dynamique même des échanges moléculaires au sein du noyau rend nécessaire d'intégrer comme un acquis la dimension intrinsèquement stochastique de l'expression génétique, qui provoque ainsi une variété phénotypique.
Le développement embryonnaire, classiquement envisagé comme la concrétisation d'un programme génétique à l'œuvre, peut être abordé selon une logique endodarwinienne. Le déterminisme y est très relatif, les devenirs cellulaires sont très plastiques, et les signaux échangés par les cellules semblent souvent pouvoir être étudiés sous l'angle d'une relation trophique.
L'apoptose, souvent appelée mort programmée, ne semble pas pourtant répondre de manière satisfaisante à cette caractérisation. Cependant il peut sembler contradictoire, dans l'approche endodarwinienne, que des cellules engagent leur propre destruction. Des observations montrent que malgré les apparences, ce phénomène avéré ne contredit pas l'approche endodarwinienne.
Pour tester les hypothèses endodarwiniennes, il faut en passer par une étude sur cellules isolées car travailler sur des populations cellulaires provoque des effets de moyenne qui masquent la variabilité intercellulaire. Des techniques d'études de l'expression génétique sur cellules isolées existent. Parmi celles-ci, une adaptation particulière de la RT-PCR, l'amplification par l'extrémité 3' (TPEA) a été simplifiée et rationalisée pour être compatible avec une approche globale. Les premiers résultats issus de ce travail de mise au point, quoique préliminaires, peuvent déjà être lus selon une grille sélective plutôt qu'instructive et confirment l'existence d'une expression aléatoire des gènes.
Une autre approche possible est celle de l'étude des variations de la méthylation des gènes, que l'on sait corrélée à des variations d'expression. Au cours d'une cinétique de différenciation, un gène est étudié sous cet angle et révèle une variabilité importante, dans le temps et de cellule à cellule. Par ailleurs, le profil de différentes sous populations au cours de cette cinétique va dans le sens d'une expression aléatoire et d'une sélection des cellules ayant le profil adapté.
Ces approches expérimentales et bibliographiques incitent à penser que la proposition théorique initiale est pertinente, économe et généralisable, même si, bien sûr, la question reste ouverte.

INTRODUCTION

I Objectif général

I. A Une proposition théorique.

Ce travail veut contribuer à promouvoir une grille de lecture renouvelée des phénomènes biologiques. Il s'intègre dans une histoire des sciences récentes, faite de propositions alternatives à un modèle dominant que l'on caractérisera sous l'appellation de paradigme déterministe.

L'ambition de ce travail n'est pas de nier les apports effectifs dont la biologie moléculaire a bénéficié au sein du cadre déterministe dans lequel elle s'est développée. Il soulève cependant des arguments tendant à accréditer la nécessité d'une vision plus globale, en redonnant une place décisive aux phénomènes stochastiques, notamment dans le cas de l'expression génétique, non pas comme du bruit de fond, comme des paramètres marginaux, mais au contraire en donnant à cette stochasticité une place prépondérante, nécessaire à l'émergence de phénomènes organisés et en apparence déterminés.

Il est tout à fait légitime qu'une discipline scientifique naissante, telle la biologie moléculaire, adopte pour accompagner son essor des modèles simplifiés, permettant de comprendre en première approximation les relations entre les différentes molécules impliquées dans l'expression génétique, principalement les acides nucléiques et les protéines. Néanmoins, il nous semble évident, face à l'évolution fulgurante de cette discipline, que la complexité croissante des réseaux de relations

entre gènes, des niveaux de régulation de la transcription et de l'expression, rendent insatisfaisante une approche dont le fondement est déterministe.

Qu'entend-on par déterminisme, en biologie, et notamment en génétique ? Le terme fut forgé en 1865 par Claude Bernard, qui le qualifie de "premier principe des sciences expérimentales" et en explique la force dans l'*Introduction à l'étude de la médecine expérimentale* : "Il faut admettre comme axiome expérimental que dans les êtres vivants aussi bien que dans les corps bruts les conditions de tout phénomène sont déterminées d'une manière absolue […] La négation de cette proposition ne serait autre chose que la négation de la science même" (Bernard, 1865, *in* Lecourt 1999). En génétique, cela se décline par l'axiome, dans un premier temps caricatural, que tout phénomène a une cause génétique, ces causes expliquant la reproductibilité des phénomènes biologiques dans des conditions similaires (chez tous les embryons, les organes se forment au même moment, tout type cellulaire soumis à une drogue y répondra de manière prévue). Jean Gayon propose la définition suivante : "le déterminisme en tant que doctrine générale, c'est la conviction que l'aspect génétique est essentiel dans tous les domaines biologiques et prime finalement sur tous les autres" (Gayon, 2003). Il souligne d'ailleurs qu'il y a là un amalgame avec le réductionnisme génétique (nous reviendrons sur cette distinction). La conséquence logique de cet axiome déterministe est, en génétique, l'instauration de la métaphore du "programme génétique", souvent confondu avec l'ADN, parfois disjoint de lui (Atlan, 1972). À cet égard, il ne faut jamais perdre de vue que programme signifie étymologiquement "écrit à l'avance", et que les versions "dures" ou "douces" de la métaphore programmatique comportent donc toutes une dimension de prédestination, qui pose un problème au rationalisme scientifique.

Pour utile que soit cette métaphore programmatique, et elle le fut effectivement, elle possède comme toute métaphore des avantages et des limites. Et s'il est devenu courant depuis quinze ans de la remettre en cause régulièrement, ou plutôt un aspect de celle-ci, s'il est de plus en plus convenu de considérer cette approche comme une première approximation, il n'en demeure pas moins qu'elle reste le cadre global de la pensée génétique contemporaine. Le déterminisme génétique a fait l'objet de nombreuses contestations dans les années 1970, mais celles-ci n'ont pas permis une remise en cause fondamentale et l'ont, au contraire, renforcées. Nous reviendrons sur ces débats. Conditionnée par un développement technique parallèle, la biologie moléculaire semble être maintenant traversée de résultats contradictoires. Nous pensons que l'interprétation de ces résultats sous le seul angle du déterminisme -ou réalisation automatisée d'un programme-, quelles que soient les nuances apportées, est de nature à mettre en péril l'ossature de la discipline et ce qu'on est en droit d'attendre d'elle, c'est-à-dire un minimum de force prédictive.

La biologie moléculaire contemporaine repose sur une connaissance détaillée des cascades de régulations, enchaînements de causes et d'effets élémentaires. Ces connaissances sont par définitions provisoires, car inévitablement limitées par l'ignorance temporaire que l'on a des effets de tout autre molécule sur l'ensemble de celles qui sont étudiées. La biologie moléculaire évolue vers une exploration systématique de l'ensemble des relations possibles entre toutes les molécules. Tendre vers l'exhaustivité devient nécessaire, car il peut suffire d'une molécule pour perturber la connaissance d'un système donné. Or l'exhaustivité demeure une limite inaccessible. Plus important, cette logique confine la biologie moléculaire à une

démarche descriptive et non pas explicative des processus du vivant. Nous développerons cette idée.

Il s'agit donc d'une panne conceptuelle. Cette panne repose sur l'axiome de la spécificité d'action des molécules (ou des sites actifs). Cet axiome, utile et souvent expérimentalement vérifié en première approximation, peut aussi devenir un handicap, ce que nous soulignerons, à une compréhension réaliste des phénomènes biologiques. Nous défendrons l'idée dans cette thèse que le concept de spécificité n'a pas de bases solides en biologie moléculaire, même s'il peut être parfois fonctionnellement utile. Nous proposerons, au cours de ce travail, un modèle qui permet de tenir compte des affinités relatives que peuvent avoir les molécules entre elles (parfois si fortes qu'elles peuvent prendre l'apparence d'une spécificité fonctionnelle), mais qui ne fait pas reposer seulement sur ces affinités les relations causales entre les événements élémentaires. Car il apparaît que, plus qu'un problème de connaissance ou d'ignorance de mécanismes précis, c'est un problème d'échelle d'observation qui explique la panne que beaucoup ressentent.

De plus, même au niveau moléculaire, un corpus cohérent d'observations récentes nous semblent de nature à remettre en cause la possibilité matérielle d'un fonctionnement déterministe du vivant.

Cette panne conceptuelle doit être abordée de front, peut-être au prix d'un saut paradigmatique. Ne pas explorer d'hypothèses alternatives, comme par exemple celles qui structurent ce travail c'est, nous semble-t-il, se résigner, à terme, à transformer la biologie moléculaire en une vaste opération d'empilement de données sans cohérence interne, au péril de la scientificité de l'ensemble.

Cette conviction n'est bien sur pas nouvelle. Le constat de la nécessité d'intégrer une dimension aléatoire dans les phénomènes inter ou sub-cellulaire, comme nous allons le proposer, est une évidence pour beaucoup, même s'il est parfois difficile de savoir si cette préoccupation n'est pas plutôt une manière de sauver le cœur de la pensée déterministe en acceptant, à la périphérie, des phénomènes parasites. Or nous pensons par ailleurs qu'il est illusoire de chercher à sauver cet édifice conceptuel en n'accordant qu'une place marginale aux phénomènes non programmés. En effet, pourquoi, alors que l'accumulation de données mettant en péril la vision déterministe ne cesse de prendre de l'ampleur, ne pas envisager que c'est le cœur même de cette grille de lecture qui est à revoir ? Pourquoi se cantonner à lui adjoindre quelques codicilles de stochasticité dans l'espoir d'en sauver l'essentiel ? Notre thèse est qu'il est difficile de renoncer à la pensée déterministe pour au moins trois raisons principales.

Premièrement, parce qu'elle a, au moins en apparence, l'immense intérêt d'expliquer la très remarquable propriété de reproductibilité des phénomènes génétiques. On peut ici illustrer cette vision par quelques exemples. Si chaque individu, à un stade donné de son développement, se met à développer tel organe, ou tel type cellulaire, c'est bien parce qu'un programme réalise de manière contrôlée cette opération. De même, du moins pour des caractères simples, les livres de génétique regorgent d'exemples où une mutation génétique modifie un caractère phénotypique, laissant entendre qu'il y a bien une relation déterministe entre le déterminant et le caractère (nous en citerons). Or nous chercherons à démontrer que cette reproductibilité ne saurait reposer uniquement sur une grande précision, une grande fiabilité des régulations, mais aussi et surtout sur leur caractère dégénéré.

Deuxièmement, parce que le déterminisme génétique a résisté à de nombreux niveaux de critiques grâce auquel il s'est progressivement nuancé en rajoutant des hypothèses *ad hoc*, ce qui l'a renforcé. En effet, ces nuances que nous détaillerons lui permettent de fournir des explications à de nombreux problèmes auparavant insolubles. Mais elles pourraient aussi contribuer à l'affaiblir, lui faisant courir le péril de pouvoir expliquer tout et son contraire, au risque de sa propre cohérence.
Troisièmement, parce qu'une critique du déterminisme génétique relativise le rôle central attribué à l'ADN dans les processus cellulaires. La critique de l'importance classiquement accordée à l'ADN n'est pas inédite, mais, dans ses formes diverses, elle va rarement de pair avec une critique radicale du déterminisme génétique. Nous évoquerons ces critiques. Nous défendrons l'idée que cette relativisation est légitime.

Nous argumenterons ici, sans rien remettre en cause des phénomènes observés ni de leur reproductibilité (les tissus se développent effectivement à des stades donnés, il y a effectivement un lien entre des mutations génétiques et des variations phénotypiques), qu'une autre grille de lecture est possible, voire qu'il est nécessaire de renoncer à la précédente, parce qu'elle est construite sur des amalgames et des imprécisions. Il y a dans les approches classiques une discordance d'échelle entre les phénomènes observés et leur interprétation moléculaire.

Nous proposerons dans ce travail un paradigme probabiliste, fondant les relations entre les molécules du vivant. Dans le cadre de ce paradigme, nous proposerons l'existence de dynamiques que nous appellerons *endodarwiniennes*. D'après ce nouveau paradigme, l'expression stochastique et différentielle d'un génome commun par des cellules clonales permet

l'adaptation d'une fraction d'entre et conséquemment leur multiplication et leur sélection, ayant pour résultat l'organisation du tissu.

Nous chercherons à démontrer que cette proposition apporte des réponses nouvelles aux critiques faites historiquement au déterminisme génétique.

Le but de cette thèse est donc d'expliciter ce que serait cette approche nouvelle faisant une part prépondérante aux phénomènes élémentaires aléatoires dans la génération d'ordre global. Elle fournira des éléments bibliographiques souvent épars, mais qui, rassemblés, semblent plaider pour la pertinence et la parcimonie de ce modèle. Elle présentera enfin des résultats expérimentaux augurant la possibilité de mesurer leur importance au sein du vivant. Cette proposition se veut testable et réfutable, et répond donc aux critères de l'investigation en sciences expérimentales. Elle pourrait déboucher sur une méthode généralisable en biologie moléculaire et cellulaire.

Elle s'intéresse plus particulièrement, sur le plan expérimental, à une étude de la différenciation cellulaire. Ce phénomène largement étudié et dont une chronologie sera présentée constitue une discipline vaste. Elle s'est développée au sein du paradigme déterministe. Nous chercherons à démontrer qu'à l'inverse, elle est toute indiquée pour tester des hypothèses probabilistes. C'est pourquoi nous avons concentré notre approche expérimentale sur une étude de la différenciation in vitro de lignées cellulaires clonales, modèles de la différenciation neuronale.

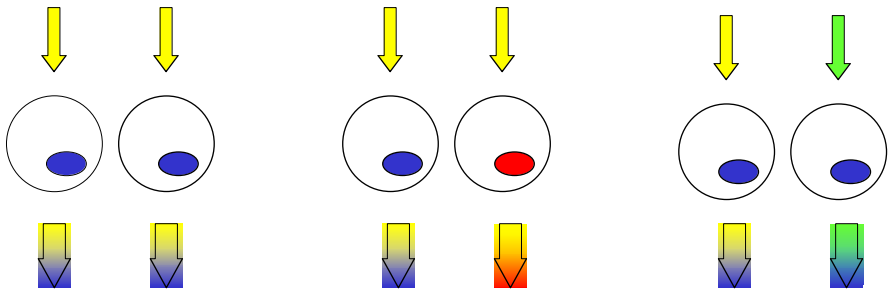


Figure 1 : Le déterminisme suppose une spécificité de réponse à un signal.
A. Deux cellules de même génome (ovale bleu) répondront de manière identique à un signal identique (flèche entrante jaune). Ce signal peut être constitué d'une/plusieurs molécule(s) ou d'une variation de paramètres extérieurs à la cellule ou d'une combinaison. Dans le cadre déterministe, la variabilité de réponse peut être due (B.) à une différence génétique et/ou (C.) à une variation de signal.

Supposer et tester un comportement aléatoire intercellulaire nécessite de mettre à jour l'hétérogénéité de la population. On montrera que pour cela, une analyse à l'échelle de la cellule individuelle (suivie d'une approche statistique) est à bien des égards indispensable.

Le but de ce travail, en remettant en cause ce lien logique et en mettant en relation une approche probabiliste de la différenciation cellulaire avec une étude de cellules isolées est de contribuer à promouvoir cette hypothèse.

I. B Qu'est ce que le déterminisme en biologie ?

Par modèle déterministe de l'expression génétique, nous caractérisons la vision classique selon laquelle les cellules reçoivent chacune des signaux émis par l'environnement ou par d'autres cellules, intègrent ces signaux et y réagissent en engageant une cascade d'expression génétique aboutissant à une réponse appropriée à ces signaux, finement régulée. La nature de cette réponse repose sur l'information génétique contenue dans le génome (Figure 1). Une population homogène de cellules de génomes identiques répondra à ces signaux de telle sorte que le profil moyen d'expression sera le reflet fidèle du profil effectif d'expression de chacune des cellules.

Dans sa version simplifiée, ci-dessus, ce modèle est évidemment rigide. L'utilisation de ce modèle l'est bien sur beaucoup moins : il a été tempéré par de nombreuses observations qui ont mis en évidence un bruit de fond dans l'expression génétique, c'est-à-dire une variation intercellulaire marginale. Il a gagné en souplesse au fur et à mesure de l'avancée des connaissances sur les mécanismes cellulaires. Néanmoins,

il conserve sa structure globale. En effet, pour expliquer l'apparition d'une hétérogénéité, par exemple la formation d'un nouveau type cellulaire au sein d'une population homogène, il suppose l'émission et la réception d'un signal, suivis d'une réponse à ce signal. Ce faisant, il présuppose une asymétrie antérieure. Il y a donc une contradiction puisque ce qui est à démontrer et ce qui est supposé se confondent. De plus, cette seconde asymétrie doit alors être expliquée, selon la même logique. De proche en proche, chaque asymétrie ainsi repose sur une asymétrie "préalable". On peut donc considérer que c'est un modèle de création d'ordre par l'ordre, puisque chaque état, supposé stable ou ordonné, repose sur un état ordonné préalable. La question de l'apparition initiale de l'ordre est sans cesse repoussée en amont.

I. C Existe-t-il une crise du déterminisme?

Nous allons tenter ici de convaincre que le déterminisme génétique est en crise. Faire ce diagnostic suppose de la mesure et de la méthode. Il ne peut se résumer à quelques exemples opportunément choisis (même si on y fera appel), le silence étant fait sur les exemples inverses.

Nous partirons donc des constats suivants :

-La critique du déterminisme génétique n'est pas une idée neuve en biologie.

-Les critiques historiques du déterminisme ont principalement porté sur sa rigidité excessive due à ses hypothèses initiales simplificatrices.

-Le paradigme déterministe tel que défini plus haut, a su intégrer un certain nombre de ces critiques et a gagné en souplesse et en complexité.

C'est sur la base de ces constats, que nous faisons nôtres, que nous chercherons cependant à montrer qu'une crise latente persiste. En effet, nous pensons qu'il est illusoire de chercher à rendre plus souple une pensée rigide par essence et que la biologie moléculaire telle que nous la connaissons permet de changer radicalement de paradigme, de gagner en simplicité, et d'évacuer des hypothèses non nécessaires à la compréhension des phénomènes biologiques.

I. C 1 Un concept robuste

Le déterminisme génétique est constitutif de la biologie moléculaire, même s'il lui préexiste. Par abus de langage (on peut considérer que cette discipline naît avant 1953), nous entendons dans ce travail par *biologie moléculaire* la discipline qui, à la suite de la découverte de la structure de l'ADN (Watson et Crick, 1953), a entrepris d'expliquer au niveau moléculaire les phénomènes impliquant la maintenance, la duplication et l'expression du génome. La découverte de James Watson et Francis Crick est une avancée fondamentale qui a permis de donner une explication mécanique et moléculaire à l'édifice de la théorie de l'évolution et plus généralement à la compréhension du monde vivant dans son unité. Trouver simultanément dans une structure moléculaire le support matériel des gènes et l'explication de leur mutabilité et de leur transmissibilité était sans conteste un pas de géant pour les évolutionnistes de toutes disciplines. Cette influence ne s'est pas restreinte au phénomène de l'hérédité mais progressivement à celui de l'expression des gènes, notamment par le modèle de l'opéron lactose de François Jacob et Jacques Monod (Jacob et Monod, 1961). Décrit chez la bactérie pour la production de β-galactosidase, ce modèle propose l'existence de deux types de gènes, les gènes

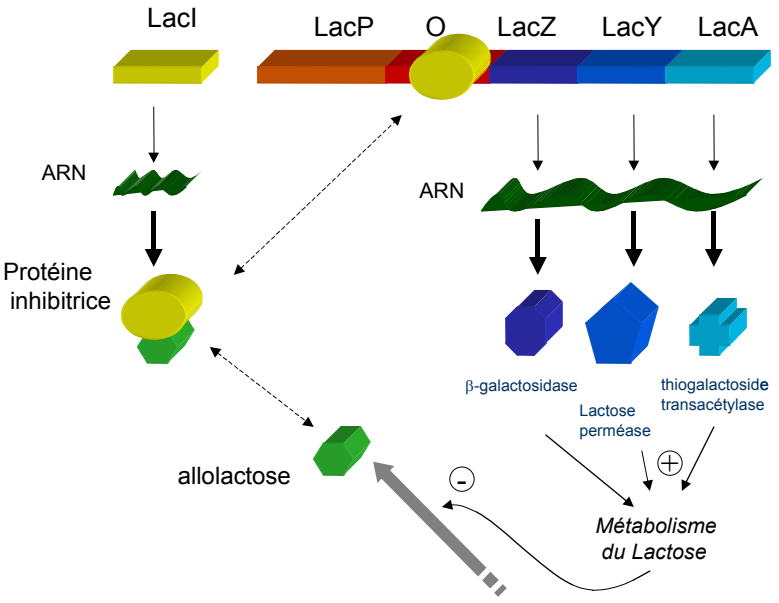


Figure 2 : Modèle de l'Opéron Lactose chez *E. coli*.
Deux types de gènes sont définis : ceux qui codent pour des molécules régulatrices (ici LacI code pour le répresseur de l'opéron) et ceux qui codent pour les protéines de structure et les enzymes (Ici LacZ, LacY et LacA). Ce modèle repose sur la spécificité d'interaction entre protéines et séquences d'ADN, ainsi que sur une stabilité de ces interactions.

structuraux qui codent pour les protéines et les enzymes, et les gènes régulateurs, codant pour les protéines régulatrices (Figure 2). Ce modèle est instructif au sens où il réduit l'ensemble de ces réactions à une mécanistique d'enchaînements élémentaires qui sont chacun la réponse adéquate à une instruction antérieure (appelée aussi information ou signal). La machinerie cellulaire s'apparente formellement à un programme informatique (l'essor de l'informatique, contemporain de la biologie moléculaire, n'est certainement pas anecdotique dans cette contamination sémantique) répondant à des commandes.

Le modèle de Monod et Jacob a donné lieu à différents modèles généraux de régulation de la transcription, notamment celui de Britten et Davidson (Britten et Davidson, 1969). Dans celui-ci, la combinatoire des différentes protéines codées par les gènes régulateurs permet d'obtenir une variété d'associations des ces protéines ; chacune de ces associations permet la régulation spécifique d'un gène (Figure 3, page suivante). L'ensemble constitue ainsi un modèle de régulation du génome entier reposant sur la spécificité. Ce dernier est aussi l'ancêtre d'un modèle de référence, ayant vocation à expliquer l'ensemble du développement par des cascades intégrées de réponses à des signaux, proposé plus de trente ans plus tard par l'un de ces deux auteurs (Davidson, *et al.*, 2002). Ainsi, ce cadre de pensée a été, soulignons-le, d'une immense fécondité pendant cette période. La possibilité de discrétiser l'ensemble des phénomènes biologiques en successions de régulations génétiques les a rendus intelligibles en termes moléculaires, et dans certains cas simples de maladies monofactorielles, a pu permettre de comprendre les relations de causalité avec une mutation génétique identifiée. De plus, cette avancée intellectuelle fut renforcée par au moins trois développements technologiques majeurs : le génie génétique, permettant de couper et/ou insérer différents éléments

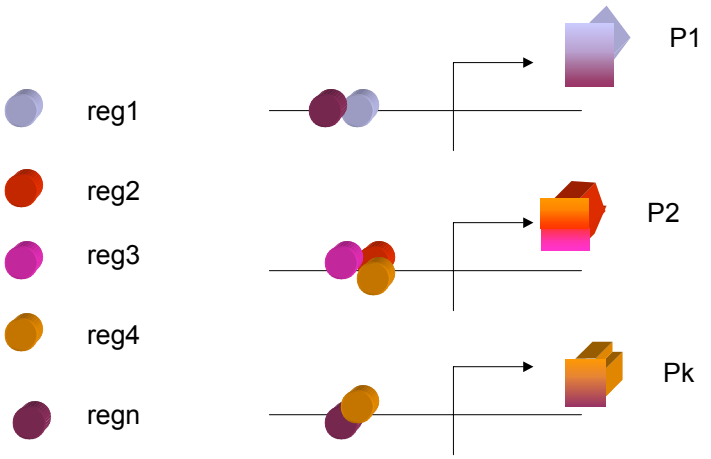


Figure 3 : Modèle de Britten et Davidson (simplifié).
Les protéines régulatrices (reg1 à regn) s'associent pour contrôler l'expression des autres gènes, codant pour les protéines P1 à Pk. Les associations entre régulateurs sont spécifiques du gène à contrôler et la combinatoire permet d'expliquer le contrôle général de l'expression génétique.

d'ADN dans des organismes, le séquençage de l'ADN permettant de caractériser le contenu génétique de n'importe quel organisme, et la réaction en chaîne à la polymérase (PCR) qui permet de détecter par amplification (et de réutiliser) une séquence d'acide nucléique donnée. Ces trois percées, qui reposent sur les propriétés structurelles propres de la molécule d'ADN, ont renforcé, par la force des résultats produits, la place centrale que l'on accorde au génome. Mais si féconds que soient ces débouchés, il reste pertinent de s'interroger sur leurs limites et sur le biais qu'ils imposent à la perception que l'on se fait des interactions intra- et intercellulaires.

Comme le souligne Michel Morange "la régulation [moléculaire] est devenue le fait biologique par excellence". Il note que dès 1961 " F. Jacob et J. Monod suggérèrent que les modèles qu'ils avaient proposés étaient susceptibles d'expliquer les variations d'activité géniques qui se produisaient au cours du développement embryonnaire et de réconcilier génétique et embryologie" (Jacob et Monod, 1961, cité in Morange, 1999). Néanmoins les critiques affluèrent : il était simpliste et peu transposable aux organismes "supérieurs", la régulation négative n'est pas le seul mode de régulation. De plus, de nouveaux niveaux de complexité, découverts progressivement, n'étaient pas pris en compte : épissage des ARN, méthylation/acétylation de la chromatine (nous y reviendrons), structure tridimensionnelle de la chromatine, ces niveaux de complexité étant progressivement regroupés sous le terme d'épigénétique comme si la génétique avait soudainement besoin de données auxiliaires pour garder ses capacités explicatives. Le génome lui-même s'avéra plus complexe, avec l'existence de séquences morcelées ou répétées, de réarrangement chromosomiques, avec aussi la proportion écrasante des séquences non codantes. Le modèle initial ne pouvait certes pas se voir reprocher

de ne pas tenir compte de phénomènes inconnus à sa conception ! Et de fait, il intégra progressivement ces données, auxquelles chacun s'accorda à n'attribuer qu'une influence marginale sur le cœur du modèle : le fait que la régulation de l'expression des gènes se fasse principalement au niveau de la transcription. Alors que certains proposaient même que le développement soit "illogique" (Lewin, 1984), la découverte des gènes homéotiques, conservés structuralement et fonctionnellement, confirma que les propositions de Jacob et Monod avaient été des "anticipations remarquables de ce que l'analyse moléculaire des gènes du développement révélait". Et de fait, les " modèles de régulation ont donc assez peu évolué durant ces trentes dernières années. On peut, sans simplifier indûment dire que la biologie moléculaire vit encore aujourd'hui dans le cadre du paradigme moléculaire établi dans les années 1960" (Morange, 1999).

I. C 2 Des bases fragiles

I. C 2 a La spécificité

Mais à trop se renforcer de nuances parfois contradictoires, n'est-on pas en droit de penser que l'édifice conceptuel est vacillant ? On peut penser que oui. En effet, quelles étaient les prédictions du modèle déterministe ? Dans le cadre du modèle de Britten et Davidson, les gènes de régulation codent pour des facteurs qui, par diverses combinaisons, contrôlent l'expression de plusieurs ensembles de gènes. Ces modèles supposent l'existence de régulateurs spécifiques, c'est-à-dire portant chacun une information correspondant à une commande du programme. Ceux-ci, ou du moins des combinaisons spécifiques de tels régulateurs de l'expression génétique, expliqueraient la différenciation des cellules en différents types cellulaires spécifiques.

C'est la recherche de ces régulateurs spécifiques qui a mobilisé pendant des années les efforts des chercheurs mais ce programme de recherche aboutit à un résultat spectaculaire : les régulateurs spécifiques n'existent pas en tant que tels et toutes les molécules régulatrices sont extrêmement communes, présentes et partagées dans les différents types cellulaires (Duboule et Wilkins, 1998). On verra que l'on peut rapprocher cette constatation d'un acquis important de l'embryologie moléculaire, à savoir que les inducteurs du développement sont non spécifiques et eux aussi très partagés (Dieterlen-Lièvre, *et al.*, 1996). Les gènes homéotiques, ceux-là même qui sont très conservés au sein du vivant, de la drosophile aux mammifères, et qui ont paru venir au secours du déterminisme génétique, sont des régulateurs non spécifiques susceptibles de reconnaître des séquences de six nucléotides extrêmement présentes et disséminées dans tout le génome (Pellerin, *et al.*, 1994). Or ils ne se fixent, in vivo, que sur un nombre restreint de séquences. Pour comprendre ce fonctionnement, dans la logique déterministe, il faut faire appel à des cofacteurs, ce qui revient à recommencer la recherche de spécificité et accepter que ces régulateurs ne sont pas la cause d'un programme, mais les partenaires moléculaires d'un ensemble complexe de régulations. Sans causes, la pensée déterministe est orpheline, et l'on peut soit en chercher de nouvelles, soit chercher à penser autrement. Cette non-spécificité a conduit à parler de bricolage évolutif (Duboule et Wilkins, 1998), pour expliquer que c'était cette non-spécificité qui permettait de dérouler un programme d'une grande adaptativité : mais ne s'agit-il pas d'une astuce sémantique, puisqu'en toute logique, un programme est constitué d'étapes prévues à l'avance ? Ce n'est pas le cas ici : pourquoi donc persister à penser qu'il existe un programme ?

Comment dans le cadre d'une acception non pas stricte, mais seulement rigoureuse de la notion de programme, accepter la possibilité de reprogrammation de certaines cellules par dédifférenciation/transdifférenciation ? Nous reviendrons sur ces aspects, mais il semble ici qu'ils illustrent le risque tautologique inhérent à la métaphore programmatique en biologie : d'une part, il existerait des phénomènes programmés, et d'autre part ceux qui ne le sont pas, parce qu'ils sont la conséquences d'aléas et d'imprévus, peuvent être considérés comme reprogrammés. Tout ressortirait donc d'un programme, même l'imprévu, c'est-à-dire même son contraire. Cette explication totalisante ne nous paraît pas compatible avec la réfutabilité scientifique.
On aura compris, encore ici, et on tentera de le démontrer avec plus de détails plus bas, qu 'il ne s'agit bien évidemment pas de remettre en cause l'existence de ces phénomènes avérés, mais de proposer de les comprendre sous un autre angle.

On voit que la notion de programme est intimement liée à celle de spécificité, qui trouve son origine dans le modèle moléculaire de stéréospécificité, parfois abusivement présenté comme un modèle "clef-serrure". C'est sur cette notion que reposent les modèles de régulations vus plus haut. La stéréospécificité est, en toute rigueur, une relation exclusive entre deux structures tridimensionnelles permettant leur association à l'exclusion de toute autre. La spécificité n'est pas une notion relative. C'est une notion absolue, non quantitative, et donc périlleuse à manipuler en sciences expérimentales. Elle est pourtant largement utilisée tant au niveau moléculaire (spécificité des interactions) qu'au niveau cellulaire (spécificité des types cellulaires). Elle s'illustre parfaitement dans le cas de maladies monofactorielles, qui furent d'ailleurs des modèles de grand intérêt pour comprendre les principes de bases de l'expression génétique. Dans un certain nombre de cas limités, en effet, la

mutation d'une base dans la séquence d'un gène peut aboutir à un phénotype malade, comme l'anémie falciforme. Cette maladie est due au fait que la molécule d'hémoglobine, mutée en un point précis de sa séquence, induit une conformation tridimensionnelle différente de cette protéine et modifie ces propriétés de fixation de l'oxygène. Mais cet exemple canonique, comme d'autres, ne doit pas faire oublier que ces cas de liaison spécifique entre un seul gène et un caractère sont rares. Que les exemples de ce type aient permis de comprendre les bases du fonctionnement des gènes est indéniable, mais ils constituent les exceptions de relations beaucoup plus complexes. En outre, ils peuvent occasionner une erreur logique. On prétend parfois démontrer la spécificité en rappelant que chez tous les malades présentant une pathologie "génétique" caractérisée, on trouve effectivement la même mutation. Mais il est alors indispensable de rappeler qu'il s'agit ici d'un rapport d'inclusion et non pas d'équivalence : si tous les malades portent la mutation, cela ne signifie pas pour autant que tous les porteurs de mutations sont malades. La structure des systèmes de santé est telle que, pour des raisons économiques évidentes, on ne fait pas la recherche, pour toute une population, de l'ensemble des maladies rares liées à un allèle dont ils pourraient être les porteurs. On effectue ces recherches dans certaines populations à risques, ou dans les familles dont un membre est malade. Ce qui ne signifie pas que tous ceux qui ne sont pas surveillés et ne présentant pas la pathologie ne sont pas porteurs de ce gène. Cet argument est-il spécieux ? S'il est certes parfaitement logique, on pourrait cependant objecter que les cas de porteurs non atteints sont rares. Et qu'il est peu pertinent de raisonner sur des phénomènes marginaux. Or de telles évaluations existent et ont été recensées (Wolf, 1997). Il apparaît, pour les rares études qui ont été entreprises, que les résultats confirment de manière spectaculaire la variété phénotypique qu'il peut y avoir dans des

populations portant les mêmes mutations connues pour être liés (pour "être la cause") de maladies "spécifiques". La spécificité n'est donc qu'apparente, biaisée par le fait que l'on n'observe que les "malades" et pas la partie immergée de l'iceberg des porteurs non atteints.

Raisonnons provisoirement sur un exemple. Nous avons dit la prudence qu'il fallait réserver à ce type d'approche, nous ne le perdons pas de vue. Mais n'oublions pas non plus que la biologie moléculaire assoit de manière générale ses démonstrations sur des exemples. Pour pouvoir la critiquer sur certain de ses aspects, il nous faut aussi utiliser sa logique propre. Nous nous intéresserons au cas de l'interleukine 2 (IL-2), une cytokine qui agit comme un facteur de croissance pour l'expansion clonale des lymphocytes T. Initialement décrite dans ce système, cette cytokine a pu un temps en paraître spécifique. Mais au fur et à mesure de la progression des connaissances, celle-ci fut identifiée comme un partenaire de la progression dans le cycle cellulaire des lymphocytes T (passage de la phase G1 à S), dans la maturation des thymocytes à différents stades chez la souris, dans la différenciation des cellules autres que les lymphocytes T, dans l'augmentation de la synthèse des immunoglobulines, et de la transcription de la chaîne J des lymphocytes B, dans l'activation de l'activité cytotoxique des cellules NK, dans l'induction des cellules killer induites par les lymphokines, dans la promotion de la prolifération, la différenciation et l'activité cytolytique des macrophages, dans l'activité tumoricide anticorps-induite des monocytes. Elle est de plus importante pour le système immunitaire en général, comme l'ont montré des études sur des souris KO. On comprend qu'un défaut de sécrétion d'IL-2 soit associé, chez l'homme, a des problèmes immunologiques. Par ailleurs, IL-2 est impliquée à des degrés divers au processus d'apoptoses. Enfin, IL-2

est décrite comme un facteur trophique des cellules du stroma du microenvironnement du thymus (Hollander, 1999).

Cet exemple illustre bien que la spécificité d'action d'IL-2 est toute relative, bien qu'on ait légitimement pu penser, initialement, que cette cytokine soit spécifique. Cela montre que la spécificité reflète l'ignorance bien compréhensible que l'on a des différents rôles d'une molécule tant qu'on ne l'a pas étudiée sous tous les angles. L'article décrivant cette multiplicité de fonctions en vient, dans un premier temps, à la conclusion que la molécule IL-2 doit être finement régulée pour permettre de jouer ses rôles. On peut rester sceptique face à cette proposition. Cette multifonctionnalité ne traduit-elle pas plutôt qu'IL-2 est un partenaire de ces régulations, sans en être la cause ? C'est ce à quoi semble se résoudre l'auteur, qui finit par plaider pour une régulation stochastique de son expression, la démontrer (y compris son expression monoallélique, elle-même stochastique), et la rapprocher de plusieurs situations comparables.

Une autre manière non satisfaisante de postuler la spécificité d'un régulateur pour une fonction est de le croire à tort, indispensable. Une approche classique en biologie moléculaire est de chercher à corréler l'expression d'un gène à une fonction particulière. Quand cette corrélation est bonne, on le déclare spécifique de cette fonction. Mais cela peut conduire à de nombreuses surprises. C'est le cas du récepteur G-CSFR au facteur de stimulation de colonies de Granulocytes (G-CSF). Le nom de ce récepteur laisse entendre qu'il est une voie nécessaire de transduction du signal aboutissant à la différenciation granulocytaire, et pourtant, on a montré récemment que le remplacement du domaine spécifique de ce récepteur, ou sa délétion, n'empêchait pas la différenciation granulocytaire (Jacob, *et al.*, 1998). Cela

ne signifie pas que G-CSFR ne joue aucun rôle dans cette différenciation, mais souligne qu'en cas de carence, les cellules semblent capables de recruter d'autres voies pour assurer la fonction.

De même, les expériences de *knock-out* fonctionnels où des gènes d'intérêt sont sélectivement inactivés ont donné des résultats surprenants. Dans la majorité des cas, les animaux présentent effectivement un phénotype modifié, voire létal. Mais dans des proportions non négligeables de cas (jusqu'à 30 %), le phénotype est peu ou pas changé (Melton, 1994), même en l'absence de gènes considérés comme indispensables, notamment la myoglobine (Garry, 1998). De toute évidence, des mécanismes de recrutements, d'autres gènes sont à l'œuvre pour suppléer cette absence (Greenspan, 2001 ; Edelman et Gally, 2001).

Ces exemples mettent en évidence qu'il peut être parfois vain de regarder de trop près une cascade de régulation, voire de la décrire parfaitement, si d'autres voies alternatives fonctionnent (peut être pas par défaut) tout aussi bien. La biologie moléculaire a pour but, à son niveau d'analyse, de nous fournir les explications de phénomènes cellulaires. Il semble qu'elle en soit souvent réduite à ne nous décrire que des cascades de régulations inévitablement partielles, plus ou moins arbitrairement choisies, quitte à faire l'impasse sur d'autres, et donc ne pas répondre à son objet initial.

I. C 2 b Le génocentrisme.

Un des aspects de la contestation du déterminisme est la critique du génocentrisme. Nous nous inscrivons dans cette critique, même si nous ne la confondons pas avec la critique du déterminisme génétique lui-même. Le déterminisme préexiste à la connaissance de l'ADN.

Nous nommerons ici génocentrisme l'importance démesurée accordée au génome dans les phénomènes vivants, et nous nous restreindrons strictement à cette définition. Puisque l'ADN est le support des gènes, il est "l'espace des causes" : "Le génome contient non seulement une série de plans de construction, mais aussi un programme coordonné de synthèse des protéines, ainsi que les moyens de contrôler son exécution" (Jacob et Monod, 1961). Pousser cette logique à l'extrême conduit à donner une place à part, centrale, organisatrice, au génome dans les phénomènes vivants. Richard Dawkins (Dawkins, 1976) a poussé cette centralité du génome dans les processus biologiques jusque dans ses extrêmes retranchements, puisqu'il va jusqu'à envisager les individus comme des artefacts inventés par les génomes pour se reproduire. Cette thèse, dont la dimension provocatrice est évidemment à prendre en compte, a été féconde puisqu'elle a fourni un cadre de recherche important, et fut utile pour l'avancée d'une discipline matérialiste et le recul des visions anthropocentriques. Néanmoins elle était plus une manière de penser stimulante, qu'une véritable démonstration scientifique. Sa généralisation abusive lui confère un statut de dogme y compris dans les processus physiologiques. De sorte que ce n'est que très sporadiquement que l'on n'envisage que la possibilité, pour des cellules qui partagent le même génome, de rentrer en compétition. Un des seuls cas de figures où une approche compétitive puisse être envisagée est celle du système immunitaire dans le cadre de la sélection clonale, mais là encore, il repose sur une variation

génétique induisant une sélection de certaines cellules par opposition à d'autres. En cela nous restons dans le cadre d'une vision dawkinsienne d'une compétition sur la base de différence génétique.

L'ADN, possédant des caractéristiques particulières, serait au centre du vivant et soumettrait l'ensemble des niveaux biologiques à sa propre dynamique de propagation. En tout point cette approche s'apparente à un essentialisme moléculaire. En effet, il apparaît rapidement que la position centrale accordée aux acides nucléiques ne l'est qu'arbitrairement. Car si l'ADN est nécessaire à la synthèse des protéines, via la séquence unidirectionnelle transcription-traduction (encore que des contre-exemples de protéines autoréplicatives existent, voire de protéines synthétisées par d'autres protéines sans intervention ribosomique et donc d'ARNm, les "Non Ribosomal Peptide Synthetases" (NRPS) présentes chez certaines bactéries et champignons (Simon, 1976)), et la mise en application du code génétique, il n'en est pas moins vrai que les protéines sont nécessaires à la synthèse d'ADN, à sa maintenance, à sa propagation. Que cette rétroaction n'ait pas la limpidité du code génétique, mais soit au contraire d'une prodigieuse complexité, ne signifie en rien qu'elle ne soit pas, du point de vue où l'on se place, à l'origine de l'ADN. Cela ne fait qu'illustrer l'élégance et la simplicité et l'universalité du code génétique mais ces paramètres ne constituent pas une justification suffisante à l'instauration d'une hiérarchie moléculaire.

En outre le code génétique est une métaphore. Elle fut forgée avant même la découverte de la structure de l'ADN, par le physicien Erwin Schrödinger dans son livre *What is Life ?* (*Qu'est-ce que la vie ?*) paru en 1944. Ce livre fut visionnaire à bien des égards. Partant du principe que les composants atomiques des gènes sont en effectifs trop réduits pour que la loi des grands nombres suffise à expliquer la

reproductibilité de leur fonctionnement, il postule que le vivant s'organise à partir d'un principe d'ordre, *codé* dans les gènes situés sur les chromosomes, qu'il envisage avec une grande clairvoyance comme un "cristal apériodique". De ces hypothèses, Schrödinger en propose d'autres, notamment "que le code en miniature puisse se trouver en correspondance exacte avec un plan de développement très complexe et très élaboré et contenir en même temps les moyens de le mettre à exécution". Il propose plus loin que cela se fasse "par un "mécanisme" entièrement différent du "mécanisme de probabilité" de la physique" (Schrödinger, 1944). Ces postulats d'une très grande élégance influencèrent la pensée des biochimistes contemporains, jusqu'à Francis Crick lui-même, qui confessa souvent que la lecture de ce livre avait été décisive dans son engagement en biologie. On voit qu'ils concentrent l'essentiel de la pensée déterministe. Néanmoins, derrière cette métaphore du code, la réalité biochimique de ce qui est appelé "code génétique" est constituée d'un nombre limité de réactions chimiques, impliquant des affinités moléculaires préférentielles mais pas absolues dont la reproductibilité est grande (grâce notamment à un rétrocontrôle enzymatique d'erreurs). Il n'y a pas lieu, dans le cadre d'une approche rationnelle, d'extrapoler quelque conclusion que ce soit de la simplicité de ces réactions. L'ADN n'est pas plus au cœur du vivant que les protéines qu'il code ou celles qui en assurent la maintenance et la réparation. Il n'existe pas plus d'organismes vivants dépourvus d'ADN, c'est-à-dire pas un seul, que d'organismes dépourvus de protéines ou même de lipides. Certes, un acide nucléique a comme propriété remarquable de se répliquer et d'évoluer, et donc d'être potentiellement à lui seul le substrat d'une évolution moléculaire. Les formes les plus simples d'ARN circulants, les viroïdes, peuvent cependant être mis en regard des protéines du prion, tout aussi autonomes pour leur pouvoir transformant.

La conséquence du fait que l'ADN n'est pas au centre du vivant (qu'il n'y a pas de centralité dans le vivant) est qu'il n'est pas pertinent de justifier des dynamiques cellulaires par leur capacité à propager leur ADN.

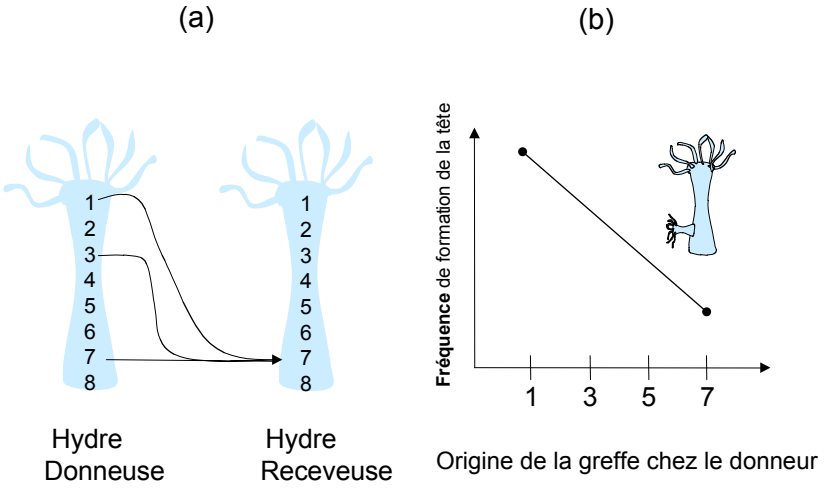
La critique du génocentrisme n'est cependant évidemment pas inédite. Elle a d'ailleurs plusieurs chefs de file, mais peu vont jusqu'à remettre en cause radicalement le déterminisme génétique lui-même. Nous ne prétendons pas, dans cette évocation limitée, réduire l'apport de ces critiques aux quelques lignes que nous leur accordons. Nous y faisons allusion en n'ignorant pas qu'elles ont suscité de nombreux et riches travaux en biologie, auquel notre travail ne peut cependant se confronter sans se détourner de son objet. Nous les présentons en mettant en avant une propriété qu'à nos yeux ils partagent : la critique qu'ils portent tend à ne pas réduire le programme aux programme génétique, en le délocalisant, sans pour autant donc nier l'existence d'un programme en tant que tel.

L'auto-organisation telle qu'a pu la proposer Henri Atlan est une proposition qui critique la centralité de l'ADN vu comme un programme, préférant la métaphore de données (Atlan et Koppel, 1990). Cette théorie peut se rapprocher par plusieurs aspects de celle que nous proposons. Notamment, par son économie en hypothèses, qui relativisent la nécessité de prévoir une régulation précise de certains phénomènes émergents. Néanmoins, dans ce cadre, le programme existe bel et bien par ailleurs, même s'il n'est pas contenu dans l'ADN, mais délocalisé dans la machinerie cellulaire (celle-ci étant d'ailleurs un produit des gènes). Il existe donc toujours, pour cette approche, un programme, dont la précision est atteinte par sensibilité aux conditions initiales.

Un autre critique du génocentrisme, sur la base de nombreuses observations et expérimentations, consista à proposer qu'un génotype ne détermine pas un

phénotype, mais plusieurs, un champ de phénotype possibles baptisé norme de réaction du génotype (Dobzhansky, 1970). L'environnement influe lui aussi sur cette norme en canalisant la réalisation d'un de ces phénotypes. Ici, le déterminisme n'est plus seulement génétique, mais alors généralisé aussi à l'environnement, ce qui est une autre forme de délocalisation. Cette approche a pour éminent mérite de démontrer à quel point le passage du génotype au phénotype est complexe et non entièrement contenu dans les gènes. Mais il n'est pas, lui non plus une critique fondamentale du déterminisme, mais plutôt un aménagement de la théorie : le programme est généralisé a tout ce qui peut influer sur l'expression des gènes, le génome lui-même comme l'environnement. Il existe un déterminisme environnemental qui répond et module le déterminisme génétique. Même si cette approche est évidemment nécessaire, puisque personne ne peut nier sérieusement l'influence de l'environnement sur l'expression génétique, elle ressort plus d'une démarche de description que d'explication. Elle est vouée à faire intervenir des hypothèses *ad hoc* pour chaque environnement. Elle est une formalisation du fait que les influences environnementales et génétiques sur le phénotype sont étroitement intriquées. Cela relativise le déterminisme génétique, mais ne le remet pas en cause radicalement.

La théorie de l'information de position, enfin, propose que les cellules d'un organisme puissent "connaître" leur position dans l'organisme, grâce à l'existence de gradient morphogénétique traversant l'organisme (Wolpert, 1969). Chacun de ces gradients définirait un axe de coordonnées, et l'intégration de ces différents actes permettrait à la cellule de "savoir" où elle se situe dans cet "espace des gradients", et par exemple, de se différencier de manière adéquate. Ces gradients d'orientations différentes permettraient, par leurs interactions, et différents effets de seuil, de créer de la



D'après Gilbert, 1991

Figure 4 : Gradients morphogènes et information de position
L'existence de gradients morphogènes peut être montré chez l'hydre par des greffes. (a) Hydres donneuse et receveuse. L'hydre est constituée d'un axe apical (1) - basal (8). Le pôle apical est la zone de l'hypostome (tête) On peut induire la formation d'un hypostome secondaire latéral en position basale par des greffes.
(b) En fonction de l'origine de la greffe le long de l'axe, la fréquence de formation de l'hypostome secondaire varie.

discontinuité, dans un espace initialement continu. Ceci serait en relation avec le nombre fini de types cellulaires connus (dont on peut noter incidemment que la définition est très arbitraire). Depuis, ces gradients ont souvent été effectivement observés, notamment dans le cadre du développement embryonnaire. Ils sont corrélés à des expressions de gènes particulières.

Mais on peut constater que si cette théorie nuance l'importance de l'information génétique dans la voie de différenciation, c'est pour la remplacer par une autre forme d'information, comme son intitulé le démontre. À bien des égards, c'est de nouveau une externalisation partielle de la programmation du développement, qui ne remet pas celle-ci en cause.

On peut d'ailleurs s'interroger sur l'interprétation des résultats expérimentaux qui ont contribué à accréditer cette théorie. Ces expériences sur des hydres (Webster et Wolpert, 1966) reposent sur des greffes. On greffe par exemple une portion de la partie apicale d'une hydre donneuse dans la partie basale d'une hydre receveuse. On observe dans quelle mesure cette greffe modifie le devenir des différentes portions de la receveuse. On constate bien un effet inductif. Mais, fait notable, cet effet se mesure non pas par son niveau, mais par *le nombre* d'implants ayant un effet inducteur, autrement dit par leur *fréquence* (Figure 4). Ce qui est décrit comme une information de position agit donc par une augmentation de fréquence des événements que par une spécification de ceux-ci. L'existence, dans chaque cas de greffe, d'une proportion non nulle (et non négligeable) de résultats non conformes, souligne la marge de manœuvre dont disposent les cellules pour s'engager ou non, avec une certaine probabilité, dans une voie de différenciation.

La critique du génocentrisme est parfois confondue à tort avec une critique du darwinisme même, voire un retour aux idées prédarwiniennes (voir l'article de Kirschner, *et al.*, 2000, fournissant plusieurs éléments de critique du génocentrisme, mais qui est curieusement intitulé "Molecular Vitalism"). Charles Darwin ne connaissait évidemment pas l'ADN, et cet amalgame ne doit pas tromper. En ce qui concerne nos propositions, il est évident qu'il s'agit d'un élargissement du darwinisme et des dynamiques de hasard-sélection au sein de populations clonales, et non pas de leur négation.

L'ensemble de ces éléments est de nature à nous faire penser que le déterminisme est en crise. On a vu les différentes critiques qui ont conduit à en nuancer le contenu. Elles ont permis un temps d'en sauver l'essentiel. Mais la spécificité, biais d'observation intrinsèque des connaissances transitoires, cède le pas à une vision dégénérée des interactions moléculaires. La notion de programme s'accommode de toutes les nuances jusqu'à des contradictions internes. La centralité du génome est largement contestée. Les résultats s'accumulent qui nécessitent, selon cette logique, de nouvelles hypothèses *ad hoc*. Celles-ci, qui se multiplieront encore, font reposer de plus en plus la pensée déterministe sur un socle d'hypothèses périphériques et fragmentaires.

Ce faisant, ce n'est plus à un assouplissement mais à une fragilisation de cette approche que nous sommes confrontés. Cela légitime la recherche d'un paradigme nouveau.

Soulignons ici la différence qu'il peut y avoir entre déterminisme et réductionnisme et les conséquences du passage progressif de l'un à l'autre. Comme le souligne Jean Gayon, "l'enjeu du déterminisme génétique est la possibilité de prédire le phénotype

d'un organisme (...) à partir de son génotype"(Gayon, 2003). On a vu les raisons différentes pour lesquelles il semble illusoire de croire à cette possibilité. La génétique est passée du déterminisme au réductionnisme dont le but n'est pas de prédire mais d'expliquer les phénomènes. Cette approche est parfaitement adaptée pour isoler une voie de régulation, étudier les relations entre les molécules qui la composent, mais ne permet pas de prévoir la manière dont cette voie va s'insérer dans le contexte plus large du génome en relation avec l'organisme, la manière dont elle va suppléer des voies similaires ou elle même être supplée par d'autres, la manière dont elle va inhiber des voies antagonistes ou être inhibée par elles. Or ce sont ces questions qui idéalement, si elles trouvaient des réponses, pourraient servir le projet déterministe. Mais plus nos connaissances sur la complexité des relations intergéniques progressent, plus nous savons que cet espoir est vain.

