

L'unité du vivant

Les techniques de la génétique moléculaire ont entièrement renouvelé nos idées dans les domaines de l'évolution, du développement embryonnaire, du cycle de division cellulaire et de ses dérèglements. Les homologues de structure et de fonction observées entre certaines molécules trouvées chez des organismes aussi différents qu'insectes et mammifères mettent une fois de plus en évidence l'unité du vivant.

François Jacob

RÉFÉRENCES

1. Van Ooyen A, Nusse R. Structure and nucleotide sequence of the putative mammary oncogene *int-1* ; proviral insertions leave the protein-encoding domain intact. *Cell* 1984 ; 39 : 233-40.
2. Wilkinson DG, Bailes JA, McMahon AP. Expression of the proto-oncogene *int-1* is restricted to specific neural cells in the developing mouse embryo. *Cell* 1987 ; 50 : 79-88.
3. Shackleford GM, Varmus HE. Expression of the proto-oncogene *int-1* is restricted to postmeiotic male germ cells and the neural tube of mid-gestational embryos. *Cell* 1987 ; 50 : 89-95.
4. Baker NE. Molecular cloning of sequences from *wingless*, a segment polarity gene in *Drosophila* : the spatial distribution of a transcript in embryos. *EMBO J* 1987 ; 6 : 1765-73.

ADRESSE

F. Jacob : professeur au Collège de France et à l'Institut Pasteur. Institut Pasteur, 25, rue du Dr Roux, 75724 Paris Cedex 15, France.

Dans toute discipline, la recherche connaît des périodes fastes et des périodes creuses. Au cours des années 1970, la biologie moléculaire s'est vue soudain dotée de moyens nouveaux, d'une extraordinaire puissance, pour l'analyse des acides nucléiques. A la suite de quoi, tous les laboratoires se sont mis à repérer des gènes, les isoler, en faire la séquence, en chercher la régulation. D'où, pour commencer, une série de travaux quelque peu fastidieux. Mais après quelques années, la situation s'est renversée. On assiste maintenant à une explosion de résultats qui, dans bien des aspects, bouleversent les idées jusqu'alors en cours. Cela notamment dans les domaines de l'évolution, du développement de l'embryon, du cycle de division cellulaire et de ses aberrations que représentent les oncogènes. Et tout converge ! Pour analyser certains phénomènes chez la souris, où la génétique est si difficile, on peut alors utiliser certaines données obtenues

chez la drosophile, où la génétique est si avancée. Je voudrais le montrer avec deux exemples.

Oncogènes et gènes de développement

Premier exemple : l'homologie récemment trouvée entre des oncogènes de mammifères et des gènes qui gouvernent certains aspects de la morphogenèse chez la drosophile. Tel est le cas de l'oncogène *int-1*. Celui-ci est un gène activé dans les tumeurs mammaires de la souris à la suite de l'insertion, dans le génome, du virus de la tumeur mammaire de la souris. Ce virus est capable de transformer, en culture, les cellules de l'épithélium mammaire mais non celles d'autres lignées. L'ADN du gène *int-1* a été cloné, sa séquence établie. De la séquence on peut prévoir que la protéine déterminée par ce gène est sécrétée [1]. Chez l'embryon, le gène semble être exprimé, entre le 9^e et le 15^e jour, dans le cerveau et dans la moelle épinière [2]. Chez l'adulte mâle, on le

trouve exprimé dans les spermatides exclusivement [3]. Curieusement, le gène ne semble pas exprimé dans le tissu mammaire ni aucun autre tissu de la femelle.

Le gène *wingless* fut à l'origine repéré, dans la drosophile, par une mutation qui transformait l'aile en une réplique du notum, autre composant du thorax. Récemment, d'autres mutations du même gène ont été isolées. Ce sont des mutations récessives létales pour l'embryon homozygote, du type « polarité segmentaire ». Chez les embryons mutants, on trouve le nombre normal de segments, mais dans chaque segment manquent certaines des structures cuticulaires. Le gène *wingless* a été cloné et son expression dans l'embryon précisée. Il est transcrit le long d'une bande de trois ou quatre cellules dans chaque segment ; soit quelque seize bandes le long de l'embryon [4]. Point important : chez les animaux chimères, formés de groupes de cellules de génotype sauvage et de groupes de cellules de génotype mutant, les ailes sont normales. En d'autres termes, le phénotype des cellules mutantes peut redevenir normal au voisinage de cellules de type sauvage. Le produit du gène *wingless* semble donc être diffusible d'une cellule à ses voisines.