II. Énoncé de l'hypothèse endodarwinienne.

Il semble nécessaire, au vu des données précédentes et de l'essoufflement de la capacité prédictive du modèle instructionniste dominant, de chercher dans les ressources propres de la biologie les outils qui nous permettraient de proposer un cadre de lecture renouvelé des phénomènes observés. La pertinence de cette proposition pourrait s'apprécier notamment par sa capacité, non pas à contredire les analyses préexistantes, mais à les englober dans un contexte plus large. Pour asseoir sa scientificité, elle devrait pouvoir être testable et réfutable, caractères qui échappent de plus en plus à la logique instructionniste.

II. A Un parallèle historique

Il apparaît qu'on peut trouver dans le corpus biologique une idée féconde qui a répondu, à une échelle différente, à un problème similaire. Cette idée fut si largement combattue, disséquée et nuancée mais finalement admise et renforcée, que cela nous fournirait en outre un avantage pour en tester la pertinence dans notre système, eu égard à l'expérience capitalisée des contestations historiques. Cette idée est celle de sélection naturelle (Darwin, 1859). Défendue par Charles Darwin et son contemporain Alfred Wallace, elle était une réponse à l'essentialisme fixiste. Pour celui-ci, et ses défenseurs, le monde vivant, résultant d'une création d'ordre divin, était immuable. Prétendre que les espèces pouvaient descendre d'autres espèces, et donc ne pas êtres des créatures initiales, était impensable. La pensée transformiste

mit à bas cette certitude qui n'était d'ailleurs pas d'ordre scientifique mais idéologique. Elle ne se limite d'ailleurs pas au darwinisme et lui préexistait. Jean-Baptiste Lamarck notamment formula avant lui, une théorie de la transformation des espèces (Lamarck, 1809). Il faisait reposer le mécanisme de celle-ci sur une transmission à sa descendance de caractères acquis par un individu au cours de son existence. En tant que tel ce mécanisme ne nécessite pas de compétition entre individus. Ce que le darwinisme a d'original, au sein du courant transformiste, c'est d'avoir pris conscience, d'une part, que des variations inter individuelles présentes à la naissance existent (ce qui conduit à ne pas considérer comme indispensable l'acquisition de caractères), et d'autre part, par l'intermédiaire des travaux de Thomas Malthus, que l'on ne pouvait mathématiquement pas fournir des ressources à tous les individus d'une même population, étant considéré son rythme exponentiel de prolifération. Ce constat impose l'idée de compétition pour les ressources, qui transforme de fait cette approche descriptive en une approche intrinsèquement dynamique.

Ainsi ce n'est pas tant l'hérédité des caractères acquis qui différencie ces deux approches, d'autant plus que Darwin la postulait comme Lamarck (même s'il n'en faisait pas le moteur de sa sélection). La différence principale est que Darwin trouve dans la compétition pour les ressources le moteur pour que la dynamique enclenchée ne cesse de s'entretenir et de se répandre, une fois la vie apparue. C'est d'ailleurs un caractère fondamental de l'approche darwinienne. Elle combine une dimension descriptive, en cohérence avec les observations, à une dimension dynamique, en fournissant des arguments décisifs pour une explication de la persistance dans le temps des phénomènes. L'approche fixiste, ou même le transformisme lamarckien, pouvaient à la rigueur décrire le monde vivant tel qu'il était

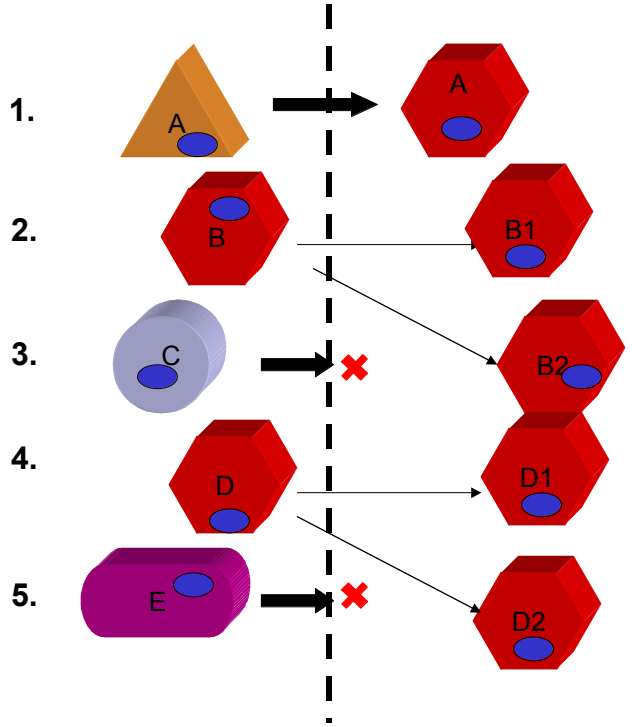


Figure 5 : L'hypothèse endodarwinienne.
Dans une population de cellules clonales (même génome schématisé par un ovale bleu), les cellules sont *a priori* hétérogènes quant à leur profil d'expression génétique. Elles se stabilisent et s'adaptent aux conditions environnementales soit parce que leur profil d'expression est déjà adapté (cas 2 et 4) soit parce que leur comportement exploratoire aléatoire leur permet d'exprimer finalement une combinaison de gènes adaptée (cas 1).D'autres cellules ne trouvent pas les combinaisons adéquates et meurent (cas 3 et 5). Il en résulte une multiplication préférentielle des cellules ayant les profils adaptés (B donne B1 et B2, D donne D1 et D2). Cela concourt à une homogénéisation de la population cellulaire.

à un instant donné, mais ne possédaient pas les ressources propres pour expliquer que face à une catastrophe par exemple, les espèces vivantes s'adaptent et surmontent l'épreuve. Nous montrerons que cette remarque a des répercussions sur nos propositions.

II. B L'hypothèse endodarwinienne

Nous proposons que cette dynamique de compétition pour les ressources qui présuppose des variations individuelles, soit transposée au sein des relations intercellulaires d'un même organisme. Quand bien même celles-ci partagent le même génome, en raison de leur origine clonale, rien ne permet d'exclure qu'elles puissent, dans l'utilisation différentielle de leur génome, entrer en compétition pour des ressources de divers ordres et de transmettre cet avantage à leur descendantes avec souplesse et adaptabilité. A la différence de l'approche classique qui propose que les cellules répondent de manière intégrée, et donc ordonnée, à des signaux, nous proposons que le hasard soit à l'origine de l'état d'expression des cellules, hasard qui est dans un second temps stabilisé, en une dynamique hasard/stabilisation analogue au hasard et à la sélection (Figure 5).

Nous proposons que par défaut, les cellules d'un même organisme, mais plus précisément d'un même tissu, expriment leurs gènes de manière stochastique et que cet aléa d'expression *a priori* entraîne une sélection des profils d'expression génétique les plus appropriés. Les cellules agissent ici dans leur intérêt immédiat, qui est l'accès aux ressources, sans prédétermination. Il s'agit donc d'une forme de sélection naturelle qui se fonde non pas sur les différences génétiques entre cellules, puisque sur ce plan elles sont clonales et donc homogènes, mais sur l'utilisation

différentielle des gènes dans le cadre de ce génome commun. C'est cette dynamique que nous appelons dynamique endodarwinienne rythmée par le hasard (expression stochastique) et la sélection (stabilisation des profils d'expression). Dans cette perspective, le génome peut être considéré comme la mémoire de la cellule. Au sein d'une cellule soumise à des variations environnementales permanentes auxquelles elle doit réagir, le génome représente l'histoire des réponses évolutives trouvées et archivées par la lignée cellulaire, ce qui lui permet à son tour de réagir de manière probabiliste.

II. C Nécessité pour l'hypothèse endodarwinienne d'expliquer la reproductibilité des phénomènes.

II. C 1 Un paradoxe apparent

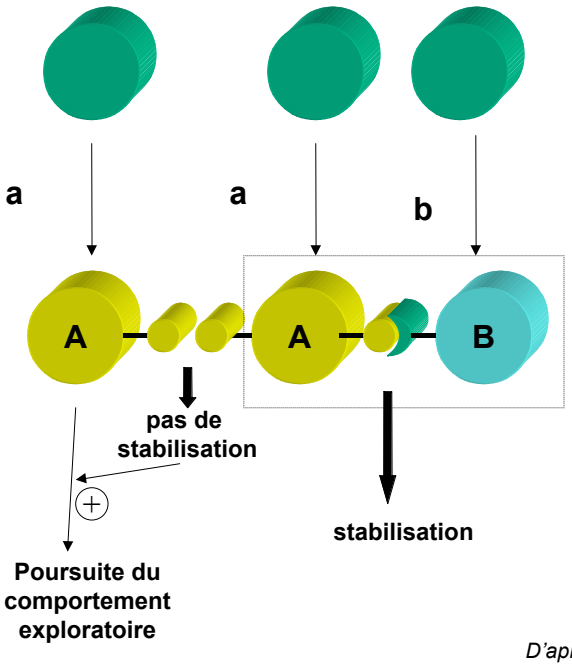
Le déterminisme génétique explique la reproductibilité et la précision des phénomènes du vivant. L'approche probabiliste doit le faire tout autant. D'un point de vue déterministe, la reproductibilité des réactions s'explique par le fait que leur support, les gènes se transmettent physiquement. Confrontées aux mêmes conditions, des cellules pourvues du même génome pourront y répondre de la même manière. Dans une population, la variabilité pour un gène est *a priori* un paramètre mesurable, au sens où il est possible de déterminer la proportion de chaque allèle dans la population. Par opposition, on doute, dans un premier temps, que la variabilité de l'expression génétique soit un paramètre suffisamment solide pour être le support d'une sélection, puisque l'on imagine mal comment elle pourrait être héritable. On conçoit difficilement qu'une "variabilité d'expression" se transmette. En

effet, il paraît surprenant d'alléguer que dans deux organismes donnés, à un même stade du développement et dans un tissu donné, l'expression génétique aléatoire intercellulaire, puisse produire le même résultat, par exemple la formation d'un organe dans le cadre de l'embryogenèse.
En d'autres termes c'est la reproductibilité d'un phénomène reposant sur le hasard qui est difficilement acceptable *a priori*.

II. C 2 Les arguments pour résoudre ce paradoxe.

II. C 2 a La stabilisation limite les choix

On peut proposer plusieurs réponses de divers ordres à ce paradoxe.
-La première, c'est que ce phénomène ne repose pas sur le hasard, mais sur le binôme hasard/sélection. La reproductibilité de ce phénomène ne doit pas être conçue comme provenant seulement de la stochasticité supposée de l'expression génétique mais de deux "forces" , la première étant cette expression stochastique et la seconde étant la stabilisation des profils d'expression qui sont adaptés au milieu et à l'instant. Cette seconde phase structure la population cellulaire en une restriction de l'espace des possibilités vers l'espace de la viabilité. Si chaque cellule est affectée d'une probabilité p d'exprimer une combinaison de gène adaptée, plus la taille de la population sera grande, plus la probabilité de voir se réaliser des profils d'expression appropriés au contexte (et permettre aux cellules concernées de coloniser le milieu) sera grande. Et donc plus la reproductibilité sera grande.



D'après Kupiec, 19 96

Figure 6 : Stabilisation d'un profil aléatoire.
Dans ce modèle, le rôle des interactions cellulaires est de stabiliser les cellules différenciées spontanément en A ou B (avec des probabilités respectives a ou b) plutôt que d'être la cause des différenciations. Les cellules A et B se stabilisent par des interactions entre molécules de surface. En effet, celles-ci déclenchent des cascades de kinases/phosphatases non spécifiques qui stabilisent l'expression des gènes et donc stabilisent le profil de chacune.

-La seconde c'est qu'il est possible d'envisager la stabilisation d'un profil aléatoire sans faire appel à des régulateurs spécifiques. Même si cette question fondamentale dépasse quelque peu le cadre de ce travail, on peut ici rappeler des propositions théoriques qui permettent d'envisager une réponse non déterministe à cette question (Kupiec, 1996). Supposons deux cellules (Figure 6). Chacune est susceptible d'exprimer au hasard un gène a ou b. Ces gènes codent pour des protéines de surface A et B. Supposons désormais que seule l'interaction moléculaire A-B (par opposition à A-A ou B-B) conduise à une stabilisation de l'expression de chaque gène dans chaque cellule, par un processus de (dé)phosphorylation non-spécifique du régulateur qui augmentera la stabilité son affinité pour le gène sur lequel il est fixé (a dans l'une, b dans l'autre). La stabilisation a donc lieu sans instruction spécifique. Cette proposition schématique pourrait servir de cadre à enrichir. L'auteur note que les homéogènes sont proches en termes de structure mais varient par leur affinité pour une séquence donnée, qui est décroissante au fur et à mesure de l'activation de ces différents homéogènes (Pellerin, *et al.* 1994).

II. C 2 b Un phénomène aléatoire est statistiquement reproductible

Penser qu'un phénomène aléatoire n'est pas reproductible est un contresens. si les événements élémentaires sont imprévisibles, le phénomène global, est reproductible *statistiquement*. Dans le contexte de l'hypothèse endodarwinienne, les cellules vont réagir à une variation environnementale (pris au sens le plus large, incluant les cellules voisines) comme à un stress, auquel elles vont répondre de manière aléatoire. Cette variété de réactions possibles va se distribuer autour d'une moyenne, et cette distribution sera d'autant plus reproductible que le nombre d'événements

élémentaires sera grand. La reproductibilité de cette distribution est la condition nécessaire et suffisante pour qu'un crible sélectif puisse se mettre en place, et il est superflu de supposer que les cellules réagissent de manière homogène. Si cette distribution des configurations observées est reproduite, il se trouvera parmi les cellules des individus qui réagiront de manière adéquate à la variation de milieu et qui pourront se répandre au détriment des autres. Le nombre d'événements élémentaires, ici le nombre de cellules réagissant à une variation environnementale est un paramètre déterminant pour prévoir si un système donné peut ou non ressortir d'une dynamique endodarwinienne. Nous y reviendrons quand nous aborderons les limites de l'hypothèse.

II. C 2 c Transmissibilité d'un profil d'expression aléatoire

Puisque nous venons de montrer à plusieurs titres comment un phénomène aléatoire peut être reproductible, il reste à envisager les modalités par lesquelles il peut être transmis. Comment envisager, dans un modèle probabiliste, que les cellules sélectionnées transmettent à leurs descendantes le profil d'expression adéquat ? On peut répondre de deux manières différentes. La première c'est de postuler que la sélection s'opère continûment génération après génération, chacune reproduisant une variabilité d'événements élémentaires. Il n'y aurait pas à proprement parler de transmissibilité directe, mais plutôt une "transmission statistique", et stabilisation continuelle, par exemple selon le modèle évoqué ci-dessus.

Mais il est possible de postuler une transmission directe, c'est-à-dire qu'une fois le profil adapté "trouvé", il pourrait être transmis concrètement aux cellules filles, en leur "épargnant" un comportement exploratoire. Ceci est en rapport avec la structure

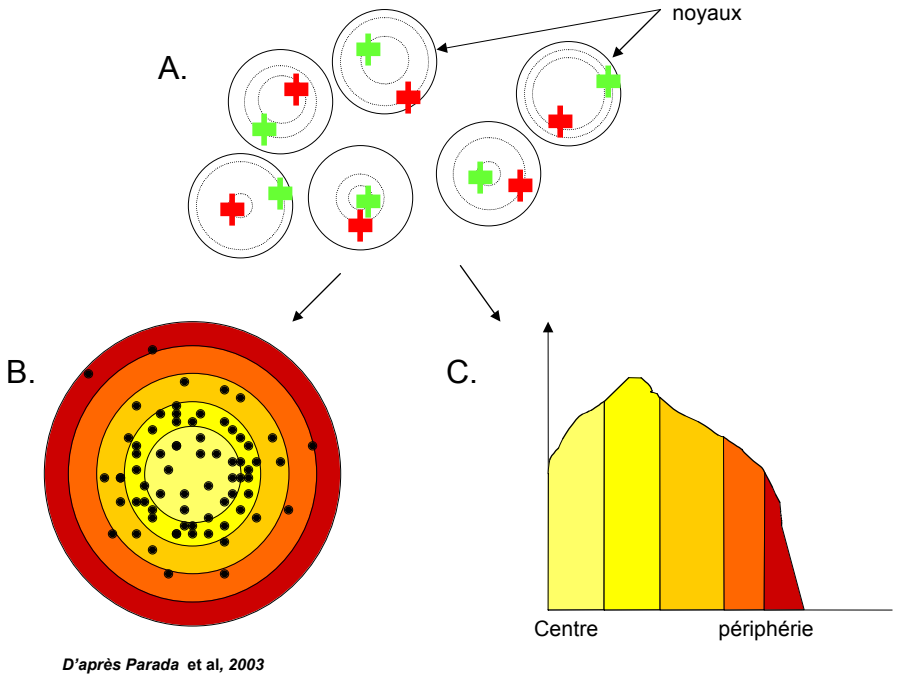


Figure 7 : Variabilité du positionnement des chromosomes.
A. Pour une paire de chromosomes donnée, on peut observer des positions radiales différentes pour chacun, d'un noyau à l'autre. Or cette position radiale est corrélée avec le niveau d'expression. Il existe une distribution radiale de ces positions que l'on peut représenter de deux manières (B et C)

tridimensionnelle de la chromatine. Une littérature récente a précisément et physiquement décrit la notion de territoires chromosomiques. Les auteurs de ce champ d'investigation récent soulignent que les chromosomes, à l'interphase, ne se retrouvent pas dans des positions aléatoires. Ils occupent des zones topographiques précises, baptisées territoires chromosomiques. Ces zones peuvent être très contournées, mais ne s'interpénètrent pas, ce qui est nécessaire à la recondensation de chacun des chromosomes. Cette répartition chromosomique n'est pas absolue, elle possède une certaine souplesse, mais elle n'est pas non plus totalement aléatoire. On note par exemple que des positions relatives entre territoires chromosomiques sont corrélées de manière statistiquement significative (mais pas absolue). Cette topographie tridimensionnelle dynamique est en relation avec l'accessibilité de certaines séquences aux facteurs de transcription. Une observation complémentaire est que la position radiale d'une portion de chromosome (distance par rapport au centre du noyau) peut être corrélée à la précocité de sa réplication (Parada, *et al.*, 2003). En résumé, on observe que les positions relatives et absolues des chromosomes ne sont pas aléatoires et qu'elles sont corrélées au niveau d'expression des gènes (Figure 7). Nous sommes donc en présence d'une explication physique potentielle de la transmissibilité individuelle et directe d'un profil d'expression d'une cellule mère aux cellules filles. Si la configuration spatiale de l'ensemble des chromosomes d'une cellule est en relation avec l'expression des gènes, elle peut être un mode de description du profil d'expression. Si en outre cette configuration est transmissible, alors, le profil d'expression est dans une certaine mesure transmissible. C'est ce qu'ont constaté deux équipes qui ont montré que de manière étonnante, ces configurations sont transmissibles même après un passage par une mitose (Walter, *et al.*, 2003 ; Gerlich, *et al.*, 2003). Il est très notable de

constater par ailleurs que cette transmissibilité est souple. On montre par exemple que ce profil d'expression se retrouve d'une génération à la suivante avec une certaine probabilité, mais que ce lien se dégrade si l'on compare des organisations nucléaires à deux générations cellulaires d'intervalle. Nous sommes donc typiquement devant un phénomène souple qui possède à la fois une variabilité et une transmissibilité, et qui est corrélé à l'expression génétique, ce qui en fait un objet d'étude désigné pour l'approche endodarwinienne. Il convient cependant ici de rester prudent. L'héritabilité des topographies chromosomiques est un sujet d'intérêt récent, et il semble trop tôt pour accorder avec confiance à cette corrélation un rôle causal dans l'expression génétique. Il n'est pas à exclure qu'elle en soit une conséquence, ou que les deux phénomènes soient des conséquences parallèles d'une cause non encore identifiée. Néanmoins, la littérature sur ce sujet fait appel à l'expression génétique stochastique comme lien possible avec ces observations (Parada, *et al.*, 2003).

Ainsi, il semble possible de faire reposer l'acquisition reproductible et la transmission de profils d'expression adaptés sur des paramètres qui n'impliquent pas de spécificité ou de programme, mais qui au contraire sont souples et adaptables. Comment concilier cette approche avec le modèle déterministe ?

II. D Relation entre paradigme déterministe et paradigme probabiliste.

II. D 1 Un déterminisme rationalisé tolère une dimension stochastique.

La dynamique endodarwinienne que nous proposons paraît s'opposer à l'approche instructionniste classique. En proposant que les cellules ne répondent pas de manière déterminée à un stimulus environnemental, mais de manière aléatoire, il pourrait sembler que nous contredisions l'approche classique, qui a longtemps été efficace, et a longtemps fourni un cadre de lecture qui permettait de résoudre de nombreux problèmes biologiques. Comment expliquer alors que nous nous posions en opposition à cette approche longtemps féconde ? Et comment alors entreprendre de réexpliquer les phénomènes antérieurement décrits à la lueur de cette nouvelle grille d'analyse ?

La réponse à ces questions légitimes est en fait assez simple : la dynamique endodarwinienne ne s'oppose pas à l'approche instructionniste. Nous allons essayer de montrer ici qu'elle ne se substitue pas à elle, mais qu'elle l'englobe : nous chercherons à démontrer que l'approche instructionniste est à la fois un cas particulier de l'approche endodarwinienne, une première approximation et une limite. Nous proposons donc ici un paradigme englobant l'approche traditionnelle.

Il faut d'abord chercher à comprendre comment le déterminisme, au-delà des spéculations théoriques initiales, s'est aussi imposé expérimentalement comme le principe de base du fonctionnement biologique. Les pionniers de la biologie moléculaire ne disposaient pas des outils discriminants qui leur auraient permis de savoir si les cellules clonales se comportaient ou non de manière hétérogène, même si certains ont pu le postuler (Lederberg, 1959). Répondre à une telle question aurait nécessité d'étudier l'expression génétique au niveau de la cellule unique, ce qui ne fut possible que tardivement. On a donc, en première approximation, extrapolé des résultats d'observations globales. Une population supposée homogène de milliers ou millions de cellules était analysée, ses ARN (reflets de l'expression génétique) ou

protéines purifiés et identifiés, et l'on en déduisait l'activité d'une cellule supposée représenter le tissu dans son ensemble. Le problème est que cette démarche ne permettait de n'évaluer que la moyenne de l'expression génétique d'un grand nombre de cellules et induisait en fait un parti pris théorique : elle présupposait que toutes les cellules d'un même tissu exprimaient les mêmes gènes, donc que le mécanisme à la base de ce phénomène était déterministe. Alors même que la nature de ce mécanisme était précisément une des énigmes à résoudre. Par ce raisonnement circulaire, L'éventualité d'une variabilité aléatoire d'expression de cellule à cellule, preuve potentielle, au contraire, d'un mécanisme probabiliste, fut d'emblée gommée par un biais d'observation.

Dans ce contexte déterministe, l'expression stochastique des gènes n'est pas inédite en tant que telle, mais on ne lui accordait pas le statut que lui donne l'approche endodarwinienne. L'étude de l'expression stochastique des gènes trouvait jusqu'ici sa justification dans l'étude du bruit de fond de l'expression génétique, étant donné que chacun s'accorde désormais à penser qu'un déterminisme strict ne serait pas crédible au vu des observations. L'expression stochastique n'était pas ignorée jusqu'ici, y compris à l'intérieur du modèle déterministe, mais son statut était radicalement différent : elle y est décrite comme un élément de souplesse injecté dans le système, voir comme la confirmation que comme les systèmes vivants ne sont pas parfaits, ils produisent inévitablement des réactions non conformes, à la marge. Mais une prise de conscience progressive de l'importance de la dimension stochastique de l'expression génétique conduisit certains auteurs à lui proposer un rôle : on constata que la structure de certaines boucles de régulations simples, au fonctionnement déterministe, peut aboutir à des équilibre bistables qui créent de la diversité. En ce sens, jusqu'à présent, l'analyse déterministe englobait la possibilité

d'une expression stochastique en tant que bruit de fond inévitable, voire sporadiquement créateur de diversité, mais toujours dans le cadre d'une régulation fine qui permet d'en limiter les excès par des contrôles spécifiques (McAdams et Arkin, 1999).

II. D 2 Le paradigme probabiliste englobe le paradigme déterministe

Notre approche est différente, nous proposons que les mécanismes d'apparence déterministe, quel que soit le niveau d'observation auquel on les étudie, ne soient que la résultante d'une observation moyenne qui masque les variabilités individuelles. On peut même envisager qu'au sein d'une même cellule, un mécanisme puisse avoir l'apparence du déterminisme, pour la simple raison que la variabilité est très resserrée autour de la moyenne des phénomènes observés. Reprenons le cas du code génétique. Les ARN de transfert sont le support physique de la correspondance du code génétique. Ils associent l'acide aminé au codon correspondant. L'explication physico-chimique de cette association est une affinité préférentielle qui pour être très forte, n'en est pas pour autant absolue. Abstraction faite du rétrocontrôle ultérieur, ces associations ont donc une probabilité non certaine de se réaliser "correctement", même si cette probabilité est très proche de p=1. La prétendue spécificité alléguée à ce genre de mécanisme n'est en fait qu'une première approximation, dans le sens où, la probabilité étant très proche de 1, il n'est pas nécessaire de rentrer dans les détails de sa valeur précise. Dans cet exemple, il y a bien un parti pris fonctionnel, même s'il reste implicite, pour se contenter de cette approximation, parce qu'elle suffit à expliquer le phénomène de la traduction. Que ce parti pris soit dans de nombreux cas justifié, qu'il soit souvent inutile de rentrer dans un questionnement sur la valeur exacte de la probabilité de survenue d'une réaction chimique, par exemple

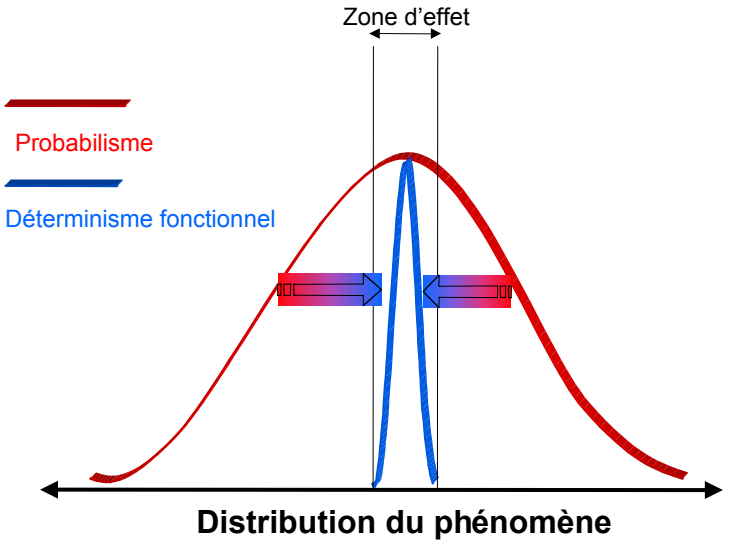


Figure 8 : Le déterminisme, cas particulier du probabilisme.
Un phénomène probabiliste se distribue autour d'une moyenne. On suppose qu'une réaction biologique donnée est enclenchée si le phénomène se place entre les valeurs de la "zone d'effet". Si la distribution du phénomène est large (courbe rouge), seule une partie des cas se situera dans la zone d'effet et donnera une réponse. Si la distribution est très resserrée, la totalité des cas pourra donner lieu à une réponse, on aura l'impression d'un déterminisme fonctionnel, même si le phénomène garde intrinsèquement sa dimension probabiliste.

dans le cas où elle est très proche de un ou très proche de zéro, cela est légitime. Ce qui ne l'est rigoureusement pas c'est de considérer que parce que ces approximations sont souvent suffisantes, elles peuvent se substituer à la réalité des faits, et servir de justification théorique à une approche générale de la biologie moléculaire reposant sur la "spécificité". Or, en toute rigueur, la spécificité n'existe pas en biologie. Et de fait, ces premières approximations sont les limites mathématiques de probabilités. Elles en sont des cas particuliers. L'exemple du code génétique était caricatural (encore qu'une "erreur" non "corrigée", au stade de la traduction, puisse bien sûr occasionner des conséquences graves). Mais on concèdera que dans le domaine de l'expression génétique, les probabilités sont souvent très éloignées de "zéro" ou de "un".

C'est pourquoi la dynamique endodarwinienne est une proposition plus large qui inclut les dynamiques apparemment déterministes. En proposant que l'expression d'un gène dans un contexte donné soit une probabilité, elle inclut l'ensemble des possibles, y compris les plus extrêmes, y compris les limites théoriques (Figure 8).

Remarquons enfin que cette approche est économe en hypothèses. Dans le cadre d'une dynamique déterministe, il est nécessaire de supposer l'existence de mécanismes contrôlant l'expression génétique dans chaque cellule de manière à ce que cette expression soit en adéquation avec les besoins de cette dernière. De fait, de tels mécanismes de contrôle de l'expression génétique existent, ou sont à tout le moins décrits comme tels. Mais dans l'approche endodarwinienne, il n'est pas nécessaire de présupposer, en théorie, de mécanismes de contrôle coûteux, en énergie, en ressources, et en précision. Le contrôle de l'adéquation de l'expression génétique aux besoins de la cellule est une caractéristique émergente de la dynamique endodarwinienne, qui se fait mécaniquement par le principe de sélection

et qui peut soit inciter la cellule "non adaptée" à poursuivre l'expression aléatoire jusqu'à ce qu'elle exprime une combinaison adaptée de gènes, soit l'éliminer. Nous verrons que les mécanismes de contrôle peuvent être compris de manière différente, comme des canalisateurs évolutifs.

APPROCHE BIBLIOGRAPHIQUE

Avant de présenter les tests expérimentaux par lesquels nous avons cherché à contribuer à la mise à disposition de techniques permettant de révéler et mesurer la stochasticité de l'expression génétique et de la méthylation de l'ADN, nous avons considéré qu'il était à indispensable de chercher dans la littérature les indices allant potentiellement dans le sens de l'hypothèse endodarwinienne. Nous proposons ici une sytnhèse des données expérimentales accréditant l'existence d'une expression stochastique généralisée des gènes, qui est la base nécessaire d'un mécanisme endodarwinien, et par ailleurs la remise en cause fondamentale de la notion de "programme génétique".

Puis à travers une approche bibliographique de deux domaines où la notion de programmation est souvent prépondérante, le développement et l'apoptose, nous tenterons de montrer qu'une autre approche de ces disciplines est possible et que nos propositions visant à critiquer le déterminisme en biologie sont compatibles avec les phénomènes observés. Proposer une grille de lecture renouvelée des phénomènes biologiques nous conduit logiquement à chercher si celle-ci propose des axes de cohérence forts dans les contextes où elle est a priori la moins attendue. Elle permettra aussi de voir dans quelle mesure certains phénomènes peuvent être expliqués de manière diverse, voire contradictoire, ce qui pourra révéler qu'il est parfois impossible de trancher entre les différents modèles en l'absence de données complémentaires. Une attitude rationnelle face à ce type d'indécision n'est alors pas d'en justifier un de manière arbitraire, mais de chercher alors lequel des deux cadres explicatifs est le plus parcimonieux. On tentera de convaincre que l'approche endodarwinienne l'est plus que l'approche déterministe classique.

I Les observations expérimentales de variabilité non génétique et de stochasticité de l'expression génétique

I. A Variété des phénomènes biochimiques stochastiques

Dans le contexte de la biologie moléculaire et de son paradigme déterministe, on trouve néanmoins dans la littérature des éléments d'approche probabiliste qui pourraient être la base de dynamiques endodarwiniennes, même si celles-ci ne sont que très peu souvent envisagées. Ce paradoxe au sein du paradigme dominant pose la question du statut de ces exemples. Sont-ils des exceptions marginales au sein d'un ensemble incroyablement plus large de données allant dans un sens déterministe ? Sont-ils la base d'une possible réfutation expérimentale de ce paradigme ? On tentera de montrer que la diversité des situations physiologiques auxquelles elles se rattachent, et leur dispersion au sein du monde vivant, semble plaider pour la seconde hypothèse, à moins de supposer de nombreuses et sporadiques apparitions, hypothèse peu parcimonieuse. C'est donc une mise en cohérence de ces observations que nous proposons ici, inédite à notre connaissance. Toutes ne ressortent pas strictement d'une dynamique endodarwinienne, mais elles ont en commun une approche probabiliste et un recul pris par rapport à l'identité génétique comme source d'homogénéité.

I. A 1 Large présence des phénomènes stochastiques

L'étude des phénomènes stochastiques au sein d'un organisme, au niveau cellulaire ou tissulaire, est très documentée. Les phénomènes stochastiques au niveau cellulaire sont nombreux et ne doivent pas être confondus entre eux. Si on s'attardera sur la problématique particulière de l'expression génétique on peut néanmoins remarquer que plusieurs travaux ont aussi porté sur l'influence de la stochasticité sur la phosphorylation et la méthylation (nous y reviendrons par ailleurs), sur certains effets de résonance moléculaire (Dean Astumian, *et al.*, 1997) et même de certaines composantes de la régulation du cycle cellulaire. Des aspects stochastiques ont été décrit dans différents aspects de la cancérogenèse : la simulation de modèles de carcinogenèse depuis cinquante ans fait appel à des composantes stochastiques (Armitage et Doll, 1954) pour décrire les dynamiques des mutations génétiques associées à l'apparition des tumeurs ou plus récemment pour décrire l'élimination des cellules cancéreuses (Michor, *et al.*, 2003). Nous sortons ici du cadre de l'endodarwinisme, puisque la tumorigenèse repose dans ces modèles sur une ou des mutations génétiques malignes, mais c'est la stochasticité du modèle, et donc sa non-programmation, que nous mettons en avant. De la même manière, les mécanismes de vieillissement cellulaire, par modification des télomères (et donc là encore, modifications génétiques) ont une composante stochastique (Rubelj et Vondracek, 1999). Enfin Gerhart et Kirschner consacrent un chapitre entier de leur ouvrage de synthèse (Gerhart et Kirschner, 1997) à ce qu'ils nomment les mécanismes exploratoires, regroupant des cas de progression par essai/erreur conceptuellement semblables aux dynamiques endodarwiniennes, mais à différents niveaux du vivant, celui de l'individu (recherche de nourriture chez les fourmis) qui nous intéresse peu ici, mais aussi celui systèmes immunitaires et nerveux déjà décrits (*Cf. infra*), de l'angiogenèse et de la trachéogenèse, et enfin au niveau

subcellulaire, dans le comportement dynamique du cytosquelette. Ce niveau subcellulaire présente aussi une autre dynamique qui pourrait présenter des similitudes avec l'endodarwinisme : il s'agit de la répartition apparemment stabilisée des territoires chromosomiques à l'intérieur du noyau, en rapport avec le niveau d'expression des gènes qu'ils contiennent. Enfin, la ségrégation aléatoire des organites cellulaires, et notamment des chloroplastes, a été théorisée (Birky et Skavaril, 1984) et observée (Hennis et Birky, 1984). Ces observations ont montré que consécutivement à une partition inégale, un mécanisme de rééquilibrage *a posteriori* existait. Notons que si ces exemples subcellulaires peuvent présenter un profil hasard/stabilisation, ils ne ressortent pas de l'endodarwinisme au sens strict car il n'y a pas de sélection d'une unité de reproduction, comme peut l'être la cellule (analogue en cela à l'unité de sélection qu'est l'individu dans le cadre classique de la sélection naturelle). Il y a seulement une stabilisation d'états particuliers.

Ces exemples variés montrent que la dimension stochastique des phénomènes biochimiques est une donnée fondamentale qui doit être prise en compte pour une meilleure compréhension des mécanismes du vivant.

I. A. 2 De nombreux systèmes où une variabilité non génétique est observée

I. A. 2 a Variabilité chez les procaryotes

Chez les procaryotes, une hétérogénéité non génétique a été étudiée dans le cadre de la variation de taille des population de bactériophages (Delbruck ,1945), dans les temps de divisions cellulaires bactériennes (Powell 1956, 1958), dans l'alternative

lyse/lysogénie des bactéries infectées par le phage λ, dans l'induction ou la répression de bactéries cultivées sous induction LacZ (Novick et Viener, 1957 ; Benzer, 1953 ; Maloney et Rotman, 1973 ; Tolker Nielsen, *et al.*, 1998), dans le comportement natatoire des bactéries flagellées (Spudich et Koshland, 1976), dans la variation aléatoire du gène (*fliC* ou *fljB*) qui code pour la flagelline de *Salmonella enterica* (Stocker, 1949), dans l'induction de la différenciation (Russo-Marie, *et al.*, 1993, Chung, *et al.* 1994, Chung et Stephanopoulos 1995).

I. A. 2 b Variabilité intercellulaire dans les voies de différenciation

En accord avec un modèle probabiliste, des expériences sur de nombreuses lignées ont démontré que leur cinétique de différenciation est mieux décrite lorsqu'on assigne à chaque cellule une probabilité de se différencier par unité de temps. C'est le cas des expériences pionnières de Till sur des cellules souches hématopoïétiques murines (Till, *et al.*, 1964) renforcés bien plus tard (Mayani, *et al.*, 1993 ; Abkowitz, *et al.*., 1996). D'autres signes d'une variabilité d'expression ont été détectés dans des lignées de mélanocytes (Bennett 1983), de myoblastes (Lin, *et al.*, 1994), de cellules de la crypte intestinale (Paulus, *et al.*, 1993), d'hépatocytes (Michaelson, 1993 ; accompagnée de sélection : Van Roon, *et al.*, 1989), de chondrocytes (Bohme, 1995). Un modèle de différenciation reposant sur une base stochastique a été proposé pour les lymphocytes T (Davis, *et al.*, 1993 ; Chan, *et al.*, 1994), les lymphocytes B (Deenick, *et al.*, 1999), les érythrocytes (Hu, *et al.*, 1997) et les cellules originaires de la crête neurale (Baroffio, *et al.*, 1992).

I. A. 2 c Hétérogénéité intercellulaire d'expression génétique dans un contexte homogène

D'autres observations expérimentales de variabilité d'expression intercellulaire ont été faites dans des lignées exprimant des gènes rapporteurs sous contrôle de différents promoteurs. C'est le cas de la lignée macrophagique RAW 264 (Ross, *et al.*, 1994) transfectée de manière stable par LTR-HIV1-LacZ. Des résultats comparables ont été obtenus concernant l'expression de lacZ dans la lignée humaine Jurkat sous contrôle de différents promoteurs ou éléments cis activateurs (Fiering, *et al.*, 1990), ainsi que pour des séquences cibles d'hormones stéroïdes (Ko, *et al.*, 1990) et dans des constructions de cellules expriment la luciférase sous divers promoteurs (White, *et al.*, 1995 ; Takasuka, 1998). Dans le cas de cellules murines musculaires multinucléées, des différences d'expression internucléaires pour des gènes donnés, dans un contexte cytoplasmique supposé homogène ont été mesurées (Newlands, *et al.*, 1998).

Plus surprenant encore est le cas l'expression hétéroallélique. Alors que les allèles baignent dans un environnement transcriptionnel commun, des observations en nombre croissant montrent que pour certains gènes les deux allèles s'expriment différemment. Cela peut être dû au phénomène d'empreinte, sur lequel nous reviendrons (méthylation différentielle des allèles en fonction de leur origine parentale, conduisant à la répression de l'expression de l'un des deux). Mais dans de nombreux cas, ce phénomène très dynamique et variant de cellule à cellule semble être dû à une expression stochastique, notamment pour les gènes d'immunoglobulines (Nemazee, 2000), des récepteurs olfactifs (Chesset al, 1994), de

la globine (Wijgerde, 1995) - où cette variabilité est décrite comme héritable (de Krom, 2002), des récepteurs T et NK (Held, *et al.*, 1999),des cytokines IL-2 et IL-4 (Hollander, *et al.*, 1998 ; Rivière, *et al.*, 1998).

Ainsi tous ces exemples mettent en évidence le phénomène de plasticité phénotypique qui permet, sur la base d'un génome commun et par des dynamiques aléatoires, de d'aboutir à une diversité qui peut permettre de répondre à un besoin de souplesse dans l'adaptabilité à des fluctuations environnementales. Reste à envisager des mécanismes qui permettent, dans un second temps, de stabiliser certaines options.

I A. 3 Variabilité et stabilisation

I. A 3 a Des précurseurs

Postuler des compétitions à l'intérieur de l'organisme et donc son préalable, de la diversité, est une idée ancienne. Elle fut proposée pour la première fois en 1881, par l'embryologiste darwinien Wilhelm Roux, dans *Der Kampf der Theile im Organismus - Le combat entre les parties de l'organisme-* (Roux, 1881), dans lequel il prévoyait que les organes et les cellules puissent rentrer en compétition. Mais cette hypothèse resta isolée, et n'eût pas de postérité immédiate. Rappelons que dans le modèle endodarwinien, la compétition est envisagée au niveau des cellules qui sont des unités capable de se reproduire avec des taux différentiels en fonction de leur niveau d'adaptation au milieu. Dans ce modèle, la compétition entre organes, n'est pas pertinente, ceux-ci ne se reproduisant pas. Cette hypothèse fut reprise

sporadiquement comme telle, au seul niveau cellulaire, par Spiegelman, qui la rebaptisa compétition physiologique (Spiegelman, 1945). Elle a enfin été remise au goût du jour expérimentalement par les travaux de l'équipe de H. F. Nijhout (Nijhout et Emlen, 1998, Klingenberg et Njhout, 1998). Ces travaux, qui portent sur des lépidoptères et des hannetons, mettent en évidence que les relations entre organes sont plus fondées sur la compétition que sur l'échange de signaux. Cette compétition pour des ressources en quantité limitée, peut être un des mécanismes de la régulation de la taille de ces organes et une explication des processus de symétrisation.

I. A 3 b Deux théories majeures

Mais ce sont deux théories majeures qui, au vingtième siècle, firent appel à des dynamiques de hasard et de sélection à l'intérieur de l'organisme. Nous les présentons ici succinctement, dans le but de mettre en évidence en quoi elle peuvent se rapprocher ou rester distantes des hypothèses endodarwiniennes.

-La théorie de la sélection clonale.

La théorie de la sélection clonale a bouleversé l'immunologie à la fin des années 1950. Elle fut une réponse audacieuse à la question suivante : Face au nombre énorme d'antigènes potentiellement présents, comment expliquer la diversité du répertoire des anticorps de reconnaissance spécifique ? Jusque dans les années 1950, l'immunologie reposait sur une théorie dite instructive de synthèse des

anticorps. L'idée centrale était que les cellules synthétisaient des gammaglobulines toutes identiques, qui étaient ultérieurement façonnées par le contact avec un antigène, pour lui devenir complémentaire, et que ce "moulage" était alors inscrit dans la mémoire génétique de la cellule, son génome. Cette théorie était un défi pour le dogme central de la biologie moléculaire, qui présuppose le transfert unidirectionnel d'information de l'ADN à la protéine. Elle semblait en effet supposer une information dans le sens rétrograde. C'est en 1955 que Niels Jerne émet le premier l'hypothèse que les anticorps préexistent à l'antigène, et sont présents en une très grande diversité dans l'organisme (Jerne, 1955). L'entrée d'un antigène n'entraîne donc pas le "moulage" d'un anticorps spécifique, mais la reconnaissance de l'antigène par son anticorps complémentaire préexistant, et l'activation consécutive de processus cellulaires de fabrication en grande quantité de cet anticorps. Burnet affinera cette théorie en 1957 en remplaçant le concept d'anticorps soluble par celui de récepteur cellulaire, dont l'activation par un antigène spécifique va activer la cellule (qui ne porte qu'un type de récepteur). Toutes les cellules produisant ce même récepteur étant supposées de même origine, l'hypothèse prend le nom de sélection clonale (Burnet, 1957). Cette hypothèse va être ardemment débattue mais finalement admise quand en 1976, Tonegawa démontre la possibilité de générer des millions d'anticorps différents avec seulement quelques gènes, par une mécanique de recombinaison génique (Tonegawa, 1976). Cette théorie, qui présuppose une variabilité très importante, puis une sélection, est clairement une illustration d'un processus darwinien à l'intérieur d'un même organisme. Elle repose néanmoins sur une différence génétique entre cellules (après la recombinaison génique) et s'éloigne en cela de notre propos qui est de supposer la possibilité d'une compétition à génome identique. Mais néanmoins, elle prouve qu'une dynamique de

hasard (avec production en surnombre) et de sélection, peut être pertinente au sein des cellules d'un même individu.

-La théorie de la stabilisation sélective des neurones

Proposée par Changeux en 1976, elle se place dans la continuité historique de spéculations non démontrées ou formalisées depuis Ramon y Cajal (1899), sur la coexistence de mécanismes de compétition et de régression au cours du développement embryonnaire. C'est au cours d'une rencontre en 1972 organisée par Edgar Morin que Changeux propose un développement des synapses sur le mode "*"darwinien"* ***mais*** *de manière épigénétique*" (nous soulignons, comme s'il était incongru qu'un phénomène darwinien ne soit pas entièrement génétique) et sous le contrôle de l'activité du réseau. Il formalisera cette proposition en 1973 et 1976, en proposant un mécanisme plausible selon lequel, à certains points critiques du développement, des contacts synaptiques surnuméraires sont formés et l'activité du réseau stabilise ultérieurement certains de ces contacts et engage la destruction des autres (Changeux, *et al.*, 1973 ; Changeux et Danchin, 1976). Des indices expérimentaux en faveur de cette théorie ont abondé dans différents systèmes neuraux dans les années qui suivirent, notamment par les travaux d'Edelman (Edelman et Mountcastle, 1978 ; Edelman, 1987) qui parla alors de "darwinisme neural". La connectivité du neurones est fréquemment décrite en termes de plasticité et de compétition, jusqu'à celle entre différents axones sur une même cible (Kasthuri, *et al.*, 2003 ; Buffelli, *et al.* 2003)

Ainsi l'organogenèse du cerveau, dont la complexité est inouïe, s'avère descriptible en des termes adaptatifs, dans un contexte d'homogénéité génétique, plutôt que par un mécanisme instructif résultat d'un programme prédéfini et précis. Si cette organogenèse si complexe peut reposer sur une dynamique de hasard stabilisation, et si celle-ci s'avère plus parcimonieuse qu'une dynamique instructive, il devient pertinent de se demander dans quelle mesure une généralisation n'est pas envisageable.

I. A 3 c D'autres systèmes reposant sur une dynamique de hasard/sélection

Chez les protozoaires, le phénomène dit de variation clonale phénotypique (Clonal Phenotypic variation, CPV) ressort d'un mécanisme de hasard/sélection en ce qu'il fait reposer une adaptation sur une variation phénotypique. Certains parasites protozoaires ne peuvent pas avoir de reproduction sexuée pendant de longues plages de leur cycle de vie, dépendant de leur hôte. La variation phénotypique peut être un des moyens de substitution pour s'adapter rapidement au milieu. Si les cellules filles activent les mêmes gènes pour répondre à cette pression de l'environnement, on aura alors un clone phénotypique (Meunier, 2001).

Chez le nématode *Caenorhabditis elegans*, connu pour le déterminisme canonique de ses destins cellulaires, certains cas de différenciation sont décrits de manière stochastique. Différentes espèces de nématodes ont pour particularité d'avoir un nombre de cellules fixes, chacune ayant une origine déterminée. On peut ainsi tracer le devenir des cellules d'un précurseur donné. Ils sont *a priori* des candidats de choix pour démontrer la rigueur et la puissance d'une dynamique déterministe. Or, il a été montré que pour certaines d'entre elles dans cet environnement déterminé, il existe

un aléa de différenciation (cellules AC et VU venant des précurseurs Z1.ppp ou Z4.aaa) (Greenwald, *et al.*, 1992).

Chez l'annélide *Helobdella robusta*, dans le contexte des premières divisions cellulaires stéréotypées de l'œuf, une phase transitoire d'expression stochastique du gène HRO-WNT-A dans l'un ou l'autre des deux types cellulaires AB ou CD est observée, suivie d'une phase dite "déterministe" d'expression (Huang, et al., 2001).

Un mécanisme dénommé "compétition cellulaire" (Morata et Ripoll, 1975) est documenté dans la croissance disques imaginaux de la drosophile, mécanisme par lequel des cellules certes mutées rentrent en compétition pour le facteur Decapentaplegic (Dpp) dit de survie (Moreno, 2002). Il suppose certes une différence génétique (une mutation pour un gène) mais propose à ce titre une explication à la disparition progressive des cellules mutées par des taux de prolifération différentiels.

Dans un article de synthèse en 1993, James Michaelson recense lui aussi un certain nombre de cas de figures faisant intervenir une hétérogénéité de l'expression suivie d'une sélection dont il rappelle le caractère structurant dans le cas du système immunitaire, et du système nerveux mais aussi dans le développement du foie. Il montre qu'au cours du développement de cet organe chez le rat, les cellules qui synthétisent l'albumine remplacent progressivement celles qui synthétisent l'α-foetoprotéine, grâce à un taux de prolifération supérieur. Un mécanisme de concurrence identique, toujours dans le foie, est aussi mis en œuvre lors d'une inflammation. Les cellules qui synthétisent le fibrinogène se multiplient aux dépens de celles qui synthétisent l'albumine. James Michaelson élargit son propos au développement du système urogénital, et plus généralement à l'embryologie, notamment dans le développement des membres de vertébrés (Michaelson, 1993).

Ses propositions et ses remarques en font sans conteste un précurseur de l'endodarwinisme.

L'ensemble de ces observations ne fait pas preuve. Mais on est en droit de faire plusieurs constats. Une importante variabilité non génétique est observée aussi bien chez les procaryotes que chez les eukaryotes. Parmi ces derniers, et notamment dans le règne animal, elle peut concerner aussi bien les protozoaires que les métazoaires. Elle semble jouer à différents niveaux hiérarchiques du vivant. Elle concerne les phases de développement ou non. Elle est pressentie dans certaines phases de différenciation. Elle intervient dans des situations normales ou pathogènes. Cette variabilité est accompagnée de sélection dans plusieurs systèmes.

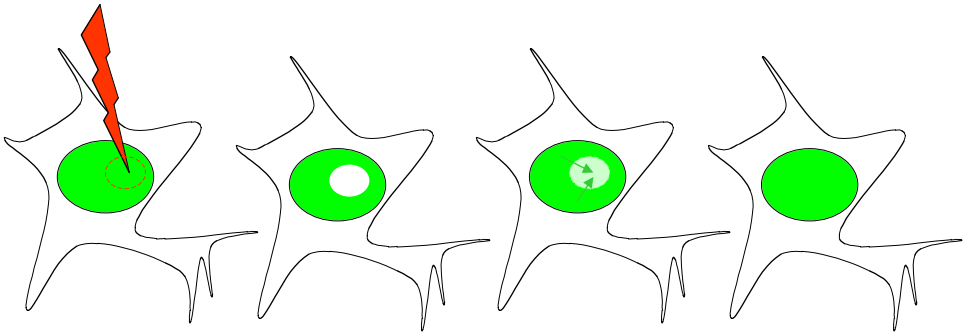
La diversité de ces cas de figures rend pertinente l'hypothèse de travail que la variabilité non génétique ait une portée plus générale que les nombreux cas sporadiques où elle est remarquée. Cette hypothèse doit bien sûr être testée. Dans ce cadre, nous aborderons ensuite une forme que peut prendre cette variabilité non génétique : l'expression stochastique des gènes.

I. B L’expression stochastique des gènes

Dans un article de synthèse proposant une théorie unifiée de l'expression génétique, il est mentionné en préambule, que chez les métazoaires composés de plusieurs centaines de types cellulaires, la "remarquable diversité dans la spécialisation cellulaire est atteinte par l'expression et la régulation finement contrôlées d'un sous-ensemble de gènes dans chaque lignée cellulaire" (Orphanides et Reinberg, 2002).