Ces deux histoires séparées, celle de l'oncogène *int-1* de la souris et celle du gène *wingless* de la drosophile, se sont récemment rencontrées. En utilisant une sonde de souris, on a recherché et on a trouvé, dans le génome de drosophile, une séquence homologue de *int-1*. Par hybridation sur les chromosomes géants de glande salivaire, cette séquence de drosophile a été localisée à la même position cytologique que le gène *wingless* [5]. Les séquences des deux ADNc de drosophile, celle homologue de *int-1* et celle de *wingless*, sont les mêmes. Et l'injection d'ARN « antisens » de la séquence *int-1* dans des œufs de type sauvage produit des embryons qui présentent un phénotype *wingless* [6].

Si le rôle du proto-oncogène *int-1* dans la morphogenèse est ainsi bien établi chez la drosophile, il reste flou dans le cas de la souris. Mais, récemment, l'étude du xénope a montré que, chez les vertébrés aussi, ce gène intervenait dans le dévelop-

pement. En utilisant comme sonde l'ADNc de souris, on a pu repérer et cloner le gène *int-1* de xénope. Si l'on injecte, dans des œufs fécondés, l'ARN messager de ce gène, celui-ci est traduit et la protéine distribuée dans l'ensemble de l'œuf. La quasi-totalité des embryons survivent et se développent en gastrulas apparemment normales. Mais au stade morula, on trouve que la plaque neurale bifurque dans la région antérieure où elle est flanquée de deux notochordes, tandis qu'elle s'épaissit dans la région postérieure. L'effet est aboli si l'on enlève de la protéine la séquence signal permettant la sécrétion [7]. Il semble ainsi que l'expression ectopique de *int-1* conduit à la formation de deux axes antéro-postérieurs. Ce qui montre bien le rôle de ce gène dans les processus de morphogenèse chez les vertébrés.

Comment rendre compte des effets divers qu'exercent ces molécules semblables dans des organismes divers ? Bien évidemment c'est en terme de biologie moléculaire qu'une explication pourra être trouvée. Selon toute vraisemblance, le gène *int-1* détermine la synthèse d'une molécule qui agit comme facteur de croissance sur certains types cellulaires à certaines étapes de la différenciation des organismes. Et les cas de ce genre s'accumulent. Le proto-oncogène *c-rel* a une longue séquence en commun avec le gène de drosophile *dorsal*, gène à expression maternelle, impliqué dans l'établissement de la polarité dorso-ventrale [8]. Le proto-oncogène *c-raf* est identique au gène de drosophile *pole-hole*, gène de segmentation de la classe dite « terminale » à expression maternelle et zygotique [9]. Le proto-oncogène *c-kit* est, chez la souris, le produit du gène dominant *W* (*white spotting*), jouant un rôle dans la migration des cellules de la lignée érythroïde, des cellules germinales et des mélanoblastes issus de la crête neurale [10]. Enfin, récemment, on a montré que le produit du proto-oncogène *c-mos* est identique au facteur cytotatique qui bloque en métaphase la méiose des œufs de xénope [11]. Il est remarquable de voir l'étude d'un proto-oncogène, au rôle souvent très difficile à définir chez les mammifères, s'éclairer soudain par des résultats obtenus dans l'étude du

développement embryonnaire chez des insectes ou des batraciens.

Homéotiques chez les vertébrés et les insectes

Autre exemple de convergence, plus surprenant peut-être encore puisqu'il s'agit de gènes qui interviennent pour établir le plan du corps, chez les insectes et chez les vertébrés. L'histoire a commencé dans les années 1930, quand un étrange mutant de drosophile apparut qui, au lieu d'une paire d'ailes, en portait deux. Cette mutation fut appelée *bithorax*. C'était la première d'une série dont l'étude connaît, depuis quelques années, un développement explosif. La morphogenèse de la drosophile est dirigée par l'action successive de groupes de gènes. Les premiers établissent les axes du corps, antéro-postérieur et dorso-ventral. Les suivants divisent le corps en segments, c'est-à-dire en modules répétés le long de l'axe antéro-postérieur. Les troisièmes définissent la localisation et précisent les caractéristiques de chaque segment. Ces derniers gènes, appelés « homéotiques » (*HOM*), sont réunis dans le génome en deux groupes : le complexe *bithorax* et le complexe *antennapedia*. Ce dernier ainsi nommé parce que la première mutation trouvée dans ce complexe remplaçait, sur la tête de la mouche, une antenne par une patte. Le complexe *antennapedia* contient les gènes qui déterminent les caractéristiques des segments antérieurs alors que le complexe *bithorax* est chargé des segments postérieurs. Si, par exemple, une délétion enlève l'ensemble du complexe *bithorax*, tous les segments postérieurs au second segment thoracique deviennent des copies du second segment thoracique [12].

Il apparaît aujourd'hui que les gènes homéotiques gouvernent la synthèse de facteurs de transcription qui règlent l'activité de groupes de gènes de structure. C'est ainsi qu'ils déterminent le destin des cellules dans les segments. La manière dont ils sont mis en action semble très compliquée et reste encore mal comprise. Un complexe comme *bithorax*, qui définit les particularités de chacun des neuf segments postérieurs, semble ne

RÉFÉRENCES

- Rijsewijk F, Schuermann M, Wagenaar E, Parren P, Weigel D, Nusse R. The *Drosophila* homolog of the mouse mammary oncogene *int-1* is identical to the segment polarity gene *wingless*. *Cell* 1987 ; 50 : 649-57.
 - Cabrera CV, Alonso MC, Johnston P, Phillips RG, Lawrence PA. Phenocopies induced with antisense RNA identify the *wingless* gene. *Cell* 1987 ; 50 : 659-63.
 - McMahon AP, Moon RT. Ectopic expression of the proto-oncogene *int-1* in *Xenopus* embryos leads to duplication of the embryonic axis. *Cell* 1989 ; 58 : 1075-84.
 - Steward R. *Dorsal*, an embryonic polarity gene in *Drosophila* is homologous to the vertebrates proto-oncogene, *c-rel*. *Science* 1987 ; 238 : 692-4.
 - Ambrosio L, Mahowald AP, Perrimon N. Requirement of the *Drosophila raf* homolog for torso function. *Nature* 1989 ; 342 : 288-90.
 - Geissler EN, Ryan MA, Housman DE. The dominant-white spotting (*W*) locus of the mouse encodes the *c-kit* proto-oncogene. *Cell* 1988 ; 55 : 185-92.
 - Sagata N, Watanabe N, Vande Woude GF, Ikawa Y. The *c-mos* proto-oncogene product is a cytosolic factor responsible for meiotic arrest in vertebrates eggs. *Nature* 1989 ; 342 : 512-8.
 - Lewis EB. A gene complex controlling segmentation in *Drosophila*. *Nature* 1978 ; 276 : 565-70.
 - Akam M. The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* 1987 ; 101 : 1-22.
 - Gehring WJ. Homeo boxes in the study of development. *Science* 1987 ; 236 : 1245-52.
 - Akam M. Hox and *HOM* : homologous gene clusters in insects and vertebrates. *Cell* 1989 ; 57 : 347-9.
 - Graham A, Papalopulu N, Krumlauf R. The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* 1989 ; 57 : 367-78.
 - Duboule D, Dollé P. The structural and functional organization of the murine *Hox* gene family resembles that of *Drosophila* homeotic genes. *EMBO J* 1989 ; 8 : 1497-505.
 - Hogan B, Holland P, Schofield P. How is the mouse segmented? *Trends Genet* 1985 ; 1 : 67-74.
 - Lumsden A, Keynes R. Segmental patterns of neuronal development in the chick hindbrain. *Nature* 1989 ; 337 : 424-8.
 - Murphy P, Davidson DR, Hill RE. Segment-specific expression of a homeobox-containing gene in the mouse hindbrain. *Nature* 1989 ; 341 : 156-9.
- contenir que trois gènes, régissant la structure de trois protéines [13]. Ces protéines sont exprimées en quantités variées et en combinaisons variées avec d'autres protéines pour donner aux cellules de chaque segment leur spécificité. Chacun des gènes de structure est entouré d'un jeu complexe d'activateurs (*enhancers*) qui en modulent l'expression selon la position des segments.
- L'analyse des gènes *HOM* a mis en évidence deux régularités remarquables. La première est la présence, dans tous les gènes homéotiques, d'une séquence très semblable de 180 paires de bases [14]. Cette séquence, appelée « homéoboîte », définit un domaine protéique qui a une haute affinité pour l'ADN. Selon toute vraisemblance, les gènes *HOM* viennent, par duplication et divergence, d'un gène ancestral commun, possédant une homéoboîte et fonctionnant comme régulateur d'autres gènes.
- Seconde particularité : un surprenant parallélisme entre la position qu'occupent, sur le chromosome, les éléments génétiques d'un complexe *HOM* et la région du corps où ces gènes exercent leur effet, le long de l'axe antéro-postérieur de l'animal. Les mutations à différents sites d'un complexe tel que *bithorax*, modifient différentes régions du corps de l'animal. Et l'ordre dans lequel ces mutations peuvent être alignées le long du chromosome est le même que celui dans lequel les régions modifiées sont disposées le long du corps [12]. Autrement dit, l'organisation des segments le long de l'axe antéro-postérieur de l'animal reflète l'organisation des éléments génétiques du complexe le long du chromosome (*figure 1*). L'activation de ces complexes *HOM* a lieu au début de la vie embryonnaire quand se forme le blastocèle. Pour chaque cellule, c'est la position de cette cellule dans le blastocèle qui entraîne l'activité d'une combinaison particulière de gènes *HOM*. Une fois que les cellules ont reçu les instructions définissant leur position sur l'axe antéro-postérieur, et donc leur avenir dans le développement, elles en conservent la mémoire de manière permanente [13].
- Depuis quelques années, on peut parler de complexes *HOM* [13], non plus seulement chez les insectes, mais