Cette assertion est en fait surprenante au vu d'un nombre conséquent de données montrant l'expression stochastique des gènes : celle-ci est en effet avérée, quelle que soit l'importance et la signification biologique qu'on lui accorde. Elle a été postulée dans le cadre de prédictions théoriques (Kupiec 1981, 1983 ; Ko 1991), elle a été modélisée (Arkin, *et al.*, 1998, Cook, *et al.*, 1998 ; Thattai et Oudenaarden, 2001 ; Kuznetsov 2001,, *et al.* 2002 ; Sasai et Wolynes, 2003), en plusieurs composantes (Swain, *et al.*, 2002), elle a enfin, été observée expérimentalement (Chelly, *et al.*, 1989 ; Ross, *et al.*, 1994 ; Hume, 2000 ; Elowitz, *et al.*, 2002, Heams 2003), dans des systèmes divers. Pour avoir passé ces étapes de l'investigation scientifique elle doit être considérée comme une donnée et non pas comme une spéculation. Elle semble néanmoins être une réalité souterraine, puisque des articles récents ressentent encore la nécessité d'en annoncer l'existence (Elowitz, *et al.*, 2002), bien qu'elle ait été caractérisée de longue date. Cela souligne la difficulté à diffuser cette idée. Comme le soulignent Sasai et Wolynes, avant d'explorer des hypothèses stochastiques, la puissance de la grille déterministe n'est pas facile à remettre en question :

« La vision déterministe a beaucoup d'atouts. Les miracles du développement embryonnaire requièrent intrication et précision. Les cycles cellulaires, un élément dynamique important et constitutif du vivant, sont souvent décrits comme des horloges. Avec la grande quantité d'informations disponibles désormais à l'ère post - génomique, il est difficile de résister à faire des analogies entre les cellules et les processeurs d'information fait par l'homme, les ordinateurs électroniques, qui rectifient leurs programme avec une détermination que Laplace aurait trouvée impressionnante. La vision stochastique n'est cependant pas sans intérêt. " (Sasai et Wolynes, 2003)



D'après Misteli, 2001

Figure 9 : Technique de FRAP.
Un marqueur fluorescent est associé à une protéine nucléaire.
A. Une zone du noyau est excitée par un laser.
B. Cette excitation entraîne une perte de la fluorescence des protéines présentes dans cette zone.
C. En fonction de leur mobilité, les protéines fluorescentes avoisinant la zone vont progressivement y diffuser.
D. Il y a un recouvrement progressif de la fluorescence dans la zone sous l'effet de cette diffusion. La vitesse à laquelle se fait ce recouvrement donne une indication sur le coefficient de diffusion de la molécule.

Non seulement cette vision ne manque pas d'intérêt mais elle est, au vu des progrès récents, le mode de description le plus rationnel des événements moléculaires. Nous allons présenter ici les données les plus récentes qui permettent de manière décisive d'affirmer que l'expression des gènes est fondamentalement stochastique.

I. B 1 Bases moléculaires de l'expression stochastique des gènes.

I B 1 a La haute dynamicité des échanges moléculaires

Nous chercherons ici à illustrer qu'un certain nombre d'acquis récents, concernant les interactions moléculaires impliquées dans l'expression des gènes, mettent en difficulté l'approche déterministe. Un certain nombre d'arguments expérimentaux donnent une cohérence fonctionnelle à une description stochastique de l'expression des gènes, par opposition à une logique instructionniste.

Tout d'abord, une découverte fondamentale de ces dernières années est la nature très dynamique du noyau cellulaire (Misteli, 2001). Cette observation a été permise notamment par la technique de recouvrement de fluorescence après exctinction (Fluorescent Recovery After Photobleaching - FRAP) qui l'a révélée de manière surprenante. Cette technique permet de calculer les coefficients de diffusion et les constantes d'associations des protéines (Figure 9). Elle a permis de montrer la diffusion des protéines dans le noyau, ou plutôt, du fait de son encombrement, leur marche au hasard dans un environnement restreint. Les protéines nucléaires sont très mobiles, Les compartiments nucléaires se désassemblent à la mitose pour se

reformer dans les cellules filles. Les associations entre molécules sont très fluctuantes : au sein même de la chromatine, on a montré expérimentalement que les histones H1 sont en association extrêmement transitoire ; leur temps de liaison est de l'ordre de grandeur de la minute, puis ils se dissocient et se réassocient ailleurs, au point que l'on a parlé de modèle "stop-and-go". Des activateurs de transcription comme les récepteurs au stéroïdes (glucocorticoïdes ou œstrogènes) qui se lient à l'ADN sont encore plus dynamiques. Il en va de même pour les complexes de préinitiation de la transcription : des études in vitro ont montré que l'efficacité d'association des éléments de ce complexe est très faible (chez la drosophile, cette durée in vitro peut être de plusieurs minutes, alors que le passage de l'initiation à l'élongation in vivo est de quelques secondes). La chromatine est en fait traversée de flux permanents de tous ses composants et partenaires moléculaires. Seul un système de concentration localisée des facteurs de transcription par la formation d'un nombre limité d'"usines à transcription" permet d'augmenter la probabilité d'association de ces molécules à des niveaux compatibles avec le temps biologique. Les constituants de ces usines à transcription peuvent être par ailleurs stockés dans des compartiments eux-mêmes dynamiques, au sens où ils sont dépendants du flux entrant et du flux sortant de molécules. Ces compartiments (de vingt à quarante par noyaux) sont eux-mêmes soumis à régulation et sont sensibles aux inhibiteurs de kinases et de phosphatases. Les usines à transcription, au nombre de 2500 par noyaux à un instant donné, limitent le nombre de gènes exprimables simultanément et laissent la porte ouverte à une expression différentielle de cellule à cellule sur la base de combinatoires. Par ailleurs il a été montré que le complexe de transcription se dissocie après chaque événement pour devoir se reformer, avec une probabilité certes augmentée de réassociation au même endroit.

Ces données donnent dans l'ensemble une vision très dynamique des échanges moléculaires dans le noyau cellulaire, et semblent peu compatibles avec un modèle instructionniste où les réactions se succèdent en cascades dirigées et organisées. Il n'apparaît ni parcimonieux, ni même réaliste qu'un quelconque programme puisse être à l'origine de ces mouvements aléatoires browniens. La logique qui se dégage des observations les plus récentes est celle d'un comportement exploratoire, d'un tâtonnement moléculaire dont la conséquence est une adaptabilité permanente de la réponse cellulaire globale aux variations environnementales. Cela n'est pas incompatible avec des processus de d'autostabilisation par rétroaction positive de certaines associations moléculaires. La quantité des interactions que peut engager une seule molécule au gré de ces échanges permanents et transitoires démultiplie, pour un nombre fixé de molécules, les combinaisons possibles, et rend crédible un comportement exploratoire de la cellule. Ces processus qui favorisent certaines associations peuvent contribuer à limiter fonctionnellement les possibilités, et à donner donc une apparence de déterminisme à des phénomènes intrinsèquement probabilistes.

Dans le cadre de cette logique, on doit ici mentionner la problématique largement ignorée de la disponibilité des molécules régulatrices en relation avec leur nombre de sites potentiels. On verra plus loin l'exemple de régulateurs présents en nombre bien moindre par cellule que celui de leurs séquences cibles. Il ne peut donc être mécaniquement présent partout et cela impose une compétition topographique pour cette molécule entre ses différents sites d'affinité.

L'ensemble de ces paramètres souligne la très haute dynamicité des échanges nucléaires et conduit à proposer que l'induction de la transcription se fasse par une augmentation de la probabilité de transcription dans chaque cellule plutôt que par une augmentation du taux de transcription dans chaque cellule (Hume, 2000). Cette observation générale s'explique au niveau moléculaire : les cellules passeraient d'un état non activé à activé par dépassement du seuil de concentration d'une molécule Cet effet de seuil, associé à une autocatalyse qui amplifierait les différences intercellulaires pourrait être le cadre explicatif de la variété de destinées au sein d'une population. Cette augmentation en nombre de cellules concernées plutôt qu'en taux d'expression par cellule a été montrée in vitro dans le cas de l'actine β dans des fibroblastes et pour LacZ sous inducteur dans différentes lignées transfectées .

I. B 1 b La dimension stochastique apparaît avec l'effet d'échantillonnage.

Les phénomènes stochastiques (ou le bruit, le terme varie) sont souvent invoqués quand, au niveau élémentaire, un effet d'échantillonnage est susceptible d'intervenir. Le point de vue déterministe modélise des boucles de rétroaction, le plus souvent sans se demander si chaque molécule qui intervient est présente en quantité suffisante. Or c'est un problème primordial. Si un gène donné est très largement exprimé, les protéines synthétisées seront présentes en grand nombre, et l'on peut envisager une approche statistique de leurs effets, et supposer par exemple qu'elles sont en quantité saturante pour la régulation dans laquelle elles interviennent. Mais si le gène est peu exprimé, et que la protéine est produite à un faible nombre de copies entre quelques dizaines et quelques centaines (Arkin et McAdams, 1999), elle ne sera pas en situation saturante. Cela pourra avoir des conséquences notamment sur

la répartition des copies à la division cellulaire (effet d'échantillonnage) ou sur sa non-disponibilité dans les régulations pour lesquelles elle est supposée intervenir.
En se fondant sur les données de P. Guptasarma (Guptasarma, 1995) chez *E. coli,* A. Arkin et H. McAdams (1999) prennent le cas d'un gène exprimé en cinquante copies d'ARNm. Ils rappellent que lors de la division cellulaire qui s'ensuivra (et qui devrait aboutir en théorie à un partage égal du nombre de copies en deux, soit une moyenne de vingt-cinq copies par cellules filles), des calculs statistiques montrent que 6 % des cellules auront moins de 19 copies et 6 % des cellules filles en auront plus de 30. Cette asymétrie, simple conséquence d'un situation statistique donnée, sans faire intervenir à un seul moment de programme de réalisation, peut suffire à donner des sous-populations de part et d'autre d'un seuil particulier à l'espèce d'ARN considérée (et à son utilisation dans une boucle de régulation) et qui peut conduire à une variation phénotypique dans un environnement commun.
Il devient donc indispensable au vu de l'exemple considéré de rechercher simultanément les régulations dans lesquelles un ARN ou une protéine donnée interviennent *et* leur disponibilité à remplir tout ou partie de ces activités.

On connaît par ailleurs une certain nombre de protéines impliquées dans la régulation de l'expression est inférieure à son nombre de sites potentiels sur l'ADN. C'est le cas par exemple du répresseur MeCP2, qui est présent à 1 million de copies par cellule (en moyenne) alors qu'il peut se fixer sur environ 40 millions de sites sur l'ADN. Cela implique mathématiquement qu'il ne puisse être présent partout simultanément, donc qu'il y a une combinatoire de distributions possibles, et donc d'effets biologiques possibles.

L'approche probabiliste commande de prendre en compte les ratios entre les éléments qui interagissent, alors que l'approche déterministe néglige en règle générale cette question pour les considérer souvent comme étant en quantité suffisante pour interagir. C'est ce type d'observation, la limitation des ressources au sein d'une population, qui permit à Charles Darwin de proposer un moteur évolutif fondé sur la sélection des variants adaptés. Une recherche des dynamiques endodarwiniennes doit impérativement les prendre en compte. Car une analyse fondant la reproductibilité d'un phénomène élémentaire sur un mécanisme déterministe serait incompatible avec une limitation ou une répartition inégale, cellule à cellule, de certaines ressources.

La disponibilité ne tient pas seulement à la quantité d'une espèce d'ARN messager mais aussi à leur répartition topographique. Or les transcrits ne sont pas produits de manière linéaire. Quand on a cherché à observer leur apparition, de manière fine, ils ont été décrits comme étant produits de manière pulsatile. Le paramètre déterminant pour la production d'ARN messager est la fréquence moyenne des pulsations, dépendant elle même de la probabilité de formation du complexe de préinitiation de la transcription. Le nombre moyen d'ARN produits est ainsi dépendant de la stabilité de ce complexe. Cela a une influence sur leur disponibilité spatiale (un temps de diffusion est nécessaire à partir de leur site de transcription) et temporelle, les deux étant imbriquées de manière évidente. Cette observation, qui plus est, a notamment été faite de manière éclatante dans les fibres musculaire de souris (Newlands, *et al.*, 1998) qui sont des syncytiums, ceci ne permettant pas d'argumenter sur la base d'éventuels environnements intracellulaires différant de l'une à l'autre. En effet, avec ce raisonnement, on aurait pu soutenir l'idée que des microenvironnements différents

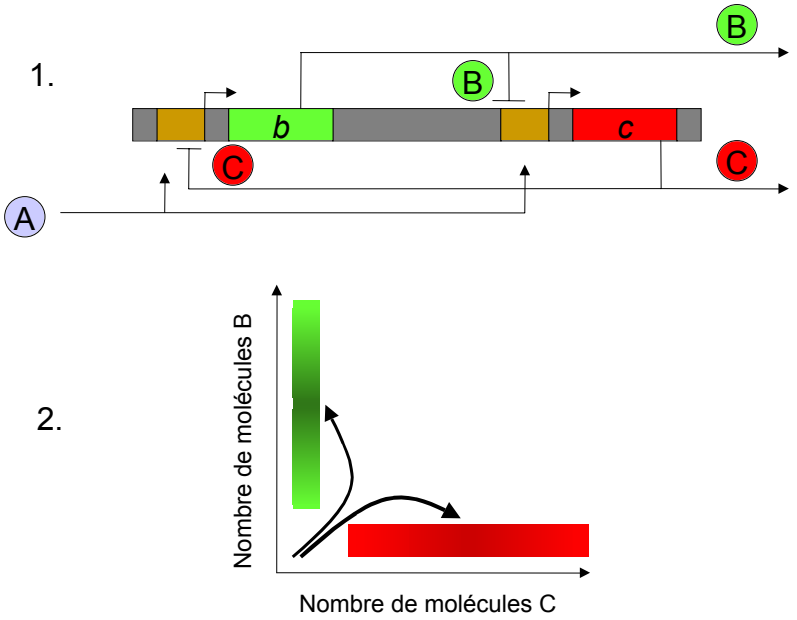


Figure 10 : Équilibre bistable.
1. Dans le modèle de régulation simple proposé ici, la protéine A active l'expression des gènes b et c. En retour, C inhibe l'expression de b et B inhibe l'expression de c. 2. Il résulte de ce modèle de régulation que seul B ou C peut être exprimé. Dans ce modèle, la sensibilité aux conditions initiales est très grande.

envoient de signaux différents, hypothèse compatible avec le déterminisme. Ici, on a constaté que les différents noyaux, au sein d'un environnement homogène, semblaient être indépendamment activés ou inactivés pour la transcription de certains ARN. Cette observation concernait aussi bien des gènes natifs que des transgènes. En outre le nombre de *loci* concernés est dynamique et varie en fonction de l'âge de la souris.

I. B 1 c La structure des boucles de régulation peut conduire à une expression stochastique.

Au sein même d'un fonctionnement déterministe, les données des simulations suggèrent que la structure même des boucles de régulation impose une dynamique stochastique au système (Figure 10). En outre, sur la base de modèles très simples de systèmes d'expression, dans lesquels, comme cela est fréquemment observé, le produit final, la protéine, a une activité répressive ou facilitatrice sur l'expression du gène, les simulations mettent en évidence que les protéines elles aussi sont produites par « paquets ». Ces paquets de molécules dont le nombre est variable sont eux-mêmes produits dans des intervalles de temps aléatoires. Quand ces différentes boucles se suivent au sein d'une cascade d'expression, l'effet en est d'autant plus amplifié. Il résulte de ces simulations que la séquence temporelle d'événements élémentaires peut largement varier d'une cellule à une autre et produire des effets de désynchronisation (McAdams et Arkin, 1999).

Ceci peut être corrélé à diverses observations empiriques chez les procaryotes (Mac Adams et Arkin, 1997) : réponses chémotactiques différentielles, variations de phases dans l'expression génétique du pili chez E. Coli, différences de nombre de

générations par unité de temps dans des cultures bactériennes, tous ces exemples étant bien évidemment considérés au sein de populations clonales. Ces phénomènes de désynchronisation semblent aussi se retrouver dans certaines études sur des systèmes eucaryotes, telles que les cellules Swiss 3T3 soumises à des variations de concentration en sérum (Brooks, 1985).

Les désynchronisations qui peuvent donc résulter de ces aléas dus à la structure même de la régulation de l'expression par ses produits finaux ou intermédiaires peuvent conférer à une certaine proportion de la population un avantage sélectif temporaire.

Par ailleurs, ces différences peuvent aussi conduire à une diversité phénotypique de la population. Les simulations montrent que des fluctuations à court terme dans la production d'une protéine peuvent survenir de manière large autour de la valeur seuil qui contrôle l'expression de certains gènes. Ces variations importantes peuvent se manifester notamment en temps nécessaire pour produire une protéine à une concentration donnée. La conséquence d'une telle variation en deçà ou au-delà d'une telle valeur seuil peut être de nature à engager explicitement le destin d'une cellule. A titre d'exemple (Arkin, *et al.*, 1998) on peut mentionner le circuit de décision par lequel les procaryotes *E. coli* infectés par le phage λ sont conduits à « choisir » entre phase lytique et lysogénie. Ce circuit est descriptible en termes de fluctuations aléatoires stabilisées dans un second temps. Les réseaux régulant l'alternative compétence/sporulation chez *Bacillus subtilis* semblent ressortir de la même logique (Grossman, 1995).

De plus, on a cherché à décomposer la stochasticité en une composante extrinsèque et une composante intrinsèque (Swain, *et al.*, 2002). La composante intrinsèque du

bruit tient au fait que même si toutes les cellules étaient parfaitement synchrones, les événements de transcription et de traduction n'interviendraient pas de la même manière dans chacune, et les gènes ne seraient pas forcément exprimés dans le même ordre (car le système n'est pas parfait, et qu'il existe des fluctuations élémentaires). Cette composante seule varie d'un gène à un autre puisqu'elle est dépendante du système d'expression d'un gène particulier. C'est en fait cette composante que nous avons évoquée plus haut. La composante extrinsèque tient à toutes les variables qui peuvent influer sur les mécanismes de transcription et de traduction : nombre d'ARN-polymérases ou de ribosomes, position de la cellule dans son cycle, disponibilité de la machinerie de dégradation de l'ARN et plus largement environnement cellulaire. De sorte que chaque composant cellulaire peut en réalité subir des fluctuations intrinsèques et à son tour être la cause d'une fluctuation extrinsèque d'un autre composant. La variation de la composante extrinsèque prise globalement a, elle, un effet sur l'ensemble des gènes. Des modèles ont récemment été mis au point pour permettre de discriminer les deux types de contribution à l'expression stochastique des gènes. Ces simulations ont permis de postuler qu'au sein de la composante intrinsèque, la transcription joue un rôle prépondérant au-delà de la valeur de deux protéines par transcrit ARNm. En deçà, les événements de traduction apportent une contribution significative au bruit.

L'ensemble de ces informations doit nous amener à relativiser les schémas traditionnels et statiques de l'expression déterministe des gènes au sein d'une cascade de réactions. Elles nous amènent à prendre conscience qu'au-delà de considérations extérieures, c'est la structure même de ces rétroactions qui aboutit à créer de la variabilité.

I. B 2 Le contrôle du bruit

Si certains auteurs semblent s'accommoder de l'idée qu'une expression aléatoire puisse être la base d'une diversité phénotypique, l'accent est souvent mis, face à cette variabilité perçue comme une perturbation, sur les mécanismes permettant de la restreindre ou de la contrôler. On peut noter que l'expression aléatoire a été tôt caractérisée comme une " transcription illégitime " (Chelly, *et al.*, 1989). Le choix de cette dénomination semble bien montrer que si l'on constate des variations par rapport à ce qui est "prévu" par le programme, il faudrait alors considérer ces variations comme anormales, et donc comme parasites et non pas comme utiles. La nécessité d'un contrôle de celle-ci devient, dans ce cadre, indispensable.

De fait, il existe un certain nombre de mécanismes de répression de ces aléas. Comme le notent Arkin et McAdams, « la conception des circuits de régulation génétique utilise la redondance, les rétroactions et d'autres mécanismes pour produire le déterminisme nécessaire des résultats à partir de circuits construits sur des éléments intrinsèquement générateurs de bruit. La fiabilité opérationnelle dans la régulation est particulièrement cruciale dans le développement des organismes métazoaires complexes ». En d'autres termes, c'est bien au niveau opérationnel, c'est-à-dire ici macroscopique que doit intervenir la fiabilité, nonobstant des événements élémentaires imprévisibles. Plutôt que de chercher à envisager une homogénéité intercellulaire de réponse, ces auteurs invoquent, au titre des mécanismes de régulation de la stochasticité, la coopération transcriptionnelle intercellulaire, qui permet de répartir des produits d'expression différentiellement produits. ils notent en outre l'existence de points de contrôle dans les cascades de

réactions, mais remarquent qu'ils n'entretiennent qu'imparfaitement le synchronisme intercellulaire. Ils s'attardent plus sur le troisième mécanisme de régulation, la redondance génétique. Dans le cadre de leur modèle, l'existence de gènes à copies multiples a tendance à diminuer drastiquement leurs aléas d'expression quand bien même ils sont intégrés à des cascades de régulations antérieurement décrites. Il est donc parfaitement envisageable que ce soit ce genre de mécanismes qui soit à l'origine de la canalisation de phénomènes aléatoires et permettent d'augmenter la probabilité d'expression de manière si proche de un qu'il aura une apparence déterministe même au niveau élémentaire.

Mais dire cela, c'est acter le fait que la base de l'expression génétique est fondamentalement stochastique, que cette stochasticité est tamponnée dans un certain nombre de cas où l'aléa est trop grand pour que sa reproductibilité soit assurée de manière satisfaisante. Ici, le génome joue bien son rôle de mémoire structurale de réponses évolutives sélectionnées car elles favorisaient certains phénotypes. L'information sur le nombre de copies d'un gène étant au moins aussi importante que la séquence du gène lui-même.

En outre, la canalisation, principalement due à la redondance génétique, est un argument qui ne doit pas être présenté de manière biaisée. Dans le cadre déterministe, la redondance est souvent invoquée à juste titre pour expliquer, en partie, l'adaptabilité d'un organisme face à par exemple, des mutations défavorables. Les copies intactes permettent d'assurer la fonction. Mais cette redondance semble parfois opportunément passée sous silence quand il faut expliquer l'évolution par mutations, et qu'elle pourrait alors être un handicap. Il en va de même avec, à l'inverse, la pléïotropie. C'est certes un principe d'économie qui permet de réaliser plusieurs fonctions avec le même gène (ce qui remet en cause le concept de

spécificité), mais une mutation sur ce gène peut mettre en péril plusieurs fonctions simultanément. Ces deux exemples, redondance et pléïotropie, dont la réalité n'est pas à contester, sont ainsi des problèmes pour le paradigme déterministe. Ils n'en sont pas dans le paradigme probabiliste, où ils sont des illustrations du caractère dégénéré des interactions moléculaires, reposant sur des affinités relatives, jamais absolues.

L'ensemble de ces données montre-il que la logique déterministe s'accommode en son sein de phénomènes stochastiques ?

Il ne nous semble pas pertinent de raisonner sur cette base. L'ensemble de ces mécanismes de contrôle présentés ici sont fondés sur le même présupposé de spécificité qui par ailleurs provoque dans ce modèle des fluctuations aléatoires. Le raisonnement semble circulaire. Par ailleurs, la redondance a aussi un coût en terme de contrôle des effets synergiques des éléments redondants qui se surajoute à celui du contrôle de ces éléments pris un à un. C'est pourquoi, même si ces propositions révèlent la stochasticité de l'expression génétique et la renforcent, il semble plus que jamais parcimonieux de préférer un système où la fiabilité est assurée de manière moins coûteuse que par des contrôles précis et coordonnés. C'est dans un tel système que peut apparaître spontanément de la variété phénotypique.

I. B 4 Génération de variété phénotypique.

Créer du différent dans un environnement initialement homogène est la grande question qui préside à la différenciation cellulaire. On a vu que dans le cadre

déterministe, il faut supposer une asymétrie entre la cellule qui reçoit le signal de différenciation et celle qui l'émet, ce qui ne fait que repousser d'un cran la question de l'origine de leur asymétrie : alors qu'on cherche à expliquer une différence, on en invoque une autre, qu'il faudra expliquer, et ainsi de suite.

Nous avons vu que le modèle probabiliste renverse cette perspective. Il présuppose que les cellules sont par défaut dans un état de désordre et donc qu'il existe une variabilité intercellulaire d'expression. Ce modèle porte, au gré des auteurs, plusieurs noms. Il est parfois qualifié de binaire par rapport à un modèle déterministe appelé rhéostatique (Fiering, *et al.*, 2000). Ici est exprimée l'idée que, face à une situation induisant une différenciation, les cellules vont y répondre soit en s'engageant dans la différenciation, soit en ne le faisant pas (avec une certaine probabilité) : c'est donc un choix binaire. Dans le cadre du modèle classique, chaque cellule y répondra de manière homogène en adaptant son taux de transcription pour un gène donné. Cette adaptation du taux de réponse est une réaction analogue à un rhéostat dont on fixe le niveau pour s'adapter à un nouvel environnement. D'autres appellations sont parfois données comme l'opposition entre réponse analogique et réponse numérique. Ces singularités lexicales recouvrent en réalité la même opposition. Et dans de nombreux cas, majoritaires dans les études où l'on cherche à discriminer entre les deux modèles, le modèle probabiliste semble s'imposer.

Ainsi dans un certain nombre d'exemples on peut envisager qu'il y a recrutement de cellules (donc hétérogénéisation des profils d'expression intercellulaire) plutôt qu'une augmentation de taux d'expression homogène dans chaque cellule (Walters, *et al.*, 1995).

-En réponse à une stimulation oestrogénique, la molécule ApoB, secrétée dans le foie (oiseaux) subit une augmentation de ce taux de sécrétion : ceci se fait par

recrutement de cellules qui étaient à l'état silencieux, plus que par augmentation du taux de sécrétion dans les cellules initialement actives (Lin, *et al.*, 1986).

-L'induction d'enzymes impliquées dans le cycle de l'ornithine chez le rat par des glucocorticoïdes (carbamoyl synthétase, phophoénolpyruvate carboxykinase et arginase) sur des hépatocytes fœtaux en culture est discontinue et semble répondre à un mécanisme stochastique : les probabilités de transcription de chacune des enzymes sont indépendantes.

-De même que l'induction des cytokines par présentation d'antigènes dans les populations clonales de lymphocytes T est stochastique et la plupart des cellules expriment seulement une des multiples cytokines exprimées par la population clonale (Bucy, *et al.* 1994,1995).

Comme le soulignent les auteurs (Fiering, *et al.*, 2000) qui réunissent ces exemples, on peut toujours répondre à ces faits que l'environnement peut varier d'une cellule à l'autre, expliquant de manière déterministe classique les différences de réactions. Cet argument, même s'il peut paraître peu convaincant (car il appelle inévitablement la question suivante : si l'on suppose des différences microenvironnementales, comment assurer alors la reproductibilité de celles-ci d'un organisme à un autre ?) doit néanmoins être étudiée. Mais comme il est souligné dans cette même étude, et comme nous l'avons indiqué plus haut, il est de moins en moins tenable quand il s'agit d'expliquer que des noyaux d'un syncytium soient différentiellement activés.

Il pourrait même être encore plus périlleux d'expliquer de manière déterministe les cas en nombre croissant de régulation hétéroallélique de l'expression génétique, ou en tout cas de faire reposer cette explication sur des variations microenvironnementales. On sait cependant que les mécanismes d'empreintes, de mieux en mieux connus, permettent d'expliquer en partie un certain nombre de ces

cas de figures (gènes des récepteurs d'antigènes des lymphocyte B et T). Néanmoins, certains autres n'ont pas (encore ?) été liés au phénomène d'empreinte génétique comme le récepteur Ly49 des cellules NK, et des récepteurs olfactifs (Höllander, 1999).

Faire reposer, en dernier ressort, la reproductibilité intercellulaire d'une expression génétique sur la conservation planifiée de différences micro environnementales à l'intérieur d'une même cellule semble au-delà de ce qu'il est raisonnable d'admettre au vu de nos connaissances sur le fonctionnement de la machinerie cellulaire.

Ainsi, par ces exemples, on a pu constater que les mises en évidence expérimentales aboutissent bien souvent à la caractérisation de sous-populations cellulaires qui se distinguent les unes des autres non pas par leur génome mais par leurs réactions différentielles dans un environnement identique. En cela, ces données rejoignent les cadres conceptuels déjà évoqués qui se sont révélés féconds pour certaines grandes disciplines de la biologie, la sélection clonale et la stabilisation neurale. Ces exemples, on l'a déjà dit, ne font pas preuve, pas plus que leur accumulation. Mais il sont déterminants en ce qu'ils ont été le fruit d'investigations qui cherchaient précisément à discriminer entre les deux modèles et qu'ils ont nettement rendu pertinent le modèle probabiliste. Ils ont même pu servir à montrer que si les deux mécanismes pouvaient être à l'œuvre (l'augmentation de la probabilité d'expression -et donc du nombre de cellules concernées- peut s'accompagner d'une augmentation du taux d'expression de ce gène dans chacune), c'est bien la composante probabiliste qui était responsable de la majorité de l'effet macroscopiquement observé. Ils sont la base expérimentale qui soutient une

proposition théorique visant à expliquer la différenciation cellulaire par un mécanisme de hasard/sélection.

I C. les limites des dynamiques endodarwiniennes.

I C 1 Des limites déjà atteintes ?

Un petit nombre de travaux décrivent des systèmes où la dimension stochastique de l'expression génétique est explicitement réfutée. Avant de les étudier, notons que le besoin de démentir une dimension stochastique semble parfois conférer à celle-ci, par ailleurs, un statut de règle plutôt que d'exception. Il semble d'ailleurs que les travaux considérés s'accordent difficilement sur ce que pourrait être une expression stochastique. Cela reflète peut-être le fait que, bien qu'abondamment documentée, l'expression stochastique des gènes est mal appréciée, mal comprise et n'aboutit que confidentiellement à des modèles cohérents qui, s'ils étaient connus de manière plus générale, éviteraient les confusions.

Dans un article récent, l'acquisition des récepteurs CD94/NKG2 par les cellules NK est présentée comme ordonnée et non stochastique (Lian, *et al.*, 2002). Les récepteurs de la famille CD94/NKG2 sont décrits comme des hétérodimères reconnaissant le complexe majeur d'histocompatibilité de classe 1 (CMH-1). Une autre famille de récepteurs du CMH-1 est décrite, les récepteurs Ly49. L'expression des deux familles de récepteurs est généralement décrite comme stochastique, et c'est dans ce contexte qu'il faut mesurer les exemples de non stochasticité décrits

qui ne seraient donc, sur le fond, que des exceptions à la règle. Mais on peut aussi s'interroger sur le sens de cette non stochasticité. Les cinétiques décrites montrent bel et bien qu'à chaque stade une fraction seulement des cellules exprime CD94 et les différents récepteurs NKG2. Ce qui est mis en évidence, c'est l'activation progressive de ces différentes expressions. Certains éléments exprimés sont détectés de manière plus précoce que d'autres, et ce de manière reproductible. C'est donc bien une expression ordonnée, mais qui ne semble pas ressortir *stricto sensu* d'une non stochasticité. Cette expression ordonnée et séquentielle est en tout cas parfaitement compatible avec l'hypothèse que des gènes aient des probabilités différentielles d'expression, qui se traduisent par des précocités variables d'expression, et qui peuvent être rétrocontrôlés de manière à amplifier le phénomène. C'est au sein de ce même système qu'un autre exemple de non stochasticité est documenté. L'expression des récepteurs de la famille Ly49 dans les cellules NK est, on l'a vu, décrite classiquement comme stochastique. Mais il en va autrement pour la coexpression de certains d'entre eux. La coexpression des récepteurs de la famille Ly49 est d'ailleurs elle-même, en général, décrite comme stochastique, ce qui revient à dire ici que les récepteurs sont exprimés de manière indépendante. mais une combinaison semble échapper à cette règle (la coexpression Ly-49D/Ly-49H) sur la base d'études de corrélations statistiques (Smith, *et al.*, 2000). Encore une fois, outre le fait que comme dans l'exemple précédent, il semble que cela soit une exception, cela est de nouveau parfaitement compatible avec notre modèle, qui n'exclut absolument pas que certains gènes exprimés puissent augmenter la probabilité d'expression d'autres gènes, y compris au niveau d'une seule cellule, ce qui peut conduire à ce qu'il soient exprimés de manière statistiquement corrélée. Et de manière générale, dans le cas de gènes partageant des homologies de séquence

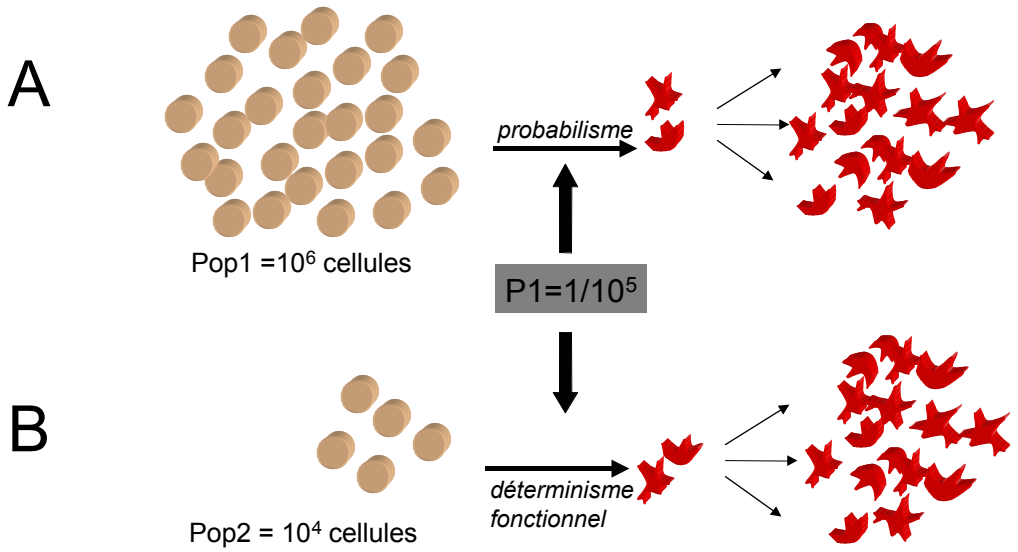


Figure 11 : Limite des dynamiques probabilistes.
Soit P1 la probabilité qu'une cellule de la population exprime la combinaison adéquate de gènes pour être adaptée à son milieu (toutes les cellules ont la même et les probabilités sont indépendantes). A. Si la taille de la population est supérieure à 1/P1, la probabilité qu'une cellule exprime la combinaison attendue est de 1. B. Si la taille de la population inférieure à 1/P1, mais qu'on observe systématiquement l'expression de la combinaison adaptée, c'est que des mécanismes ont permis d'augmenter P1 de tel sorte que l'on observe un déterminisme fonctionnel.

importante, cela se comprend d'autant mieux si l'on met en évidence que cette séquence est impliquée dans la régulation des gènes étudiés.

I C 2 Des limites prévisibles

Une théorie doit proposer le cadre de validité dans laquelle elle se déploie, et notre proposition a bien sur un tel cadre de validité.

En effet, l'hypothèse d'une dynamique endodarwinienne repose sur le fait qu'il existe une population cellulaire suffisamment grande pour qu'un phénomène de probabilité p puisse se produire, être sélectionné, et se répandre. Il est naturel ici d'admettre que tous les cas de figures ne se prêtent pas à une telle description. Pour le raisonnement, considérons que l'on parvienne à affecter une probabilité p1 à l'expression génétique d'un gène g1, pour une cellule donnée (Figure 11). Supposons par ailleurs que dans toutes les observations faites, l'expression de g1 soit détectable. Si la population cellulaire pop1 est de taille largement supérieure à 1/p1, on peut expliquer l'expression de ce gène en terme purement probabiliste. Supposons désormais que la taille pop1 de la population observée soit au contraire bien inférieure à 1/p1 : il devient illusoire de faire reposer l'expression de ce gène sur un mécanisme probabiliste. Pour le dire autrement, notre approche n'est applicable que dans le cas ou la population cellulaire est suffisamment large pour que la probabilité qu'un gène (ou une combinaison de gène) s'exprime soit égale à un au niveau de la population, quelle que soit sa probabilité au niveau d'une seule cellule. Dans les cas où cette condition n'est pas réunie, on est alors, mais alors seulement, légitimement fondé à faire appel à des modes de reproductibilité non pas statistiques mais déterministes, en rappelant bien qu'ils ne le sont jamais complètement et qu'il

s'agit d'un abus de langage fonctionnel. C'est bien pour cela que les deux approches sont complémentaires, pour peu qu'on définisse clairement leur domaine de validité. Pour cela, on peut soit augmenter la taille de la population, mais c'est justement le paramètre fixe ici, ou alors augmenter la probabilité que chaque cellule puisse exprimer la combinaison de gènes adéquate. On peut envisager l'évolution biologique comme un moteur permettant de répondre à cet impératif, en sélectionnant les individus sur ce critères les individus adaptés., en sélectionnant des gènes qui auront une action de contrôle sur les mécanismes permettant de canaliser les réactions. Cette canalisation peut avoir une apparence de déterminisme, sans être pour autant réductible à cette apparence (Figure 8, p. 54). Il est à noter que certaines atrophies d'organes sont irréversibles une fois que la taille de ce dernier passe un seuil plancher. Cela est compatible avec la problématique évoquée ici.

Ainsi nous avons cherché ici à montrer que la remise en cause du paradigme dominant était nécessaire au vu de son incapacité croissante à expliquer le réel sinon à le décrire. Nous avons par ailleurs proposé une grille de lecture, et montré que celle-ci s'appuie sur une littérature abondante et variée, constituée aussi bien d'expérimentation que de modélisation. Nous avons enfin tenté d'en borner des contours prévisibles.

Il reste maintenant à confronter cette proposition à des expérimentations discriminantes, ainsi qu'à des disciplines pour lesquelles il peut sembler aberrant de convoquer le hasard comme force structurante.

II. Une lecture non déterministe du développement.

II. A Introduction

L'embryologie, et plus particulièrement le domaine de la différenciation cellulaire, ont historiquement été des champs de recherche particulièrement marqués par les explications déterministes. La forte reproductibilité apparente de phénomènes qui se déroulaient dès les phases les plus précoces du développement, plaidait en faveur d'une explication déterministe. Dans un ouvrage auquel cette section doit beaucoup, et qui comporte de nombreux arguments potentiellement en faveur d'un développement de type probabiliste, Nicole le Douarin se fixe comme objectif de répondre à la question suivante : "Comment se peut-il que de cette cellule unique, "isolée", surgissent les constituants du corps de l'adulte, faits de plusieurs milliards de cellules harmonieusement ordonnées (...) ? ". Alors même qu'elle souligne à juste titre, on le verra plus bas, la compétition féroce qui existe entre les cellules pour les ressources au cours du développement, alors qu'elle démontre les relations trophiques entre les cellules, elle le fait dans le cadre de la persistance de la théorie du programme de développement (Le Douarin, 2000).

Cette rémanence de l'approche programmatique fut renforcée par le fait que les découvertes de la génétique et la théorisation de la biologie moléculaire accompagnèrent l'essor de l'embryologie expérimentale, et l'entraînèrent avec elle dans leur cadre conceptuel. Néanmoins, il convient de ne pas en rester aux

apparences et de voir comment, par un retour au sources à certaines étapes des progrès du domaine, l'apparente évidence du déterminisme peut être questionnée, et peut être vue sous un angle différent. Cette section n'a bien sûr pas pour but de reprendre l'intégralité de l'histoire de l'embryologie comme discipline scientifique. Cela dépasserait et de loin le cadre de notre travail. Elle souhaite mettre en évidence que le corpus d'expériences fondatrices en embryologie fait de cette discipline une théorie de la *plasticité régulée* beaucoup plus que du *déterminisme*. Et surtout, que l'interprétation qui en est faite classiquement, pour pertinente qu'elle soit, peut être modifiée. Les données expérimentales de l'embryologie ne font pas sens à elles seules : elles ne se suffisent pas en tant que telle pour discriminer un modèle déterministe d'un modèle probabiliste. Elles laissent la place à une variété d'interprétations.

Comme précédemment, il ne saurait, ici, s'agir de convaincre par simple accumulation d'exemples. Cela dit, on remarque là encore que la biologie du développement, s'est elle-même constituée autour d'exemples qui, quand ils ont été suffisamment convaincants, ont été progressivement généralisés à des degrés divers du monde vivant. Cette méthodologie est rendue nécessaire par la structure des expérimentations entreprises, inévitablement fondées sur des situations précises, et dans des espèces précises. Elle n'est donc pas mise en cause ici, mais elle ne saurait, en conséquence servir à mettre en cause une démarche de relecture critique de ces expériences et du cadre conceptuel qui les a justifiées.

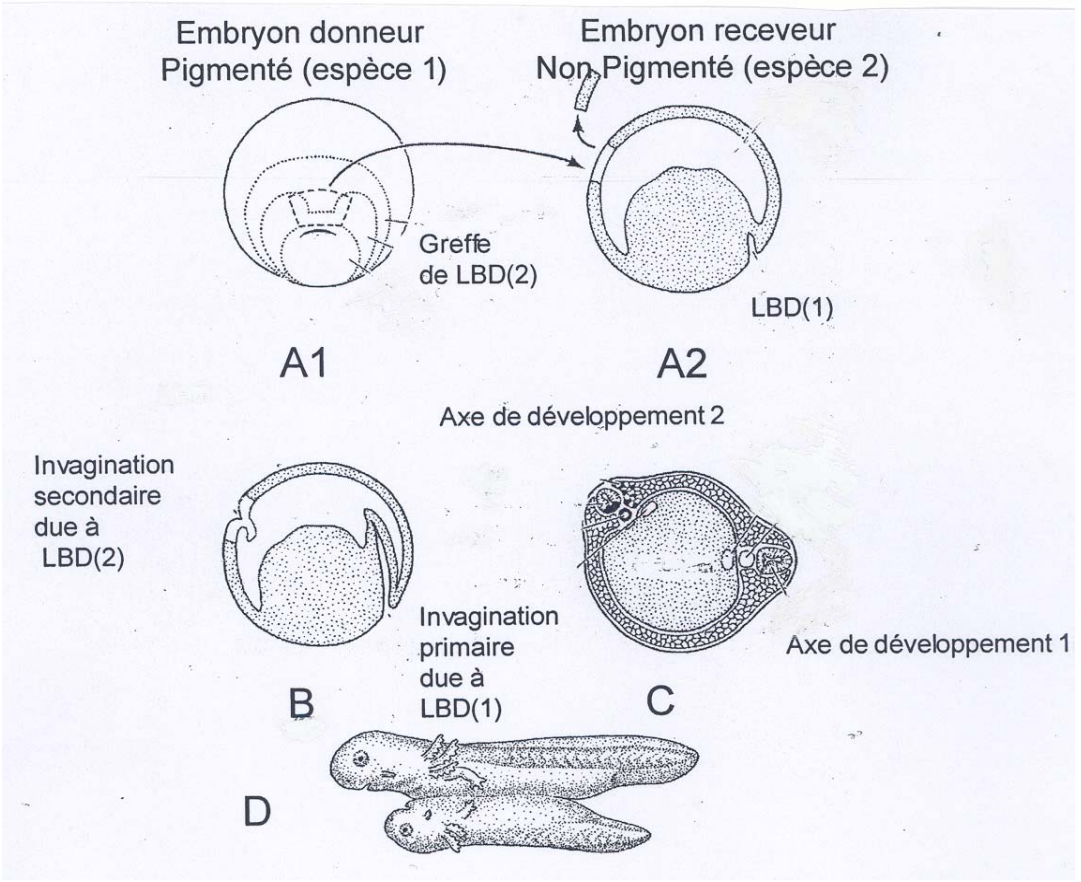


Figure 12 : Expérience de Spemann démontrant l 'existence d 'un organisateur.
Au stade Blastula, la greffe d 'une lèvre blastoporale d 'un embryon (A1) en substitution d 'une portion de l 'épiderme présomptif sur embryon receveur d 'une autre espèce (A2) aboutit, pour ce dernier, à la formation d 'une seconde lèvre blastoporale (B) puis d 'un deuxième axe de développement (C) et enfin de deux embryons siamois (D)
LBD : lèvre blatoporale dorsale.

II. B le déterminisme en embryologie : un choix d'observation contestable

II. B 1 Le déterminisme apparent des expériences fondatrices pose des difficultés d'interprétation

Si l'on considère les expériences de Spemann qui ont fondé les concepts modernes de l'embryologie expérimentale on peut, dans un premier temps, être fasciné par l'apparent déterminisme qui régit les étapes du développement. Dans son expérience célèbre publiée 1924, menée avec Hilde Mangold sur des espèces de tritons, il réalise un greffe de la lèvre blastosporale dorsale d'une gastrula de *T. taeniatus* sur le flanc d'une gastrula de *T. cristatus* (Spemann et Mangold, 1924*)*. Ces deux espèces de tritons diffèrent par la pigmentation des œufs fécondés et des cellules qui en descendent. Des greffes interespèces sont viables à ce stade très précoce. De cette sorte, on peut observer, dans l'organisme obtenu, l'origine des tissus par la différence de pigmentations.

Cette greffe aboutit à la formation d'un axe embryonnaire secondaire, en conséquence duquel deux embryons siamois se développent (Figure 12). Ces expériences et celles qui les ont complétées ont mis en évidence l'existence, dans la lèvre blastosporale, d'une région dite organisatrice, que l'on appela par la suite organisateur de Spemann. Ces expériences sont à la base du concept d'induction, l'idée que les cellules s'influencent pendant le développement, qui a été d'emblée interprété dans un cadre déterministe.

L'organisateur de Spemann fut décrit comme responsable de la mise en place de l'organisation de l'embryon et de ses organes qui matérialisent un axe

antéropostérieur. Elle ouvra la voie à une typologie de la surface de l'embryon en " territoires présomptifs " susceptibles de donner, par le jeu des interactions cellulaires du développement, des tissus ou des organes bien précis, en fonction de leur localisation par rapport à cet organisateur. Néanmoins Vogt, qui réalisa la première carte des territoires présomptifs, fit prendre conscience, dans une communication de 1928, que le rôle de l'organisateur de Spemann ne pouvait être compris si l'on ne présupposait pas, en complément, une hétérogénéité cellulaire telle que celle que l'on retrouve de fait dans l'ovocyte fécondé. On voit bien qu'ici, la théorie en cours d'élaboration, visant à comprendre comment les cellules se différencient, présente le handicap de présupposer une différence initiale pour y répondre.

L'hypothèse de Vogt fut d'ailleurs expérimentalement confirmée par la mise en évidence, en 1969, de l'existence d'un organisateur primaire, par Nieuwkoop, et qui porte son nom (Nieuwkoop, 1973). Encore fallut-il comprendre sa mise en place et présupposer, là encore une hétérogénéité préalable. Et de fait celle-ci est observable expérimentalement, puisque l'on sait que dès la fécondation de l'ovocyte un gradient de distributions des différents composants du zygote se met en place, et que pour certaines protéines, ces gradients sont mis en place pendant l'ovogenèse.

Mais on est alors confronté à un problème. Supposer que tout repose sur un gradient initial, à partir duquel une mécanique programmatique se met en place, est peu pertinent au regard de la complexité des phénomènes qui se succèdent au cours du développement. Le rôle des gradients dans la morphogenèse et le métabolisme est un sujet qui a été longtemps controversé, mais les travaux les plus récents sur les gradients métaboliques montrent qu'ils sont à la fois un élément de modulation de l'expression de gènes et un produit de celle-ci, ce qui permet difficilement de leur

attribuer clairement un strict rôle causal (Blackstone, 2003). Si l'on reconnaît qu'à certaines étapes clés, des gradients sont nécessaires, cela n'autorise pas à en déduire que tout le développement repose sur l'existence du gradient morphogénétique de l'œuf.

La réponse classique qui est faite à ce stade est généralement qu'il existe une combinaison entre la mise en place des gradients permettant une hétérogénéité préalable et l'épigenèse développementale qui expliquent le développement de l'embryon. L'hétérogénéité permettrait la mise en route d'un programme, et l'épigénèse lui apporterait une souplesse de réalisation. Mais l'on voit clairement ici qu'il s'agit bien d'une interprétation des phénomènes plutôt qu'une explication consensuelle : Tant que l'influence réciproque de chacune de ses composantes n'est pas déterminée, comment prétendre, sinon dogmatiquement, que l'une ne sert qu'à assouplir la rigueur de l'autre ? Pourquoi ne pas prétendre l'inverse ? Ce pourrait être une discussion sans fin si, comme on l'a déjà souligné, la vision probabiliste ne permettait pas réconcilier les deux approches.

De plus, si l'on en revient aux gradients de l'ovocyte, ceux-ci ne sont pas spontanés. Ils sont mis en place, quel qu'en soit le mécanisme, à une étape précédente qui impose de remonter de génération en génération, afin de trouver une cause dans une différence préexistante. Ce qui revient à abdiquer du pouvoir explicatif d'une théorie embryologique, ou d'accepter de la restreindre à une fonction de simple paraphrase du réel : ce qui existe est causé par ce qui précède. Ce que personne ne conteste, mais cela n'offre que peu d'intérêt pour un projet de compréhension du développement.

II. B 2 Un programme dégénéré et non spécifique est-il un cadre explicatif satisfaisant ?

On peut trouver des éléments qui contredisent l'idée de déterminisme dans le corpus même des expériences fondatrices de l'embryologie. Un certain nombre d'expériences, qui sont généralement moins diffusées que les spectaculaires résultats de Spemann et de ses continuateurs, tend à relativiser fortement la dimension spécifiquement organisatrice de la lèvre blastosporale. On pensait à l'époque, au vu des expériences de Spemann, que cette zone possédait des facteurs spécifiques, moléculaires, capables d'induire la neuralisation du tissu ectodermique. Des expériences ont contribué à mettre en évidence que cela était de l'ordre de l'interprétation. L'un des élèves de Spemann, Johannes Holtfreter, mit au point une technique de culture permettant d'entretenir des fragments prélevés sur le pôle animal d'une gastrula d'amphibien. Dans ce milieu, ces cellules se maintiennent et forment une vésicule de type ectodermique qui ne se différencie pas. Ce sont ces vésicules que Holtfreter va mettre en contact avec différentes structures dont on pense a priori qu'elles n'ont pas de pouvoir organisateur : lèvre blastosporale chauffée, tuée par l'alcool, ou séchée. Il les mit aussi en contact avec des tissus sans pouvoir inducteur supposé, comme des fragments d'intestin embryonnaire, auxquels il faisait subir une dessiccation. Il chercha enfin à les mettre en contact avec d'autres tissus inducteurs de groupes zoologiques très éloignés (mammifères, mais aussi reptiles, poissons, voire insectes ou annélides), dont on pouvait imaginer qu'ils ne seraient pas compatibles avec la gastrula d'amphibien. Et dans toutes ces expériences, il constata que l'induction de neuralisation pouvait avoir lieu. Comme le note Nicole le Douarin : " aucun dénominateur commun ne semblait exister entre les

nombreux tissus vivants ou morts et même les extraits acellulaires avec lesquels un effet neuralisant pouvait être obtenu ". Ceci fut confirmé de manière définitive quand Waddington et ses collaborateurs induisirent la neuralisation d'un tissu ectodermique de gastrula grâce à du bleu de méthylène. Ainsi la faculté de neuralisation semblait tenir plus à la plasticité et l'adaptabilité des cellules receveuses du " signal " que de la nature de ce signal (Le Douarin, 2000).

Ces résultats nuancent ceux de Spemann sans nier leur apport. Ils permettent de modifier l'interprétation que l'on se faisait initialement, sur la base des expériences pionnières du rôle de l'inducteur. En effet, avec ces expériences complémentaires, il apparaît que le rôle de l'organisateur de Spemann n'est pas d'apporter une information spécifique à tel ou tel territoire présomptif. La quête des régulateurs spécifiques s'est avérée être un échec. Or elle a été le programme de l'embryologie pendant des décennies.

Bien au contraire les molécules qui ont été identifiées sont à la fois extrêmement répandues et très banales, agissent à différentes étapes du développement, et leur rôle ne peut être prédit sans que l'on connaisse le contexte dans lequel elles interviennent (Dieterlen-Lièvre, *et al.*, 1996). Comment expliquer alors le pouvoir inducteur que malgré tout, Spemann a démontré, sans en donner les mécanismes, comme il le reconnaissait lui-même ? C'est en fait vers les cellules cibles que s'est alors tournée l'attention des chercheurs : les inducteurs quels qu'ils soient, semblaient être capables de déclencher une mécanique interne qui aboutit à la neuralisation de certains tissus. Mais dire cela, c'est aussi reconnaître que ces mécanismes et les incertitudes qui les entourent ne nous permettent pas de

discriminer entre une mécanique probabiliste ou déterministe de la neuralisation des cellules ectodermiques.

Ceci est encore accentué par d'autres expérimentations qui valent d'être mentionnées. Le concept d'induction neurale évolua sensiblement quand on remarqua de manière expérimentale que certaines cellules précurseurs des futures cellules nerveuses, prélevées tôt au stade blastula et mises en culture se différenciaient spontanément en neurones.

Ces observations combinées avec d'autres renversèrent radicalement la perspective du développement embryonnaire : elles installèrent l'idée que les substances inductrices avaient pour rôle d'inhiber des différenciations spontanées plutôt que d'en déclencher. Ainsi le rôle de l'organisateur de Spemann est d'inhiber la neuralisation dans les tissus ventraux plutôt que d'activer celle des tissus ectodermiques.

Or cette approche ressort typiquement d'une grille de lecture endodarwinienne. Elle peut s'inscrire dans une logique où les cellules ont spontanément tendance à explorer différentes voies susceptibles de les conduire à se différencier en un type cellulaire en fonction des ressources disponibles. C'est seulement dans un second temps que cette tendance spontanée, ce comportement exploratoire, est inhibé, ou stabilisé *a posteriori*, et peut ici expliquer l'homogénéisation de la population en neurones. Une dynamique de hasard/stabilisation est apte à décrire des mécanismes comme ceux-ci, où il est vraisemblable de penser que ces cellules ne reçoivent pas de signal particulier, mais explorent spontanément les différentes voies qui les conduisent progressivement à un phénotype adapté au milieu.

II. B 3 Les gènes du développement, nouvel avatar du déterminisme ?

La découverte de gènes impliqués dans la mise en place du plan de l'embryon (Lewis, 1978) a permis de redonner de la vigueur à la conception déterministe de l'expression génétique, notamment le constat de leur action reproductible sur des populations de cellules ciblées très précisément, qui semble ne laisser que peu de place au hasard. La colinéarité des gènes *hox*, qui les place, sur le chromosome, dans le même ordre que celui de leur activation temporelle, et ce de manière très conservée dans le règne animal, est de nature à conforter un fonctionnement déterministe. Il n'est pas inutile de se demander dans quelle mesure leur fonctionnement n'est pas un contre-exemple, voire une réfutation des dynamiques proposées par l'approche endodarwinienne.

Il convient pour cela de rappeler que si le rôle de ces gènes est crucial dans les phases de développement, il n'est pas pour autant strictement causal. Le rôle des gènes est souvent étudié par mutation ou délétion, et observation des conséquences de sa perte de fonction. Par exemple, on peut ainsi montrer que, dans de très diverses espèces, des gènes orthologues sont impliqués dans la formation de l'œil (*pax6* chez la souris, *eyeless* chez la drosophile) et que l'on peut même les substituer l'un à l'autre pour restaurer ce rôle. Cela incite à en faire les "gènes de l'œil". Cette simplification est cependant abusive. Leur surexpression dans des tissus différents, chez des individus adultes, ne donne pas d'œil (Schedl, *et al.*, 1996). Cela souligne que ces gènes ne sont pas à eux seuls les *déterminants* de l'organogenèse de l'œil. Ils sont des partenaires nécessaires de sa formation, mais sont intégrés dans un réseau d'interactions complexes, dépendant en outre du contexte épigénétique. De nouveau, on n'explique pas l'apparition d'un caractère par

l'expression d'un gène, et il est indispensable de faire appel à d'autres variables pour en comprendre le fonctionnement. De surcroît, notre connaissance actuelle du fonctionnement des gènes homéotiques ne nous permet que d'établir des corrélations entre leur expression et l'apparition des structures dans lesquelles ils sont impliqués. Les mécanismes cellulaires correspondant ne sont pas connus. Rien, par exemple, ne permet actuellement de conclure que leur mode d'action est univoque.

Il importe donc de ne pas confondre l'importance relative d'un gène et le déterminisme de son expression. Evelyn Fox Keller note, à l'appui de cette nuance que "la complexité de la dynamique du développement a sapé complètement la pertinence des gènes comme *causes* du développement" (Fox Keller, 2000). Le paradigme probabiliste ne s'oppose absolument pas à l'idée qu'un gène particulier puisse avoir un rôle cardinal dans certaines fonctions. Il postule justement que dans ce cas, les mécanismes évolutifs auront précisément permis de maximiser sa probabilité d'expression dans le contexte où il agit, de sorte de parvenir à un déterminisme fonctionnel de son expression (figure 8, p. 54).

Dans le cadre de cette probabilité d'expression relative, même si en apparence très proche de 1, on peut ici souligner une propriété notable des gènes *hox*. Nous avons souligné la colinéarité séquentielle et temporelle de leur organisation. Il a été démontré qu'une troisième forme de colinéarité se surajoutait à celles-ci. L'affinité des protéines Hox correspondantes pour leur séquence cible diminue selon cette même colinéarité. Pour le dire autrement, plus un gène *hox* est en 5' dans le complexe homéotique (colinéarité séquentielle), plus il est exprimé tardivement (colinéarité temporelle) et moins l'affinité de la protéine Hox correspondante pour sa séquence cible est forte (Pellerin, *et al*., 1994). Cette séquence est le motif C/GTAAT.

Cette décroissance d'affinité doit avoir un sens biologique. On comprendrait en effet mal que la sélection naturelle ait favorisé une affinité maximale pour certaines protéines Hox (codées par des gènes en 3' de l'homéodomaine, exprimés précocement) et pas pour d'autres, qui plus est codées par des gènes voisins des premiers, si cela était également nécessaire. Il est pertinent de penser ici qu'une affinité forte est nécessaire dans les premiers stades du développement embryonnaire, où les populations cellulaires sont petites et où la reproductibilité statistique d'un phénomène pourrait ne pas être vérifiée. Cette affinité pourrait sans dommage être plus relative dans les phases tardives, permettant en outre aux protéines d'adopter un comportement exploratoire leur permettant de balayer un spectre plus large de séquences. Nous sommes peut-être ici devant la matérialisation explicite d'un mécanisme permettant d'aboutir graduellement à un déterminisme fonctionnel à partir d'un fonctionnement probabiliste. A une autre échelle, on peut d'ailleurs remarquer que certaines étapes clefs du développement ont eux aussi une variabilité étonnante : c'est le cas de certains gradients (mais pas tous), comme le gradient de la protéine bicoïd, qui est structurant dans la théorie de l'information de position, et qui varie cependant d'un embryon à l'autre. Des mécanismes de correction d'erreur *a posteriori* sont postulés pour expliquer cette diversité (Houchmandzadeh, *et al.*, 2002). Cela ressort typiquement d'un fonctionnement adaptatif par essai/erreur, plus que d'un programme finement régulé. Cette souplesse peut être rapprochée des récents résultats sur la réactivité d'embryons de drosophile aux variations de pression mécanique expérimentalement exercée sur eux. Ces pressions peuvent suffire à induire l'expression ectopique de gènes du développement comme le gène Twist, impliqué localement dans le changement de forme des cellules du mésoderme lors de l'invagination de celui-ci au

cours de la gastrulation (Farge, 2003).Ces résultats nuancent eux aussi fortement la notion de programme de développement. Pour reprendre les mots de Frederik H Nijhout, "les gènes ne fournissent pas les instructions pour le développement, mais ils aident à fournir les bases matérielles du développement" (Nijhout, 1990).

Dans ce même article où il tente de se défaire des métaphores dont tout généticien du développement est tributaire, il remarque au passage que certaines expériences montrent que des mutations dramatiques comme *Antennapedia*, dans des conditions particulières, n'ont jamais une pénétrance complète et qu'il existe une variété de réponses possibles à ces mutations (Schneuwly, et *al.*, 1987).

Cette dernière précision doit nous amener à rappeler, comme pour tout phénomène d'apparence déterministe, qu'un biais d'observation peut venir de notre habitude à regarder les succès et négliger les taux d'échecs. Le déterminisme apparent du fonctionnement des gènes homéotiques doit là encore être mis en balance avec le fait que, quelles que soient les espèces, tous les embryons ne donnent pas des adultes viables, et que ces pertes sont peut-être le prix à payer pour que des mécanismes complexes puissent dans certains cas s'enchaîner harmonieusement et aboutir à un développement complet.

II. B 4 La plasticité des devenirs cellulaires pose un problème à l'interprétation programmatique

-Dédifférenciation - Transdifférenciation

La capacité de régénération n'est pas une observation récente. Il y a plus de 230 ans Spallanzani observait que des salamandres adultes pouvaient régénérer leurs

membres, leur queue, ou leur mâchoire, et espérait dans la conclusion de son traité qu'un jour l'homme puisse être doté de pareilles capacités. On sait aujourd'hui que ces capacités régénératives sont principalement le fait de deux mécanismes qui ne sont pas indépendants l'un de l'autre : l'existence de cellules souches adultes, et la capacité de certaines cellules dans certaines espèces à se dédifférencier et transdifférencier pour remodeler un organe ou un membre. Nous nous intéresserons principalement au second mécanisme dans cette section sur la plasticité cellulaire.

La dédifférenciation est un mécanisme qui peut survenir, notamment chez les salamandres à la suite d'une perte ou d'une amputation d'un membre. Autour de cette zone, les myofibres, les cellules de Schwann, les cellules périostéales et le tissu conjonctif se dissocient pour (re)former des cellules mononucléées qui migreront et se transdifférencieront de sorte que le membre soit reconstitué. Ce phénomène est particulièrement frappant pour les cellules musculaires striées, hautement différenciées et multinucléées qui se donnent en cellules mononucléées à proximité du site d'amputation.

Bien que limité en apparence, ce mécanisme est un fait biologique de portée générale puisqu'elle concerne le règne végétal dans sa grande majorité, les animaux diblastiques (Cnidaires, notamment l'hydre) comme les triblastiques (Planaires). Elle concerne aussi certains vertébrés comme les amphibiens. Chez les mammifères, les capacités de dédifférenciation ne semblent pas être nulles (Sanchez Alvarado, 2001), même si elles ne sont pas observées in vivo. Des résultats expérimentaux récents tendent à prouver que des facteurs synthétiques peuvent provoquer la dédifférenciation de myotubes de souris (Rosania, *et al.*, 2000), ainsi que certains facteurs de transcription introduits par transgénèse (Odelberg, *et al.*, 2000) qui permettent la dédifférenciation d'une fraction des cellules qui retrouvent alors leur

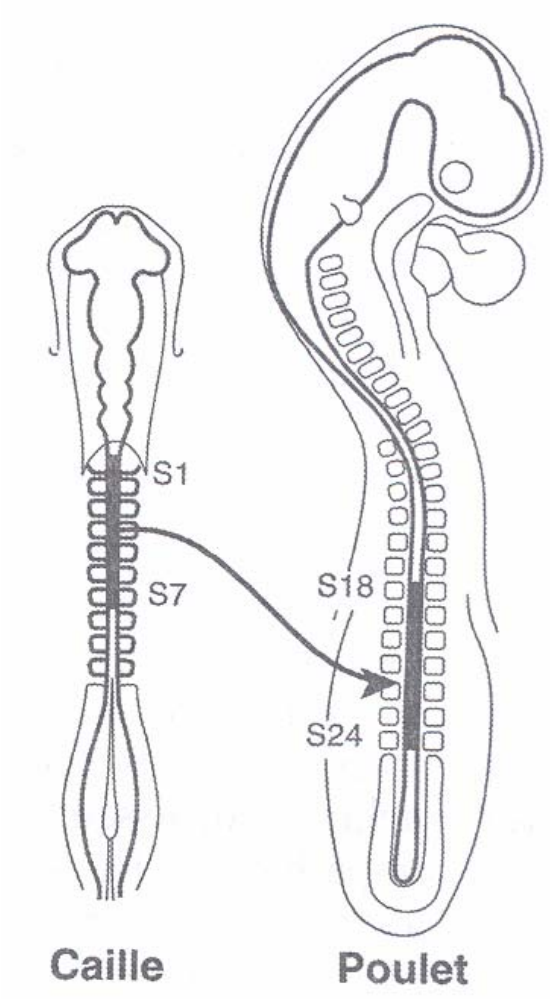
pluripotence. Enfin des myotubes hybrides triton/souris ont permis de montrer que lorsqu'on active en culture la dédifférenciation, le noyau murin reprend son cycle cellulaire et sort de son état différencié post-mitotique (Velloso, *et al.*, 2001), renforçant l'hypothèse que ce n'est pas le noyau, mais plutôt l'incapacité globale de la cellule murine à réagir au sérum qui est responsable du maintien de l'état cellulaire différencié. Cela est à mettre en relation avec le fait que toutes les cellules ne sont pas capables de dédifférenciation / transdifférenciation. Chez les amphibiens ce sont les cellules du système nerveux, de l'intestin, de l'œil et du cœur qui sont concernées, sans savoir si cette liste est définitive.

Comment interpréter cette dédifférenciation, d'un point de vue endodarwinien ? On peut d'abord remarquer que cette propriété est un défi aux typologies cellulaires, qui sont indissociables d'une approche déterministe. En effet, on explique souvent que telle cellule réagit à une variation du milieu en tant qu'elle appartient à un type cellulaire particulier, possédant des marqueurs caractéristiques de ce type. Une position typologique précise permettra ainsi de comprendre une réponse cellulaire précise. Cette catégorisation sert de grille de réponse pour expliquer la dynamique de réaction de la cellule. L'exemple de la dédifférenciation montre que la cellule peut aussi répondre en échappant à ces catégories voire en passant de l'une à l'autre, ce qui ne justifie plus une approche cloisonnée des typologies cellulaires. Celles-ci doivent être considérées, nous semble-t-il, comme des profils moyens, et provisoires. Et l'on ne saurait fonder avec certitude la réaction présumée d'une cellule sur son appartenance à tel ou tel type cellulaire.

La dédifférenciation/transdifférenciation peut-elle être vue comme un programme ? Là encore, la littérature nous montre que les facteurs de dédifférenciation sont non seulement communs à tous les types cellulaires susceptibles de subir le phénomène,

mais sont en outre des molécules non spécifiques de cette fonction. Ce sont principalement des facteurs de transcription de la famille des facteurs de croissance fibroblastique (FGF), ou d'autres comme le facteur de transcription Msx-1 et Msx-2. L'ensemble des études disponibles ne permet pas de déterminer un gène, parmi ceux-ci ou d'autres, qui soit le déterminant de la dédifférenciation, et des cofacteurs sont recherchés. Ces gènes sont impliqués, leur expression est corrélée au phénomène, mais il semble difficile, là encore, de comprendre la logique d'un hypothétique programme de dédifférenciation. En outre, observation importante, les myotubes soumis à induction par Msx-1, qui apparaît comme un acteur majeur du phénomène, ne sont que partiellement stimulées à devenir des cellules mononucléées : seules 9% d'entre elles retrouvent leur potentiel pour se transdifférencier ultérieurement en ostéoblastes, chondrocytes, ou adipocytes. Soulignons ici que cette action sur une partie seulement des cellules, est typique d'un fonctionnement probabiliste de l'expression génétique (Echeverri et Tanaka, 2002).

Il semble donc difficile, d'un point de vue expérimental, de parler de reprogrammation. D'autant que nous avons souligné (cf. supra) que cela est aussi très périlleux, d'un point de vue logique. Si l'on accepte de dire que tout ce qui n'est pas programmé peut être considéré comme reprogrammé, on dévoile la nature assez rhétorique de l'ensemble. Car à considérer que tout phénomène physiologique soit une perpétuelle reprogrammation, on finit par vider le mot de son sens.



Tiré de le Douarin, 2000

Figure 13 : Exemple de greffe interespèces.
Greffe des somites 1 à 7 (niveau vagal) de caille en substitution des somites 18 à 24 (niveau brachial) de poulet.

- Le cas des cellules de la crête neurale

Les cellules de la crête neurale, étudiées notamment sur les embryons de poulet, proviennent d'une région précise de l'ectoderme et possèdent un nombre de propriétés particulières. En nombre limité au départ, elles sont pluripotentes, et nous en verrons ci-dessous les conséquences. Elles ont une capacité migratoire très importante qui leur permet de quitter leur localisation d'origine pour gagner de nombreuses régions de l'embryon, dans lesquels elles ont des devenirs très différents. Elles participent à la formation du système nerveux périphérique, où elles peuvent jouer le rôle de neurones ou de cellules annexes (cellules gliales, cellules de Schwann). D'autres donnent des cellules endocrines, d'autres des cellules pigmentaires, d'autres participent à la formation du squelette, de la cornée, de muscles ciliaires. Cet inventaire non exhaustif veut illustrer que la pluripotence de ces cellules est très large. Elles sont pour ces raisons un sujet d'étude classique des mécanismes du développement embryonnaire.

On peut dresser des cartes des territoires présomptifs de ces cellules dans la crête neurale. Le terme de territoire présomptif ne doit pas nous induire à penser que chaque cellule de ces territoires est déterminée dans un destin particulier. Le terme de "présomptif" exprime bien, subtilement, que l'appartenance à un territoire est corrélée à un devenir "présumé", mais celui-ci n'est aucunement fixé de manière rigide. Les cartes donnent une topographie des devenirs qui pourrait laisser penser, à tort, que ces devenirs sont gravés dans le marbre. Mais on doit interpréter ces cartes à la lueur des nombreuses expérimentations complémentaires qui les accompagnent et qui, elles, mettent en évidence la plasticité de ces devenirs.

Des expériences de ce type reposent sur des xénogreffes qui sont compatibles entre deux espèces proches, par exemple caille et poulet. De plus on est capable, par des marquages avec des anticorps spécifique de la caille, de tracer le devenir des cellules de cailles greffées sur un poulet. Pour évaluer la plasticité des cellules de la crête neurale, on procède de la manière suivante. On greffe un fragment de crête neurale de caille dans un embryon de poulet, à un niveau de l'axe nerveux autre que celui dont il provient chez la caille. Par exemple, on remplace la région des somites 1 à 7 d'un embryon de poulet par celle des somites 18 à 24 d'un embryon de caille. On parle alors de transplantation hétérotopique (Figure 13). La question qui se pose alors est de savoir si, dans l'embryon de poulet greffé, les cellules du greffon vont se comporter en fonction de leur localisation initiale ou de leur nouvelle localisation. On constate alors que les cellules se comportent en fonction de leur nouvelle localisation. Non seulement elles vont migrer dans les territoires qui correspondent à cette nouvelle localisation, mais en outre elles y assureront les mêmes fonctions (si elles deviennent des neurones, elles secréteront les mêmes neurotransmetteurs) que les cellules qu'elles remplacent. Ceci montre que ces cellules, pourtant localisées dans une zone précise dans la crête neurale de la caille, présentent une certaine plasticité quand on les change de zone. Et l'on peut interpréter cette plasticité non pas comme l'exécution d'un programme, mais plutôt une sensibilité de ces cellules au conditions environnementales et tout particulièrement trophiques qui motivent leur migration. Cette plasticité, encore une fois, semble mal s'accommoder avec la prédétermination qui est sous-jacente dans la notion de programme. La possibilité de greffe elle-même contredit l'approche programmatique, ou du moins n'est compatible qu'avec une approche très nuancée du "programme", où la principale faculté de celui-ci serait son adaptabilité à en mettre un autre en route...

II. C Une relation trophique au cœur du développement

II. C 1 Les sous entendus de la métaphore informationnelle

On pourrait ici objecter que cette plasticité est justement une des conditions qui permettent de répondre à des signaux et réaliser le programme génétique. Outre le fait qu'il est toujours ambigu de vouloir faire coïncider la notion de programme avec l'adaptabilité manifeste des cellules, il semble nécessaire de regarder dans le détail la nature réelle des signaux échangés. Si ceux-ci n'apportent aucun bénéfice à la cellule qui les reçoit, les intègre, et y répond de manière appropriée, on peut effectivement les considérer comme de l'information. Mais si l'on met en évidence que ces "signaux" représentent une ressource pour la cellule receveuse, qu'elle les métabolisent et qu'elles réagissent de manière aléatoire à ce "signal" alors la vision programmatique n'a plus de pertinence.

En effet, le modèle déterministe repose sur la notion de transmission d'information. Une cellule receveuse reçoit une signal, ou encore instruction. Ce vocabulaire apparemment neutre est en fait lourd de sens. Présupposer qu'une substance reçue ait un rôle de signal n'est pas anodin, et en tout cas ancré dans une métaphore informationnelle qui a une dimension historique.

Le premier présupposé est celui de la spécificité du signal lui-même. Si la substance reçue n'est pas qu'une substance mais un signal, c'est qu'elle véhicule une unité d'information dont la spécificité va déclencher une réponse appropriée de la part de

la cellule réceptrice. Nous avons soutenu que la spécificité n'existait pas en tant que telle, en biologie. Les régulations transcriptionnelles sont dégénérées, les voies de transductions sont dégénérées (Edelman, 1978 ; Duboule, 1998), mais l'on persiste à chercher les signaux spécifiques de chaque situation. Et quand on en trouve ils sont souvent multiples, et interviennent souvent dans des régulations contradictoires. En résumé, ils ne sont pas en tant que tel des signaux.

Par ailleurs, recevoir des signaux et y répondre, même si l'on nuance ces termes, c'est en substance obéir à un dessein extérieur auquel la cellule serait asservi. C'est être la partie d'un tout au bénéfice duquel on agit de manière appropriée et potentiellement altruiste puisque l'intérêt du tout peut ne pas coïncider avec l'intérêt de la cellule. Or rien ne vient jamais expliquer comment une cellule peut faire primer l'intérêt à long terme de la population sur son intérêt immédiat. En terme évolutif, cela s'appelle de la sélection de groupe et l'on sait que les sélections de groupe ne sont acceptables que dans la mesure ou l'on découvre que chaque individu (ici, chaque cellule) tirait un bénéfice immédiat de ce comportement. Bénéfice non évident au premier abord.

II. C 2 La relation trophique au cœur des échanges intercellulaires

- Portée et limite d'une métaphore de substitution

Comment alors, si l'on met de côté la métaphore programmatique, donner un cadre cohérent à l'activité de chaque cellule prise isolément, dans la mesure où l'on

constate que globalement ces actions individuelles assurent généralement la survie de l'ensemble ici, l'organisme ?

Peut-être en regardant de plus près ce que nous avons l'habitude de qualifier de signal. Et en remarquant que ces " signaux " sont souvent, des facteurs trophiques. De nombreux exemples existent et permettent de donner une lecture trophique aux phénomènes qui président à la différenciation cellulaire, notamment dans les expériences les plus spectaculaires. Une illustration est donnée par les travaux de Hamburger et Levi-Montalcini.

Viktor Hamburger, dans les années 1940, travaillait sur le développement des neurones moteurs de la moelle épinière. Son modèle expérimental était le poulet. Cherchant à étudier l'adaptabilité des centres nerveux en fonction des changements d'environnement, il greffa, sur un embryon de poulet, un bourgeon supplémentaire qui permettait d'envisager d'obtenir un membre surnuméraire. C'est effectivement ce qu'il obtint, mais il remarqua en outre que cela avait provoqué là une hyperplasie, c'est-à-dire une augmentation du nombre de neurones en relation avec une hyperactivité fonctionnelle, qui avait été induite par la greffe de nouveaux tissus cibles de l'innervation. Comme le note Nicole le Douarin, " il émit l'hypothèse que les tissus cibles de l'innervation produisaient probablement des substances de type trophique qui transportées de manière rétrograde de la synapse au corps cellulaire du neurone, étaient responsables de l'hyperplasie constatée ". La collaboratrice d'Hamburger, Rita Levi-Montalcini, obtint des résultats sur les neurones des ganglions sensoriels similaires à ceux d'Hamburger sur les motoneurones. Ces résultats ainsi que d'autres, complémentaires, leur permirent de valider progressivement l'hypothèse d'un rôle trophique de l'innervation neurale (Le Douarin, 2000). On voit qu'ici, ce que nous serions tentés d'interpréter classiquement comme

un signal est en réalité une ressource pour laquelle la cellule à un intérêt immédiat. Cela active une réaction exploratoire du tissu neuronal dont la masse et le nombre de branchements s'accroît à la mesure de ce que la nouvelle source permet.

Un autre exemple de relation trophique peut conduire celle-ci à être abusivement présentée comme un signal. Dans le cadre des études de dédifférenciation, on peut entretenir des myotubes de tritons, indéfiniment, en culture. Ce sont, on l'a vu, des cellules multinucléées. Quand on enrichit le milieu en sérum de veau, de mouton ou de poule ces cellules repassent à l'état mononucléé (Tanaka, *et al.*, 1999). On pourrait donc ici dire que les cellules par la fourniture ou la privation de sérum, reçoivent des signaux leur indiquant la voie à suivre. N'est-il pas plus simple de considérer que la différenciation est ici un stress correspondant à la privation de sérum, entraînant une adaptation sélective lors du passage d'un milieu nutritif à un milieu non nutritif et pour laquelle le phénomène inverse se produirait donc par enrichissement du milieu ?

Dans ces deux exemples, la métaphore du signal apparaît superflue voire trompeuse. Le comportement de la cellule peut s'interpréter en tant que tel et non pas dans le dessein d'un organisme supérieur. Que les intérêts à tous les niveaux coïncident, c'est ici évident. La cellule se nourrit et prolifère, et l'organisme se développe. Mais il y a bien un intérêt immédiat pour la cellule individuelle. Et rien n'autorise à penser que l'intérêt global prévaut. Bien plus logiquement, il semble parcimonieux de proposer que ce soit au contraire quand les cellules, individuellement, trouvent leur intérêt immédiat, dans la limite des contraintes imposées par leur génome, que peuvent se développer les conditions qui rendent possible un organisme multicellulaire. C'est sur la base de ces hypothèses que des travaux théoriques ont

pu proposer une place prépondérante pour la relation trophique dans la construction de la multicellularité (Kupiec, 1986, 1997).

Il faut bien sûr apporter de la nuance à cette proposition. Une cellule n'est pas réductible à une entité réagissante dans l'immédiateté. Elle est aussi le produit d'une histoire évolutive, dont le génome est l'archivage. La cellule répond aux variations environnementales avec le stock de réponses évolutives mémorisées que sont les gènes. Que ceux-ci aient été sélectionnés parce qu'ils ont permis le succès évolutif des individus qui les portaient, cela n'est pas contesté ici. Il n'en demeure pas moins qu'ils forment un ensemble par lequel chaque cellule doit pouvoir elle-même trouver des moyens de survie, et que faire l'impasse sur ce niveau d'organisation n'est pas cohérent dans l'approche endodarwinienne. Nous verrons à ce titre que même les gènes impliqués dans l'apoptose peuvent être soumis à cette grille de lecture.

Par ailleurs il importe de préciser ce que nous entendons par "ressource". Il est parfaitement légitime d'opposer à cette proposition qu'un certain nombre de "signaux" échangés n'ont pas de valeur énergétique ou métabolique particulière. A tout le moins, la survie des cellules n'est pas menacée par l'absence d'un grand nombre de signaux (les cellules du système immunitaire peuvent attendre un vie durant l'arrivée hypothétique de l'antigène sans mourir pour autant). Passons nous donc alors d'une métaphore "programme" à une métaphore "écosystème" qui possède autant d'approximations que la précédente ? Nous ne le pensons pas. Certes, cela reste une métaphore, une tentative allusive d'unifier nos propositions, mais cela n'implique pas une correspondance totale des deux éléments comparés. Toute métaphore a ses limites. Dans le cas de celle de l'"écosystème", une de ces limites est qu'un écosystème ne se reproduit pas. Un autre est donc celle de la définition de ce qu'est une "ressource". Dans le modèle endodarwinien, les signaux

sont fonctionnellement perçus comme des ressources en ce qu'ils permettent à la cellule de se multiplier, quel que soit le mécanisme déclencheur de cette multiplication. Cela n'implique pas formellement qu'ils aient en tant que tels une valeur métabolique particulière (Atamas, 1996).

- Relation trophique et vascularisation

Un autre exemple que l'on peut donner ici, sur la réinterprétation de la différenciation en terme d'adaptation aux ressources, est le cas des cellules de la crête neurale qui interviennent dans la vascularisation. Certaines de ces cellules sont décrites comme " destinées " à constituer l'endothélium vasculaire, c'est-à-dire la partie la plus interne, des vaisseaux sanguins irriguant les centres nerveux. La différenciation de ces cellules s'accompagne de la mise en surface d'un récepteur à un facteur de croissance vasculaire dit Vascular Endothelial Growth Factor (VEGF). LE VEGF est produit par le cerveau en formation et il induit alors les cellules endothéliales issues de la crête neurale à pénétrer dans l'épithélium neural pour constituer l'endothélium vasculaire. On semble donc avoir ici un exemple classique de signal émis provoquant une différenciation. Mais il est instructif de prêter attention au contexte de la sécrétion de ce facteur. Celle-ci a lieu dans le tube neural, mais dans un premier temps, le gène codant pour VEGF n'est pas exprimé. C'est seulement quand les cellules se retrouvent en situation d'anoxie, due à l'épaississement progressif de l'épithélium et l'impossibilité croissante pour l'oxygène de diffuser au travers, qu'elles vont secréter ce facteur qui va induire la vascularisation et favoriser la livraison d'oxygène (Le Douarin, 2000). Vu de cette manière, VEGF n'est pas un signal de différenciation mais la simple conséquence d'un stress cellulaire. Les cellules ne cherchent pas ici à communiquer entre elles, mais l'une émet des molécules

particulières à son état de stress et les cellules receveuses métabolisent cette molécule et vont à sa source. Il y a donc une relation trophique simple entre ces deux types cellulaires, au sens où elles ont un intérêt immédiat dans ce qui s'apparente à l'échelle de l'organisme comme une coopération intercellulaire mais qui est en fait une relation de symbiose. Ces résultats concernant la vascularisation et l'angiogenèse sont généralisables. L'hypoxie ou l'anoxie sont souvent impliqués dans le déclenchement des cascades menant à la vascularisation.

- Relation trophique et apoptose

Dans un ordre d'idée similaire, on peut ainsi et de manière symétrique envisager l'apoptose dans le contexte d'une vision trophique des relations intercellulaires. L'apoptose est un phénomène actif qui engage les cellules vers la mort, et se distingue en cela des nécroses qui sont des phénomènes passifs. Il s'agit donc dans l'acception classique qui est faite du phénomène, d'un suicide cellulaire qui est peu compatible avec une vision trophique : si les cellules ne réagissent qu'en fonction de leur intérêt immédiat, on comprend mal que certaines se suicident, pour le seul bénéfice de l'organisme. Mais si l'on s'intéresse aux causes qui président à l'apoptose, notamment aux " signaux " qui sont impliqués dans l'initiation du programme de " suicide ", on peut remarquer qu'il s'agit souvent d'une carence pour des molécules nécessaires au développement des cellules et qui doivent être nécessairement apportées par l'alimentation, comme le rétinol (ou vitamine A). La carence pour cette vitamine au cours de l'embryogenèse de poussins de cailles est reconnue comme une source d'apoptose qui obère un bon développement.

Par ailleurs, l'exemple souvent présenté comme canonique de l'apoptose, la perte des palmures entre les doigts au cours de l'embryogenèse des mammifères, peut

s’interpréter en terme d’accès aux ressources. En effet, l’apoptose suit la formation des cinq axes vasculaires qui sont l’ébauche de la vascularisation des futurs doigts. Les cellules qui rentrent en apoptose sont celles qui sont les plus éloignées de ces axes. Pour celles-ci, le problème de l’accès à l’oxygène devient problématique au fur et à mesure de l’épaississement des épithéliums. Les cellules les plus proches des vaisseaux, elles, profitent en priorité des ressources et progressivement la main se sculpte en épousant le contour de ces axes.

Grâce à ces exemples que l’on pourrait multiplier, on peut comprendre qu’il est possible de considérer les formes de différenciations cellulaires (et même les plus extrêmes comme la mort) comme une réaction locale à des conditions trophiques au sens large (alimentation, oxygène, espace).

II.D Des surproductions transitoires.

Revenons aux travaux de Hamburger et Levi Montalcini, sur la mécanique de l'innervation des tissus cibles. On vient de voir qu'ils envisagèrent la relation trophique pour expliquer le comportement exploratoire des neurones. Mais ils allèrent plus loin. En effet, ils démontrent que dans une phase où les neurones s’adaptent à la présence d’un champ d’innervation supplémentaire, il y avait " production " en surnombre de neurones, suivie d’une mort cellulaire des neurones non adaptés. Ces résultats furent généralisés à l’ensemble des centres et ganglions nerveux.
On voit que ce que nous serions tentés d’interpréter classiquement comme un signal est en réalité une ressource pour laquelle la cellule à un intérêt immédiat. La métaphore du signal est en l’occurrence entièrement superflue et donc trompeuse ici.

Le comportement de la cellule peut s'interpréter en tant que tel et nous pas dans le dessein d'un organisme supérieur. Que les intérêts à tous les niveaux coïncident, c'est ici évident. La cellule se nourrit et prolifère, et l'organisme se développe. Mais il y a bien un intérêt immédiat pour la cellule individuelle. Citons ici encore N. le Douarin :

" Loin d'assister à un développement selon un plan linéaire, *on voit alterner des phases de production intensive (en fait*, *de surproduction) avec des moments de restructuration massive de certaines cellules*, *la réassociation ou la transformation profonde de bien d'autres permettent (seules) la " sculpture " du vivant. Le développement et la vie d'un individu ne reposent pas uniquement sur la genèse puis sur la préservation de ses cellules comme le laisserait supposer un célèbre aphorisme de Bichat :* " la vie est l'ensemble des fonctions qui résistent à la mort "*. La construction même de l'organisme puis son fonctionnement impliquent aussi un modelage permanent par la destruction ou par le " suicide " d'un grand nombre de cellules ".*

On verra que la notion de suicide cellulaire peut être diversement interprétée, et en tout cas en d'autres termes que celui d'un "altruisme cellulaire" qui amènerait au sacrifice de certaines pour le bénéfice des autres, il n'en reste pas moins que la synthèse qui est faite ici du développement correspond trait pour trait à une dynamique endodarwinienne. Comment comprendre sinon, la surproduction transitoire des cellules ? S'il existe une signalisation classiquement présentée comme incroyablement fine et régulée, comment envisager que l'on passe systématiquement par un trop plein de production qui est ensuite systématiquement dégradé ? Comment l'évolution n'aurait-elle pas mis en place des régulations moins coûteuses en énergie, en ressources et en temps ? Les dynamiques

endodarwiniennes permettent de répondre à cette question fondamentale. Il y a surproduction parce qu'il n y a pas de programme, mais des phases où le comportement exploratoire des cellules est stimulé de sorte que certaines d'entre elles puissent adopter le comportement adapté à la situation, comportement dont elles retirent un bénéfice immédiat. Dans le cas des neurones qui métabolisent les substances trophiques des tissus cibles *" plusieurs motoneurones de la moelle épinière envoient leurs axones jusqu'aux fibres musculaires*, *mais c'est celui qui établit le contact le premier ou le plus efficacement (en terme d'absorption ou de facteurs trophiques) qui survit. Les autres sont éliminés de la compétition : leur cône de croissance et leur axone se rétractent ; finalement*, *ils meurent ".*

La logique de surproduction "transitoire" peut être transposée à d'autres échelles. Dans tous les cas elle relativise l'impression de fine régulation que l'on retire de l'observation des "succès", puisqu'on les met en regard de tous les "échecs" qui ont été nécessaires. Une échelle différente de celle de la mortalité cellulaire peut être par exemple celle de la mortalité embryonnaire.

Il existe des données mal connues et peu souvent évoquées, concernant l'importance des cas d'avortement précoces. En terme médical, plus ils sont précoces, plus ils sont ignorés, la mère n'ayant même pas eu le temps de prendre conscience de son démarrage de grossesse. Au plan fondamental, il est pourtant nécessaire de documenter ce taux d'échecs. S'il est faible, les échecs pourront être vu comme l'exception à la règle d'un programme finement régulé. Si ce taux est élevé, il devient pertinent de comprendre ces nombreux échecs comme le prix à payer (la base aléatoire d'une dynamique endodarwinienne) pour assurer le succès des autres démarrages. Et dans ce cas, admettre que l'évolution biologique a sélectionné une dynamique d'essai /erreur plutôt que celle d'un programme finement

régulé. Il convient de rester prudent mais il semble que les chiffres disponibles plaident plutôt en faveur de la deuxième branche de l'alternative (Scott Forbes, 1997). Les taux globaux de d'avortement spontanés humains sont estimés, selon les études, de 18,5 % à 74 % (Simpson et Golbus, 1992). Insatisfaisant écart de résultats, impropre à une conclusion définitive et traduisant l'absence d'un consensus sur la méthode d'évaluation, conséquence vraisemblable du peu d'importance accordée cliniquement à cette problématique. Les études les plus précises (Wilcox, *et al.*, 1988) fixent à 31 % l'estimation basse du taux global d'avortements suivant fécondation, toutes causes confondues, tous stades confondus. Au sein de cette estimation, les deux tiers de ce taux concernent des grossesses qui s'arrêtent avant d'être cliniquement repérées. Ces résultats ne semblent pas être compatibles avec un processus évolutif qui sélectionnerait progressivement les combinaisons génétiques les plus aptes à assurer un développement précis et reproductible.

II E Des greffes qui dévoilent des coopérations cellulaires sur une base plus fonctionnelle que génétique

L'intérêt à court d'une cellule terme peut être détecté dans d'autres systèmes expérimentaux. Et là encore, un retour aux sources est utile. Dans les expériences de Spemann publiées en 1924, que nous avons évoquées plus haut, les structures qui apparaissent sont pour partie hybrides et pour partie spécifique d'une des deux espèces. Rappelons ici pour mémoire qu'en greffant une deuxième lèvre blastosporale à une gastrula de triton, Spemann et Mangold obtenaient deux embryons siamois. Le cas des structures hybrides est ici à considérer de près. Que

des cellules provenant d'organismes différents se rassemblent pour constituer des structures hybrides semble indiquer que ce qui les guide est un intérêt à court terme, fondé sur des relations fonctionnelles et topographiques, relations qui sont elles-mêmes conditionnées par des considérations trophiques. Ces relations semblent prévaloir ici, dans le cas des structures hybrides, sur l'origine génétique. En cela, elles semblent ne pas être compatibles avec le schéma classique de cellules au service de l'organisme dans le but d'assurer sa survie et la propagation de son génome. En effet, dans le contexte d'un organisme ou l'ensemble des cellules agissent de telle manière qu'elles favorisent la propagation de leur génome (Dawkins, 1976) par rapport à celui des autres, on comprend mal la coopération entre cellules ayant la même fonction mais provenant d'individus différents, qui plus est d'espèces différentes. Sans même aller jusqu'à là, si l'on envisage un organisme comme un ensemble où chaque cellule réagit de manière à assurer sa propre survie sans considération sur son patrimoine génétique, alors l'aspect fonctionnel de la coopération entre cellules d'origines éloignées se conçoit mieux au sein de cet organisme chimérique. Les cellules des somites ont ici un intérêt immédiat à être rassemblées et cet intérêt immédiat prime et de loin sur l'intérêt à long terme de la propagation du génome de chacune. Cette remarque permet d'illustrer que la coopération intercellulaire peut tout à fait s'intégrer à une théorie endodarwinienne. De même que la biologie des populations a depuis bien longtemps intégré les phénomènes de symbiose dans son cadre explicatif, et que personne ne songerait à affirmer qu'elles vont à l'encontre du principe général de compétition dans la sélection naturelle, on peut tout à fait postuler une coopération fonctionnelle entre cellules comme une relation structurante de l'organisme, à supposer que l'on puisse définir des intérêts bien compris pour chacune des cellules prises isolément.

II F Conclusion

Dans cette approche endodarwinienne du développement nous avons voulu montrer qu'il était possible de poser des bases pour une lecture non "programmatrice" des observations faites. Nous avons souligné la plasticité des devenirs, la non-spécificité des régulateurs, le fonctionnement probabiliste sous l'apparence du déterminisme. Nous avons proposé qu'il serait utile d'envisager les étapes du développement sous l'angle d'une relation trophique.

Si les cellules, prises individuellement, ont un comportement aléatoire, cela remet en cause le cadre général déterministe dans lequel on décrit usuellement les phénomènes biologiques. On a vu que l'on pouvait alors proposer un autre cadre, autour de la notion d'accès différentiel aux ressources, et qui remplace la vision programmatique. Insistons ici, au travers de cette remarque, sur le fait que les cellules émettent et reçoivent bien des molécules ; que les cellules réagissent à un certain nombre de stimuli et cela, indépendamment du fait de savoir si elles y réagissent de manière homogène ou non. Notre cadre conceptuel ne nie pas les interactions cellulaires à courte ou longue distance. De manière générale, ce ne sont pas les observations qui sont contestées ici mais plutôt l'échelle à laquelle on les interprète, et aussi la conception des expérimentations qui fait balayer un spectre incomplet d'hypothèses.

Dans le cadre du développement, il est cependant prévisible que les dynamiques à l'œuvre sur des grandes populations de cellules ne puissent être reproduites, notamment dans les phases précoces. L'embryologie est donc, le contexte idéal pour tester les limites de prédictibilité de nos hypothèses et mettre en évidence les

mécanismes qui, bien que probabilistes, ont l'apparence d'un déterminisme fonctionnel.

Nous allons désormais porter notre attention sur un autre champ d'étude qui semble marquée par une approche déterministe : l'apoptose.

.

III L'apoptose contredit-elle les dynamiques endodarwiniennes ?

III. A Introduction

Il existe une typologie précise et consensuelle des différentes morts cellulaires observables dans un organisme. Ces différentes morts se situent entre deux cas canoniques extrêmes : la nécrose qui est décrite comme une mort passive et désordonnée, et l'apoptose, décrite comme une mort cellulaire programmée, active et ordonnée. Des figures intermédiaires ont été décrites (cf. infra). Ici encore, il ne s'agira pas de traiter intégralement du champ d'étude immense qui s'intéresse à la mort cellulaire. Alors que, quelle que soit l'approche que l'on puisse avoir sur la structuration et le développement d'un organisme, approche déterministe ou probabiliste, l'existence d'une mort cellulaire comme facteur structurant, permettant la " sculpture du vivant " (Ameisen, 1999), apparaît comme intelligible et nécessaire, nous chercherons à montrer qu'au sein des différents types de morts cellulaires, le phénomène d'apoptose peut dans un premier temps être contre-intuitif dans le cas d'une société de cellules régie par des dynamiques endodarwiniennes. En effet, en première approximation, le concept de mort cellulaire programmée, souvent qualifié de "suicide cellulaire", est caractérisé par une apparence d' "altruisme cellulaire" radical au cours duquel certaines cellules sont présentées comme se sacrifiant pour le bénéfice de la communauté. Quels qu'en soient les déclencheurs, l'apoptose est une dynamique interne à la cellule concernée, laquelle met en jeu des mécanismes génétiques précis qui conduisent à son autodestruction. Concevable dans un cadre conceptuel où l'ensemble des cellules est au service de l'organisme dans la finalité

de protéger et transmettre son patrimoine génétique, cette approche parait en revanche s'opposer de manière frontale, à l'approche endodarwinienne. En effet, pour celle-ci, l'organisme n'est pas une finalité (même formelle) pour chaque cellule, mais la résultante d'un état d'équilibre entre elles. Cette résultante, cet état de fait non programmé ne justifie pas le "suicide" de certains individus cellulaires au profit d'autres. Les cellules ne réagissent aux variations de leur environnement immédiat que dans leur intérêt propre et instantané, certes dans la limite des solutions permises par le génome qu'elles possèdent et qui peuvent nuancer, comme on l'a souligné auparavant, cette immédiateté. Un modèle endodarwinien prédit une mort cellulaire à l'issue d'une sélection, et non d'une programmation. Il y a donc une contradiction potentielle entre notre modèle et le phénomène d'apoptose, central dans de nombreux processus biologiques, et il est nécessaire de chercher à lever cette contradiction.

III B La typologie des morts cellulaires a des contours flous qui mettent partiellement en cause la spécificité des processus apoptotiques

Comme mentionné initialement, il existe une variété de morts cellulaires, distinguables par des caractéristiques observables qui ont conduit à en dresser une classification.

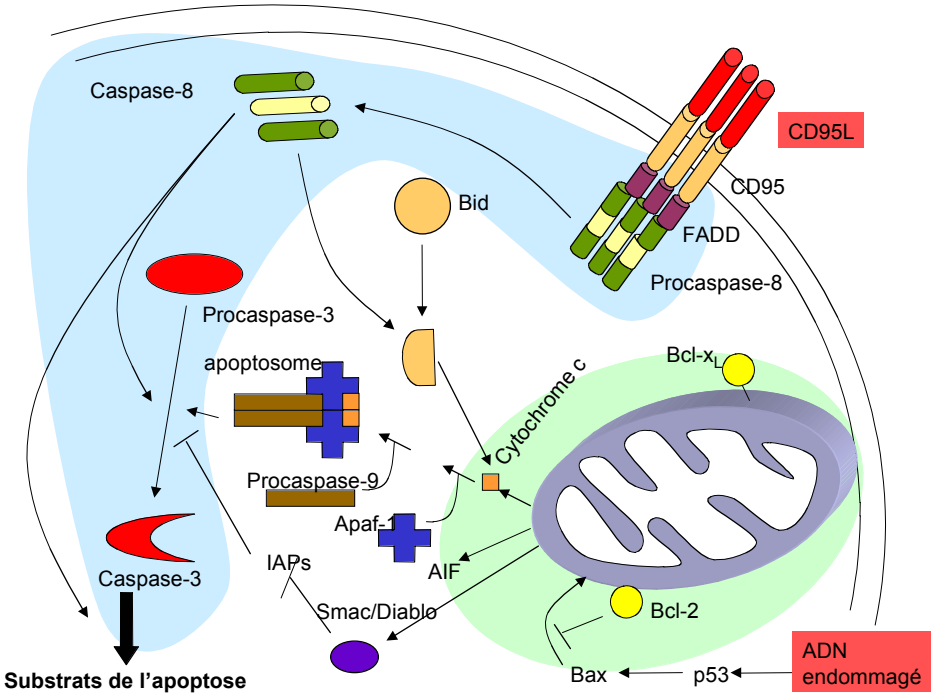
III. B 1 La nécrose

La nécrose est présentée comme un phénomène passif, souvent associé à une mort pathologique, au cours duquel une entrée d'eau entraîne la perte de l'intégrité des

membranes nucléaires et cellulaires. En conséquence, le contenu cytoplasmique est relargué dans l'espace extracellulaire. L'ADN est fragmenté de manière aléatoire. Les agents responsables de la nécrose, s'ils sont encore présents dans la cellule, peuvent alors se répandre et éventuellement menacer les cellules avoisinantes.

III. B 2 L'apoptose

A l'opposé, l'apoptose est un processus de mort qui induit une diminution du volume cellulaire, et une désolidarisation des cellules alentour par ruptures des liaisons intercellulaires, une condensation tant du cytoplasme que du noyau, une fragmentation nucléaire, une condensation chromatinienne puis un clivage de celle-ci en fragments réguliers, et la formation de corps apoptotiques, qui aboutissent à une mort de la cellule dans ses propres limites, facilitant ainsi l'élimination et/ou le recyclage des composants. Cette mécanique qui permet de circonscrire la mort, repose sur des cascades de réactions internes à la cellule et activées par elle. C'est pour cette raison que l'on parle de suicide cellulaire, même si ce terme est problématique. Il est évidemment métaphorique, mais néanmoins lourd de sens et de présupposés. On peut lui préférer le terme plus neutre de " mort propre " qui est en fait la définition minimale de l'apoptose par opposition aux autres mécanismes de mort cellulaire. On peut d'ailleurs noter que certaines cellules T cytotoxiques procèdent à la destruction de leurs cellules cibles par l'utilisation provoquée de ces mécanismes internes de "suicide". Cela peut être envisagé soit comme un recrutement d'une fonction préexistante (en l'occurrence la capacité d'autodestruction) par la cellule tueuse, soit à l'inverse comme l'intériorisation de processus de destruction exogène, liée par exemple à la machinerie de lutte contre



D'après Hengartner, 2000

Figure 14 : Schéma simplifié des deux grandes voies d'apoptose.
Sur fond bleu, la voie dite du signal de mort, faisant intervenir un récepteur transmembranaire CD95 et des procaspases 8 et 3. Sur fond vert la voie mitochondriale, qui fait intervenir des protéines de la famille Bcl2. Ces deux voies influent sur l'apoptosome.

ceux-ci. Cette seconde approche pourrait fournir une piste d'explication de l'apparition contre-intuitive de l'apoptose, nous y reviendrons.

Les mécanismes de l'apoptose sont extrêmement conservés et répandus de manière généralisée dans le monde vivant. Elle a été caractérisée dans tous les chez tous les animaux multicellulaires où elle a été cherchée, notamment les cnidaires, les nématodes, les mammifères, les insectes, les amphibiens, les oiseaux. L'apoptose, en tant que phénomène actif, a été caractérisée de manière moléculaire par la présence quasi généralisée d'acteurs comme une famille de cystéines protéases, les caspases, ainsi que d'autres protéines comme APaf-1 et la famille de protéines Bcl-2 associées à la mitochondrie.

On trouve des homologues des gènes impliqués dans l'apoptose dans une grande variété d'espèces. Des mécanismes analogues sont aussi repérés dans le règne végétal et dans certaines espèces de champignons. Même s'il est vain de prétendre faire une synthèse ici de la complexité des mécanismes d'apoptose identifiés, on peut néanmoins présenter un schéma des deux grandes voies d'apoptose décrites pour les cellules de mammifères (Hengartner, 2000) dont certains éléments seront discutés plus loin. La figure 14 illustre ces voies et l'on s'y réfèrera à plusieurs reprises au cours du texte.

-La voie dite "du signal de mort". Si un ligand se fixe sur le récepteur transmembranaire CD95, il provoque le recrutement via FADD de procaspases-8 qui s'assemblent transitoirement pour donner des caspases actives, elles mêmes inductrices de caspase-3. La caspase-3 active est un effecteur décrit des phénomènes d'apoptose.

-La voie dite mitochondriale.

Suite à des dommages dans l'ADN, une cascade de réactions conduit à l'activation de protéines de la famille Bcl2. Parmi elles, certaines ont des effets agonistes sur l'apoptose ; Bax provoque notamment la perméabilisation de la membrane mitochondriale externe, et le relargage du cytochrome C. Associé à Apaf-1, le cytochrome C permet l'activation de la procaspase 9 au sein de l'apoptosome, qui conduira à l'activation de la caspase-3. D'autres membres de la famille Bcl-2, comme Bcl-2 elle-même et Bcl-xL ont un rôle antagoniste en inhibant directement ou indirectement le relargage du cytochrome C. De nombreux intervenants viennent complexifier cette voie.

III. B 3 Des formes intermédiaires présentent des caractéristiques mixtes

Entre ces deux formes extrêmes de mort cellulaire, des formes intermédiaires ont été décrites, notamment dans le cadre de la mort neuronale, dont on a vu qu'elle était impliquée dans une dynamique de structuration du système nerveux. En plus des deux types évoqués ci-dessus, on trouve des figures de mort cellulaire neuronale dite autophagique, qui fait intervenir des vacuoles autophagiques cytoplasmiques d'origine lysosomale et une dilatation mitochondriale. Par certains aspects, la mort neuronale autophagique, qui a elle-même plusieurs variantes, micro- ou macroautophagique, s'apparente à chacune des deux figures extrêmes.

Si l'on en reste à la schématisation d'une mort passive et d'une mort contrôlée, on peut comprendre qu'elles ne soient pas intégrées de la même manière dans un schéma endodarwinien. Une mort passive peut être considérée comme la conséquence d'une contre sélection de la cellule dans le cadre par exemple d'une compétition pour les ressources. On a vu que cette mort cellulaire est indispensable,

dans ce schéma, en tant qu'il est la conséquence logique de la surproduction transitoire de cellules permettant un tri sélectif des cellules adaptées. En revanche, l'apoptose, qui implique une mise en place interne, individuelle, d'une machinerie d'autodestruction, défie en apparence l'approche endodarwinienne, puisque la mort d'individus cellulaires n'est plus seulement subie, mais décrite comme provoquée par les cellules elles-mêmes.

Néanmoins, l'intérêt de l'ébauche de typologie des morts cellulaires présentées plus haut, comprenant des intermédiaires entre les extrêmes, est de nous faire prendre conscience que les limites entre ces phénomènes ne sont pas étanches. Si une typologie est une construction nécessaire pour ordonner la connaissance que l'on a de phénomènes apparentés (et en cela, elle est la première et légitime étape de la connaissance scientifique), elle demeure une construction subjective, qui comportera toujours, même de manière résiduelle, une part d'arbitraire. En l'espèce, les différentes morts cellulaires sont regroupées selon qu'elles donnent lieu à certaines combinaisons de phénomènes (cf. supra), mais on constate qu'ils sont individuellement partagés par différents types. Ainsi, la forme intermédiaire, dite de mort autophagique (Klionsky et Emr, 2000), notamment décrite dans la mort neuronale, elle-même une catégorie subdivisable, comporte des traits de caractères propres soit à l'apoptose, soit à la nécrose. Or il apparaît qu'elle est un élément fondamental de la mort cellulaire neuronale. Elle est décrite par exemple dans la mort neuronale au sein du noyau isthmo-optique du poulet (Clark, 1990), et dans la mort par privation en NGF de neurones sympathiques, ainsi que de cellules PC12, et potentiellement dans la mort de cellules de Purkinje (Ohsawa, *et al.*, 1998 ; Xue, *et al.*

1999). Ceci est de nature à relativiser la prépondérance de l'apoptose dans la mort structurante au cours du développement du système nerveux.

Il est ici utile de porter notre attention sur les figures caractéristiques de l'autophagie. Le critère principal est l'apparition de vacuoles intracellulaires bimembranaires contenant des fragments de reticulum endoplasmique ou mitochondrial et des hydrolases lysosomales. Cela est cohérent avec l'interprétation évolutive de ce mécanisme : l'autophagie est la trace de réponses anciennes à des conditions de stress externes (déficit de métabolites, hypoxie) ou internes (accumulation d'organites ou de composants cytoplasmiques) endommagés, par laquelle la cellule se protégeait des désordres internes et luttait contre les conditions néfastes en réduisant son métabolisme et en améliorant ses capacités de recyclage (Yuan, *et al.* 2003). C'est sur la base de ce mécanisme de protection qu'aurait pu se développer une mort cellulaire. On commence à comprendre ici que des mécanismes anciens de protections peuvent êtres détournés au profit de fonctions qui en utilisent les mécanismes mais en modifient la finalité. L'autophagie pourrait, dans les maladies neurodégénératives, et en particulier la maladie de Parkinson (Stefanis, *et al.*, 2001), être activée comme un mécanisme compensatoire à une insuffisance du protéasome, mécanismes qui dégénérerait progressivement vers une mort cellulaire fatale à l'organisme. L'autophagie a par ailleurs un rôle important dans les métamorphoses d'insectes dans lesquels de grandes quantités de cellules sont détruites en l'absence du nombre suffisant de cellules spécialisées dans leur évacuation : par ses points communs avec la nécrose, elle apparaît plutôt comme un mécanisme d'auto évacuation de cellules mortes qu'un mécanisme de mort cellulaire en tant que tel

(Yuan, *et al.*, 2003). Sa position intermédiaire en fait un candidat à la modélisation d'un passage par étape d'un mécanisme de nécrose à un mécanisme d'apoptose.

En outre, même dans les cas où l'on peut caractériser et discriminer clairement des figures d'apoptose et de nécroses, une coexistence de ces figures au sein de mêmes cellules individuelles a pu être décrite (Kaminski, *et al.*, 2002) sur des lignées humaines.

En résumé, la mort structurante à l'intérieur d'un organisme n'est pas du seul ressort de l'apoptose et fait appel à des mécanismes qui peuvent avoir un apparentement avec des morts passives. Il est intéressant à noter, même si cela semble tenir de l'exception plutôt que de la règle que des figures proprement nécrotiques sont mises en évidence dans certains cas de mort neuronale développementale.

Cela n'évacue pas la nécessité de donner une lisibilité endodarwinienne à l'apoptose, mais cela complexifie passablement le paysage de la mort cellulaire comme outil d'édification de l'organisme.

III. C La non-spécificité des molécules impliquées dans l'apoptose

III. C 1 Certains acteurs de l'apoptose ne sont pas toujours indispensables.

Les deux voies principales de l'apoptose sont constituées d'acteurs impliqués dans des cascades de réactions dont la détermination apparente est le socle permettant de caractériser l'ensemble comme un " programme d'autodestruction ". Néanmoins les rôles connus de ces acteurs ne sont pas toujours en cohérence avec leur

intervention supposée dans la voie de l'apoptose. En particulier, l'absence d'un certain nombre de ces acteurs décrits comme cruciaux n'empêche pas l'apoptose, mais la retarde seulement. Cela ne signifie pas que l'ensemble du phénomène d'apoptose doive être pour autant intégralement repensé. On observe en effet un certain nombre d'invariants dans les processus apoptotiques qui sont plus nombreux que les exceptions à la règle. Néanmoins, observer les éléments de variations, les partenaires non indispensables, cela permet d'avoir une approche plus probabiliste de la mort cellulaire " programmée ".

Parmi les acteurs moléculaires qui jouent un rôle central, on peut noter le cas du cytochrome C, dont la liaison avec Apaf-1 est cruciale dans la voie d'apoptose mitochondriale (Figure 14, p.124). Le relargage du cytochrome C suite à la perméabilisation de la membrane mitochondriale externe, est décrit comme un épisode fondamental de cette voie (même si le processus de perméabilisation reste non déterminé). Il s'avère que des expériences réalisées en l'absence de cytochrome C n'ont pas aboli la fonction apoptotique. Celle-ci a été retardée, mais a bien eu lieu (Li, *et al.*, 2000).

De même le partenaire Diablo/Smac qui inhibe une famille de protéines inhibitrices d'apoptose (IAPs) ; en son absence, l'apoptose n'est pas modifiée (Okada, *et al.*, 2002).

Par ailleurs des expériences de mutants caspases-9 -/- ou d'Apaf-1-/- dans certains cas d'apoptoses stress dépendantes, n'ont montré aucune variation dans leur taux de morts cellulaires (Marsden, *et al.*, 2002). Il est à noter que cela ne concerne que certains types de telles apoptoses. D'autres expériences vont dans le sens inverse et plaident en faveur de la nécessité de ces acteurs. Il faut ajouter à ces nuances, que les effets de ces mutants avaient des pénétrances variables, qui allaient jusqu'à la

possibilité d'obtenir une certaine proportion (minoritaire) d'adultes bien portants (Honarpour, *et al.*, 2000). Ceci tendant à relativiser l'univocité des voies d'apoptoses tout comme on a pu précédemment relativiser l'importance des voies de signalisation spécifiques en mettant en évidence des nombreuses voies de recours.

Certaines figures d'apoptose, comme la dégradation et la fragmentation du génome, pourtant canonique de ce phénomènes semblent parfois n'être pas observées et ne sont donc pas strictement indispensables au bon déroulement du phénomène.

Ces divers éléments épars ne sont pas de nature, répétons le, à remettre en cause le phénomène d'apoptose, dans sa dimension de voie organisée conduisant à l'autodestruction. Mais pourraient changer la conception qu'on a du contrôle de ce phénomène en introduisant une plasticité peu compatible avec la notion de programme, rigide par définition.

On peut par exemple envisager que certains acteurs, même absents, puissent être remplacés par des homologues structuraux recrutés *ad hoc*. De même certaines caspases pourraient en suppléer d'autre dans les expériences de mutants double récessifs. Ces hypothèses de travail doivent être testées. Cela étant, en parallèle aux investigations complémentaires suggérées on est aussi en droit de proposer une réinterprétation partielle des rôles de ces acteurs. En effet, les observations ci-dessus forment un ensemble de données qui ont troublé certains auteurs qui tendent désormais, pour certains types de mort apoptotiques, notamment stress induites, à redéfinir le rôle de l'apoptosome. Puisqu'en son absence, l'apoptose peut avoir lieu, avec retard et avec un taux de caspases moindre, il semble que celui-ci puisse être interprété comme un amplificateur du phénomène apoptotique plutôt que comme sa véritable cause (Marsden, *et al.*, 2002).

Cela peut être corrélé au fait que l'un des déclencheurs de la voie d'apoptose mitochondrie dépendante est la dégradation de l'ADN. Or la dégradation d'ADN est aussi la conséquence de l'activation de cette voie, en tant qu'elle est une des figures des phases finales de l'apoptose. Il peut paraître paradoxal que la conséquence d'une voie d'apoptose en soit le déclencheur. Ceci se comprend peut-être mieux à la lueur de ce qui vient d'être suggéré : tout se passe en fait comme si l'apoptosome était pris dans une classique boucle de rétroaction positive, ici parfaitement cohérent avec le rôle proposé dans la littérature d'apoptosome comme amplificateur.

Il a été noté et démontré que, si, chez *C. elegans*, la voie d'apoptose semblait, univoque, elle pouvait inversement, chez les mammifères, procéder selon de multiples voies, différentes pour chaque type cellulaire, mais aussi pour un type cellulaire donné, différentes pour chaque "signal" déclencheur (Ameisen, 2002).

Cela n'abolit pas son rôle effectif, mais cela le replace dans un ensemble plus vaste, où, dans certains cas, et notamment, notons-le ici, dans certains cas de stress, les cellules peuvent puiser dans leur génome différents outils, différentes voies permettant d'aboutir au phénotype de mort. Sans remettre en cause les voies classiques d'apoptoses, on peut imaginer qu'il en existe un certain nombre d'alternatives. Encore une fois il n'y a donc pas de spécificité à proprement parler, Et cela aboutit encore une fois à remettre en cause la notion de programme, puisqu'on peut replacer ces phénomènes dans un contexte d'expression aléatoire due au stress aboutissant à l'une ou l'autre des voies qui enclenchent une dynamique dont la fin est commune. Cela ne donne toujours pas de rationalité à cette fin, nous y viendrons, mais cela rationalise le mécanisme et le replace dans un contexte où la

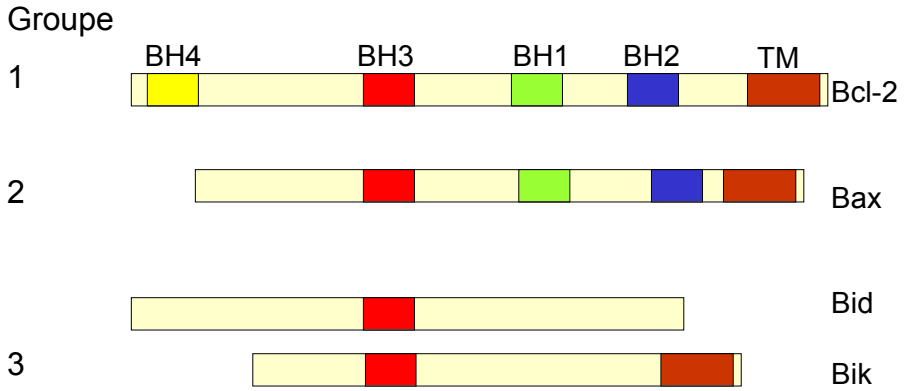


Figure 15 : La famille des protéines Bcl-2.
Ces protéines interviennent notamment dans la voie mitochondriale. On voit que les protéines Bcl-2 et Bax ont une séquence très proche. C'est cette homologie qui leur permet de se neutraliser en s'hétérodimérisant.

détermination n'est qu'apparente et due à un effet d'entonnoir où plusieurs mécanismes différents, et aléatoirement déclenchés produisent le même effet.

III. C 2 Des molécules antagonistes ont des structures très proches.

Il apparaît par ailleurs que de nombreux partenaires des voies d'apoptose ont des structures très proches, indépendamment de leur fonction agoniste ou antagoniste dans l'apoptose. C'est notamment le cas des membres de la famille Bcl2. Ces protéines sont, on l'a vu des acteurs de la voie d'apoptose mitochondriale. Ils interviennent notamment dans les processus liés au relargage du cytochrome C par la mitochondrie, qui participe à l'apoptosome. Mais ils n'interviennent pas tous de manière convergente. Bax est une protéine du cytosol qui permet la solubilisation du cytochrome C. Bcl-2 et Bcl-xL , deux protéines liées à la membrane externe de la mitochondrie par queue hydrophobe C-terminale et appartenant à cette même famille, ont un rôle antagoniste. Bcl-2 réprime l'activité de Bax. Bcl-xL réprime l'activité de Bid , un autre membre de la famille, mais qui lui a, une fois tronqué une activité agoniste sur le relargage de cytochrome C, comme Bax (Figure 14, p.124).

Quelles sont les relations structurales et évolutives entre ces membres ? Même si certaines, on va le voir, sont suffisamment proches pour que l'on puisse légitimement postuler une origine évolutive commune, certaines semblent n'avoir de relations structurales que limitées, qui procède peut-être plutôt d'une évolution convergente. Notons donc que cette famille de protéines est caractérisée avant tout par des homologies de séquences qui sont plus à envisager sous un angle fonctionnel que phylogénique. Les membres de cette famille sont classés en trois groupes (Antonsson et Martinou 2000) (Figure 15). On voit que des groupes contenant des

acteurs antagonistes partagent une forte homologie de structure. En outre, les études de relations structure/fonction ne semblent pas mettre en évidence que la séquence non commune au deux soit réellement impliquée dans la différence de fonction. Cette homologie structurale est au contraire expliquée par la capacité de ces acteurs antagonistes à se neutraliser par hétérodimérisation. Il semble donc difficilement soutenable qu'il existe une information, un programme codé génétiquement qui enclenche une voie de mort puisque ce qui ressort de la génétique est partagé par des acteurs antagonistes.

III. C 3 L'apoptose comme une rupture d'équilibre

Il semble donc qu'il faille plus réfléchir en termes de rupture d'équilibre entre différents acteurs dans des chaînes de réactions chimiques, ce qui nous conduit à considérer la cellule comme un espace traversé par des tendances contradictoires et non pas comme une entité prenant unidirectionnellement une décision (même si l'on utilise ce terme au sens métaphorique) d'autodestruction. À l'appui de cette approche en termes d'équilibre instable on peut noter qu'au moins chez *C. elegans*, les premiers stades de l'apoptose sont réversibles, mais pas les phases finales (Metzstein, *et al.*, 1998). Comme s'il était possible à la cellule de résister un temps à ces déséquilibres internes, jusqu'à des phases où les dommages, notamment génomiques sont trop importants et s'autoamplifient. Ainsi, la vie serait donc bien l'ensemble des fonctions qui résistent à la mort, selon l'adage de Bichat. La mort cellulaire ne serait donc, en première approximation, qu'une probabilité de survenue, probabilité qui finirait statistiquement par se réaliser (Klarsfeld et Revah, 1999). On pourrait en outre attribuer aux mécanismes déclencheurs le rôle de canaliser les

mécanismes favorisant la rupture de l'équilibre instable. Notons ici que les caspases, qui sont des cystéines protéases, peuvent donc être considérées comme des enzymes du métabolisme cellulaire. Qu'un stress induise la nécessité de recyclage et de nouvelles synthèses protéiques n'est pas illogique. Par ailleurs, le rôle des caspases peut être différent d'une espèce à l'autre si l'on considère les homologues dans chaque espèce. Ce n'est pas tant leur structure que leur niveau de recrutement qui leur confère leur rôle. Enfin, le rôle des caspases n'est pas limité à la voie d'apoptose, même dans les cas où l'on pensait ce rôle univoque (Ameisen, 2004). L'activation de certaines caspases dans certaines circonstances, semble être nécessaire à la différenciation cellulaire ou au contrôle du cycle cellulaire (Ameisen, 2002). Dans ce contexte, le rôle des caspases est intelligible, même si la voie dans laquelle elles interviennent aboutit à une fin contre intuitive. De manière plus générale, la réflexion pionnière de Jean Claude Ameisen sur l'acquisition évolutive des mécanismes de l'apoptose a largement contribué à proposer une approche non finaliste de la mort cellulaire (Ameisen, 1996, 1998, 1999, 2002, 2004 ; Arnoult, *et al*.,2001).

III. C 4 Les déclencheurs de l’apoptose sont souvent des " signaux par défaut "

Les mécanismes de déclenchement de l'apoptose sont souvent nommés signaux, dans le sens où ils sont décrits comme convoyant de l'information, en l'occurrence une consigne d'autodestruction. Mais ces signaux sont de natures fort diverses, et peuvent être perçus comme tels par défaut ou par excès. Si un certains nombre de molécules sont connues pour activer par leur présence l'apoptose, c'est pour d'autres

leur absence ou leur diminution qui est le déclencheur apparent du phénomène. Nous verrons que cela peut en changer les perspectives d'approche.

Il existe un certain nombre de signaux dont la présence déclenche l'apoptose. C'est notamment le cas du Fas-ligand, qui se lie au récepteur transmembranaire CD95 dit récepteur de mort et en enclenche la voie d'apoptose. De même le Tumor necrosis factor–alpha (TNF-α), peut se fixer sur un récepteur TNF-R des macrophages et induire la mort cellulaire.

Mais un grand nombre de signaux d'apoptose sont bien souvent, à l'inverse, des phénomènes de carence et d'agression. L'apoptose neuronale est notamment enclenchée par une carence en facteurs de croissance NGF (Neuron Growth Factor) pour les neurones (Yuan, *et al.*, 2002), ou pour les lymphocytes l'interleukine 2 IL-2 qui est un élément nécessaire à la mitose. L'absence de rétinol, c'est-à-dire de vitamine D, peut être une cause d'apoptose. De manière générale, l'hypoxie est un déclencheur d'apoptose. Des phénomènes agressifs comme les rayonnements UV ou les rayons X peuvent aussi en être à l'origine.

Ces remarques permettent de nuancer le caractère "altruiste" qui sous tend la logique prêtée à l'apoptose dans le développement de l'organisme. Quand bien même un mécanisme actif de destruction est mis en place, on voit dans les cas ci-dessus que celui-ci est une conséquence de stress importants, qui provoquent des dégâts intracellulaires, notamment dans l'intégrité de l'ADN. Nous ne sommes pas dans le cas d'une cellule bien portante qui se donnerait en sacrifice à l'organisme car elle serait topographiquement située à un endroit où l'organisme aurait intérêt à la détruire.

Un des exemples classique de l'apoptose peut illustrer cette logique. On sait qu'elle est à l'œuvre dans le processus de disparitions des palmes interdigitales des

embryons de mammifères. Les cellules qui constituent ces zones meurent au cours du développement par un phénomène actif. Cela est présenté classiquement comme une contribution de ces cellules à l'édification de l'organe. D'autres hypothèses, peut-être plus parcimonieuses, ont cependant été avancées. Notamment on peut remarquer que dans la séquence des événements qui conduit à la disparition de ces structures, l'apparition des cinq axes de vascularisation précurseurs des doigts précède l'apoptose des cellules des palmes. Cette vascularisation induit un gradient de disponibilité des ressources en fonction de la distance de chaque cellule à ces axes. On constate donc que ce sont les cellules qui sont les plus éloignées de ces sources, et donc qui ont l'accès minimal aux ressources qui sont défavorisées et meurent. Tout se passe ici comme si la structure résultante etait le produit d'une disponibilité différentielle des ressources dans un environnement compétitif pour celles-ci. Ceci n'explicite pas encore pourquoi la cellule met en place une dynamique individuelle d'autodestruction, mais illustre cependant le contexte dans lequel cela est réalisé, qui est un contexte d'adversité. Ceci étant une nouvelle fois de nature à nuancer la qualification de signaux que l'on prête aux molécules déclencheuses, ou à leur absence. Il semble bien ici que ces carences ou ces agressions ne sont pas tant porteuses d'information (auquel cas, il est nécessaire de renoncer à la notion de signal car on la vide de sens) qu'inductrices de stress provoquant des dommages internes destructurants.

III. D Une approche écosystémique peut fournir des clefs de réinterprétation de la mort cellulaire.

L'approche écosystémique des relations intercellulaires doit être pratiquée avec précaution. L'ensembles des caractéristiques d'un organisme ne se retrouve pas dans un écosystème (en particulier, un écosystème ne se reproduit pas) et il est donc clair que ce rapprochement a une dimension métaphorique non négligeable. Néanmoins cela n'invalide pas en soi la pertinence de ce rapprochement, pour peu que l'on garde à l'esprit que toute métaphore a sa limite. En outre, l'approche informationnelle a, elle aussi, une forte charge métaphorique. En somme il reste à exercer une vigilance permanente sur la pertinence ponctuelle de chaque analogie, quelle qu'elle soit, pour ne pas se laisser dépasser par des rapprochements superflus, qui n'auraient d'autres nécessité que celle de coller trop près aux implications de la métaphore utilisée.

On a vu plus haut, dans le contexte de l'embryogenèse de la main et des doigts de mammifères, que l'on pouvait réfléchir en termes d'accès aux ressources pour conceptualiser la causalité de certaines morts cellulaires, quels qu'en soient les mécanismes (Sonigo, *in* Kupiec et Sonigo, 2000). On peut aussi, dans un autre contexte, faire appel à des relations type proies/prédateurs entre les cellules et leurs ressources. On voit par exemple lors de l'importante mort des cellules T dans le thymus celles qui demeurent sont celles qui n'éradiquent pas complètement leur antigène. Il y a là une analogie possible avec les courbes classiques de relations entre proies et prédateurs où l'on sait que si le prédateur est superefficace, il risque de contribuer à épuiser son milieu en proies et donc à mettre en péril sa propre existence, si ce n'est celle de son espèce. La relation qui peut assurer une survie maximale aux prédateurs est, toute chose étant égales par ailleurs, un taux de prédation modéré, qui permet de maintenir par oscillations autour d'une moyenne, un effectif constant de proies et de prédateurs. On retrouve cette logique dans les

cellules T. Celles qui demeurent ne sont pas celles qui sont les plus efficaces pour la capture des ressources (ici, l'antigène) mais celles qui entretiennent à bas bruit un taux d'antigène bas mais constant.

Par ailleurs, les relations entre individus dans un écosystème ne se résument pas à de la compétition. Le réseau de relations envisageable est complexe et comporte, entre autres, des dynamiques de coopération, de symbiose, de parasitisme, comme stratégies d'accès aux ressources et de propagation de la descendance. Dans ce contexte, il apparaît que même dans un écosystème, la mort provoquée d'un individu, même si elle est bien sur contingente, s'insère dans une dynamique plus large qui ne fait pas pour autant appel à un programme ou à une finalité qui serait la survie de l'écosystème. En particulier la mort provoquée d'origine interne, par un parasite par exemple. Nous allons donc voir que ce genre d'approche va permettre de penser la mort cellulaire apoptotique.

Au cours de cette section nous avons abordé plusieurs arguments, qui nous semblent plaider en faveur d'une relativisation forte de l'apoptose en tant que programme d'autodestruction. Nous avons vu qu'elle n'était pas radicalement différente des autres types de morts cellulaires, qu'elle n'était pas la seule impliquée dans le développement des organes, qu'elle faisait appel à une biochimie précise mais qui semblait douée d'une certaine souplesse de réalisation, et dont on ne comprenait qu'imparfaitement le déterminisme génétique (puisque des motifs structuraux communs semblaient avoir des rôles antagonistes), que cette biochimie semblait souligner un rôle d'amplification du phénomène plutôt que strictement déclencheur, que l'apoptose semblait souvent être une réponse à des carences ou à

des conditions agressives, qu'elle était couplée avec des mécanismes antiapoptiques qui traduisait l'issue mortelle comme une rupture d'équilibre plus que comme une finalité, au sein de relation partiellement comparables à certaines modélisations de dynamiques de population. L'ensemble de ces éléments nuancent la dimension "altruiste" du phénomène apoptotique. Mais il reste à se demander quelle est l'origine de ce phénomène, et si les hypothèses quant à cette origine ne changent pas, de manière encore plus fondamentale, la perception contemporaine que l'on peut avoir de l'apoptose dans le contexte de la multicellularité.

Les différents types de mort cellulaires, au premier rang desquels l'apoptose, sont ainsi considérés comme une propriété de l'état multicellulaire, par lequel celui-ci s'organise et met en place des structures qui ont pour conséquence de maintenir sa structure et créer les conditions de sa reproduction. La mort cellulaire est centrale dans le développement des organismes multicellulaires, tant animaux que végétaux. Qu'en est-il des organismes unicellulaires ?

III. E L'apoptose chez les unicellulaires.

III. E 1 Le paradoxe de l'apoptose chez les organismes unicellulaires

A première vue, il semble incongru de penser que la mort cellulaire puisse être une stratégie pour les organismes unicellulaires. Dans le contexte d'une approche dawkinsienne, on admet que, dans le cadre de la propagation d'un génome, certaines cellules d'un organisme se détruisent, ou du moins ne s'opposent pas indéfiniment à leur propre destruction, si cela maximise l'adaptation de l'organisme

au milieu et par conséquent favorisent les chances de propagation de ses gènes. Mais dans le cas de populations bactériennes, il semble difficilement acceptable d'intégrer la mort cellulaire, puisqu'à la différence du cas précédent, chaque cellule, constituant en soi un organisme, possède son propre génome et tend à le propager, au détriment des autres. La non-homogénéité génétique d'une population bactérienne, impliquant inexorablement une compétition pour la survie, rend *a priori*, en première approximation, incompréhensible une mort cellulaire adaptative, même dans le cas d'une approche dawkinsienne. En outre, on comprend mal comment la mise en place d'une capacité de "suicide" puisse avoir d'autre conséquence que la contre-sélection directe.

Si l'on met en évidence l'existence d'une mort cellulaire dans ce contexte, cela serait de nature à modifier radicalement la perception de l'apoptose dans les organismes multicellulaires. Tout d'abord, cela ferait de la mort cellulaire un caractère primitif, et non pas une acquisition contemporaine du passage à l'état multicellulaire. Cela aurait comme conséquence directe d'abolir la dimension "altruiste" de ces morts observées, à quelque degré de sens que l'on utilise ce terme. En effet, la mort cellulaire et notamment l'apoptose ne serait qu'un recrutement, une adaptation d'une fonction ancienne (auquel nous n'avons pas encore conféré de rationalité, cf. infra) dans le cadre de relation trophiques de type pseudo symbiotique des cellules d'un organisme. Ainsi l'apoptose qui ne serait pas la trace de la réalisation harmonieuse d'un plan de développement programmé (écrit à l'avance) et donc ne serait plus un défi à l'entendement d'une approche endodarwinienne du développement.

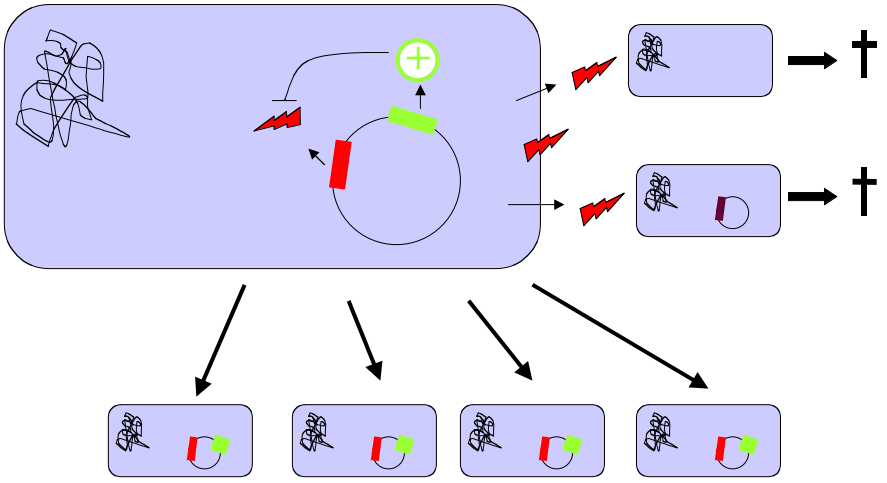


Figure 16 : Le module d'addiction.
Les bactéries qui possèdent la toxine (rouge) et l'antidote vont se répandre dans la population de deux manières : en tuant les bactéries non pourvues de l'antidote, et en se multipliant. L'ensemble de la population va donc être contraint d'exprimer les deux gènes pour survivre (d'après Ameisen, 2002).

III. E 2 Deux modèles rendent intelligible une apoptose chez les unicellulaires.

Or il se trouve que la mort cellulaire programmée est observée dans des organismes unicellulaires, et que l'on peut comprendre les mécanismes qui permettent, ou plutôt rendent indispensable, la propagation de cette propriété. La mort cellulaire programmée, ou du moins ses prémisses est observable tant chez les bactéries que chez les eukaryotes unicellulaires (Ameisen, 2002). Elle se comprend dans ces deux cas, si l'on admet qu'il peut, au sein d'une cellule, exister des conflits. Dans le cas d'une bactérie, le conflit peut survenir entre le génome et un plasmide, tandis que chez un eukaryote unicellulaire celui-ci peut intervenir entre le noyau, et la mitochondrie, trace d'une ancienne endocytose.

Deux modèles ont été proposés pour comprendre l'apparition et la propagation de cette propriété (Ameisen, 2002). Ils ne s'opposent pas.

-Le premier modèle est celui dit de l'addiction.

Il repose sur le constat que dans leur compétition pour la survie, les bactéries peuvent mettre au point des stratégies qui font intervenir des toxines. Pour cela, un binôme moléculaire entre en jeu : une toxine, qui détruit les autres cellules, et un antidote qui protège la cellule émettrice de la toxine. Evolutivement, ces modules sont souvent portés par des plasmides. La bactérie a un avantage sélectif à posséder ces modules, qui, une fois une fois introduits, confèreront à celle-ci la

capacité de se multiplier aux dépens des autres , et donc assureront leur propre propagation (Figure 16).

On comprend que l'intrusion d'un tel plasmide va conduire rapidement à sa propagation généralisée : l'ensemble des cellules ne le possédant pas, soit qu'elles ne le possédaient pas initialement, soit qu'elles ne l'ont pas reçu lors d'une division. Ce faisant, les bactéries intériorisent un module poison/antidote qui a beaucoup à voir avec le mécanisme de mort cellulaire programmée. Si l'on constate que l'antidote à une demi-vie inférieure à celle du poison (par ce qu'il est plus susceptible d'être détruit par des protéases cellulaires), on peut alors accepter que la bactérie hôte soit prise dans un engrenage addictif qui lui fait produire du poison et son antidote en permanence, et dont elle ne peut se passer, à la fois pour survivre aux dépens des autres, et à la fois pour se protéger.

La théorie endosymbiotique postule l'apparition des mitochondries par domestication d'une bactérie par endocytose incomplète. Les hypothèses évoquées ci-dessus peuvent donc s'appliquer à une relation conflictuelle non plus tant entre génome bactérien et plasmide qu'entre le noyau et la mitochondrie. "Dans ce contexte, les mécanismes par lesquels la cellule eukaryote induit la perméabilisation de la membrane mitochondriale peuvent être vus comme un mécanisme par lequel la cellule provoquer son autodestruction par destruction du module d'addiction (ici la mitochondrie)." (Ameisen, 2002)

Dans le cadre de ce modèle, on peut aussi comprendre qu'une portion seulement de cellules en viennent à s'autodétruire dans un protoorganisme. La bactérie *Bacillius subtilis*, dans des conditions défavorables, peut s'engager dans des divisions asymétriques, qui répartissent différentiellement les facteurs protéiques dans les deux cytoplasmes non complètement étanches. L'une des deux proto-cellules filles a

un rôle de protospore et l'autre de cellule somatique. C'est le cas le plus simple de ce qui peut s'apparenter à la différenciation d'une lignée germinale. Par les jeux des interactions entre une famille de facteurs de transcription inégalement répartis, une cascade de réactions provoquerait, par une série d'allers et retours, la destruction finale de la cellule "somatique". Le non-cloisonnement total permet de penser la cellule soit comme une entité propageant son génome, soit comme deux protoentités entrant en conflits pour des ressources. C'est dans cet entre-deux ambigu et sur la base des mécanismes décrits ci-dessus, que l'on peut penser à la fois la mort cellulaire au niveau unicellulaire et à la fois l'apparition d'une forme simple d'apoptose dans un contexte multicellulaire et dans un environnement de stress.

-Le second modèle est celui du "péché originel"

Il s'appuie sur l'idée qu'un gène codant pour une fonction intervenant dans le métabolisme, la division, la différenciation du cycle cellulaire ne peut être sélectionné que dans la mesure ou son inhibiteur est lui aussi simultanément sélectionné, dans la logique de ne pas perturber par excès le fonctionnement de la cellule dans laquelle il se met à fonctionner. Or on peut constater que dans les cellules de mammifères comme chez les levures, des dérèglements de tels couples de gènes impliqués dans la coordination du cycle cellulaire (cyclines) peuvent provoquer des phénotypes associés à de l'apoptose. Ces couples de molécules peuvent représenter les ancêtres de gènes impliqués dans l'apoptose. Dans ce modèle, la mort cellulaire ne serait pas une acquisition fonctionnelle, mais le produit d'une continuelle adaptation à des équilibres instables pouvant mener à la destruction, comme on le postulait plus haut. L'autodestruction pourrait ici résulter, initialement, des différents usages dans

différentes voies de produits de l'expression génétique. Cela rendrait illusoire la recherche de molécules spécifiques de l'apoptose.

À travers ces deux modèles, non mutuellement exclusifs, on voit qu'il est possible de penser l'apparition de l'apoptose autrement que par la nécessité d'assurer la survie d'un organisme. L'apoptose n'est pas une acquisition de l'état multicellulaire, elle est la résultante de fonctions anciennes reposant sur des équilibres instables et/ou d'acquisitions subies de modules ambivalents qui ont été recrutés ultérieurement dans le cadre de nouvelles fonctions. Que l'organisme soit capable de canaliser un certain nombre de ces fonctions pour en faciliter la reproductibilité dans des conditions particulières, ceci est fort probable, mais ne justifie pas de présenter fondamentalement l'apoptose comme un suicide altruiste.

III. F Conclusion

Ainsi, et nous nous accordons en cela à la position de Jean-Claude Ameisen, l'apoptose ne nous semble pas pouvoir être appelée à proprement parler mort cellulaire programmée.

Elle semble être un phénomène évolutif lié à l'intériorisation de mesures de protections. Les gènes et protéines de l'apoptose ne forment pas un groupe indépendant, mais participent à diverses activités du métabolisme cellulaire. Ces activités multiples montrent que la programmation de la mort ne peut donc leur être strictement associée, et pourrait une fois de plus pousser à chercher de nouveaux facteurs. Nous pensons que cette recherche est illusoire. Nul doute que les voies d'apoptose seront dans l'avenir décrites de plus en plus précisément, et que d'autres

partenaires seront impliqués. Mais quand on constate la diversité des voies qui peuvent aboutir à l'apoptose, il semble que la recherche d'une connaissance précise de chacune d'entre elles soit moins importante qu'une approche plus globale et plus dynamique. Les mécanismes déclencheurs ressortent plus de ruptures d'équilibres instables que d'une programmation de destruction.

Ainsi il ne nous semble pas que les phénomènes apoptotiques s'opposent à l'approche endodarwinienne pour laquelle la survie de la cellule est, sinon prépondérante, du moins indissociable de celle de l'organisme. Bien au contraire, la continuité des différents types de morts cellulaires et leur non cloisonnement, le caractère dégénéré des régulations, semblent plaider en faveur d'un modèle adaptatif au niveau cellulaire.

Cela n'est évidemment pas en contradiction avec le fait que les organismes qui ont conservé et adapté ces mécanismes ont pu être sélectionnés. Il ne s'agit pas, ici encore d'opposer la sélection naturelle, qui joue sur les organismes, aux mécanismes endodarwiniens. Ces deux phénomènes d'échelles différentes coexistent et ne peuvent durablement être en opposition.

APPROCHE EXPERIMENTALE

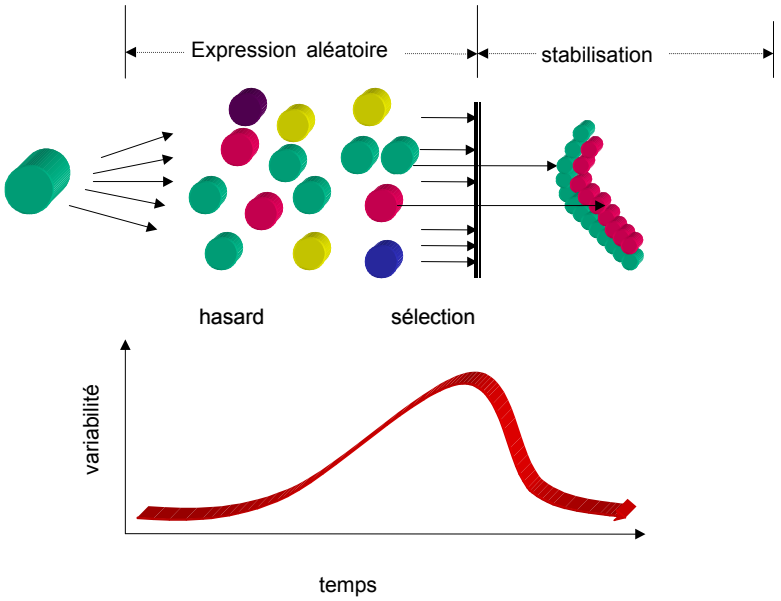


Figure 17 : Une prédiction de l'hypothèse endodarwinienne.
La différenciation cellulaire est causée par une variation dans l'environnement des cellules. Elles y réagissent par une expression aléatoire suivie d'une sélection des profils appropriés. Cela se traduit par une augmentation transitoire de la variabilité intercellulaire de l'expression génétique, maximale pendant la phase d'expression aléatoire.

I Introduction : Une prédiction du modèle endodarwinien nécessite d'être testée par une approche sur cellules isolées.

Nous nous plaçons dans l'hypothèse d'une dynamique endodarwinienne. Dans ce cadre, les cellules sont censées avoir des profils d'expression génétique aléatoires. Nous proposons que, dans ce cadre, on étudie des cellules clonales au cours d'une cinétique de différenciation. Dans un premier temps les cellules clonales sont adaptées à l'environnement dans lequel elles sont cultivées. On introduit un facteur de différenciation dans le milieu, et l'on constate alors une différenciation progressive de la population cellulaire. Cherchons dans ce cadre un critère qui nous permettrait de vérifier si ce processus ressort d'une dynamique endodarwinienne ou déterministe. Dans le premier cas, nous postulons l'apparition d'une augmentation transitoire de la variabilité intercellulaire de l'expression génétique (Figure 17). En effet, si les cellules clonales réagissent de manière aléatoire à un stress environnemental (ici, l'apparition d'un facteur de différenciation), ces réponses individuelles vont être différentes les unes des autres et la variabilité intercellulaire de l'expression génétique va augmenter. Mais dans un second temps, les cellules qui expriment les combinaisons de gènes qui leur permettent d'être adaptées à ce nouveau milieu vont être sélectionnées. A l'inverse, les cellules qui n'expriment pas les combinaisons de gènes nécessaires à la survie dans le nouvel environnement vont soit être amenées à poursuivre leur comportement exploratoire, soit à être éliminées. Cela a pour conséquence une homogénéisation progressive du profil d'expression génétique des cellules adaptées. C'est la raison pour laquelle l'augmentation de la variabilité n'est supposée que transitoire, suivie d'une restriction reflétant l'homogénéisation évoquée. Si les cellules clonales suivent au contraire une dynamique déterministe, on peut postuler qu'elle répondront, à chaque instant, de manière homogène (les unes vis-à-

vis des autres) aux variations environnementales. Cette réponse peut évoluer au cours du temps, mais il n'y a pas de raison de postuler qu'à un instant quelconque, les cellules réagissent différemment entre elles, étant entendu qu'elles possèdent le même génome, et donc les mêmes outils de réponse. On ne devrait pas constater d'augmentation transitoire de la variabilité de l'expression génétique au cours d'une phase de différenciation dans un contexte de dynamique déterministe.

Pour pouvoir réaliser cette expérimentation, il faut accéder à une approche sur cellules isolées. En effet, comme on l'a déjà évoqué plus haut, si l'on travaille à l'échelle d'une population cellulaire, on ne peut pas discriminer les deux hypothèses. L'observation du profil d'expression génétique d'une population entière, c'est-à-dire les espèces d'ARNm synthétisées en un point donné de la cinétique de différenciation peut soit représenter l'image fidèle de chaque profil d'expression individuel, soit n'être qu'un profil moyen qui masquera par définition la variabilité intercellulaire de l'expression génétique. C'est pourquoi il apparaît nécessaire, dans le cadre de cette expérimentation, de pouvoir accéder simultanément au profil d'expression génétique de plusieurs cellules individuelles et de constater si oui ou non on observe une augmentation transitoire de la variabilité définie. Cela passe par des études sur cellules isolées mais, nous le verrons dans un second temps, cela peut aussi être réalisé en s'intéressant aux variations des profils de méthylation de l'ADN.

II. Les techniques d'étude de l'expression génétique sur cellules isolées

II. A Contexte de l'apparition de ces techniques

Hormis quelques cas sporadiques où il était possible de mettre en évidence une variabilité intercellulaire de l'expression génétique (Spudich et Koshland, 1976), cette recherche n'a pu trouver sa faisabilité que par la mise au point de techniques qui permettaient d'étudier cette expression sur des cellules isolées, et de comparer ensuite les résultats.

Les raisons qui ont conditionné l'arrivée de ces techniques ne tiennent pas aux problématiques présentées de ce travail, et notamment la prédiction dont nous venons de présenter les grandes lignes. Elles tiennent à des questionnements particuliers ou l'on pouvait pressentir qu'une variabilité serait détectée. C'est le cas notamment de l'étude du système immunitaire et du système nerveux, dont on a vu qu'ils ont été étudiés dans une perspective de hasard/stabilisation, et pour lesquels il était légitime de présupposer une variabilité d'expression.

Ce fut ensuite, la volonté technique d'améliorer les seuils de détection des techniques d'amplification des acides nucléiques, principalement la réaction en chaîne à la polymérase (PCR) pour amplifier l'ADN et, depuis 1989 ; la combinaison de la transcription inverse de l'ARN en ADN et de la PCR (RT-PCR). Ultérieurement d'autres champs d'étude ont tiré profit de ces avancées techniques et de cette

amélioration des seuils de détection, comme l'étude des phases précoces des cancers. Là encore, l'idée qu'une mutation génétique puisse être à l'origine d'une variabilité d'expression rendait pertinente l'approche sur cellules isolées.

Ces approches sur cellules isolées ont longtemps été longtemps freinées par la difficulté qu'il y avait à récupérer de si faibles quantités de cellules. Les mises en suspension sont possibles pour les cellules en culture, ce qui permet même éventuellement une automatisation de l'aliquotage de cellule isolées. Cependant, la mise en suspension dans des solutions neutres empêche de connaître la localisation initiale de la cellule recueillie, ce qui peut poser des problèmes pour les études à base de sections tissulaires. D'autres techniques étaient donc nécessaires pour répondre à ces besoins. L'utilisation du laser a permis de révolutionner, dans les années récentes, les techniques de microdissection et a donné accès aux études sur cellules isolées (Busche, *et al.*, 2000). Ces dernières ont aussi bénéficié de la technique du patch-clamp (cf. infra).

II. B Deux grands types de méthodes pour analyser les variations intercellulaires

II. B 1 Les méthodes permettant de visualiser l'expression génétique par des marqueurs

Une des premières études ayant mis en évidence une expression aléatoire des gènes dans de cellules eukaryotes a utilisé un gène rapporteur LacZ sous contrôle d'un promoteur LTR-HIV1. Cette construction a été de manière stable dans la lignée

macrophagique murine RAW264. Des stimuli externes ont alors été appliqués pour stimuler l'expression de ce gène : le constat fut alors fait que seules certaines cellules, parmi une population dont toutes les cellules possédaient la construction exprimaient LacZ (Ross, *et al.*, 1994).

D'autres types de marquages existent (Todd et Margolin, 2002).

Ces marquages peuvent intervenir au niveau de l'ADN :

-agents fluorescents intercalants sur l'ADN global, permettant d'en déterminer la quantité dans chaque cellule et séparer celles ci sur la base de leur ploïdie.

-l'hybrydation in situ chromosome spécifique, région spécifique, ou séquence spécifique, peut permettre de discriminer les sous population encore plus finement.

Elle a notamment été utilisée pour étudier la coexpression de gènes de cytokines différentes dans des lymphocytes T CD4+ murins (Bucy, *et al.* 1995). Une autre illustration de l'utilisation de cette technique est l'étude de la transcription biallélique de deux gènes soumis à empreinte (cf. infra, pour plus de développements sur cette problématique) Igf2 et H19. La méthylation différentielle de ces gènes est telle que seul l'allèle d'origine paternelle de H19 et l'allèle d'origine maternelle de Igf2 sont supposés s'exprimer. Des études utilisant simultanément deux sondes d'hybridation, chacune propre à un de ces gènes, ont permis d'observer dans des proportions minoritaires mais significatives, l'expression simultanée des 2 allèles d'un ou des deux gènes. Cela va dans le sens d'une expression probabiliste où, autour du cas le plus probable qui est effectivement le plus représenté, toutes les combinaisons possibles d'expression sont marginalement observées (Jouvenot, *et al.*, 1999).

D'autres approches qui combinent jusqu'à 11 marqueurs fluorescents permettent de détecter des expressions simultanées de gènes et même de donner les coordonnées tridimensionnelles des sites d'expression dans le noyau (Levsky, *et al.*, 2002). Ce

genre d'approche est extrêmement prometteur, même si le nombre de marqueurs ne semble pas pouvoir être très extensible.

Les marquages fluorescents peuvent aussi concerner les protéines. L'une des plus récentes et des plus spectaculaires adaptations de ces techniques reposant sur des marqueurs fluorescent concerne des procaryotes. Dans le génome bactérien, deux gènes codant pour des protéines fluorescentes de différentes couleurs ont été intégrés de manière équidistante de l'origine de réplication, de part et d'autre sous le contrôle de promoteurs identiques. Ces précautions évitent le biais d'expression d'une protéine par rapport à l'autre. La fluorescence moyenne mesurée (en fait, une couleur intermédiaire) permet d'estimer si les deux gènes sont exprimés en même quantité et si cela est reproductible de cellule à cellule. Des modifications covalentes de ces protéines permettent de leur conférer une fluorescence beaucoup plus transitoire, ce qui permet d'éviter les phénomènes de persistance de la fluorescence et d'avoir ainsi une perception beaucoup plus dynamique. Les résultats de l'expérience montrent clairement la présence d'un fort bruit intrinsèque qui signe une expression génétique stochastique de cellule à cellule. Ces très élégants travaux sur des modèles procaryotes gagneraient à être adaptés aux eucaryotes.

Ces expérimentations peuvent permettre, par visualisation directe de la préparation, d'avoir une première approximation de la variabilité intercellulaire de l'expression génétique. Elles nécessitent un appareillage de tri cellulaire automatisé qui, sur la base du critère retenu, pourra isoler les cellules voulues. Elles sont néanmoins limitées par le nombre de gènes exprimés simultanément observables.

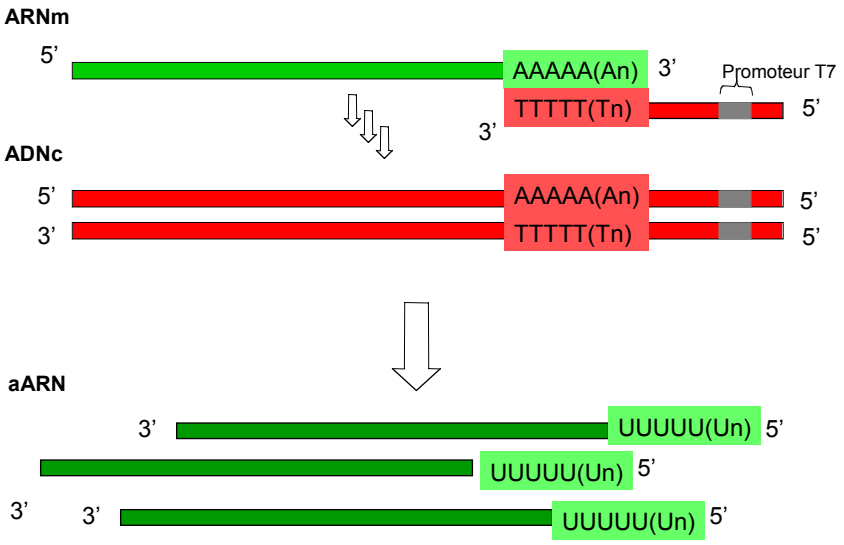


Figure 18 : l'amplification ARN.
L'ARNm est rétrotranscrit en ADN complémentaire avec une amorce comportant un promoteur pour une ARN polymérase. On réalise la synthèse du second brin et c'est ce dernier qui servira de matrice à une amplification linéaire de l'ARN antisens par rapport au brin initial. Tous les ARN messagers peuvent virtuellement être amplifiés.

II. B 2 Les méthodes reposant sur l'amplification du matériel génétique sur cellules isolées comportent plusieurs variantes.

Bien que ceci soit une moyenne, on estime à 10 000 le nombre moyen de gènes exprimés dans une cellule donnée. Une manière évidente de détecter de la variabilité est de s'intéresser à ces transcrits. Deux grands types de techniques peuvent être présentés (Monyer et Lambolez, 1995, Freeman, *et al.*, 1999). Elles répondent à des besoins souvent complémentaires, mais ont aussi permis de faire des constats convergents, par exemple la détection de l'épissage alternatif, qui permet, à partir des 10 000 gènes exprimés, d'expliquer en partie la présence estimée de 30 000 molécules différentes dans une cellule donnée (Eberwine, 2001).

Les méthodes reposant sur l'amplification linéaire de l'ARN (aARN).

Cette technique a été mise au point en 1990 sur des cellules individuelles (Van Gelder, *et al.*, 1990). A partir de l'ARNm total, de l'ADNc simple brin est synthétisé grâce à une amorce oligo-dT incorporant la séquence d'un promoteur d'ARN polymérase. Le brin d'ADNc complémentaire est ensuite synthétisé, séparé du premier, et c'est sur cette nouvelle matrice d'ADNc que l'ARN polymérase va permettre la synthèse de multiples copies d'ARN, qui seront antisens par rapport à l'ARNm initial, et seront donc nommés aARN (pour antisens autant que pour amplifiés) (Figure 18). L'intérêt principal de cette méthode est qu'elle permet potentiellement d'amplifier tous les ARNm. L'amplification linéaire permet aussi de s'affranchir en théorie des biais inhérents aux méthodes d'amplification exponentielle (cf. ci-dessous). Néanmoins, cette linéarité implique intrinsèquement une limitation

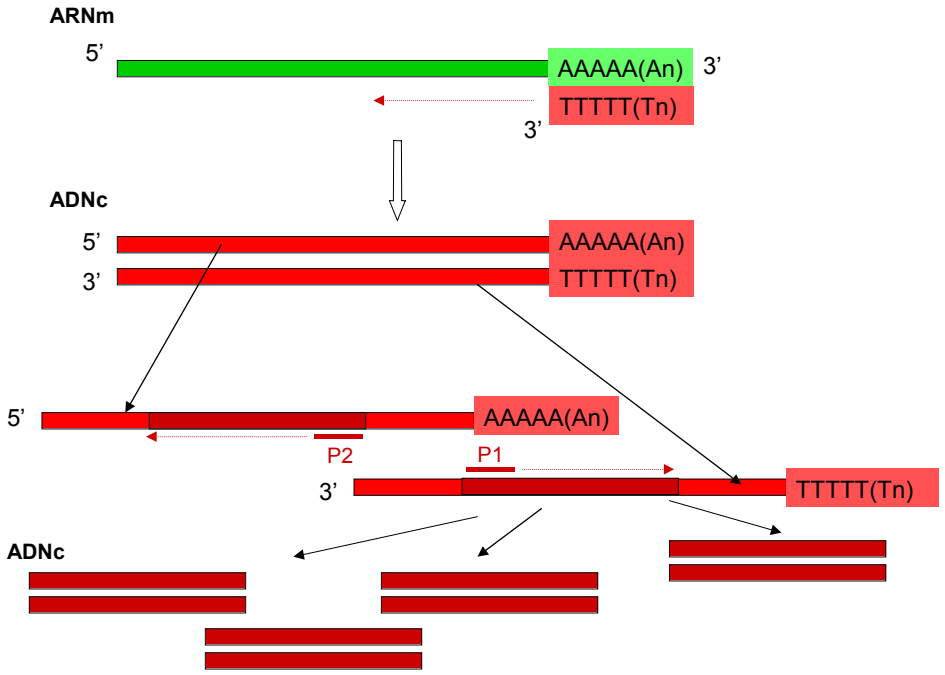


Figure 19 : La RT-PCR.
L'ARN messager est rétrotranscrit en ADNc, dont on réalise la synthèse du second brin. Les deux brins, puis chaque brin synthétisé, serviront de matrice ce qui permet une amplification exponentielle. Deux amorces spécifiques de la séquence cible sont nécessaires. A moins d'utiliser des amorces dégénérées, seules des séquences prédéfinies sont amplifiables.

en quantité de matériel amplifié (x10^6 dans le meilleur des cas), ce qui a des conséquences sur la détection des ARNm les plus rares. Cela explique probablement en partie pourquoi la méthode ne correspond qu'imparfaitement aux problématiques de ce travail, et s'est d'ailleurs relativement peu répandue dans les études sur cellules isolées. Mais, au sein d'au moins un groupe pionnier, cette technique s'est avérée capable de fournir des détections au niveau d'une cellule (Eberwine, *et al.* 1992, 2001a, 2001b), et même de très spectaculaires résultats subcellulaires (Kacharmina, *et al.*, 1999). La technique de présentation différentielle *(differential display*) qui peut être combinée avec l'aARN, nous semblait une piste prometteuse. Néanmoins son application à des cellules isolées n'a pas paru concluante, et n'est d'ailleurs apparue que très sporadiquement dans la littérature (Miyashiro, *et al.*, 1994 ; Thakker-Varia, *et al.*, 2001). Cette méthode souffre aussi de biais important, comme sa capacité à générer des faux positifs.

-Les méthodes reposant sur des adaptations de la RT-PCR.

La RT-PCR a été mise au point en 1989 dans la foulée des différentes adaptations techniques de la PCR qui ont permis la généralisation fulgurante de celle-ci dans les études de biologie moléculaire au cours des années 1990. Elle consiste en une étape de transcription inverse qui utilise une enzyme d'origine virale, la transcriptase inverse, et qui permet d'obtenir un ADN complémentaire (ADNc) du brin d'ARNm rétrotranscrit qui peut ensuite subir une PCR (Figure 19). Cette technique d'amplification exponentielle du matériel génétique était prometteuse pour qui s'intéressait aux espèces d'ARN en très petite quantité, ce qui est évidemment le cas de figure des travaux sur cellules isolées (on estime la quantité d'ARNm dans une

cellule à environ 0,5-1pg). Elle a d'ailleurs été présentée dans cette perspective (Rappolee, *et al.*, 1989). L'application de ces méthodes de RT-PCR sur cellules isolées présente l'avantage, par rapport à l'aARN, d'être globalement plus simple et plus rapide. Cette technique a été principalement appliquée par le groupe de Bernard Lambolez et Jacques Rossier, dans leurs études sur les récepteurs de l'α-amino-3-hydroxy-5-méthylisoxazole-4-propionate (AMPA) et du kainate dans les cellules neurales (Lambolez, *et al.*, 1992 ; Bochet, *et al.*, 1994 ; Ruano, *et al.*, 1995 ; Porter, *et al.*, 1999 ; Tsuzuki, *et al.*, 2001). Dans ces travaux, l'ARN est préalablement récupéré par "patch clamp", en utilisant une micropipette, permettant d'isoler le matériel génétique d'une cellule choisie et de commencer la réaction dans la pipette. La RT-PCR sur cellules isolées a été utilisée depuis dans de nombreux travaux qui ont trait principalement aux études des récepteurs neuronaux (Sucher, *et al.*, 2000 ; Dixon, *et al.*, 2000), et qui ont montré avec force la variabilité de l'expression (Kimoto, 1997).

Plusieurs limites cependant sont inhérentes à cette technique. La première tient à la très grande labilité in vitro des molécules d'ARN qui sont monobrins et, partant, beaucoup plus fragiles que l'ADN double brin. Hors de la cellule, ou de multiples enzymes interviennent dans sa maintenance et sa dégradation, l'ARN est notoirement fragile et doit être manipulé dans des conditions de haute protection. Par ailleurs, PCR comme RT-PCR sont des méthodes très sensibles aux contaminations par du matériel génétique exogène. Ce risque s'accroît exponentiellement avec le nombre de cycles d'amplification et est donc prépondérant dans les manipulations à partir de matériel génétique d'une seule cellule. Il justifie de nouveau donc de travailler dans des conditions très protégées.

Une autre limite est que les techniques d'amplifications fondées sur la PCR sont connues pour présenter un biais. Les séquences abondantes sont amplifiées préférentiellement, au détriment des séquences les plus rares. Ce problème pourrait paraître crucial dans les approches sur cellules isolées. Il le serait effectivement s'il s'agissait de prétendre que le profil amplifié est le reflet fidèle de la population initiale d'ARNm, ce qu'il n'est pas. L'objectif final est bien, dans notre approche, de comparer l'évolution d'un profil d'expression génétique au cours du temps. Ce sont donc les variations au différents temps d'une cinétique qui à terme sont signifiante. Qu'un profil d'amplification à un instant t ne soit pas homothétique du profil d'expression réel n'est pas problématique, à la réserve près que ces biais ne masquent pas excessivement la variabilité. En résumé, ces méthodes peuvent être un handicap quand on cherche à détecter un gène précis, mais ne posent pas de problème fondamental quand il s'agit de détecter des variations intercellulaires.

Une autre limite, du moins dans les premiers âges de la technique, fût qu'elle reposait sur une connaissance préalable des séquences à amplifier pour concevoir de manière séquence spécifique les amorces oligonucléotidiques. Cela a pour conséquence de limiter les approches globales et aussi de limiter le nombre de gènes amplifiables à partir d'une même cellule. La possibilité de synthétiser à façon des oligonucléotides dégénérés de séquences aléatoires (ou partiellement aléatoire) a permis de surmonter cet obstacle et d'amplifier simultanément des séquences variées, qu'il faut dans un second temps séparer.

C'est une méthode fondée sur le principe de la RT-PCR et utilisant des amorces partiellement dégénérées que nous avons contribué à améliorer dans la perspective de mettre en évidence une variabilité intercellulaire de l'expression génétique. Dans une certaine mesure, nous verrons que cette technique peut combiner les avantages

des deux approches, aARN et RT-PCR. D'autres existent, comme l'amplification globale d'ADNc proposée de manière très précoce par le groupe de Norman Iscove (Brady, *et al.*, 1990, 1995 ; Billia, *et al.*, 2001 ; Iscove, *et al.* 2002)

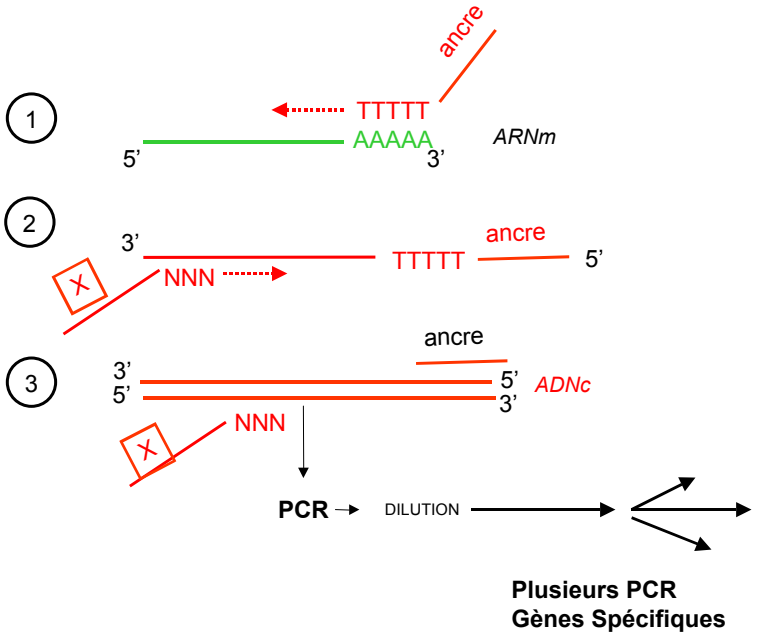


Figure 20 : Le principe du TPEA.
La première phase est une amplification non spécifique décomposée en trois temps.
1 La rétrotranscription des ARNm par une amorce polyT comportant une séquence ancre en amont.
2 La synthèse du second brin par une amorce contenant une séquence (X) absente des bases de données mammifères et une séquence partiellement dégénérée (NNN).
3 Une amplification sur le mode PCR est ensuite réalisable avec l'amorce ancre et la séquence X, ce qui conduit à une amplification exponentielle de tous les ARN rétrotranscrits.
Une dilution est possible ensuite, suivie de PCR gènes spécifiques classiques.

II. C Une modification du protocole de TPEA

Pour atteindre nos objectifs, nous avons en effet utilisé la technique dérivée de la RT-PCR par amplification de l'extrémité 3' (*3 Prime End Amplification* - TPEA). Cette technique récente s'est révélée être un outil d'étude efficace de l'expression génétique sur cellules isolées dans diverses lignées cellulaires d'origine mammifère (Dixon, *et al.*, 1998). Quand bien même cette technique n'est pas quantitative, elle a pour intérêt principal d'amplifier potentiellement toutes les séquences d'ARNm présents dans la cellule, y compris ceux présents en très petit nombre de copies. Une majorité des ARNm présents dans une cellule est, rappelons le, présent à moins de 10 copies en moyenne. Cette RT-PCR repose sur deux amplifications successives, et fonctionne sur le principe d'amorces partiellement dégénérées (Figure 20). Dans un premier temps, le lysat cellulaire contenant les ARNm est rétrotranscrit, étape à laquelle succède immédiatement une amplification non spécifique grâce à des amorces permettant l'amplification de tous les ARN polyadénylés. Dans un second temps, le produit de cette amplification non spécifique peut être dilué et fractionné, chaque fraction étant susceptible d'être alors soumise à une PCR gène spécifique. Grâce à ce fractionnement, les conditions proposées dans la technique initiale permettaient d'obtenir en théorie 40 amplifications simultanées à partir d'une même cellule, soit la détection potentielle de l'expression de 40 gènes (20 en duplicats), et ainsi de suivre les profils d'expression et de détecter des variations de cellule à cellule. Cependant, on a vu que les contaminations sont un problème critique des techniques d'amplification de matériel génétique, tout particulièrement quand il s'agit de quantité extrêmement limitées d'ARN comme c'est toujours le cas pour les techniques de RT-PCR sur cellules isolées. Ces problèmes

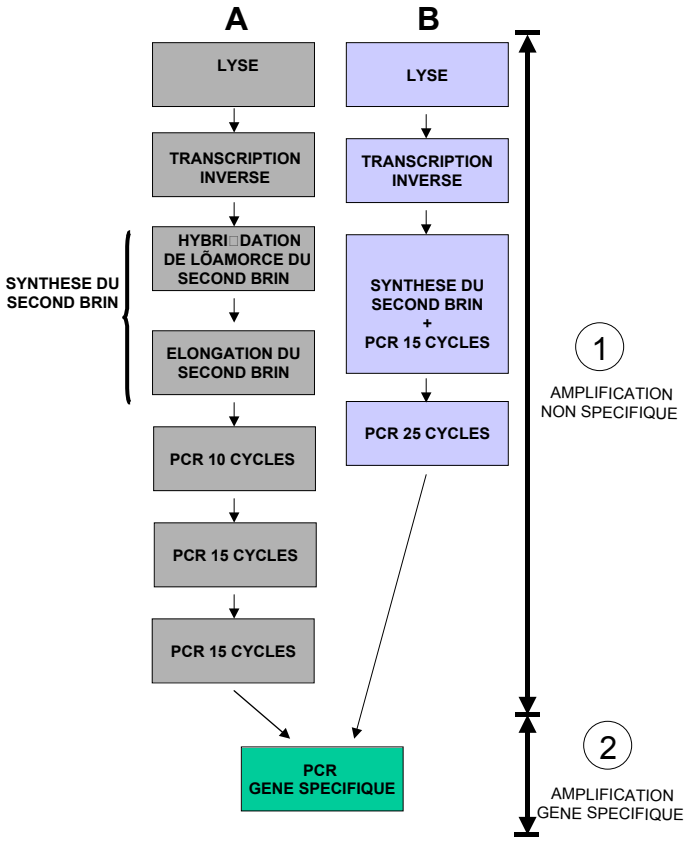


Figure 21 : Les simplifications du protocole de TPEA.
Le protocole est passé de 7 à 4 étapes, ce qui permet de diminuer de moitié les manipulations intermédiaires. Les étapes initialement fragmentées de la synthèse du second brin ont été groupées entre elles ainsi qu'aux 15 premiers cycles de PCR non spécifiques. Le nombre total de cycles est inchangé.

potentiels sont démultipliés quand il existe une phase d'amplification non spécifique, à laquelle n'importe quel ARNm exogène contaminant pourrait participer. Or la description du TPEA dans l'article original, pour séduisante qu'elle pouvait être sur le papier, était néanmoins celle d'une technique très fractionnée, et donc sensible aux contaminations. Après une acclimatation de notre technique au laboratoire, il apparaissait donc nécessaire d'apporter des modifications à cette technique pour diminuer ces risques. Par ailleurs, toute amélioration de la sensibilité de la méthode, permettant une augmentation du nombre d'aliquots possibles après fractionnement, allait dans le sens d'une plus grande finesse d'analyse des profils d'expression. Les deux questions ne sont d'ailleurs pas indépendantes, puisque la multiplication des étapes intermédiaires dans cette technique causait concrètement la perte progressive, étape après étape d'une fraction non négligeable du matériel de départ. Diminuer le fractionnement de la méthode était donc, à plusieurs titres, une question stratégique. Nous proposons donc et présentons donc ici de telles modifications de la technique du TPEA, qui maintiennent la qualité des résultats. Nous avons réduit de sept à quatre le nombre des sous-étapes de la phase d'amplification non spécifique de la technique du TPEA (Figure 21). Plus précisément, nous avons cherché et déterminé les conditions pour lesquelles un regroupement des deux étapes, initialement fractionnées, de synthèse du second brin avec une PCR non spécifique de 15 cycles. Ce regroupement est suivi d'une PCR non spécifique de 25 cycles. Il aurait pu paraître logique de chercher alors à regrouper tous ces cycles de PCR non spécifiques (qui font intervenir exactement les mêmes réactifs, notamment les mêmes amorces) en un seul bloc de 40 cycles, mais les essais dans ce sens n'ont pas permis de conclure à une fiabilité maintenue de la méthode. Le fractionnement, même rationalisé comme ici, reste nécessaire pour pouvoir, semble-

t-il, permettre à la polymérase d'agir dans les conditions optimales, en la réintroduisant à mi parcours.

Les simplifications introduites réduisent largement les risques de contaminations et permettent en outre un gain de temps significatif. Ce dernier point est crucial dans des conditions où l'individualisation des cellules n'est pas automatisé et où l'on cherche à étudier simultanément plusieurs cellules. En effet, les premières étapes, celles de l'amplification non spécifique, ne peuvent être différées une fois initiées, en raison de la fragilité de l'ARN et du faible nombre d'ADN amplifié. Il faut donc mener de front les analyses sur des cellules prélevées aux même point de la cinétique. Le gain de temps estimé lors de l'amplification non spécifique est, pour une analyse simultanée de 30 cellules isolées, d'une heure et demie sur une durée initiale jamais inférieure cinq heures.

Nous avons mené notre étude sur des lignées immortalisées murines de neuroblastome (lignée N1E-115). Il s'agit de cellules adhérentes. Cette lignée est un modèle classique de la différenciation neuronale. Celle-ci peut être provoquée notamment par une privation de sérum de veau fœtal (SVF) dans le milieu de culture, qui entraîne l'apparition de phénotypes neuronaux dans les 48 heures, accompagnés d'une mort cellulaire importante (Clejan, *et al.*, 1996). Ces cellules sont cultivées dans du DMEM, enrichi de glucose à 4,5 g/L (Invitrogen, Carlsbad, CA, USA) et de sérum de veau fœtal décomplémenté a 10 %, et de pénicilline-streptavidine 10mL/L (Invitrogen).

Malgré une dérive génétique notoirement faible, elles ont été clonées par mesure de précaution : les expériences décrites ont été systématiquement conduites à partir

d'une population d'origine clonale que nous avons réalisée, ceci pour diminuer autant que possible le -faible- risque d'hétérogénéité génétique.

Les transcriptions inverses et les réactions fondées sur la PCR ont été réalisées sur un appareil GeneAmp PCR 9700 (Applied Biosystems, Foster City, CA, USA). L'amorce de transcription inverse (5'-CTCTCAAGGATCTTACCGCTTTT-TTTTTTTTTTTTT-3'), celle du second brin (5'- CTGCATCTATCTAATGC-TCCNNNNNCGAGA-3') où N représente C, G, T ou A, et l'amorce à ancre 3' (5'-CTCTCAAGGATCTTACCGC-3'), ont été synthétisées à façon par Invitrogen, et sont celles qui ont été décrite par ans la technique initiale (Dixon, *et al.*., 1998).

Après remise en suspension des cellules N1E-115 par trypsinisation et rinçage, les cellules subissent une dilution séquentielle, de manière à obtenir la possibilité d'avoir une cellule par microlitre en théorie. Un microlitre de cette dilution est alors déposé dans chaque puits d'une plaque de Terasaki. Tous ne contiendront pas exactement une cellule. C'est une observation au microscope qui permet de ne retenir que les puits ne contenant qu'une seule cellule pour la suite de l'expérience.

Ces puits reçoivent alors un mélange permettant une lyse cellulaire dans des conditions telles que le noyau n'est pas lysé. On pourra donc se débarrasser de ce noyau et donc des contaminations en ADN génomique. Ce mélange de lyse d'un volume final de 7 microlitres contient du tampon de synthèse du simple brin (50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$; Invitrogen), 5 mM de Nonidet P-40 (Sigma, St. Louis, MO, USA), 1.5 U d'inhibiteur de Rnase RNase-OUT(Invitrogen). Après cinq minutes sur glace, les noyaux sont évacués par centrifugation (8000 g pendant 5 minutes a 4°C). Chaque surnageant est alors transféré dans un nouveau tube où il est rétrotranscrit dans un volume final de 10 microlitre après ajout de tampon de

synthèse du simple brin, de transcriptase inverse MMLV (Moloney murine leukemia virus) -Invitrogen-, et de 0,5 ng de d'amorce transcription inverse. Cette transcription inverse est réalisée à 37°C pendant 60 minutes.

La synthèse du second brin, et 15 cycles de PCR, initialement fractionnés, ont été regroupés en une seule étape, (Figure 21, p.159). Le produit de la transcription inverse est intégralement reversé dans un volume supplémentaire de réaction de PCR de 10,4 µL. Ce mélange de 20,4 µL contient aux concentrations finales : $MgCl_2$ à 4.5 mM, Tris-HCl (pH 8.3) à 67 mM, dNTP à 0.5 mM (Amersham Biosciences UK, Buckinghamshire, UK), 1 ng d'amorce de second brin, 0.4 ng d'amorce à ancre 3', l'ADN polymérase AmpliTaq à 0.35 U (Applied Biosystems). Ce mélange subit les étapes suivantes en un bloc : 50°C pendant 15 min (incubation), 72°C pendant 10 minutes (extension), 92°C pendant 5 minutes (dénaturation), 15 cycles de PCR (2.5 minutes à 92°C, 1.5 minute à 60°C, et 1.5 minute à 72°C), suivie d'une élongation à 72°C pendant 10 minutes. Ce mélange de 20,4 µL est porté à 30.4 µL de manière à ne pas modifier les concentrations des différents composants, à l'exception de l'ajout de 125 ng de l'amorce de second brin et 50 ng de l'amorce à ancre 3'. Il est alors soumis à une dénaturation de 5 minutes à 92°C, à 25 cycles additionnels de PCR (comme décrits ci-dessus) et à une élongation finale de 10 minutes à 70°C.

De plus nous avons augmenté le taux de PCR gène-spécifiques en parvenant à n'utiliser, après dilution, que 3 µL de ce volume de préamplification au lieu de 5 microlitres. Le nombre potentiel de PCR gènes-spécifiques est donc passé de 40 à 50 (20 à 25 en duplicats). Pour cela, le volume 30,4 µL d'amplification non spécifique a été dilué à un volume final de 150 µL de Tris-HCl 10 mM, EDTA 0,1 mM. Des

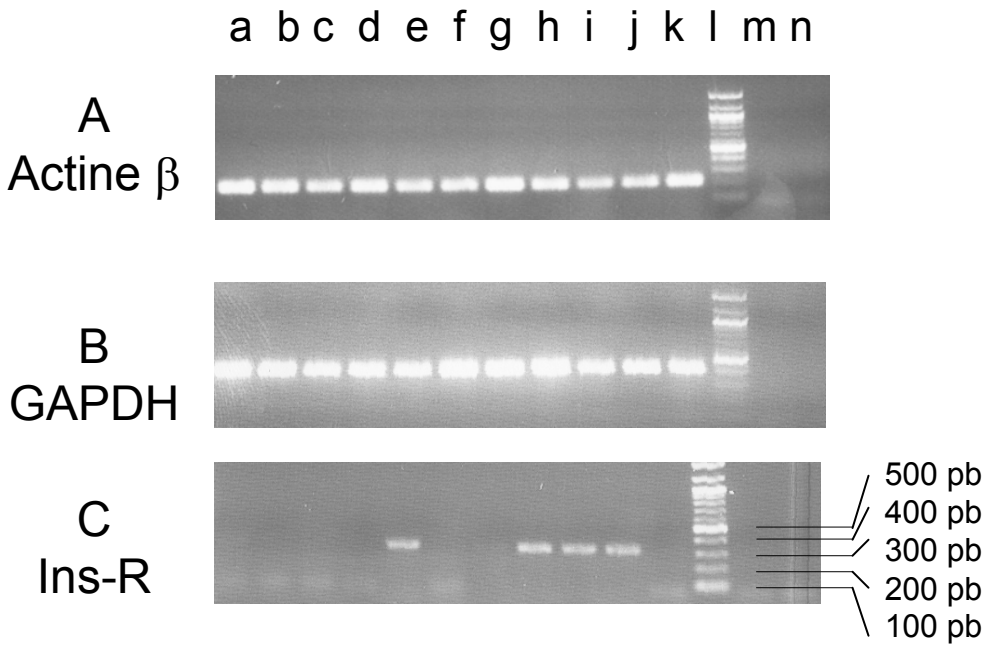


Figure 22 : Résultats d'amplification pour trois gènes par TPEA modifié.
Onze cellules individuelles clonales (a-k) ont subi une procédure de TPEA. l : marqueur de taille (100-bp DNA leader ; New England Biolabs) ; m : témoin négatif sans transcriptase inverse ; n : témoin négatif sans cellule. On détecte l'expression des gènes de l'actine-b et du GAPDH dans onze cellules sur onze. En revanche, le récepteur de l'insuline n'est exprimé que dans quatre cellules sur onze. Amorces : actine β (209 pb; GenBank ® accession no. V01217) 5'-CATCCATGCCCT-GAGTCC- 3'(forward) et 5'-ACACCTCAAACCACTCCCAG-3'(reverse); GAPDH (408 pb; accession no. X02231) 5'-GAC CCCTTCATTGACCTCAACTAC-3'(forward) et 5'-GATGCCAAAGTTGTCA TGGATGAC-3'(reverse) ; et récepteur de l'insuline (369 pb; accession no. NM 010568) 5'-TCACAAAGATGGCGTATT-3'(forward) et 5'-GACTGGAGCAA ACACAAC-3'(reverse).

aliquots de 3 µL de cette solution ont été soumis à des PCR utilisant des amorces séquences spécifiques. Chaque échantillon a été ajouté à un mélange réactif pour obtenir un volume total de 50 µl contenant du tampon PCR II 1X (100 mM Tris-HCl, 500 mM KCl ; Applied Biosystems), du $MgCl_2$ à 3,5 mM (Applied Biosystems) 12,5% de saccharose, du β-mercaptoéthanol à 12 mM, dNTP à 0,5 mM, amorces oligonucléotidiques (100 ng chaque) et ADN polymérase AmpliTaq (0,6 U).Ces concentrations sont celles préconisées par le protocole d TPEA initialement publié, à l'exception du *cresol red*, dont l'absence n'a pas empêché d'obtenir des résultats et qui a en conséquence été omis. Les températures d'hybridation des séquences ont été adaptées pour chaque couple, mais le reste des conditions de ces PCR gènes spécifiques est là encore identique à celles décrites initialement. A l'issue de la PCR, les produits d'amplification sont classiquement observés après dépôt et migration sur un gel d'agarose à 0,7% (masse/volume) contenant du bromure d'éthidium, agent intercalant, qui révèle les bandes d'ADN.

Ce protocole modifié de TPEA a donné des résultats sensibles et reproductibles. La figure 22 montre ceux qui peuvent être obtenus en utilisant les couples d'amorces pour l'actine-β, le GAPDH et le récepteur de l'insuline. Il faut d'abord noter qu'aucune détection n'est observée pour les expériences en absence de reverse transcriptase, ce qui démontre l'absence de contamination par de l'ADN génomique. Cela est renforcé par le fait que les amorces pour GAPDH sont de part et d'autre d'un intron.
Comme on pouvait s'y attendre, les gènes constitutifs d'actine-β et de GAPDH sont détectés comme exprimés dans chacune des cellules isolées. Ils constituent d'ailleurs un marqueur qui témoigne de l'existence d'une activité d'expression dans les cellules considérées. Mais par contraste, on constate que le récepteur de

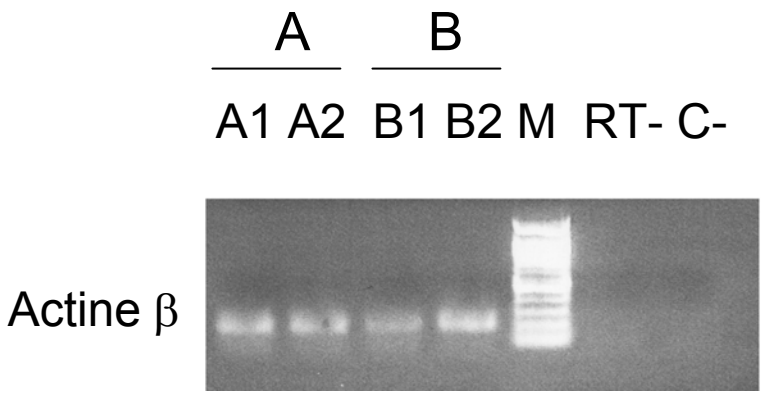


Figure 23 : **Contrôle de TPEA.**
Deux cellules (A, B) sont traitées par TPEA; chaque contenu cytoplasmique est divisé en 2 moitiés (Resp. A1 et A2, et B1 et B2) avant transcription invers ; chaque moitié subit alors l'ensemble des étapes. On ne détecte pas de biais d'expression entre les deux moitiés

l'insuline, qui n'est pas constitutivement exprimé dans ces cellules, l'est cependant dans quatre d'entre elles sur onze. A titre de contrôle additionnel, nous avons par ailleurs divisé en deux moitiés les ADN complémentaires obtenus après l'étape de transcription inverse et comparé les signaux obtenus après qu'un protocole complet de TPEA modifié avec des amorces d'actine-β a été appliqué à chacune de ces moitiés. Aucun biais d'amplification entre ces moitiés n'a été remarqué (Figure 23)

.

L'ensemble de ces résultats permet donc de conclure que les améliorations proposées rendent la technique du TPEA plus rapide, plus sensible, et plus simple. En ceci, elle semble être un compromis digne d'intérêt entre les deux principales méthodes que sont l'aARN et la RT-PCR. Elle est compatible avec un tri automatisé des cellules par cytométrie en flux. Elle permet d'envisager une véritable approche globale des profils d'expression et devraient permettre qu'elle soit plus largement utilisée dans le cadre des études sur cellules isolées.

On observe une variabilité d'expression qui soutient l'existence d'une expression stochastique. L'expression du gène du récepteur de l'insuline varie de cellule à cellule au sein d'un milieu homogène. Ceci ne peut donc être le fait de signaux spécifiques, et minimise par ailleurs la nécessité d'existence de tels signaux.

Deux gènes constitutifs, sur lesquels doit exister une forte pression de sélection pour leur expression, sont exprimés dans toutes les cellules. Cela contraste avec le gène du récepteur à l'insuline, dont l'expression n'est pas attendue dans des conditions basales comme celles de notre expérience et pour lequel une expression aléatoire est possible. Même si ces résultats sont préliminaires, et devront être multipliés (nombre de gènes, conditions différentes) il sont déjà compatibles avec une grille de lecture sélective. Ils sont le socle potentiel d'une dynamique endodarwinienne.

II. D Conclusion sur les études d'expression génétique sur cellules isolées.

Nos résultats vont dans le sens de toutes les études sur les cellules isolées : elles confirment la dimension stochastique de l'expression génétique d'au moins certains gènes quelles que soient les techniques employées. Il est hautement significatif que dans leur grande variété, toutes ces méthodes aboutissent à la même conclusion. La question du statut de cette stochasticité évoqué en introduction n'a pas été tranché dans le temps de la thèse par le biais de cette technique. Il gagnerait à l'être au cours de des temps à venir. La pluralité des méthodes incite naturellement à proposer de combiner les approches disponibles, adaptées à chaque système expérimental pour enrichir et consolider les résultats obtenus. En effet, il est toujours possible, malgré les efforts entrepris pour une s'assurer de la plus grande homogénéité possible de la population de départ, de mettre les variabilités observées sur le compte d'une différence génétique (mutation) imprévue. Pour cela, l'approche combinée d'amplification de l'ADN et de l'ARN d'autre part d'une même cellule isolée est une voie à suivre qui n'est pas incompatible avec le protocole actuel du TPEA. Une généralisation des approches sur cellules isolées serait de nature à rendre systématique la nuance sur le rôle du programme génétique. Une évaluation précise du bruit de fond, suivie d'une évaluation de l'augmentation transitoire de variabilité de l'expression génétique suite à une modification environnementale permettra à n'en pas douter de donner des réponses concrètes sur les marges de manœuvre permises par un génome à sa cellule.

Nous avons pu néanmoins mettre en évidence une variabilité transitoire d'un paramètre corrélé à l'expression génétique, de nature à rendre pertinente une approche endodarwinienne de la différenciation. Ce paramètre est la méthylation de l'ADN, et nous le présentons dans la seconde partie des résultats expérimentaux

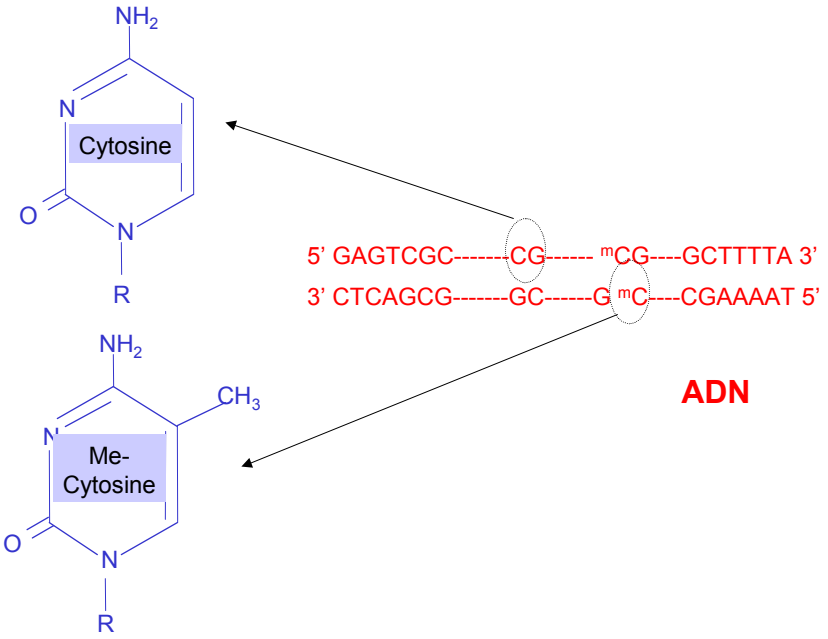


Figure 24 : Comparaison entre les formules de la cytosine non méthylée et méthylée des îlots CpG.
Le groupement méthyle est sur le carbone 5.

III L'étude de la variabilité épigénétique au cours de la différenciation.

III A La problématique de l'étude de la méthylation

Les processus épigénétiques ont suscité un intérêt croissant au cours des années récentes. Dans le contexte déterministe, ils ont été envisagés comme des paramètres de nuance du rôle du génome.

La méthylation de l'ADN, chez les mammifères, est une modification covalente sur le cinquième carbone des cytosines des dinucléotides CpG contenues dans le génome (Figure 24). La plupart des CpG sont méthylés dans le génome humain. Mais ces dinucléotides sont sous-représentés et ne représentent qu' 1% seulement du génome total. Cette sous-représentation est attribuée à un taux supérieur de mutation des résidus cytosines méthylés par rapport aux non méthylés.

Au sein d'un génome globalement méthylé sur ses CpG, 1 à 2 % de sa séquence ne l'est pas. Les zones concernées sont appelées îlots CpG car il sont enrichis en CpG (déméthylés, donc), par rapport au reste du génome. Ces îlots sont longs d'1 kb en moyenne, sont non seulement riches en CpG mais aussi riches en G/C, et sont trouvés dans les extrémités 5' des gènes. Les gènes "de ménage" connus et certains gènes exprimés préférentiellement dans certains tissus possèdent de tels îlots CpG.

La méthylation des gènes est corrélée positivement à la répression de l'expression génétique. Les gènes de ménages, qui ont un îlot CpG associé à leur promoteur, sont exprimés continûment et en parallèle on n'observe pas de méthylation

transitoire de leur îlot associé. Des études globales ont montré une corrélation entre l'activité génétique et l'hypométhylation des promoteurs. Cette corrélation est plus faible dans les CpG à l'intérieur du gène.

La méthylation des gènes subit des changements complexes pendant le développement et la différenciation chez les mammifères. Après la fertilisation, une déméthylation globale survient puis est suivie d'une reméthylation globale. Après la gastrulation, l'activation de gènes est corrélée à une perte de méthylation dans les gènes tissus-spécifiques de différentes lignées somatiques. Le cas de la lignée germinale est différent. Celle-ci présente globalement une hypométhylation à mi gestation ; suggérant soit qu'elle a échappé à la vague de méthylation globale *de novo* soit qu'elle a subit une seconde vague de déméthylation.

La méthylation différentielle de certains gènes en fonction de leur origine paternelle ou maternelle explique le phénomène d'empreinte génétique. Les gènes soumis à empreinte sont sujets dans la lignée germinale, à des modifications particulières transmises par les gamètes et nécessaires au développement embryonnaire. Les mécanismes précis de cette méthylation différentielle sont encore débattus.

Entretien de l'état de méthylation

Une propagation de l'état de méthylation d'un îlot CpG est assurée après la réplication de l'ADN par une réaction enzymatique de transfert de groupement méthyle sur les résidu cytosine du brin d'ADN néosynthétisé et initialement non méthylé. Cette activité qui utilise le CpG hémiméthylé comme substrat est considérée comme une activité de maintenance : la perte de la méthylation peut survenir suite à un empêchement de cette activité enzymatique.

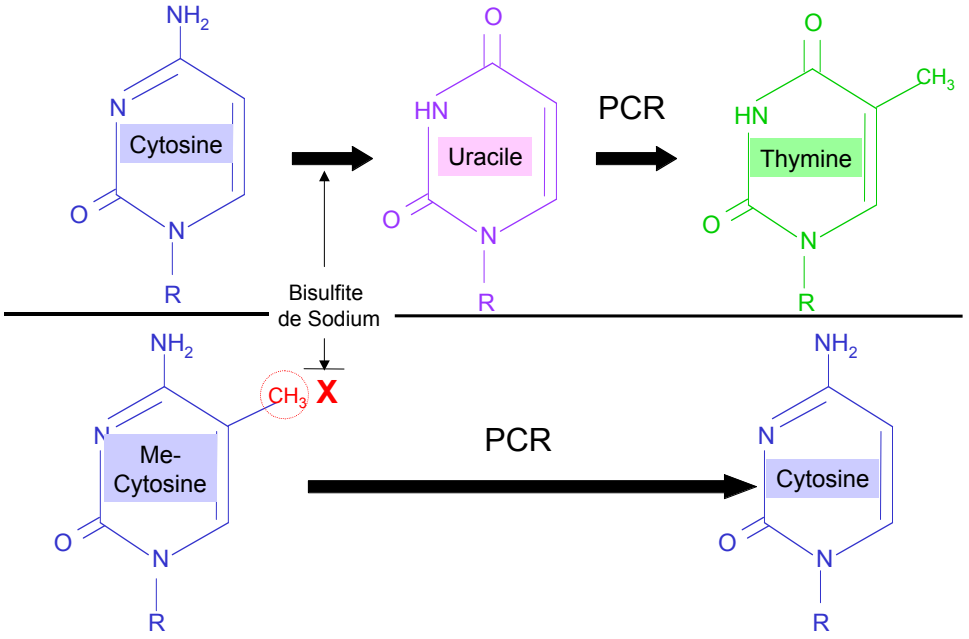


Figure 25 : Différence d'action du Bisulfite de Sodium sur la cytosine et la méthyl-cytosine.
La cytosine non méthylée est convertie en uracile. Après PCR, elle sera remplacée par une thymine. La cytosine méthylée n'est pas modifiée par le traitement au bisulfite. La résultante de ces effets est une variation de séquence entre sites méthylés et non méthylés.

Plusieurs enzymes ont été progressivement identifiées comme ADN méthyltransférases (DNMT). Si elles se virent attribuer un rôle dans la maintenance dans la méthylation, leur capacité à provoquer une méthylation *de novo*, du moins *in vivo*, a été un temps discutée, mais est maintenant admise.

La déméthylation, elle est considérée comme survenant principalement par le biais de deux cycles de réplication sans maintenance. Des exemples de déméthylation enzymatique réplication-indépendante ont été avancés. La régulation de ce processus est peu documentée.

III B Méthodologie

Participant à des travaux parallèles (voir annexe) sur l'empreinte génétique et dans ce contexte, à l'étude des variations de la méthylation de certains gènes différentiellement méthylés, il nous a semblé utile d'appliquer ces techniques d'étude à notre modèle de différenciation in vitro, sur des gènes soumis à empreinte, dont on n'avait pas de raison *a priori* de les supposer variables au cours de la différenciation.

L'étude de la méthylation après traitement de l'ADN au bisulfite de sodium repose sur la capacité de ce dernier à convertir chimiquement les cytosines non méthylées d'un dinucléotide CpG en uraciles (qui seront transformées en thymidine au cours de la PCR qui suit), sans pouvoir convertir les méthylcytosines de ces CpG. Après une PCR avec des amorces spécifiques des séquences modifiées, la méthylation de chaque CpG peut être établie en sous clonant et en séquençant le produit

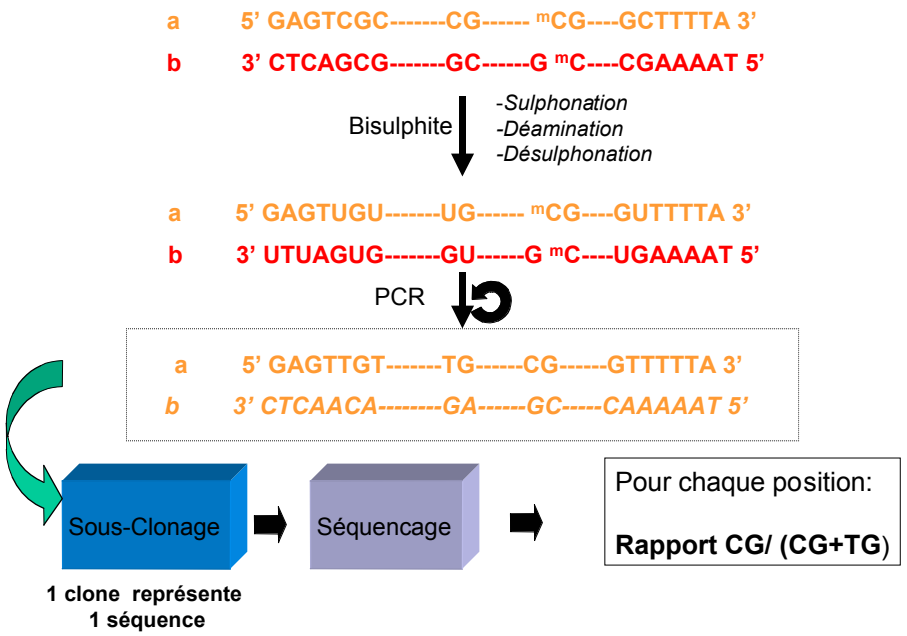


Figure 26 : Principe du traitement des cellules au bisulfite.
Celui-ci modifie la séquence en fonction des mécanismes décrits. La séquence d'intérêt subit une PCR. Le produit de PCR est sous-cloné et séquencé. On analyse plusieurs séquences pour déduire, à chaque position, le pourcentage de méthylation initial.

d'amplification. Chaque clone représente une molécule unique d'ADN de l'échantillon initial.

Toutes les cytosines n'appartenant pas à un îlot CpG auront été converties chimiquement en uraciles puis après la PCR en thymidines. Cela constitue un témoin interne. En ce qui concerne les cytosines non méthylées des CpG, elles auront étés converties en uracile puis thymidine après PCR à l'inverse des 5-méthylcytosines (Figure 25, page précédente). Il est ainsi possible, pour chaque CpG, d'établir un pourcentage de méthylation/déméthylation et d'établir sur ce plan une variabilité intercellulaire de la méthylation (Figure 26). Cette méthode répond ainsi à nos besoins, sans qu'il soit nécessaire de passer par une approche sur cellules isolées.

Un nombre de clones approprié fournira une estimation de la proportion étudiée et l'on pourra suivre cette proportion au cours d'une cinétique de différenciation.

Nous avons donc soumis les cellules N1E-115 à une cinétique de différenciation. Pour ce faire nous avons mis en culture une population de cellules obtenues à partir d'une seule cellule mère pour s'assurer de l'origine clonale de la population. La cinétique de différenciation a été conduite en privant les cellules de SVF. A intervalles donnés, trois types de populations cellulaires ont été simultanément traitées au bisulfite selon un protocole publié (Kerjean, *et al.*, 2001)

-Les cellules en cours de différenciation, adhérentes

-les cellules mortes dans le surnageant.

-Des cellules non soumises à la différenciation.

Les principales étapes du protocole de traitement de l'ADN au bisulfite de sodium sont les suivantes :

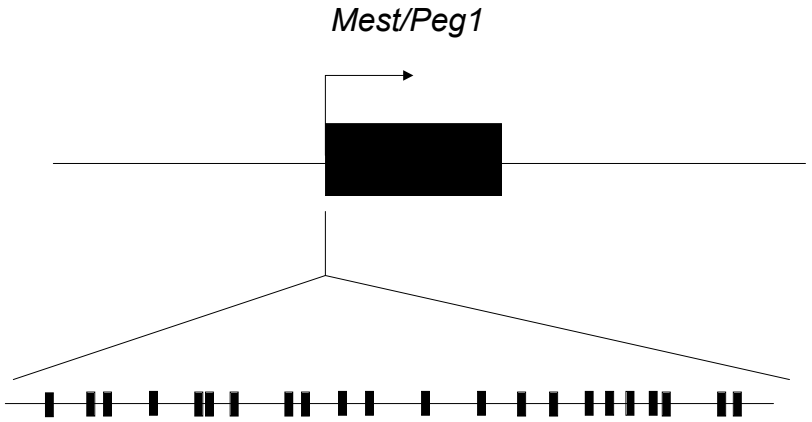


Figure 27 : Structure génomique de la zone de méthylation différentielle amplifiée.
22 positions CpG sont amplifiées dans la région du promoteur de Mest/Peg1 (accession No AF017994; -201, +116 par rapport à l'origine de transcription). Flèche : site d'initiation de la transcription ; Rectangle noir : exon ; petits carrés noirs : CpGs de la région amplifiée.

-Les cellules subissent une lyse pendant une nuit par action de protéinase K dans une incubation a 50°C.

-L'ADN est ensuite dénaturé par traitement NaOH pendant 30 minutes.

Puis la solution de bisulfite de sodium-hydroquinone est appliquée pendant quatre heures à 50°C.

-L'ADN modifié est ensuite purifié et élué, il subit un dernier traitement à la soude pendant quinze minutes à 37°C, suivi d'un ajout d'acétate d'ammonium. L'ADN modifié est alors soumis à une précipitation dans de l'éthanol 100% à -20°C pendant 30 minutes et rincé une fois à l'éthanol 70%

-Il est enfin resuspendu dans de l'eau, prêt à être utilisé ou à être immédiatement conservé à -20°C

-L'ADN est ensuite soumis à une PCR gène spécifique dans les conditions décrites (Kerjean, *et al.*, 2001).

Dans la phase exploratoire de ce travail, nous nous sommes restreints à l'étude de la région de méthylation différentielle (RMD) du gène Mest/Peg1. Cette RMD est constituée du promoteur, du premier exon et d'une partie du premier intron. Elle est différentiellement méthylée sur chaque allèle en fonction de la provenance maternelle (méthylé) ou paternelle (déméthylée - transcriptionnellement active) de celui ci. La région que nous amplifions est une séquence de 537 paires dans le promoteur (-201, + 116 vis à vis du site d'initiation de la transcription). Elle contient 22 sites CpG soumis à méthylation différentielle. Seuls les 10 derniers sont situés dans la séquence transcrite (Figure 27).

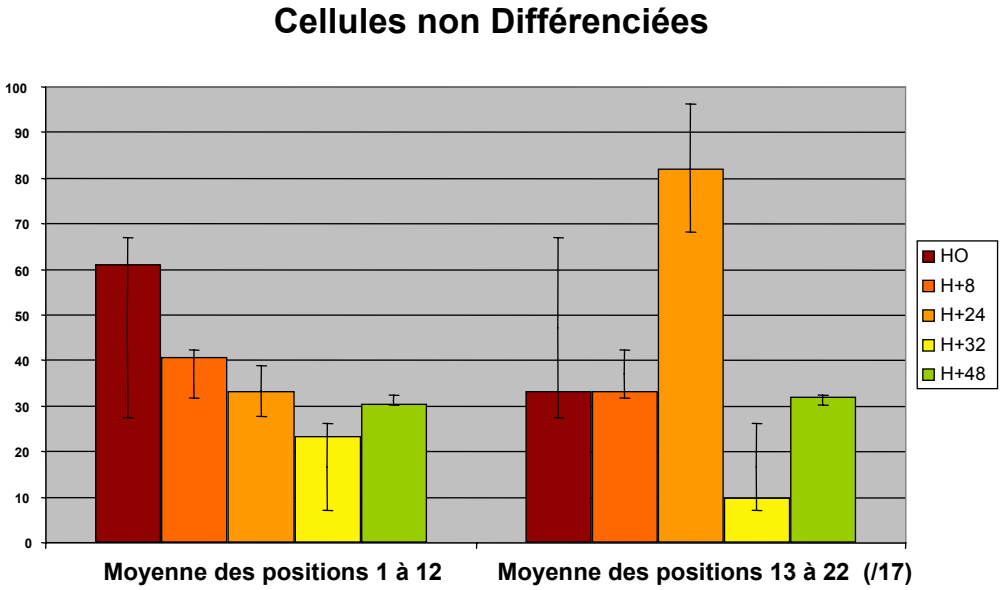


Figure 28 : Evolution du taux de méthylation au cours d'une cinétique de différenciation - Cellules non différenciées (cinétique témoin).
Gauche : moyenne des positions de 1 à 12, correspondant à la séquence non transcrite. Droite : moyenne des positions 13 à 22 (excepté la position 17, aberrante), correspondant à la séquence transcrite. Résultats complets en Annexe.

Le produit d'amplification est ensuite sous cloné dans une construction commerciale CloneAmp® pAMP1 (Life technologies) et séquencé.

III C Résultats

Dans la séquence, toutes les cytosines hors CpG ont bien été converties en thymidine. Cela constitue le témoin interne. La variabilité observée ne serait donc être imputée à une action incomplète du bisulfite de sodium.

Les résultats complets sont présentés en annexe. Par commodité, nous en présentons une version simplifiée ici. En effet, plutôt que de présenter les évolutions des variations à chacune des 22 positions CpG au cours du temps, nous en dégageons ici une tendance, sur la base d'une première observation :

-Les positions CpG 1 à 12 (non transcrites) présentent une évolution temporelle homogène de leur taux de méthylation : nous en présentons ici la moyenne.

-De même les positions CpG 13 à 22, potentiellement transcrites, présentent un profil homogène entre elles : là encore nous en présentons la moyenne.

Ainsi, nous pouvons présenter, pour chaque cas de figure (cellules différenciées, cellules mortes, cellules non différenciées témoin), deux séries de résultats.

Cellules non différenciées

L'étude de la cinétique du témoin non différencié apporte une première surprise (Figure 28). La variabilité du taux de méthylation observée au cours de cette cinétique n'était pas attendue (Résultats complets : Annexe 1).

Cellules Différenciées

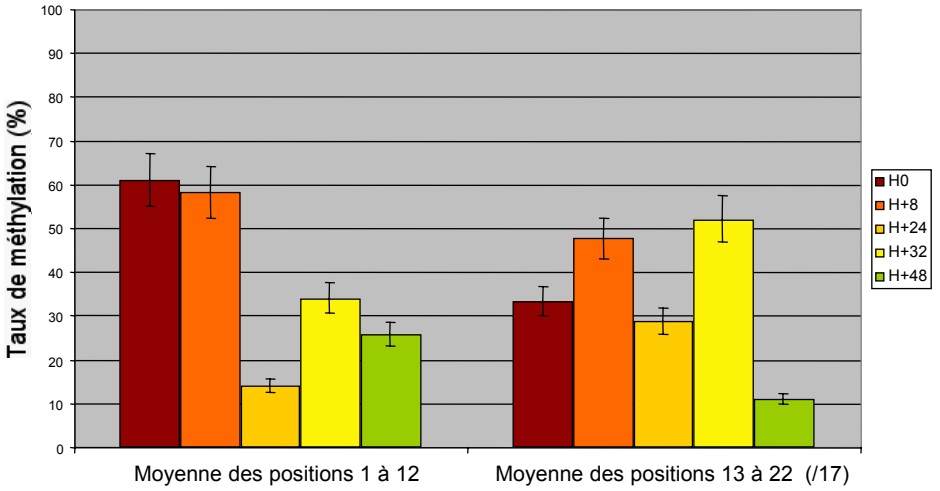


Figure 29 : Evolution du taux de méthylation au cours d'une cinétique de différenciation - Cellules différenciées.
Gauche : moyenne des positions de 1 à 12, correspondant à la séquence non transcrite. Droite : moyenne des positions 13 à 22 (excepté la position 17, aberrante), correspondant à la séquence transcrite. Résultats complets en Annexe.

Cellules Mortes

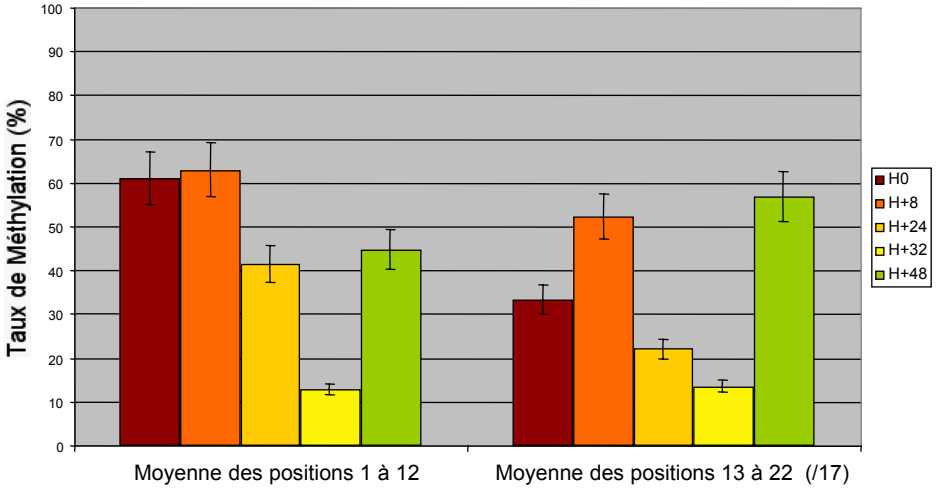


Figure 30 : Evolution du taux de méthylation au cours d'une cinétique de différenciation - Cellules mortes.
Gauche : moyenne des positions de 1 à 12, correspondant à la séquence non transcrite. Droite : moyenne des positions 13 à 22 (excepté la position 17, aberrante), correspondant à la séquence transcrite. Résultats complets en Annexe.

Les phases précoces de la cinétique sont très hétérogènes, mais la séquence de 24 à 48 heures est assez homogène une diminution transitoire de la méthylation, suivie d'une augmentation.

Ces variations sont peu significatives dans la première partie de la séquence, mais beaucoup plus dans la séquence transcrite. Dans cette dernière il y a transitoirement une méthylation quasi complète à 24 h (90%).

Cellules différenciées

Le profil de la cinétique de différenciation est le suivant (Résultats complets : Annexe 2) :

Diminution du taux de méthylation entre 8 et 24 heures.

Augmentation du taux de méthylation entre 24 et 32 heures.

Diminution du taux de méthylation entre 32 et 48 heures.

L'ensemble de la séquence possède ce profil, même s'il est plus marqué dans la première partie de la séquence (Figure 29).

Cellules mortes

Les cellules mortes ont un profil de méthylation qui diminue de 8 à 24 heures et qui augmente de 32 à 48 heures (Résultats complets : Annexe 3).

Il semble y avoir une grande homogénéité des profils de variation de méthylation dans la première partie, notamment aux positions 4 à 11, et une plus grande hétérogénéité dans les positions correspondant à la séquence transcrite (Figure 30).

Ces résultats doivent être interprétés avec prudence en raison de la logique qui présidé à leur obtention. Le gène Mest/Peg1 a été choisi non pas parce qu'il était susceptible de varier au cours de la différenciation cellulaire, mais parce qu'il était

connu pour avoir une méthylation différentielle au cours de la gamétogenèse. Il serait excessif de lui faire jouer un rôle sur la base de ces seuls résultats. Néanmoins, et justement parce qu'ils manifestent une variabilité inattendue, ceux-ci posent une question biologique ouverte.

Les constats que l'on peut faire sont les suivants.

-Une variabilité complexe de la méthylation est observée au cours d'une phase de différenciation cellulaire.

-Cette variabilité se surajoute à une variabilité de base, non négligeable, de la méthylation au cours d'une cinétique témoin.

-Les variations de méthylation sont très hétérogènes au cours des 24 premières heures et ne permettent pas de dégager une tendance précise.

-Entre 24 et 48 heures, les modifications de la méthylation suivent *une tendance inversée* chez les cellules en cours de différenciation, par rapport à la cinétique témoin.

-Entre 24 et 48 heures, les modifications de la méthylation *suivent une tendance comparable* chez les cellules mortes par rapport à la cinétique témoin.

-Les CpG de la séquence non transcrite présentent une profil de variation homogène, par rapport aux CpG de la séquence transcrite comparés entre eux.

Indépendamment de l'interprétation que l'on peut en faire, ces résultats mettent néanmoins indubitablement en évidence une variabilité intercellulaire de la méthylation. De plus, Ils montrent que cette variabilité évolue au cours de la différenciation cellulaire. On constate que les cellules mortes suivent une tendance comparable à celle des cellules différenciées. On peut proposer ainsi l'hypothèse qu'à chaque point de la cinétique, la population de cellules mortes s'enrichit en

cellules n'ayant pas suivi la tendance que l'on observe chez les cellules en cours de différenciation. L'amplitude notable des variations observées est compatible avec l'hypothèse que plusieurs options sont ouvertes à chaque instant à la cellule et que seules certaines lui permettent d'êtres sélectionnées pour la phase suivante. Cela l'est moins avec l'hypothèse que ces cellules reçoivent toutes, au cours d'une phase de différenciation, des signaux appropriés qui leur permettent d'adopter un comportement conforme auquel elles se tiennent, aux erreurs marginales près.
Les évolutions transitoires complexes du profil de méthylation au cours de la différenciation semblent en outre ici être explicables de manière plus pertinente par une fluctuation stochastique que par une régulation précise. Celle ci devrait par ailleurs expliquer qu'à chaque étape et pour chaque position, les deux situations, méthylée ou non méthylée, existent. La régulation n'aurait donc pas le même effet dans chaque cellule, au sein d'un une population et d'un milieu pourtant homogène.

Ces résultats doivent être replacés dans un contexte plus général où la méthylation des gènes n'est pas soupçonnée de subir de si larges variations sur des temps courts. Ils rendent pertinente l'idée d'appliquer plus généralement cette technique sur une batterie de gènes au cours des cinétiques de différenciation.

III D Discussion

La variabilité observée d'état de méthylation suggère que celle-ci est un phénomène dynamique au cours de la cinétique. En soit, les variations de méthylation des témoins non soumis la différenciation est une information étonnante.

Il semble que cette variabilité soit aléatoire. Cela pourrait être mis en relation avec un comportement exploratoire permettant de déréprimer au hasard l'expression de certains gènes pendant la différenciation cellulaire, avec pour conséquence qu'une fraction des cellules trouve par ce biais les réponses adéquates, sans prédétermination. Ces résultats sur des cellules ne sont cependant que préliminaires, et doivent être pris avec les précautions d'usage. Il est à noter que des vagues de méthylation *de novo* peuvent être constatées spontanément dans les cellules en culture, à la différence de ce qui est observé in vivo. Cela est associé à l'idée que dans les cultures cellulaires, des gènes non nécessaires sont rendus silencieux et donc que la méthylation peut être considérée un avantage sélectif pour la croissance. Ainsi, la prudence qui préside à la généralisation de résultats obtenus sur des cultures cellulaires n'empêche pas de remarquer que nos résultats sont complémentaires de ceux évoqués ici, et que la culture cellulaire permet peut-être de dévoiler des phénomènes physiologiques inhibés dans l'organisme.

Il est à noter que la technique de traitement au bisulfite est néanmoins assez lourde et nécessite notamment un important travail de séquençage. Il semble qu'au vu des résultats présentés et de ce qu'ils laissent percevoir elle soit néanmoins adaptée à nos problématiques.

Vu sous l'angle déterministe, les phénomènes épigénétiques permettent de seconder le génome dans la réalisation du programme, en lui apportant de la nuance et de l'adaptabilité. Néanmoins tout niveau supplémentaire de complexité nécessite, dans cette optique, une fine régulation, sur ses mécanismes propres mais aussi sur ses interactions avec les autres niveaux. De sorte que dans ce paradigme, on ne comprend plus très bien au final ce que le génome gagne à conserver ces mécanismes (qui permettraient d'expliquer que le génome n'explique pas à soi seul)

s'il doit par ailleurs mettre en place et conserver les contrôles adéquats (et donc expliquer en retour le contrôle des mécanismes épigénétiques). Dans l'approche endodarwinienne, l'ensemble de ces mécanismes a un rôle unifié : apporter une tolérance aux comportements exploratoires et permettre de fixer, même transitoirement les options adaptées.

DISCUSSION

Dans ce travail, nous avons fait le choix de mener de front plusieurs démarches. Reprendre une littérature préexistante pour souligner qu'elle n'était pas prisonnière du cadre conceptuel dominant. Améliorer des techniques existantes pour mieux les adapter à des problématiques émergentes. Importer des techniques dans des domaines où elles étaient peu pressenties. Pour chacune d'entre elles nous avons bien sûr le sentiment d'avoir apporté une contribution encore modeste. Mais le programme de recherche qui s'ouvre est vaste, et n'aurait pu être défriché dans le temps limité imparti à la préparation d'une thèse. Souhaitons ici que ce programme puisse être entrepris résolument.

Les résultats présentés dans ce travail vont dans le sens d'une expression stochastique des gènes et d'une variabilité transitoire au cours de la différenciation cellulaire. Ils ne prétendent bien sûr pas à eux seuls à constituer une démonstration de ces phénomènes. Ceux-ci sont déjà parfois partiellement avérés et nos résultats les renforcent. Ils fournissent aussi des méthodes qui pourraient permettre de faciliter ces études et de placer ces questions au cœur des préoccupations des biologistes moléculaires. Dans le cadre de cette thèse, il n'a malheureusement pas été possible d'aller plus loin, ce qui est frustrant. Les étapes suivantes d'un programme de mise en évidence de dynamiques endodarwiniennes seraient, quelles que soient les méthodes utilisées, une mesure de la variabilité au cours de cinétiques de différenciation (amorcée ici dans le travail sur la méthylation), et sa quantification de manière à obtenir un index précis de variabilité. Il faudrait ainsi la corréler avec des paramètres biologiques pour voir si cette variabilité est un bruit de fond ou elle-même

un paramètre. Grâce à nos hypothèses et nos mises au point, parmi d'autres, il est désormais possible de mettre en place de tels dispositifs expérimentaux.

A supposer que cela devienne réalité et que des dynamiques endodarwiniennes soient mises en évidence, quelles en seraient les conséquences pour la discipline ?
Dans le contexte sensible qu'est la contestation du déterminisme en biologie, il nous semble nécessaire de préciser tout d'abord clairement plusieurs éléments. La contestation du déterminisme biologique peut provenir d'horizons très différents. Certaines critiques, celles dont nous pouvons nous sentir partiellement proches, émanent de constats scientifiques et s'inscrivent elles-mêmes dans une démarche scientifique. Ces critiques reposent sur des observations documentées mettant en évidence que le déterminisme n'est pas toujours compatible avec une explication rationnelle et économe en hypothèses. Nous avons expliqué que de notre point de vue, elles avaient contribué à nuancer le déterminisme, alors que nous pensons que ses bases mêmes suggèrent que c'est une remise en cause plus générale qu'il faudrait adopter. Néanmoins, nous nous sentons pleinement tributaires de ces critiques, souvent pionnières, qui furent audacieuses en leur temps et ont contribué à faire de la biologie une discipline qui accepte de s'interroger sur elle-même et se remet en cause régulièrement.

D'autres critiques, qu'il ne s'agit pas de recenser ici, sortent du champ scientifique et procèdent par approximations et amalgames. Profitant du constat que le déterminisme n'est plus aussi triomphant qu'à l'aube des découvertes de François Jacob et Jacques Monod, car la progression des connaissances apporte inévitablement de la nuance, ou que le rôle du génome n'est plus vu comme aussi

central qu'initialement, elles ont déjà pratiqué des amalgames en avançant que la biologie ne serait donc décidément pas accessible à une approche rationnelle et expérimentale, ou que l'essentiel ne serait pas dans ce qui est accessible à l'expérience. Alors que le génome était apparu comme le support matériel de la sélection, son importance revue à la baisse ne lui permettrait plus ce rôle, délégitimerait cette sélection et entrouvrirait à nouveau la porte aux conceptions antérieures à la pensée darwinienne.

Il va sans dire que nous refusons expressément toute proximité avec les critiques de cette nature. Nous nous inscrivons clairement dans le champ de la rationalité et avons proposé plusieurs hypothèses qui sont accessibles à l'expérimentation et qui pourront éprouver la validité de nos hypothèses.

C'est aussi parce que nous souhaitons ne pas sortir de ce champ que nous ne portons pas de jugement sur le sens à donner à nos propositions au-delà du domaine strict de l'expérimentation en biologie. On a vu que la critique du déterminisme peut heurter car les phénomènes reproductibles semblent difficilement être expliqués par le hasard. Plus généralement, il peut paraître frustrant qu'un organisme soit construit sur du hasard par "opposition" à un programme. D'un point de vue sociologique, ces présupposés ont peut-être contribué à la prédominance de l'approche programmatique. Mais nous n'avons pas d'éléments qui permettent d'accréditer ou d'infirmer cette supposition et, en tout état de cause, ce ne sont pas ces questions qui ont justifié nos recherches. A la réserve près que si ces présupposés étaient avérés, le biais d'interprétation qu'ils occasionneraient serait un handicap intrinsèque à une évaluation contradictoire des hypothèses alternatives. Sorti de cette réserve, nous n'accordons pas plus de valeur en soi à chacune des

approches. Nous nous contentons de penser que l'une est plus générale, plus économe en hypothèses, et compatible avec l'ensemble des observations expérimentales qui lui préexistent. Ce faisant nous n'avons d'autre préoccupation que de contribuer à mieux expliquer les dynamiques observables. La grille de lecture que nous proposons et les outils que nous présentons pour en mesurer l'intérêt ne sauraient servir de prétexte à une extrapolation ascientifique.

Comme le montre le néologisme que nous proposons, l'endodarwinisme, nous nous plaçons d'ailleurs sans ambiguïté du côté de la synthèse néodarwinienne. Seule une interprétation abusive pourrait nous placer en opposition avec elles, parce que nous relativiserions la dimension "informationnelle" du génome. Nous avons dit à plusieurs reprises que le génome est une mémoire des réponses évolutivement archivées qui permettent à une cellule de réagir de manière probabiliste à une variation du milieu. En cela, le rôle de la sélection naturelle, qui est une des composantes de la synthèse néodarwinienne, est fondamental pour les organismes. Cela n'est pas contradictoire avec un mode d'utilisation du génome qui ne soit pas programmé. Bien plus, nos propositions visent à ouvrir aux approches évolutionnistes le fonctionnement interne des organismes. Cette proposition théorique est aussi un appel à plus de transdisciplinarité entre les évolutionnistes et les biologistes moléculaires. Les premiers testent depuis longtemps des modèles qui pourraient peut-être, après adaptation, apporter aux seconds des éléments de réponses beaucoup plus nombreux qu'ils ne le soupçonnent actuellement.

Il est vrai cependant que notre approche nous amène à relativiser l'importance des travaux dont la logique est de parfaire la connaissance des séquences génétiques

pour faire progresser les connaissances en biologie. A l'heure où l' on reconnaît que la distance entre deux gènes, les différentes modifications épigénétiques, l'épissage alternatif, la position du gène par rapport à l'origine de réplication, le degré de redondance, le degré d'accessibilité du site de transcription, la structure tridimensionnelle de la chromatine, la compétition pour les facteurs de transcription, la position de la cellule dans l'organisme, la température, le pH, la pression, les champs électriques, l'état d'expression d'autres gènes, entre autres, sont de nature à modifier l'expression d'un gène, il peut sembler que sa séquence, même si elle conditionne bien évidemment pour partie la structure tridimensionnelle de la protéine qui en est issue, n'est qu'une donnée parmi tant d'autres. A rebours de cette vision intégrée, le projet de séquençage du génome repose sur le primat d'une de ces caractéristiques sur toutes les autres. Cela ne serait pas alarmant si, au-delà des investissement humains et financiers nécessaires et pas toujours extensibles, il ne contribuait pas au formatage de la pensée en biologie. Comme le note justement Evelyn Fox Keller, "l'invocation des gènes s'est révélée manifestement efficace, non seulement pour assurer le financement et la promotion des programmes de recherches, mais également (et peut-être même tout spécialement) pour trouver des débouchés aux produits d'une industrie biopharmaceutique en rapide expansion" (Fox Keller, 2000). Les vingt dernières années, la biologie moléculaire, on l'a vu, a été traversée de débats riches sur les rôles respectifs des différentes molécules et des différents niveaux d'observation sur le fonctionnement du vivant. Et pourtant, ces vingt ans de débats n'auront pas empêché l'annonce réductrice du séquençage humain comme la découverte du "grand livre de la vie", quand il n'était pas ajouté "seul connu de Dieu auparavant" (Francis Collins, directeur du National Human Genome Research Institute, cité *in* Weigmann, 2004). Certes, il s'agit d'un propos

d'industriel non désintéressé, mais qui peut nier sérieusement qu'au-delà de la caricature, cette vision réductrice continue à irriguer la biologie contemporaine ? Comment expliquer, sinon, que l'on cherche désormais à décoder le transcriptome, le protéome, et à transformer ainsi la biologie en une pile de données ?

Dans ce contexte flottant, où les certitudes d'hier sont sans cesses remises en cause, le danger se profile d'une science génétique qui serait prisonnière d'un discours et d'un vocabulaire de plus en plus inadéquat et de moins en moins rigoureux. La notion de gène par exemple, pourtant socle commun de tous les généticiens, n'est même plus consensuelle. Les définitions de ce mot sont multiples, ce qui n'est pas forcément problématique, mais le devient quand ces définitions sont contradictoires. Information, molécule, voire combinaison des deux, les acceptions sont nombreuses. Il est possible de s'en réjouir, de noter que le flou du concept est le signe de sa richesse. Il est aussi permis de penser, et c'est notre point de vue, que cette indécision est un péril. C'est semble-t-il un avis partagé : "Comme toujours, les forces antagonistes qui travaillent à déstabiliser un ensemble de termes et de concepts sont issues (...) de la rencontre de la science avec le monde réel ; elles proviennent de l'accumulation des insuffisances du lexique existant devant les nouvelles découvertes expérimentales. Et c'est cela, je l'affirme, qui s'est passé dans le cas du gène" (Fox Keller, 2000).

Une manière de sortir de cette indécision, c'est de considérer que le rôle des gènes ne peut s'apprécier à l'échelle individuelle. Continuer de supposer, même fonctionnellement, que les gènes soient en première approximation les déterminants des caractères, ce qui en était la définition minimale morganienne, nous impose de renoncer à l'équivalence trompeuse entre discrétisation des gènes et discrétisation

de leur action. En d'autres termes, dissocier l'échelle de la définition et celle de la fonction. En terme de programme de recherche, cela se traduit concrètement par la nécessité de s'éloigner résolument des approches parcellaires, et ne plus les justifier par un souci de progression graduelle vers la complexité. Ces approches ont été fécondes pour comprendre en première approximation le fonctionnement des gènes, mais ne le sont plus pour la complexité des problèmes que la communauté des biologistes moléculaires a l'ambition de résoudre.

Le fonctionnement des gènes ne peut s'apprécier qu'en réseau. En soit, cette affirmation est tout sauf une nouveauté. C'est même, nous semble-il, la colonne vertébrale d'une nouvelle pensée génétique qui trouverait dans ce discours une manière délocaliser le déterminisme. Citons une dernière fois Evelyn Fox Keller qui au terme d'une critique radicale de la pensée génétique, semble se résoudre à cet entre-deux : "j'ai affirmé que l'hypothèse d'un programme inscrit dans l'ADN doit, elle aussi, être réévaluée et j'ai proposé pour la remplacer, le concept plus dynamique de programme réparti, dans lequel les différents composants (..) fonctionnent alternativement comme instructions et comme données" (Fox Keller, 2000). Il ne suffit cependant pas de dire qu'un ensemble fonctionne en réseau, ou que ce réseau est dynamique, ou que le programme est "distribué" pour résoudre les problèmes auxquels nous sommes confrontés. Ces critiques font échos à celles, plus anciennes, et que l'on a abordées, de la contestation du déterminisme. Or celles-ci, bénéfiques dans la nuance qu'elles ont apportées à ce concept, n'ont pas pour autant réussi à apporter une grille d'analyse de substitution. Les réactualiser conduit peut-être à tourner en rond. Si ce discours n'est pas étayé par une prospective sur les dynamiques à l'œuvre dans le réseau, il n'est qu'une manière de brouiller les pistes et de se condamner à l'inintelligibilité des phénomènes observés.

Néanmoins, accepter ce fonctionnement en réseau conduit à plusieurs conséquences, au premier rang desquelles se trouve la possibilité se passer de la notion de spécificité. A supposer même qu'elle ait un fondement, ce que nous ne pensons pas, elle n'est pas fonctionnellement intéressante dans ce cadre. Peu importe que la spécificité existe si par ailleurs, des gènes peuvent en remplacer d'autres quand ces derniers font défaut. Peu importe la précision d'une régulation si d'autres peuvent la suppléer.

Et quant à comprendre les dynamiques qui sont à l'œuvre dans ces réseaux, la question reste ouverte. Il nous a semblé, au cours de notre thèse, que les dynamiques endodarwiniennes sont une piste prometteuse. Elles sont économes en hypothèses, englobent les observations antérieures, et sont compatibles avec ce que l'on connaît des interactions moléculaires. Elles ont enfin le mérite de mettre de côté toute notion d'information et de programme, importées avec une certaine artificialité en biologie, et dont on a vu qu'elles sont de plus en plus délicates à manier, et parfois dénuées de sens.

Nous ne prétendrons pas ici, même si nous sommes convaincus qu'elles sont à l'œuvre dans de nombreux processus, que les dynamiques endodarwiniennes sont la réponse universelle au fonctionnement du vivant. L'expérimentation montrera la portée et les limites de cette hypothèse. Nous espérons seulement avoir convaincu qu'il est utile, avant de se lancer dans l'étude d'un phénomène reproductible, de vérifier dans quelle mesure le hasard ne peut pas suffire à expliquer cette reproductibilité.

Convenons cependant que cette hypothèse et ses éventuelles déclinaisons ont des visées plus larges. En proposant que la stochasticité puisse être à la base d'un mécanisme structurant, selon le principe darwinien du hasard et de la sélection, nous suggérons que cette approche économe en hypothèses, et intrinsèquement dynamique, puisse être fructueuse dans de nombreux cas de figures où l'approche déterministe est supposée sans être démontrée.

On voit ici que notre proposition dépasse le cadre de la biologie cellulaire. Tout événement élémentaire aléatoire susceptible d'être stabilisé dans le cadre d'une compétition pour une ressource peut ressortir d'une approche endodarwinienne. Nous avons mis en évidence que la théorie de l'évolution peut potentiellement s'appliquer à des situations très variées pour peu qu'elles remplissent certains critères : une variabilité spontanée des phénomènes élémentaires, une "ressource" (pris dans un sens très extensif) en quantité limitante, une adaptation sélective. Il est ainsi scientifiquement pertinent de chercher à savoir si cette théorie ne pourrait pas s'étendre à des niveaux inférieurs à celui de l'individu. Cela ouvrirait un nouveau champ d'investigation à la communauté des biologistes évolutionnistes. Cela inciterait les biologistes moléculaires à acquérir de nouveaux réflexes : déterminer une probabilité d'expression par cellule d'un gène dans des conditions données et le rapporter à la taille de la population, identifier la "ressource" en facteur limitant (plutôt que de chercher le "signal"), caractériser la capacité d'une cellule à s'adapter (plutôt qu'à déployer un "programme"). Ces questions forment pour chaque problème un programme de travail conséquent.

Si cela s'avère fructueux, il se pourrait que l'on ait, à un niveau modeste, contribué à un rendre pertinente l'hypothèse d'une importation, mais plus encore d'une généralisation des dynamiques darwiniennes dans des situations où elles n'étaient préalablement pas soupçonnées. On pourrait aussi avoir contribué à convaincre que l'on peut penser la biologie sans programme, sans spécificité, et sans information. Ou du moins fourni quelques éléments rendant ces propositions moins aberrantes qu'il n'y paraît en première analyse. C'est la notre seule ambition.

BIBLIOGRAPHIE

A

Abkowitz, J. L., Catlin, S. N. and Guttorp, P. (1996). Evidence that hematopoiesis may be a stochastic process in vivo. Nat. Med. *2,* 190-7.

Ameisen, J.C. (1996). The origin of programmed cell death. Science. 272,1278-9.

Ameisen, J. C. (1998). The origin of programmed cell death in the flow of evolution and its role in host-pathogen interactions. C R Seances Soc Biol Fil. 192, 1095-8.

Ameisen, J. C. (1999).La Sculpture du vivant. Paris, Le Seuil.

Ameisen, J. C. (2002). On the origin, evolution, and nature of programmed cell death: a timeline of four billion years. Cell Death Differ. *9*, 367-93.

Ameisen, J. C. (2004). Looking for death at the core of life in the light of evolution. Cell Death Differ. 11, 4-10.

Antonsson, B. and Martinou, J. C. (2000). The Bcl-2 protein family. Exp Cell Res *256*, 50-7.

Aranyi, T., Kerjean, A., Toth, S., Mallet, J., Meloni, R. and Paldi, A. (2002). Paradoxical Methylation of the tyrosine hydroxylase Gene in Mouse Preimplantation Embryos. Genomics *80*, 558-63.

Arkin, A., Ross, J. and McAdams, H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. Genetics *149*, 1633-48.

Armitage, P. and Doll, R. (1954). The age distribution of cancer and a multi-stage theory of carcinogenesis. Br J Cancer *8*, 1-12.

Arnoult, D., Tatischeff, I., Estaquier, J., Girard, M., Sureau, F., Tissier, J. P., Grodet, A., Dellinger, M., Traincard, F., Kahn, A., Ameisen, J. C. , Petit, P. X. (2001). On the evolutionary conservation of the cell death pathway: mitochondrial release of an apoptosis-inducing factor during Dictyostelium discoideum cell death. Mol Biol Cell. 12, 3016-30.

Astumian, R. D., Adair, R. K. and Weaver, J. C. (1997). Stochastic resonance at the single-cell level. Nature *388*, 632-3.

Atlan, H. (1972). L'organisation biologique et la théorie de l'information. Paris, Herman.

Atlan, H. and Koppel, M. (1990). The cellular computer DNA: program or data. Bull Math Biol *52*, 335-48.

B

Baroffio, A. and Blot, M. (1992). Statistical evidence for a random commitment of pluripotent cephalic neural crest cells. J Cell Sci *103*, 581-7.

Bennett, D. C. (1983). Differentiation in mouse melanoma cells: initial reversibility and an on-off stochastic model. Cell *34*, 445-53.

Benzer, S. (1953). Induced synthesis of enzymes in bacteria analyzed at the cellular level. Biochim Biophys Acta *11*, 383-95.

Bernard, C. (1865). Introduction à l'étude de la médecine expérimentale. Paris, J. B. Baillière et Fils 1865 ; Paris, Flammarion 1966.

Billia, F., Barbara, M., McEwen, J., Trevisan, M. and Iscove, N. N. (2001)., Resolution of pluripotential intermediates in murine hematopoietic differentiation by global complementary DNA amplification from single cells: confirmation of assignments by expression profiling of cytokine receptor transcripts. Blood *97*, 2257-68.

Birky, C. W. and Skavaril, R. (1984). Random partitioning of cytoplasmic organelles at cell division: the effect of organelle and cell volume. J Theor Biol *106*, 441-7.

Blackstone, N. W. (2003). Redox signaling in the growth and development of colonial hydroids. J Exp Biol *206*, 651-8.

Bochet, P., Audinat, E., Lambolez, B., Crepel, F., Rossier, J., Iino, M., Tsuzuki, K. and Ozawa, S. (1994). Subunit composition at the single-cell level explains functional properties of a glutamate-gated channel. Neuron *12,* 383-8.

Bohme, K., Winterhalter, K. H. and Bruckner, P. (1995). Terminal differentiation of chondrocytes in culture is a spontaneous process and is arrested by transforming growth factor-beta 2 and basic fibroblast growth factor in synergy, Exp Cell Res *216*, 191-8.

Brady, G., Barbara, M. and Iscove, N. (1990). Representative in vitro cDNA amplification from individual hemopoietic cells and colonies. Methods Mol Cell Biol *2*, 17-25.

Brady, G., Billia, F., Knox, J., Hoang, T., Kirsch, I. R., Voura, E. B., Hawley, R. G., Cumming, R., Buchwald, M. and Siminovitch, K. (1995). Analysis of gene expression in a complex differentiation hierarchy by global amplification of cDNA from single cells. Curr Biol *5*, 909-22.

Britten, R. J. and Davidson, E. H. (1969). Gene regulation for higher cells: a theory. Science *165*, 349-57.

Brooks, R. F. and Shields, R. (1985). Cell growth, cell division and cell size homeostasis in Swiss 3T3 cells. Exp Cell Res *156*, 1-6.

Bucy, R. P., Panoskaltsis-Mortari, A., Huang, G. Q., Li, J., Karr, L., Ross, M., Russell, J. H., Murphy, K. M. and Weaver, C. T. (1994). Heterogeneity of single cell cytokine gene expression in clonal T cell populations. J Exp Med *180,* 1251-62.

Bucy, R. P., Karr, L., Huang, G. Q., Li, J., Carter, D., Honjo, K., Lemons, J. A., Murphy, K. M. and Weaver, C. T. (1995). Single cell analysis of cytokine gene coexpression during CD4+ T-cell phenotype development. Proc Natl Acad Sci U S A *92*, 7565-9.

Buffelli, M., Burgess, R. W., Feng, G., Lobe, C. G., Lichtman, J. W. and Sanes, J. R. (2003). Genetic evidence that relative synaptic efficacy biases the outcome of synaptic competition. Nature *424*, 430-4.

Burnet, F. M. (1957). A modification of Jerne's theory of antibody production using the concept of clonal selection. Austr J Sci *20*, 67-69.

Busche, S., Gallinat, S., Bohle, R. M., Reinecke, A., Seebeck, J., Franke, F., Fink, L., Zhu, M., Sumners, C. and Unger, T. (2000). Expression of angiotensin AT(1) and AT(2) receptors in adult rat cardiomyocytes after myocardial infarction. A single-cell reverse transcriptase-polymerase chain reaction study. Am J Pathol *157* 605-11.

C

Changeux, J. P., Courrege, P. and Danchin A. (1973). A theory of the epigenesis of neuronal networks by selective stabilization of synapses. Proc Natl Acad Sci U S A *70*, 2974-8.

Changeux, J. P. and Danchin, A. (1976). Selective stabilisation of developing synapses as a mechanism for the specification of neuronal networks. Nature *264*, 705-12.

Chelly, J., Concordet, J. P., Kaplan, J. C. and Kahn, A. (1989). Illegitimate transcription: transcription of any gene in any cell type. Proc Natl Acad Sci U S A *86*, 2617-21.

Chess, A., Simon, I., Cedar, H. and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. Cell *78*, 823-34.

Chung, J. D., Conner, S. and Stephanopoulos, G. (1995). Flow cytometric study of differentiating cultures of Bacillus subtilis. Cytometry *20*, 324-33.

Chung, J. D., Stephanopoulos, G., Ireton, K. and Grossman, A. D. (1994). Gene expression in single cells of Bacillus subtilis: evidence that a threshold mechanism controls the initiation of sporulation. J Bacteriol *176*, 1977-84.

Clejan, S., Dotson, R. S., Wolf, E. W., Corb, M. P. and Ide, C. F. (1996). Morphological differentiation of N1E-115 neuroblastoma cells by dimethyl sulfoxide activation of lipid second messengers. Exp Cell Res *224*, 16-27.

D

Darwin, C. (1859). On the Origin of Species. London, John Murray. *Trad. Fr*. (1999) l'Origine des espèces, Paris, Flammarion.

Davis, C. B., Killeen, N., Crooks, M. E., Raulet, D. and Littman, D. R. (1993). Evidence for a stochastic mechanism in the differentiation of mature subsets of T lymphocytes. Cell *73*, 237-47.

Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C. H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C. T., Livi, C. B., Lee, P. Y., Revilla, R., Rust, A. G., Pan, Z., Schilstra, M. J., Clarke, P. J., Arnone, M. I., Rowen, L., Cameron, R. A., McClay, D. R., Hood, L. and Bolouri, H. (2002). A genomic regulatory network for development. Science *295*, 1669-78.

Dawkins, R. (1976). The Selfish Gene. Oxford, Oxford University Press. *Trad. fr.* Le gène égoïste. Paris, Menges 1978 ; Paris, Odile Jacob, 1996.

Dieterlen-Lièvre, F., Duprat, A. M. and Le Douarin, N. (1996). La biologie du développement : sources et perspectives. Médecine/sciences *12*, 67-75.

Dixon, A. K., Richardson, P. J., Lee, K., Carter, N. P. and Freeman, T. C. (1998). Expression profiling of single cells using 3 prime end amplification (TPEA) PCR. Nucleic Acids Res *26* , 4426-31.

Dixon, A. K., Richardson, P. J., Pinnock, R. D. and Lee, K. (2000). Gene-expression analysis at the single-cell level. Trends Pharmacol Sci *21,* 65-70.

Duboule, D. and Wilkins, A. S. (1998). The evolution of 'bricolage'. Trends Genet *14*, 54-9.

E

Eberwine, J. (2001). Single-cell molecular biology. Nat Neurosci *4*, (Suppl):1155-6.

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M. and Coleman, P. (1992). Analysis of gene expression in single live neurons. Proc Natl Acad Sci U S A *89*, 3010-4.

Eberwine, J., Miyashiro, K., Kacharmina, J. E. and Job, C. (2001a). Local translation of classes of mRNAs that are targeted to neuronal dendrites. Proc Natl Acad Sci U S A *98*, 7080-5.

Eberwine, J., Kacharmina, J. E., Andrews, C., Miyashiro, K., McIntosh, T., Becker, K., Barrett, T., Hinkle, D., Dent, G. and Marciano, P. (2001b). mRNA expression analysis of tissue sections and single cells. J Neurosci *21*, 8310-4.

Echeverri, K. and Tanaka, EM. (2002). Mechanisms of muscle dedifferentiation during regeneration. Semin Cell Dev Biol *13*, 353-60.

Edelman, G. M. and Mountcastle, V. B. (1978). The Mindful Brain: Cortical Organization and the Group Selective Theory of Higher Brain Function. Cambridge, MIT Press.

Edelman, GM. (1987). Neural Darwinism: The Theory of Neuronal Group Selection. New York, Basic Books.

Edelman, G. M. and Gally, JA. (2001). Degeneracy and complexity in biological systems. Proc Natl Acad Sci U S A *98*, 13763-8.

Elowitz, M. B., Levine, A. J., Siggia, E. D. and Swain, PS. (2002). Stochastic gene expression in a single cell. Science *297*, 1183-6.

F

Fiering, S., Northrop, J. P., Nolan, G. P., Mattila, P. S., Crabtree, G. R. and Herzenberg, L. A. (1990). Single cell assay of a transcription factor reveals a threshold in transcription activated by signals emanating from the T-cell antigen receptor. Genes Dev *4*, 1823-34.

Fiering, S., Whitelaw, E. and Martin, D. I. (2000). To be or not to be active: the stochastic nature of enhancer action. Bioessays *22*, 381-7.

Forbes, L. S. (1997). The evolutionary biology of spontaneous abortion in humans. Trends Ecol Evol *12*, 446-450.

Fox Keller, E. (2000). The century of the gene. Harvard, Harvard University Press *Trad. Fr.* (2003). Le siècle du gène. Paris, Gallimard.

Freeman, T. C., Lee, K. and Richardson, P. J. (1999). Analysis of gene expression in single cells. Curr Opin Biotechnol *10*, 579-82.

G

Garry, D. J., Ordway, G. A., Lorenz, J. N., Radford, N. B., Chin, E. R., Grange, R. W., Bassel-Duby, R. and Williams, R. S. (1998). Mice without myoglobin. Nature *395*, 905-8.

Gayon, J. (2003). Prédire ou expliquer. Sciences et Avenir Hors Série. *136.*

van Gelder, R. N., von, Zastrow, M. E., Yool, A., Dement, W. C., Barchas, J. D. and Eberwine, J. H. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci U S A *87*, 1663-7.

Gerhart, J. and Kirschner, M. (1997). Cells, embryos, and evolution. Malden, Blackwell Science.

Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R. and Ellenberg, J. (2003). Global chromosome positions are transmitted through mitosis in mammalian cells. Cell *112*, 751-64.

Gilbert, S. F. (1991). Developmental Biology. 3rd ed. Sunderland, Sinauer Associates.

Greenspan, R. J. (2001). The flexible genome.Nat Rev Genet *2,* 383-7.

Greenwald, I. and Rubin, G. M. (1992). Making a difference: the role of cell-cell interactions in establishing separate identities for equivalent cells. Cell *68*, 271-81.

Grossman, A. D. (1995). Genetic networks controlling the initiation of sporulation and the development of genetic competence in Bacillus subtilis. Annu Rev Genet *29*, 477-508.

Guptasarma, P. (1995). Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia coli ? Bioessays *17*, 987-97.

H

Heams, T. and Kupiec, J. J. (2003). Modified 3'-end amplification PCR for gene expression analysis in single cells. Biotechniques *34*, 712-4.

Held, W., Kunz, B., Ioannidis, V. and Lowin-Kropf, B. (1999). Mono-allelic Ly49 NK cell receptor expression. Semin Immunol *11*, 349-55.

Hengartner, M. O. (2000). The biochemistry of apoptosis. Nature *407*, 770-6.

Hennis, A. S. and Birky, C. W., Jr. (1984). Stochastic partitioning of chloroplasts at cell division in the alga Olisthodiscus, and compensating control of chloroplast replication. J Cell Sci *70*, 1-15.

Höllander, G. A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D. E., Bhan, AK. and Burakoff, S. J. (1998). Monoallelic expression of the interleukin-2 locus. Science *359*, 2118-21.

Höllander, G. A. (1999). On the stochastic regulation of interleukin-2 transcription. Semin Immunol *11*, 357-67.

Honarpour, N., Du, C., Richardson, J. A., Hammer, R. E., Wang, X. and Herz, J. (2000). Adult Apaf-1-deficient mice exhibit male infertility. Dev Biol *218*, 248-58.

Houchmandzadeh, B., Wieschaus, E. and Leibler, S. (2002). Establishment of developmental precision and proportions in the early Drosophila embryo. Nature *415*, 798-802.

Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C. and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. Genes Dev *11*, 774-85.

Huang, F. Z., Bely, A. E. and Weisblat, D. A. (2001). Stochastic WNT signaling between nonequivalent cells regulates adhesion but not fate in the two-cell leech embryo. Curr Biol *11,* 1-7.

Hume, D. A. (2000). Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. Blood *96* , 2323-8.

I

Iscove, N. N., Barbara, M., Gu, M., Gibson, M., Modi, C., Winegarden, N. (2002). Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. Nat Biotechnol *20*, 940-3.

J

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol *3*, 318-56.

Jacob, J., Haug, J. S., Raptis, S., Link, D. C. (1998). Specific signals generated by the cytoplasmic domain of the granulocyte colony-stimulating factor (G-CSF) receptor are not required for G-CSF-dependent granulocytic differentiation. Blood *92,* 353-61.

Jerne, N. (1955). The natural selection theory of antibody formation. Proc Natl Acad Sci U S A *41*, 849-857

Jouvenot, Y., Poirier, F., Jami, J. and Paldi, A. (1999). Biallelic transcription of Igf2 and H19 in individual cells suggests a post-transcriptional contribution to genomic imprinting. Curr Biol *20*, 1199-202.

K

Kacharmina, J. E., Crino, P. B. and Eberwine, J. (1999). Preparation of cDNA from single cells and subcellular regions. Methods Enzymol *303*, 3-18.

Kaminski, M., Karbowski, M., Miyazaki, Y., Kedzior, J., Spodnik, J. H., Gil, A., Wozniak, M. and Wakabayashi, T. (2002). Co-existence of apoptotic and necrotic features within one single cell as a result of menadione treatment. Folia Morphol *61*, 217-20.

Kasthuri, N. and Lichtman, J. W. (2003). The role of neuronal identity in synaptic competition. Nature *424,* 426-30.

Kerjean, A., Couvert, P., Heams, T., Chalas, C., Poirier, K., Chelly, J., Jouannet, P., Paldi, A. and Poirot, C. (2003). In vitro follicular growth affects oocyte imprinting establishment in mice. Eur J Hum Genet *11*, 493-6.

Kerjean, A., Vieillefond, A., Thiounn, N., Sibony, M., Jeanpierre, M., Jouannet, P. (2001). Bisulfite genomic sequencing of microdissected cells. Nucleic Acids Res *29*, E106-6.

Kimoto, Y. (1998). A single human cell expresses all messenger ribonucleic acids: the arrow of time in a cell. Mol Gen Genet *258*, 233-9.

Kirschner, M., Gerhart, J. and Mitchison, T. (2000). Molecular "vitalism". Cell *100,* 79-88.

Klarsfeld, A. and Revah, F. (1999). Biologie de la mort. Paris, Odile Jacob.

Klein, C. A., Seidl, S., Petat-Dutter, K., Offner, S., Geigl, J. B., Schmidt-Kittler, O., Wendler, N., Passlick, B., Huber, R. M., Schlimok, G., Baeuerle, P. A. and Riethmuller, G. (2002). Combined transcriptome and genome analysis of single micrometastatic cells. Nat Biotechnol *20*, 387-92

Klingenberg, C. P., Nijhout, H.F. (1998). Competition among growing organs and developmental control of morphological asymmetry. Proc R Soc Lond B *265*, 1135-1139.

Klionsky, D. J. and Emr, S. D. (2000). Autophagy as a regulated pathway of cellular degradation. Science *290*, 1717-21.

Ko, M.S. (1991). A stochastic model for gene induction. J Theor Biol *153*, 181-94.

Ko, M. S., Nakauchi, H. and Takahashi, N. (1990). The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. EMBO J *9*, 2835-42.

de Krom, M., van de Corput, M., von, Lindern, M., Grosveld, F. and Strouboulis, J. (2002). Stochastic patterns in globin gene expression are established prior to transcriptional activation and are clonally inherited. Mol Cell *9,* 1319-26.

Kupiec, J. J. (1981). Théorie probabiliste de la différenciation cellulaire, XIIème rencontre de Méribel, 161-163.

Kupiec, J. J. (1983). A probabilist theory for cell differentiation, embryonic mortality and DNA C-value paradox. Specul Sci Technol *6*, 471-478.

Kupiec, J. J. (1986). A probabilist theory for cell differentiation : the extension of Darwinian principles to embryogenesis. Spec Sci Techn *9*, 19-22.

Kupiec, J. J. (1996). A chance-selection model for cell differentiation. Cell Death Diff *3*, 385-390.

Kupiec, J. J. (1997). A Darwinian theory for the origin of cellular differentiation. Mol Gen Genet *255*, 201-8.

Kupiec, J. J. and Sonigo P. (2000). Ni Dieu ni gène. Paris, Le Seuil.

Kuznetsov, V. A., Knott, G. D. and Bonner, R. F. (2002). General statistics of stochastic process of gene expression in eukaryotic cells. Genetics *161*, 1321-32.

L

Lamarck, J.B. (1809). Philosophie Zoologique. Paris, Dentu, 1809 ; Paris, Garnier Flammarion, 1994.

Lambolez, B., Audinat, E., Bochet, P., Crepel, F., Rossier, J. (1992)., AMPA receptor subunits expressed by single Purkinje cells. Neuron *9*, 247-58.

Lecourt, D. (1999). Déterminisme, *in* **Lecourt D.** (Dir.), Dictionnaire d'histoire et philosophie des sciences. Paris Presses Universitaires de France.

Lederberg, J. (1959). Genes and antibodies. Science *129*, 1649-53.

Le Douarin, N. (2000). Des chimères, des clones et des gènes. Paris, Odile Jacob.

Levsky, J. M., Shenoy, S. M., Pezo, R. C. and Singer, R. H. (2002). Single-cell gene expression profiling. Science *297*, 836-40.

Levsky, J. M. and Singer, R. H. (2003). Gene expression and the myth of the average cell. Trends Cell Biol *13*, 4-6.

Lewin, R. (1984). Why is development so illogical? Science *224*, 1327-9.

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565-70.

Li, K., Li, Y., Shelton, J. M., Richardson, J. A., Spencer, E., Chen, Z. J., Wang, X. and Williams, R. S. (2000). Cytochrome c deficiency causes embryonic lethality and attenuates stress-induced apoptosis. Cell *101*, 389-99.

Lian, R. H., Maeda, M., Lohwasser, S., Delcommenne, M., Nakano, T., Vance, R. E., Raulet, D. H. and Takei, F. (2002). Orderly and nonstochastic acquisition of CD94/NKG2 receptors by developing NK cells derived from embryonic stem cells in vitro. J Immunol *168*, 4980-7.

Lin, C. T., Palmer, W., Wu, J. Y. and Chan, L. (1986). Estrogen induction of very low density apolipoprotein II synthesis, a major avian liver yolk protein, involves the recruitment of hepatocytes. Endocrinology *118*, 538-44.

Lin, Z., Lu, M. H., Schultheiss, T., Choi, J., Holtzer, S., DiLullo, C., Fischman, D. A. and Holtzer, H. (1994). Sequential appearance of muscle-specific proteins in myoblasts as a function of time after cell division: evidence for a conserved myoblast differentiation program in skeletal muscle. Cell Motil Cytoskeleton *29*, 1-19.

M

Maloney, P. C. and Rotman, B. (1973). Distribution of suboptimally induces -D-galactosidase in Escherichia coli. The enzyme content of individual cells. J Mol Biol *73*, 77-91.

Marsden, V. S., O'Connor, L., O'Reilly, L. A., Silke, J., Metcalf, D., Ekert, P. G., Huang, D. C., Cecconi, F., Kuida, K., Tomaselli, KJ., Roy, S., Nicholson, D. W., Vaux, D. L., Bouillet, P., Adams, J. M. and Strasser, A. (2002). Apoptosis initiated by Bcl-2-regulated caspase activation independently of the cytochrome c/Apaf-1/caspase-9 apoptosome. Nature *419*, 634-7.

McAdams, H. H. and Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale.Trends Genet *15*, 65-9.

McAdams, H. H. and Arkin, A. (1997). Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A *94*, 814-9.

Melton, D. W. (1994). Gene targeting in the mouse. Bioessays *16*, 633-8.

Metzstein, M. M., Stanfield, G. M. and Horvitz, H. R. (1998). Genetics of programmed cell death in C. elegans: past, present and future. Trends Genet *14,* 410-6.

Meunier, L. (2001). Clonal variation of gene expression as a source of phenotypic diversity in parasitic protozoa. Trends Parasitol *17*, 475-9.

Michaelson, J. (1993). Cellular selection in the genesis of multicellular organization, Lab Invest *69*, 136-51.

Michor, F., Nowak, M. A., Frank, S. A. and Iwasa, Y. (2003). Stochastic elimination of cancer cells. Proc R Soc Lond B Biol Sci *270*, 2017-24.

Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. Science *291*, 843-7.

Miyashiro, K., Dichter, M. and Eberwine, J. (1994). On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning. Proc Natl Acad Sci U S A *91*, 10800-4.

Monyer, H. and Lambolez, B. (1995). Molecular biology and physiology at the single-cell level. Curr Opin Neurobiol *5*, 382-7.

Morange, M. (1999). Régulation moléculaire*, in* **Lecourt, D.**, (Dir.), Dictionnaire d'histoire et philosophie des sciences. Paris, Presses Universitaires de France.

Morata, G. and Ripoll, P. (1975). Minutes: mutants of drosophila autonomously affecting cell division rate. Dev Biol *42*, 211-21.

Moreno, E., Basler, K. and Morata, G. (2002). Cells compete for decapentaplegic survival factor to prevent apoptosis in Drosophila wing development. Nature *416*, 755-9.

N

Nan, X., Campoy, F. J. and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell *88*, 471-81.

Nemazee, D. (2000). Receptor selection in B and T lymphocytes. Annu Rev Immunol *18*, 19-51.

Newlands, S., Levitt, LK., Robinson, C. S., Karpf, A. B., Hodgson, V. R., Wade, R. P. and Hardeman, EC. (1998). Transcription occurs in pulses in muscle fibers. Genes Dev *12*, 2748-58.

Nieuwkoop, PD. (1973). The organization center of the amphibian embryo: its origin, spatial organization, and morphogenetic action. Adv Morphog *10*, 1-39.

Nijhout, H. F. (1990). Metaphors and the role of genes in development. Bioessays *12*, 441-6.

Nijhout, H. F. and Emlen, D. J. (1998). Competition among body parts in the development and evolution of insect morphology. Proc Natl Acad Sci U S A *95*, 3685-9.

O

Odelberg, S. J., Kollhoff, A. and Keating, M. T. (2000). Dedifferentiation of mammalian myotubes induced by msx1.Cell *103*, 1099-1109.

Ohsawa, Y., Isahara, K., Kanamori, S., Shibata, M., Kametaka, S., Gotow, T., Watanabe, T., Kominami, E. and Uchiyama, Y. (1998). An ultrastructural and immunohistochemical study of PC12 cells during apoptosis induced by serum deprivation with special reference to autophagy and lysosomal cathepsins. Arch Histol Cytol *61*, 395-403.

Okada, H., Suh, WK., Jin, J., Woo, M., Du, C., Elia, A., Duncan, G. S., Wakeham, A., Itie, A., Lowe, S. W., Wang, X. and Mak, T. W. (2002). Generation and characterization of Smac/DIABLO-deficient mice.Mol Cell Biol *22*, 3509-17.

Orphanides, G. and Reinberg, D. (2002). A unified theory of gene expression. Cell *108*, 439-51.

P

Parada, L. A., Roix, J. J. and Misteli, T. (2003). An uncertainty principle in chromosome positioning. Trends Cell Biol *13*, 393-6.

Paulus, U., Loeffler, M., Zeidler, J., Owen, G. and Potten, C. S. (1993). The differentiation and lineage development of goblet cells in the murine small intestinal crypt: experimental and modelling studies. J Cell Sci *106*, 473-83.

Pellerin, I., Schnabel, C., Catron, K. M. and Abate, C. (1994). Hox proteins have different affinities for a consensus DNA site that correlate with the positions of their genes on the hox cluster. Mol Cell Biol *14*, 4532-45.

Porter, J. T., Cauli, B., Tsuzuki, K., Lambolez, B., Rossier, J., Audinat, E. (1999). Selective excitation of subtypes of neocortical interneurons by nicotinic receptors. J Neurosci *19*, 5228-35.

Powell, E. O. (1956). Growth rate and generation time of bacteria, with special reference to continuous culture. J Gen Microbiol *15*, 492-511.

Powell, E. O. (1958). An outline of the pattern of bacterial generation times. J Gen Microbiol *18*, 382-417.

R

Rappolee, D. A., Wang, A., Mark, D. and Werb, Z. (1989). Novel method for studying mRNA phenotypes in single or small numbers of cells. J Cell Biochem *39*, 1-11.

Riviere, I., Sunshine, M. J. and Littman, D. R. (1998). Regulation of IL-4 expression by activation of individual alleles. Immunity *9*, 217-28.

Rosania, G. R., Chang, Y. T., Perez, O., Sutherlin, D., Dong, H., Lockhart, D. J. and Schultz, P. G. (2000). Myoseverin, a microtubule-binding molecule with novel cellular effects. Nat Biotechnol *18*, 304-8.

Ross, I. L., Browne, C. M. and Hume, D. A. (1994). Transcription of individual genes in eukaryotic cells occurs randomly and infrequently. Immunol Cell Biol *72*, 177-85.

Roux, W. (1881). Der Kampf der Theile im Organismus. Leipzig, Engelmann.

Ruano, D., Lambolez, B., Rossier, J., Paternain, A. V. and Lerma, J. (1995). Kainate receptor subunits expressed in single cultured hippocampal neurons: molecular and functional variants by RNA editing. Neuron *14*, 1009-17.

Rubelj, I. and Vondracek, Z. (1999). Stochastic mechanism of cellular aging--abrupt telomere shortening as a model for stochastic nature of cellular aging. J Theor Biol *197*, 425-38.

Russo-Marie, F., Roederer, M., Sager, B., Herzenberg, L. A. and Kaiser, D. (1993). Beta-galactosidase activity in single differentiating bacterial cells, Proc Natl Acad Sci U S A *90*, 8194-8.

S

Sánchez, Alvarado, A. (2001). La régéneration tissulaire, Biofutur *213*, 48-51.

Sasai, M., Wolynes, PG. (2003). Stochastic gene expression as a many-body problem. Proc Natl Acad Sci U S A *100*, 2374-9.

Schedl, A., Ross, A., Lee, M., Engelkamp, D., Rashbass, P., van Heyningen, V. and Hastie, N. D. (1996). Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities.Cell *86*, 71-82.

Schneuwly, S., Klemenz, R. and Gehring, W.J. (1987). Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia. Nature *325*, 816-8.

Schrödinger, E. (1944). What is Life? The Physical Aspect of the Living Cell. Cambridge, University Press. *Trad. Fr.* (1986). Qu'est-ce que la vie ? Paris, Christian Bourgois Editeur.6.

Simon, R. D. (1976). The biosynthesis of multi-L-arginyl-poly(L-aspartic acid) in the filamentous cyanobacterium Anabaena cylindrica. Biochim Biophys Acta *422*, 407-18.

Simpson, J. L. and Golbus, M. S. (1992). Genetics in Obstetrics and Gynecology. 2nd ed. Philadephia, Saunders Company.

Smith, H. R., Chuang, H. H., Wang, L. L., Salcedo, M., Heusel, J. W., Yokoyama, W. M. (2000). Nonstochastic coexpression of activation receptors on murine natural killer cells. J Exp Med *191*, 1341-54.

Spemann, . H. and Mangold, H. (1924). Über Induktion von Embryonanlagen durch Implatation artfremder Organisatoren. Roux' Archiv für Entwicklungsmechanik *100*, 599-638.

Spudich, J. L. and Koshland, D. E., Jr. (1976). Non-genetic individuality: chance in the single cell. Nature *262*, 467-71.

Stefanis, L., Larsen, K. E., Rideout, H. J., Sulzer, D. and Greene, L. A. (2001). Expression of A53T mutant but not wild-type alpha-synuclein in PC12 cells induces alterations of the ubiquitin-dependent degradation system, loss of dopamine release, and autophagic cell death. J Neurosci *21*, 9549-60.

Sucher, N. J., Deitcher, D. L., Baro, D. J., Warrick, R. M. and Guenther, E. (2000). Genes and channels: patch / voltage-clamp analysis and single-cell RT-PCR. Cell Tissue Res *302*, 295-307.

Swain, P. S., Elowitz, M. B. and Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci U S A *99*, 12795-800.

T

Takasuka, N., White, M. R., Wood, C. D., Robertso W. R. and Davis, J. R. (1998). Dynamic changes in prolactin promoter activation in individual living lactotrophic cells. Endocrinology *139*, 1361-8.

Tanaka, E. M., Drechsel, D. N. and Brockes, J. P. (1999). Thrombin regulates S-phase re-entry by cultured newt myotubes. Curr Biol *9*, 792-9.

Thakker-Varia, S., Alder, J., Crozier, R. A., Plummer, M. R. and Black, I. B. (2001). Rab3A is required for brain-derived neurotrophic factor-induced synaptic plasticity: transcriptional analysis at the population and single-cell levels. J Neurosci *21*, 6782-90.

Thattai, M. and van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. Proc Natl Acad Sci U S A. *98*, 8614-9.

Till, JE., Mc Culloch, E. A. and Siminivitch, L. (1964). A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming. Proc Natl Acad Sci U S A *51,* 29-36.

Todd, R. and Margolin, D. H. (2002). Challenges of single-cell diagnostics: analysis of gene expression. Trends Mol Med *8*, 254-7.

Tolker-Nielsen, T., Holmstrom, K., Boe, L. and Molin, S. (1998). Non-genetic population heterogeneity studied by in situ polymerase chain reaction. Mol Microbiol *27*, 1099-105.

Tonegawa, S. (1976). Reiteration frequency of immunoglobulin light chain genes: further evidence for somatic generation of antibody diversity. Proc Natl Acad Sci U S A *73*, 203-7.

Tsuzuki, K., Lambolez, B., Rossier, J. and Ozawa, S. (2001). Absolute quantification of AMPA receptor subunit mRNAs in single hippocampal neurons. J Neurochem *77*, 1650-9.

V

Velloso, C. P., Simon, A. and Brockes, J. P. (2001). Mammalian postmitotic nuclei reenter the cell cycle after serum stimulation in newt/mouse hybrid myotubes. Curr Biol *11*, 855-8.

W

Walter, J., Schermelleh, L., Cremer, M. and Tashiro, S. and Cremer, T. (2003). Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. J Cell Biol *160*, 685-97.

Walters, M. C., Fiering, S., Eidemiller, J., Magis, W., Groudine, M. and Martin, D. I. (1995). Enhancers increase the probability but not the level of gene expression. Proc Natl Acad Sci U S A *92*, 7125-9.

Watson, J. D. and Crick, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-8.

Webster, G. and Wolpert, L. (1966). Studies on pattern regulation in hydra. I. Regional differences in time required for hypostome determination. J Embryol Exp Morphol *16*, 91-104.

Weigmann, K. (2004). The code, the text and the language of God. EMBO Rep *5*, 116-8.

Wijgerde, M., Grosveld, F. and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. Nature *377*, 209-13.

Wilcox, A. J., Weinberg, C. R., O'Connor, J. F., Baird, D. D., Schlatterer, J. P., Canfield, R. E., Armstrong, E. G. and Nisula, B. C. (1988). Incidence of early loss of pregnancy. N Engl J Med *319*, 189-94.

Wolf, U. (1997). Identical mutations and phenotypic variation. Hum Genet *100*, 305-21.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J Theor Biol *25*, 1-47.

X

Xue, L., Fletcher, G. C. and Tolkovsky, A. M. (1999). Autophagy is activated by apoptotic signalling in sympathetic neurons: an alternative mechanism of death execution. Mol Cell Neurosci *14*,180-98.

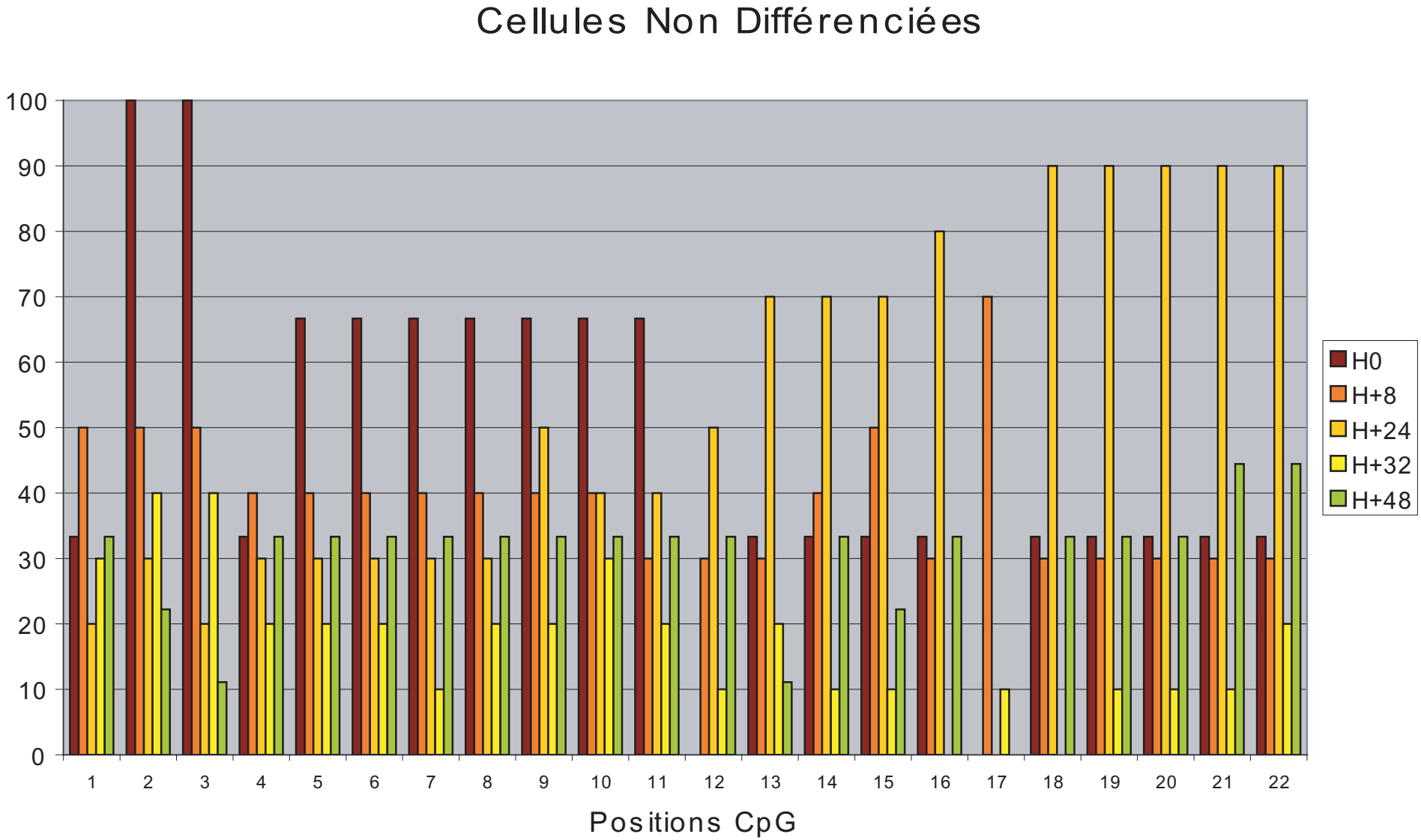
Y

Yuan, J., Lipinski, M. and Degterev, A. (2003). Diversity in the mechanisms of neuronal cell death. Neuron *40*, 401-13.

ANNEXES

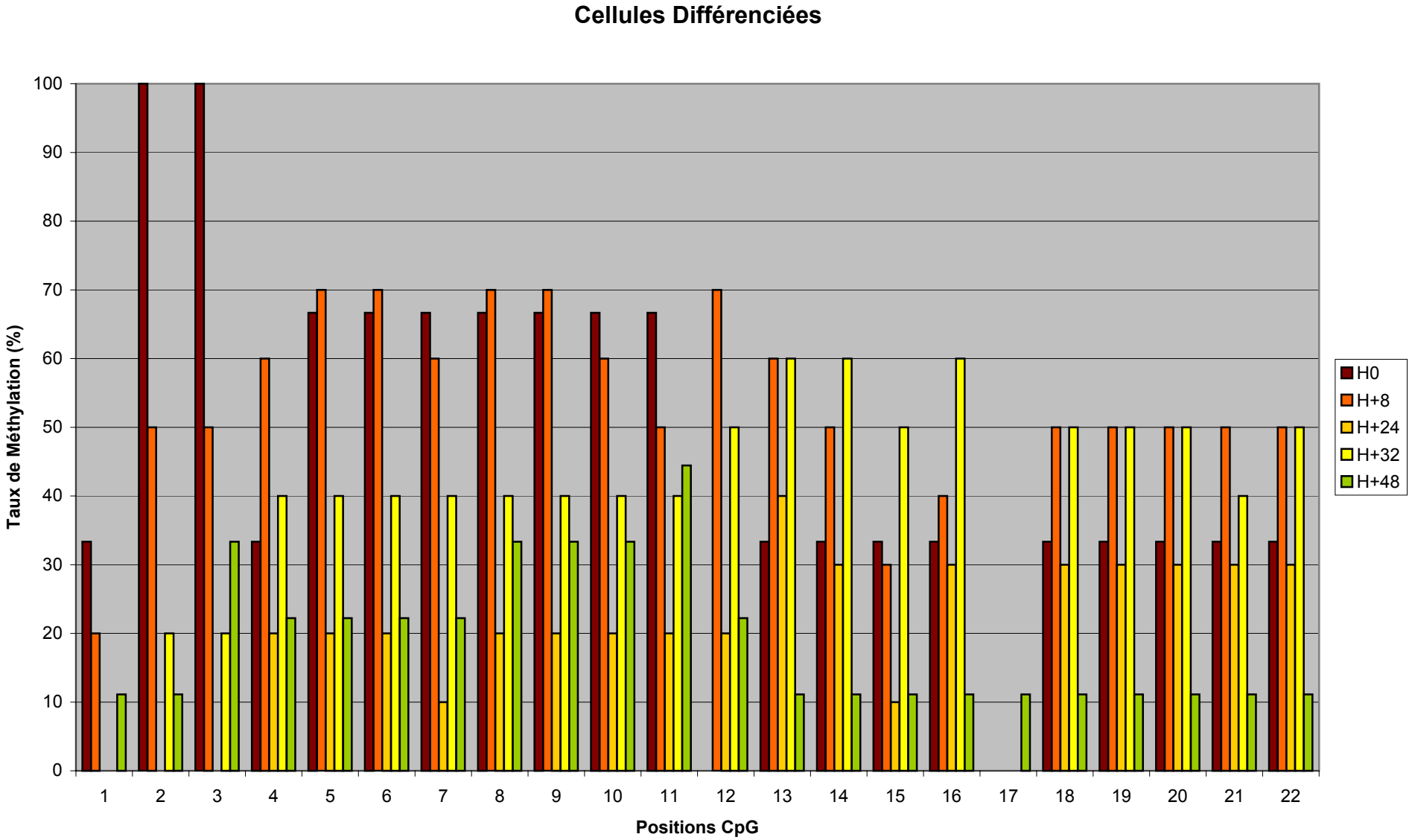
ANNEXE 1

Variation du taux de méthylation du gène Mest/Peg1. Données totales.
Population témoin de cellules murines N1E-115 en milieu riche en sérum de veau fœtal


Cellules Non Différenciées
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Positions CpG
H0
H+8
H+24
H+32
H+48

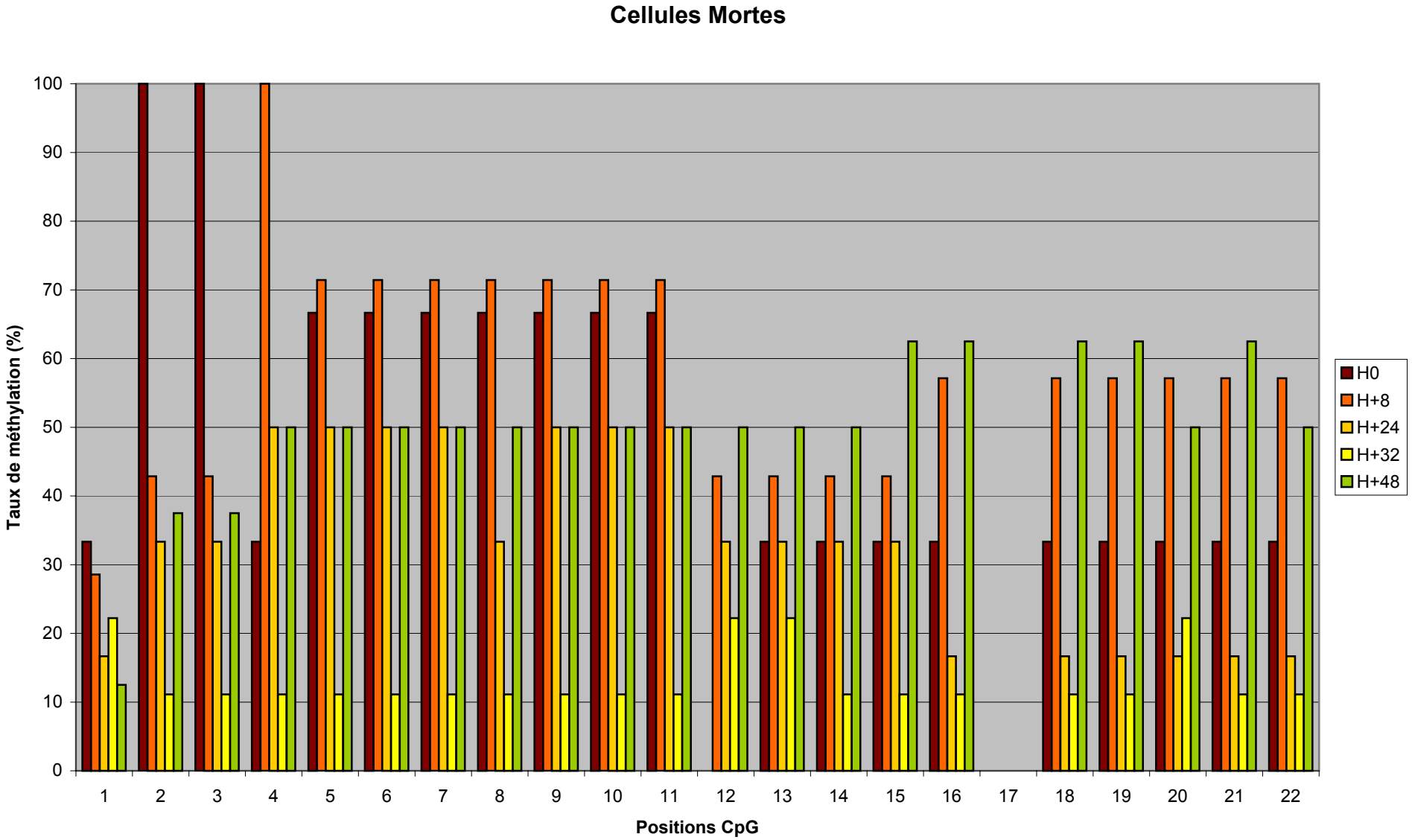
ANNEXE 2

Variation du taux de méthylation du gène Mest/Peg1. Données totales.
Population de cellules murines N1E-115 en milieu sans sérum de veau fœtal, subissant la différenciation.


Cellules Différenciées
Taux de Méthylation (%)
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Positions CpG
H0
H+8
H+24
H+32
H+48

ANNEXE 3

Variation du taux de méthylation du gène Mest/Peg1. Données totales.
Surnageant de cellules mortes de la population de cellules murines N1E-115 en milieu sans sérum de veau fœtal.


Cellules Mortes
Taux de méthylation (%)
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Positions CpG
H0
H+8
H+24
H+32
H+48

ANNEXE 4

Données brutes totales de l'étude des variations du taux de méthylation du gène Mest/Peg1

Nombre de séquences méthylées

		0	0 ind	T8+	T8 + ind	T8-	T8- ind	T8 M	T8 M ind	T24 +	T24 + ind	T24-	T24- ind	T24 M	T24 D ind	T32 +	T32+ ind	T32-	T32- ind	T32 M	T32 M ind	T48 +	T48+ ind	T4 8-	T4 8- ind	T48 M	T48 M ind
Nombre de séquences		*3*		*10*		*10*		*7*		*10*		*10*		*6*		*10*		*10*		*9*		*9*		*9*		*8*	
POSITION CpG	1	1		5	1	2		2		2		0		1		3		0		2		3		1		1	
	2	3		5	1	5		3		3		0	2	2		4		2		1		2		1		3	
	3	3		5	1	5		3		2		0	1	2		4		2		1		1		3		3	
	4	1	1	4		6		7		3		2		3		2		4		1		3		2		4	
	5	2		4		7		5		3		2		3		2		4		1		3		2		4	
	6	2		4		7		5		3		2		3		2		4		1	2	3		2		4	
	7	2		4		6	1	5	1	3		1	1	3		1		4		1		3		2		4	
	8	2		4		7		5	1	3		2		2	1	2		4		1		3		3		4	
	9	2		4		7	1	5		5		2		3	1	2		4		1		3		3		4	
	10	2		4		6		5		4		2		3		3		4		1		3		3		4	
	11	2		3		5		5		4		2		3		2		4		1		3		4		4	
	12	0		3	1	7		3		5		2		2		1		5		2		3		2		4	
	13	1		3	1	6		3		7		4		2		2		6		2	1	1		1		4	
	14	1		4	1	5		3		7		3		2		1		6		1	1	3		1		4	
	15	1		5	1	3	1	3		7		1		2		1		5		1	1	2	1	1		5	
	16	1		3		4		4		8		3		1		0		6		1	1	3	1	1		5	
	17	0		7		0		0		0		0		0		1		0		0	1	0	1	1		0	
	18	1		3		5		4		9		3		1		0		5		1	1	3	1	1		5	
	19	1		3		5		4		9		3		1		1		5		1	1	3	1	1		5	
	20	1		3		5		4		9		3		1		1		5		2	1	3	1	1		4	1
	21	1		3		5		4		9		3		1		1		4		1	1	4	1	1		5	
	22	1		3		5		4		9		3		1		2		5		1	1	4	1	1		4	1

TO : initiation de la cinétique
T8 : point de cinétique H+8
T24 : point de cinétique H+24
T32 : point de cinétique H+32
T48 : point de cinétique H+48

\+ : cellules cultivées en milieu SVF +, ne subissant pas la différenciation
\- : cellules cultivées en milieu SVF-, subissant la différenciation
M : celules mortes au cours de la différenciation (surnageant des cultures

ind : séquence C ou T indéterminable (par défaut =0)

ANNEXE 5

Publications

tration and measured fluorescence of the data shown in Figure 2 was obtained [Bartlet's nonparametric Model II regression analysis (7)] and was represented by the equation: Log(DNA) = -13.50 + 4.25 Log(Fluorescence). This equation can be used to estimate the concentration of DNA in a solution from fluorescence measured, following the protocol outlined above.

Although estimating DNA concentrations might look trivial, the outcome of several experiments depends on the accurate determinations of DNA in the samples or stock DNA preparations. PCRs can be inhibited by excessive DNA concentrations, and no amplification product will be obtained if DNA is below the amplification threshold. Concentrations of genomic DNA are tricky to estimate because of the viscosity of the concentrated solutions and the unevenness of long DNA in more diluted solutions. This easy and accurate protocol will facilitate the estimation of DNA concentrations. In our laboratory, we perform serial dilutions of genomic DNAs over several orders of magnitude to obtain an accurate estimation of DNA concentrations in stock solutions. The use of a 96-well plate or a multitube approach facilitates this strategy, and the iQ iCycler optical thermal cycler can easily be used for this purpose. The protocol outlined in this study avoids the need to purchase costly fluorescence plate readers, learn how to use an extra apparatus, and simultaneously saves precious bench space.

REFERENCES

1. **Morrison, T.B., J.J. Weis, and C.T. Wittwer.** 1998. Quantitaion of low copy transcripts by continuous SYBR Green I monitoring during amplification. BioTechniques *24*:954-962.
2. **Akey, J.M., D. Sosnoski, E. Parra, S. Dios, K. Hiester, B. Su, C. Bonilla, L. Jin, et al.** 2001. Melting curve analysis of SNPs (McSNP): a gel-free and inexpensive approach for SNP genotyping. BioTechniques *30*:358-367.
3. **Ririe, K.M., R.P. Rasmussen, and C.T. Wittwer.** 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal. Biochem. *245*:145-160.
4. **Gonzalez, J. M. and C. Saiz-Jimenez.** 2002. A fluorimetric method for the estimation of G+C mol% content in microorganisms by thermal denaturation temperature. Environ. Microbiol. *4*:770-773.
5. **Molecular Probes.** 1996. SYBR Green I nucleic acid gel stain. Molecular Probes Application Sheet MP-7567.
6. **Sambrook, J., E.F. Fritsch, and T. Maniatis.** 1989. Molecular Cloning: A Laboratory Manual. CSH Laboratory Press, Cold Spring Harbor, NY.
7. **Sokal, R.R. and F.J. Rohlf.** 1981. Biometry, 2nd ed. W.H. Freeman and Co., New York, NY.

We acknowledge support from the Spanish Ministry of Science and Technology through the Ramón y Cajal Program and projects REN2002-00041 and BTE2002-04492-C02-01. *Address correspondence to Dr. Juan M. Gonzalez, IRNAS-CSIC, Apdo 1052, 41080, Sevilla, Spain. e-mail: jmgrau@irnase.csic.es*

Received 19 December 2002; accepted 24 January 2003.

Juan M. Gonzalez and Cesareo Saiz-Jimenez
Instituto de Recursos Naturales y Agrobiología
CSIC, Sevilla, Spain

Modified 3′-End Amplification PCR for Gene Expression Analysis in Single Cells

BioTechniques 34:712-716 (April 2003)

For years, gene expression has been widely studied in most fields of biology. It has been used to address molecular mechanisms controlling basic cell functions and their alterations, including cell division, cell differentiation, apoptosis, and various infectious processes. In the vast majority of such studies, carried out by conventional techniques such as Northern blotting or RT-PCR, only the mean gene expression within large populations of cells was analyzed, and cell-to-cell variability was not taken into account. Remarkably, several independent studies using in situ detection techniques have recently pointed to the existence of important intercellular variability in gene expression, both in prokaryotic and mammalian cells. Altogether, the data strongly argue for a stochastic-based process of gene expression (1–4). To further our understanding of cell physiology, there is an increasing need to evaluate intercellular variability of gene expression in a large number of cell types and physiological situations. To this end, several techniques have recently been developed, each with various advantages and limitations (5).

In the course of our work, we used the 3′-end amplification PCR (TPEA-PCR) procedure, which has proven to be an efficient tool for studying gene expression in various mammalian cell lines at the single-cell level. Although this technique is not quantitative, it allows virtually all the mRNAs sequences present in a single cell to be amplified, including those present at low copy number (6). This RT-PCR relies on two successive amplifications. First, cell lysate-containing mRNAs are reverse-transcribed, followed by a nonspecific amplification, using primers designed to amplify all polyadenylated mRNAs. Second, there is a gene-specific amplification of up to 40 aliquots of the first amplification product. Thus, it is theoretically possible to detect up

to 40 different transcripts (20 in duplicates) in a single cell and to follow gene expression patterns and their variations from cell to cell.

However, contamination remains critical, especially when the quantity of RNA in the starting material is limited, as is always the case with single-cell RT-PCR. Decreasing the number of manipulations carried out during the procedure will reduce the risk of contamination. Here we propose such a modification of the TPEA technique (schematically shown in Figure 1) that maintains the quality of the result. We reduced the successive substeps originally required for the nonspecific amplification from seven to four. Specifically, we have combined the two second-strand synthesis substeps with a round of 15 nonspecific PCR cycles. This simplification largely reduces the contamination risk, results in a significant gain of time, and becomes particularly convenient when a large number of cells are to be screened concomitantly. For example, the amplification time required for 30 single cells is 1.5 h shorter than that required for the classical TPEA technique (8 h including cell harvest and lysis).

In our experiments, murine neuroblastoma cells (clone N1E-115) were cultured in DMEM with 4.5 g/L glucose (Invitrogen, Carlsbad, CA, USA) supplemented with 10% decomplemented FBS and 10 mL/L penicillin-streptavidin (Invitrogen).

All the reverse transcription- and PCR-based reactions were carried out in a GeneAmp® PCR 9700 apparatus (Applied Biosystems, Foster City, CA, USA). Reverse transcription primer (5′-CTCTCAAGGATCTTACCGCTTTTTTTTTTTTTTTT-3′), second-strand primer (5′-CTGCATCTATCTAATGCTCCNNNNNCGAGA-3′, where N represents C, G, T, or A), and 3'-heel primer (5′-CTCTCAAGGATCTTACCGC-3′) were those described by Dixon et al. (6).

After cell harvest and serial dilutions in Terasaki plates, single cells (1 µL) were identified under a light microscope and lysed in situ in a final volume of 7 µL containing 1× first-strand buffer (50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$; Invitrogen), 5 mM Nonidet P-40™ (Sigma, St. Louis, MO, USA), 1.5 U RNase Inhibitor RNaseOUT™ (Invitrogen). After 5 min on ice, nuclei were removed by centrifugation (8000× *g* for 5 min at 4°C). Each supernatant was then transferred to a new tube and reverse transcribed (first strand) in a final volume of 10 µL containing 1× first-strand buffer, 200 U Moloney murine leukemia virus reverse transcriptase (Invitrogen), and 0.5 ng reverse transcription primer for 60 min at 37°C.

cDNA second-strand synthesis, primer extension, and 15 rounds of PCR were combined as a single step instead of being carried out separately (Figure 1). The 10-µL reverse transcription volume was added to a 10.4-µL PCR mixture and was subjected to 50°C for 15 min (incubation), 72°C for 10 min (primer extension), 92°C for 5 min (denaturation), 15 PCR cycles (2.5 min at 92°C, 60°C for 1.5 min, and 72°C for 1.5 min), and 72°C for 10 min. Final concentrations of this 20.4-µL reaction volume were 4.5 mM $MgCl_2$, 67 mM Tris-HCl (pH 8.3), 0.5 mM dNTPs (Amersham Biosciences UK, Buckinghamshire, UK), 1 ng second-strand primer, 0.4 ng 3′-heel primer, and 0.35 U AmpliTaq® DNA polymerase (Applied Biosystems).

The 20.4-µL RT-PCR volume was then increased to 30.4 µL without changing reagent concentrations, except for the addition of 125 ng second-strand primer and 50 ng 3′-heel primer, and subjected to denaturation at 92°C for 5 min, 25 additional PCR cycles (as

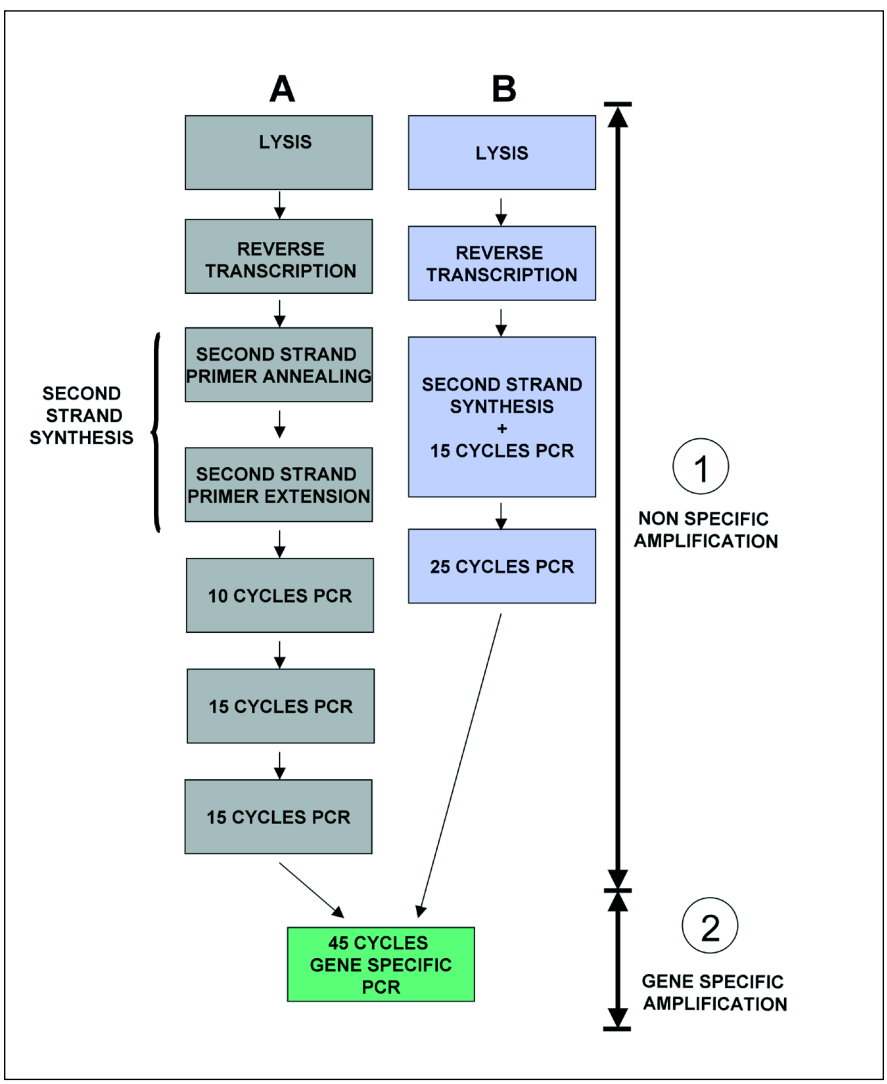


Figure 1. Original (A) and modified (B) TPEA-PCR protocols.

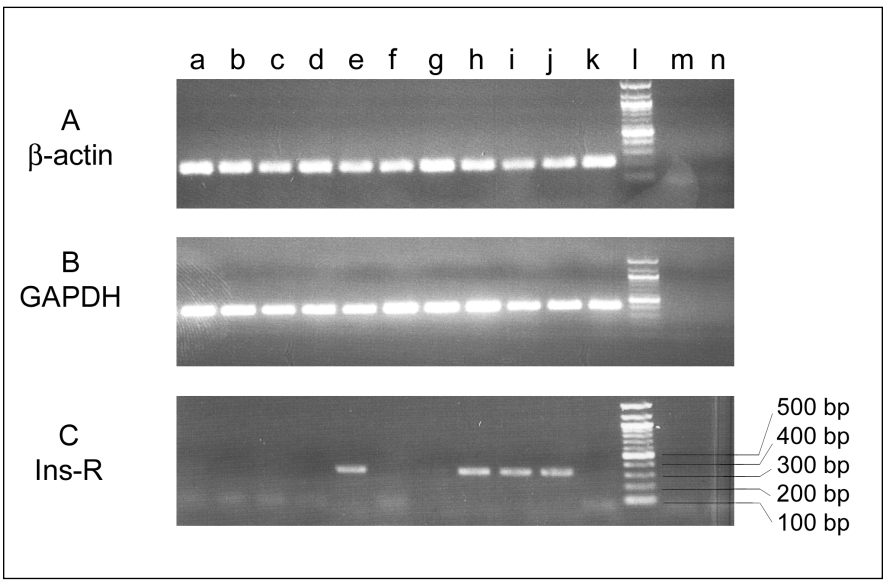


Figure 2. Agarose gel electrophoresis of β-actin (A), GAPDH (B), and insulin receptor (Ins-R; C) cDNAs. The modified TPEA-PCR protocol was used to detect the expression of the three genes in 11 single NE1-115 cells (lanes a–k). Lane l, molecular weight marker (100-bp DNA ladder; New England Biolabs). Lane m, negative control (no reverse transcriptase). Lane n, negative control (no cell). Primers were the following: β-actin (209 bp; GenBank® accession no. V01217) 5′-CATCCATGCCCT-GAGTCC-3′ (forward) and 5′-ACACCTCAAACCACTCCCAG-3′ (reverse); GAPDH (408 bp; accession no. X02231) 5′-GACCCCTTCATTGACCTCAACTAC-3′ (forward) and 5′-GATGC-CAAAGTTGTCATGGATGAC-3′ (reverse); and insulin receptor (369 bp; accession no. NM 010568) 5′-TCACAAAGATGGCGTATT-3′ (forward) and 5′-GACTGGAGCAAACACAAC-3′ (reverse).

described above), and a final elongation at 72°C for 10 min.

Additionally, we improved the yield of the second gene-specific PCR by using only 3 µL of the nonspecific amplification product, instead of 5 µL. The potential number of gene-specific PCRs is thereby increased from 40 to 50. Thus, the 30.4-µL final reaction volume was diluted to a 150-µL final volume in (10 mM Tris-HCl, 0.1 mM EDTA), and 3 µL of this solution were used for the second PCR using gene-specific primers. Each sample was subjected to PCR in a 50-µL final volume reaction containing at final concentrations: 1× PCR Buffer II (100 mM Tris-HCl, 500 mM KCl; Applied Biosystems), 3.5 mM $MgCl_2$ (Applied Biosystems), 12.5% sucrose, 12 mM β-mercaptoethanol, 0.5 mM dNTPs, 0.6 U AmpliTaq, and primers at 100 ng/reaction. Cresol red was found to be unnecessary and was therefore omitted. Annealing temperatures were adjusted when necessary, but otherwise PCR conditions were those used by Dixon et al. (6).

The modified TPEA protocol gave sensitive and reproducible results. In Figure 2, we show the results we obtained using three different sets of primers specific for β-actin (A), GAPDH (B), and the insulin receptor (C) genes. Importantly, no signal was detected when reverse transcriptase was omitted, demonstrating the absence of genomic DNA amplification. This result was strengthened by the fact that the GAPDH primers used for this amplification were intron spanning.

As expected, the two constitutive genes (β-actin and GAPDH) were found to be expressed in all the single cells. In contrast, expression of the insulin receptor gene, which is not constitutively produced in N1-E115 neuroblastoma cells, was found to be variable and was detected in four of 11 cells.

As an additional control, we have also divided cDNA pools in two after the reverse transcription step and compared the signals obtained after the full TPEA procedure was performed with β-actin primers. No bias in the amplification of transcripts was detected (data not shown). Altogether, these modifications result in a faster, more efficient, and more reliable TPEA-PCR technique.

REFERENCES

1. **Elowitz, M.B., A.J. Levine, E.D. Siggia, and P.S. Swain.** 2002. Stochastic gene expression in a single cell. Science *297*:1183-1186.
2. **De Krom, M., M. van de Corput, M. von Lindern, F. Grosveld, and J. Strouboulis.** 2002. Stochastic patterns in globin gene expression are established prior to transcriptional activation and are clonally inherited. Mol. Cell *9*:1319-1326.
3. **Hume, D.A.** 2000. Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. Blood *96*:2323-2328.
4. **Kupiec, J.J.** 1996. A chance-selection model for cell differentiation. Cell Death Differ. *3*:385-390.
5. **Todd, R. and D.H. Margolin.** 2002. Challenges of single-cell diagnostics: analysis of gene expression. Trends Mol. Med. *8*:254-257.
6. **Dixon, A.K., P.J. Richardson, K. Lee, N.P. Carter, and T.C. Freeman.** 1998. Expression profiling of single cells using 3 prime end amplification (TPEA) PCR. Nucleic Acids Res. *26*:4426-4431.

We are grateful to Anne Tyvaert for technical assistance. We thank Drs. Sitbon, Taylor (Montpellier), and Tardieux (Paris) for support and critical reading of this manuscript. We are grateful to Dr. Sonigo for constant support and helpful advice. Address correspondence to Dr. Jean-Jacques Kupiec, Equipe de Génétique des virus, Institut Cochin, INSERM U567, CNRS UMR 8104, 22 rue Méchain, 75014 Paris, France. e-mail: kupiec@cochin.inserm.fr

Received 23 December 2002; accepted 7 February 2003.

Thomas Heams and Jean-Jacques Kupiec
Institut Cochin
Equipe de Génétique des virus
INSERM U567 and
CNRS UMR 8104
Paris, France

European Journal of Human Genetics (2003) 11, 493–496
© 2003 Nature Publishing Group All rights reserved 1018-4813/03 $25.00
www.nature.com/ejhg

ARTICLE

In vitro follicular growth affects oocyte imprinting establishment in mice

Antoine Kerjean*[,1,2], Philippe Couvert[1], Thomas Heams[1], Céline Chalas[1,2], Karine Poirier[1], Jamel Chelly[1], Pierre Jouannet[2], Andras Paldi[3] and Catherine Poirot[4]

[1]*GDPM, Institut Cochin, INSERM, CNRS-Université Paris V and CHU Cochin, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France;* [2]*GREFH, Université Paris V and CHU Cochin, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France;* [3]*Institut Jacques Monod, 2 place Jussieu, 75005 Paris, France;* [4]*Laboratoire de Biologie de la Reproduction, CHU La Pitié-Salpétrière, 75013 Paris, France*

***In vitro* folliculogenesis of cryopreserved ovarian tissue could be an effective method for insuring fertility for patients who receive gonadotoxic treatment. Although several culture systems have been described for growing female gametes *in vitro*, the production of competent oocytes for further development remains a considerable challenge. The purpose of our study was to determine whether maternal primary imprinting progresses normally during mouse oocyte growth *in vitro*. We analysed the DNA methylation status of differentially methylated regions of the imprinted genes *H19*, *Mest/Peg1* and *Igf2R* using fully grown germinal vesicle-stage oocytes (fg oocytes) produced by *in vitro* folliculogenesis from early preantral follicles. When compared to fg oocytes removal from control females, we observed after *in vitro* development, a loss of methylation at the *Igf2R* locus in six out of seven independent experiments and *Mest/Peg1* locus (one out of seven), and a gain of methylation at the *H19* locus (one out of seven). These results provide insight into the dysregulation of the process of primary imprinting during oocyte growth *in vitro* and highlight the need for effective new biomarkers to identify complete nuclear reprogramming competence after *in vitro* folliculogenesis.**
European Journal of Human Genetics (2003) **11,** 493–496. doi:10.1038/sj.ejhg.5200990

Keywords: bisulfite; culture; DHPLC; imprinting; methylation; oogenesis

Introduction

Various culture systems have been described for the maturation of female gametes *in vitro*.[1] However, the development of primordial follicles to maturity, in which oocytes acquire *in vitro* competence to undergo maturation, fertilization and embryonic development, remains a considerable challenge. Recent evidence suggests that mechanisms regulating oocyte growth underlie the establishment of maternal primary imprinting during mouse oogenesis.[2] This lead to monoallelic expression/repression of imprinted genes involved in embryonic growth, placentation and behavioral development.[3] During oogenesis, expression and/or methylation analysis of imprinted genes provide evidence for the erasure of the parental imprint in primordial germs cells.[4] New imprints are then later initiated during the development of nongrowing oocytes to fully grown (fg) oocytes as indicated by a marked change of maternal germline-specific imprints.[2,5] Thus, maternal primary imprinting may be particularly susceptible to methylation changes that occur at imprinted loci during oogenesis. Since previous studies in mice have suggested that culture conditions could affect the epigenetic regulation of genomic imprinting,[6,7] the aim of our study was to determine the consequence of *in vitro* folliculogenesis on the progress of maternal primary imprinting during germinal vesicle-stage oocyte growth.

*Correspondence: Dr A Kerjean, Epigenix Research Group, GDPM, Institut Cochin, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France. Tel: +(33) 1 44 41 24 56; Fax: +(33) 1 44 41 24 62;
E-mail: kerjean@cochin.inserm.fr
Received 8 November 2002; revised 3 February 2003; accepted 11 February 2003

Materials and methods

Preantral follicles containing early growing germinal vesicle-stage oocytes (eg oocytes) were isolated from ovaries of 11- and 12-day old F1 hybrid female mice (C57bl6 × CBA). Early preantral follicles were selected according to the following criteria: intact follicle structure with one or two layers of granulosa cells, visible round and central oocyte, adhering thecal cells, follicle diameter between 100 and 130 μm. Follicles were placed individually into 10 μl drops of culture medium, under pre-equilibrated mineral oil (Sigma-Aldrich) at 37°C in an atmosphere of 5% CO_2 in air, using a system similar to the one previously described.[8] Culture medium consisted of α minimal essential medium (Life Technologies) supplemented with 5% heat inactivated FCS, 10 μg/ml transferrin (Boehringer Mannheim, France), 5 μg/ml insulin (Boehringer Mannheim) and 100 mIU/ml r-FSH (Serono, France). Healthy oocyte–cumulus complexes consisting of compact granulosa cells with regular borders surrounding the entire mass were used in these experiments. On day 11 of culture, fg oocytes were freed of their surrounding cumulus cells by repeated pipetting in 1 mg/ml hyaluronidase (Sigma-Aldrich) and washed in serum-free modified HTF medium (Irvine Scientific, USA). As control, we used antral follicle grown *in vivo* collected on the surface of the ovary of adult mice.

Bisulfite treatment of 5–8 fg oocytes imbedded in agarose beads was performed as previously reported.[9] Special care was taken to avoid contamination from granulosa cells. DNA methylation analysis of specific differentially methylated regions (DMRs) of *H19*, *Mest/Peg1* and *Igf2R* was performed on independent pools of 5–8 fg oocytes by denaturing high-performance liquid chromatography (DHPLC) analysis as recently reported.[10]

Results and discussion

We have analysed the methylation status of DMRs of imprinted genes using a newly developed DHPLC-based procedure.[10] Analysis of control fg oocytes from different PCR products indicated that *H19* alleles remained essentially all unmethylated (Figure 1c), whereas *Mest/Peg1* alleles (Figure 2c) and *Igf2R* alleles (Figure 3c) were almost all methylated when compared to curves representing 100 and 0% methylated/unmethylated control alleles. These methylation profiles were characteristic of the maternal imprint and provided evidence that the maternal methylation imprint was already established in *in vivo* fg oocytes as previously described.[2,5,11]

After *in vitro* folliculogenesis, DHPLC profiles of PCR products from *in vitro* fg oocytes showed almost unmethylated alleles in six out of seven samples at the *H19* locus (Figure 1d) and methylated alleles in six out of seven samples at the *Mest/Peg1* locus (Figure 2d). In one (one out of seven) PCR product, a gain of allelic methylation and a

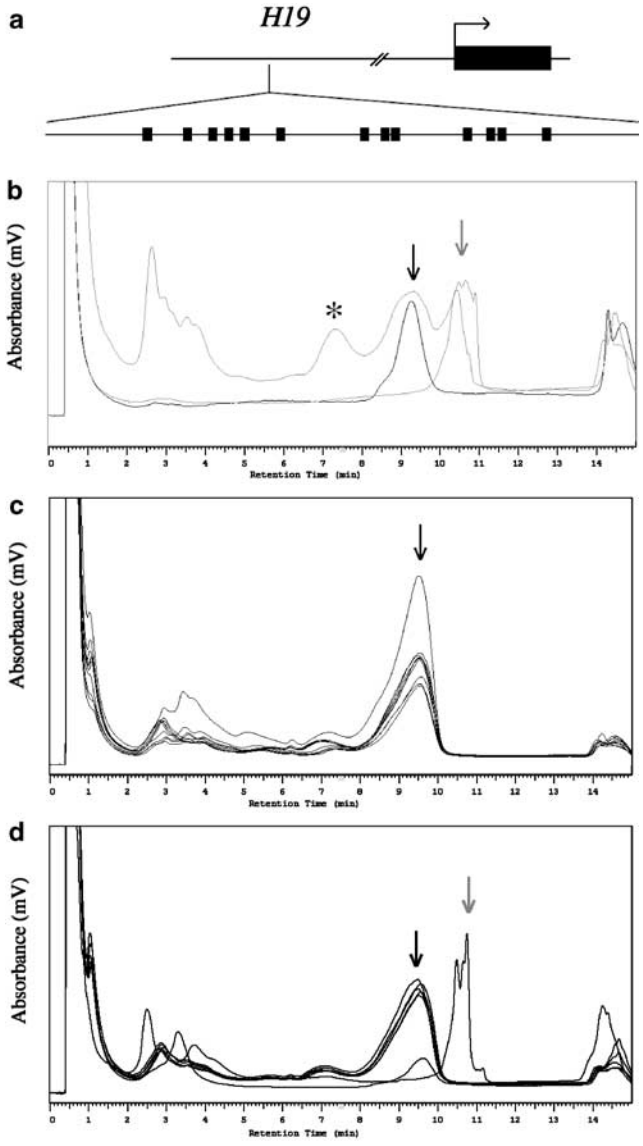


Figure 1 Methylation analysis of imprinted *H19*, (**a**) Genomic structures of the DMR studied here. We amplified 13 CpGs of the *H19* imprinting control region (accession no. U19619.1; nt −2975, −3204), Arrow, transcription start site of the gene; black box, exons; small black squares, individual CpGs within the areas amplified. (**b**) DHPLC chromatograms of PCR products of control samples. Methylation profiles were studied by newly developed DHPLC-based method, as recently described.[10] Single peaks consisting of homoduplexes characterized homogeneous samples with methylated or unmethylated alleles, whereas multiple peaks consisting of homoduplex and heteroduplex formations characterized heterogeneous samples: (*), heteroduplexes; granulosa cells; control methylated alleles; control unmethylated alleles. (**c**) DHPLC chromatograms of PCR products of *in vivo* grown fg oocytes. $n=9$ chromatograms are superposed for *H19*, $n=8$ for *Mest/Peg1* and $n=7$ for *Igf2R*. (**d**) DHPLC chromatograms of PCR products of *in vitro* grown fg oocytes $n=7$ chromatograms are superposed: (→), unmethylated alleles; (↓), methylated alleles; (*), heteroduplexes.

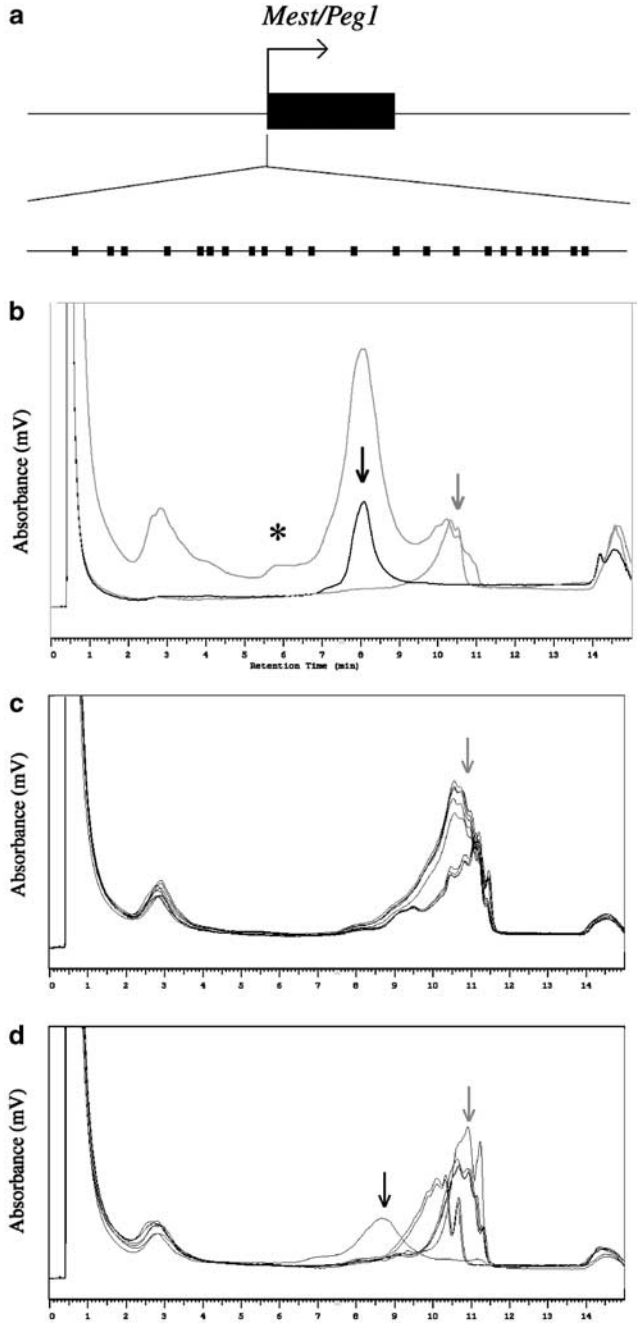


Figure 2 Methylation analysis of imprinted *Mest/Peg1* gene. (**a**) Genomic structure of the DMR studied. We amplified 22 CpGs of the *Mest/Peg1* promoter region (accession no. AF017994; nt −201, +116 relative to transcriptional start). See Figure 1(a) for details. (**b**) DHPLC chromotograms of PCR products of control samples. See Figure 1(b) for details. (**c**) DHPLC chromotograms of PCR products of *in vivo* grown fg oocytes $n=8$ chromotograms are superposed. (**d**) DHPLC chromotograms of PLR products of *in vitro* grown fg oocytes. $n=7$ chromotograms are superposed. See Figure 1(d) for details.

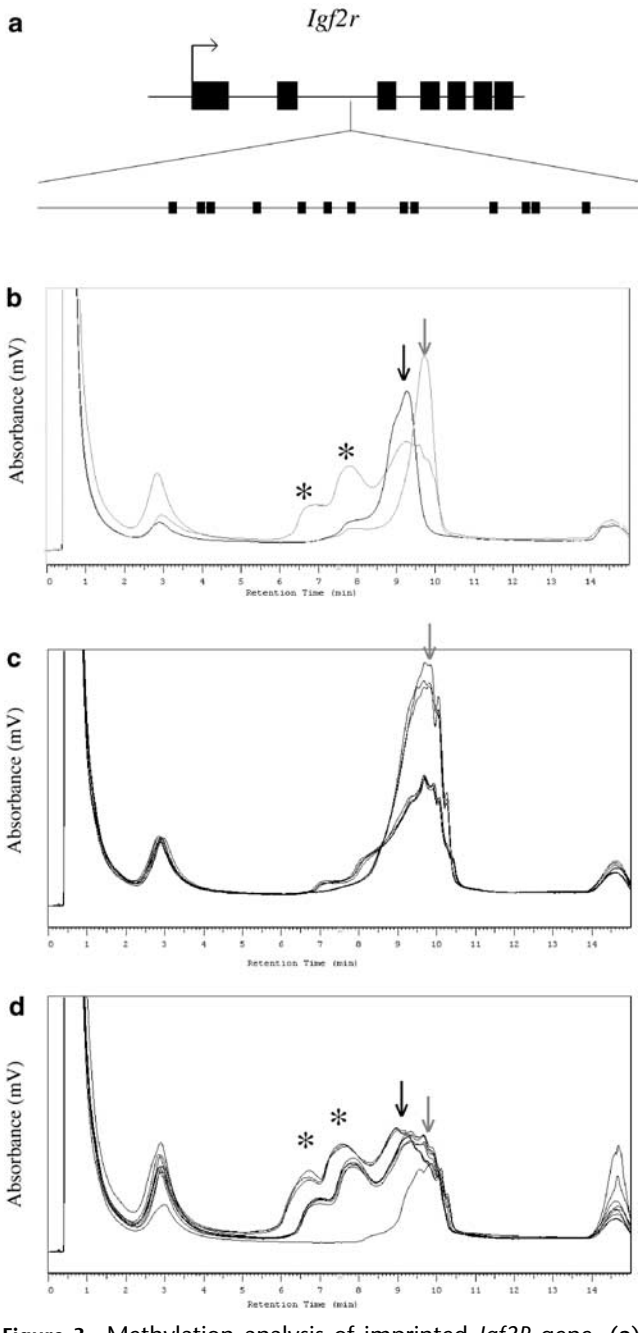


Figure 3 Methylation analysis of imprinted *Igf2R* gene. (**a**) Genomic structure of the DMR studied. We amplified 13CpGs of the *Igf2R* within intronic DMR2 (accession no. L06446; nt 1857–2101). See Figure 1(a) for details. (**b**) DHPLC chromotograms of PCR products of control samples. See Figure 1(b) for details. (**c**) DHPLC chromotograms of PCR products of *in vivo* grown fg oocytes. $n=7$ chromotograms are superposed. (**d**) DHPLC chromotograms of PCR products of *in vitro* grown fg oocytes. $n=8$ chromotograms are superposed. For details see Figure 1(d).

loss or absence of allelic methylation were unexpectedly observed at the *H19* and *Mest/Peg1* loci, respectively. Such abnormal monoallelic profiles for *H19* and *Mest/Peg1* were different from multiallelic patterns that have been observed in granulosa cells (Figure 1, 2 and 3b). In contrast, at the *Igf2R*-R2 locus, DHPLC chromatograms showed profiles corresponding to almost methylated alleles in only one out of eight PCR products, whereas multiple peaks were found in seven out of eight samples (Figure 3d). These multiple peaks, consisting of homoduplex and heteroduplex formations were characteristics of heterogeneous samples with methylated alleles and unmethylated alleles, as confirmed by genomic bisulphite sequencing (not shown). However, since PCR amplifications were performed on pools of five to eight *in vitro* fg oocytes, we could not establish whether this heterogeneity affected individual fg oocytes (methylated and unmethylated alleles in the same fg oocyte) or the entire population of fg oocytes (fg oocytes with methylated alleles, fg oocytes with unmethylated alleles). Nevertheless, these results indicated, for the first time, the disruption (loss, absence or delay) of maternal methylation imprint establishment during *in vitro* folliculogenesis. The mechanisms that could affect maternal primary imprinting establishment as a consequence of *in vitro* culture remain elusive. We hypothesize that these dysregulations could be stochastic events at some imprinted loci as illustrated by improper *de novo* methylation at the *H19* locus and improper demethylation at the *Mest/Peg1* locus. Alternatively, these imprinting errors could be enhanced by *in vitro* cultivation at other specific imprinted loci, as exemplified by results at the *Igf2R*-R2 locus. The consequence of imprinting methylation errors for incomplete oocyte competence requires further investigation. However, it is tempting to speculate that such epigenetic abnormalities could be involved in subsequent abnormal development.[12–14] Recently, complete oogenesis supporting development to full term after nuclear transfer and *in vitro* fertilization was successfully and efficiently accomplished by *in vitro* folliculogenesis using a 28-day *in vitro* culture system.[15] This complete nuclear reprogramming was consistent with establishment of normal methylation imprinting at the *Igf2R* locus and illustrates the critical role of culture procedures on subsequent epigenotype and phenotype. Although, *in vitro* folliculogenesis will enable important progress in different fields of biology and medicine, these results highlight the need for improving *in vitro* culture conditions and developing biomarkers to identify complete nuclear competence.

Acknowledgements

This work was supported by a grant from the 'la Société Générale'. We thank Susan Saint-Just for reading the manuscript.

References

1 Smitz JE, Cortvrindt RG: The earliest stages of folliculogenesis *in vitro*. *Reproduction* 2002; **123**: 185–202.
2 Obata Y, Kono T: Maternal primary imprinting is established at a specific time for each gene throughout oocyte growth. *J Biol Chem* 2002; **277**: 5285–5289.
3 Reik W, Dean W, Walter J: Epigenetic reprogramming in mammalian development. *Science* 2001; **293**: 1089–1092.
4 Surani MA: Imprinting and the initiation of gene silencing in the germ line. *Cell* 1998; **93**: 309–312.
5 Kono T, Obata Y, Yoshimzu T, Nakahara T, Carroll J: Epigenetic modifications during oocyte growth correlates with extended parthenogenetic development in the mouse. *Nat Genet* 1996; **13**: 91–94.
6 Dean W, Bowden L, Aitchison A *et al*: Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. *Development* 1998; **125**: 2273–2282.
7 Khosla S, Dean W, Brown D, Reik W, Feil R: Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. *Biol Reprod* 2001; **64**: 918–926.
8 Cortvrindt R, Smitz J, Van Steirteghem AC: *In-vitro* maturation, fertilization and embryo development of immature oocytes from early preantral follicles from prepuberal mice in a simplified culture system. *Hum Reprod* 1996; **11**: 2656–2666.
9 Kerjean A, Vieillefond A, Thiounn N *et al*: Bisulfite genomic sequencing of microdissected cells. *Nucleic Acids Res* 2001; **29**: e106.
10 Couvert P, Poirier K, Carrié A *et al*: A DHPLC-based method for DNA methylation analysis of differential methylated regions from imprinted genes. *Bio Techniques*, 2003; **34**: 356–362.
11 Lucifero D, Mertineit C, Clarke HJ, Bestor TH, Trasler JM: Methylation dynamics of imprinted genes in mouse germ cells. *Genomics* 2002; **79**: 530–538.
12 DeBaun MR, Niemitz EL, McNeil DE *et al*: Epigenetic alterations of H19 and LIT1 distinguish patients with Beckwith–Wiedemann syndrome with cancer and birth defects. *Am J Hum Genet* 2002; **70**: 604–611.
13 Young LE, Sinclair KD, Wilmut I: Large offspring syndrome in cattle and sheep. *Rev Reprod* 1998; **3**: 155–163.
14 Young LE, Fernandes K, McEvoy TG *et al*: Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. *Nat Genet* 2001; **27**: 153–154.
15 Obata Y, Kono T, Hatada I: Maturation of mouse fetal germ cells *in vitro*. *Nature* 2002; **418**: 497.