aussi chez les mammifères. Cela grâce à l'existence de la boîte homéotique [14]. La remarquable conservation de structures observée à travers le monde vivant permet en effet d'utiliser les séquences homéoboîtes d'insectes pour « pêcher », en quelque sorte, par hybridation des séquences homologues dans le génome d'autres espèces. Chez les mammifères, on peut ainsi isoler de nombreux gènes contenant une homéoboîte. Beaucoup de ces gènes sont disposés en petits groupes formant chacun un complexe *HOM*. Chez la souris et chez l'homme, on connaît actuellement quatre complexes désignés de *Hox-1* à *Hox-4*. Et comme chez la drosophile, les différents gènes du complexe *HOM* le mieux analysé chez la souris sont exprimés dans différentes régions du corps, le long de l'axe antéro-postérieur [15]. Là encore, au moins dans la deuxième moitié de la gestation, on observe une corrélation entre la position d'un gène dans le complexe sur le chromosome et la région du corps dans laquelle ce gène est exprimé [16, 17] (*figure 1*). Plus un gène se trouve placé loin du début du complexe sur le chromosome, plus son domaine d'expression se situe vers l'arrière de l'animal. En outre, entre certains gènes homéotiques de mammifère et de drosophile, on relève des analogies de séquence qui ne se limitent pas à la seule homéoboîte.

De telles similitudes d'organisation et de séquence ont des implications bien inattendues. La première a trait à l'évolution : il faut bien admettre que des gènes à homéoboîte existaient avant la séparation des chordés et des arthropodes. Un ancêtre commun aux insectes et aux mammifères devait déjà posséder une sorte de complexe *HOM*, avec les équivalents de certains gènes trouvés chez la drosophile et permettant de définir les régions successives du corps le long de l'axe antéro-postérieur.

Autre implication, plus troublante encore : à l'homologie des éléments génétiques des complexes *HOM* trouvée entre insectes et mammifères doit bien répondre quelque homologie des domaines d'expression dans les régions du corps. Ce qui pose problème car, à première vue, le corps

d'un insecte et celui d'un mammifère semblent bâtis selon des principes totalement différents. Le corps de l'insecte est fait de segments, c'est-à-dire de modules répétés, initialement identiques. La fonction des gènes homéotiques paraît être d'y établir les différences, de préciser la structure propre à chaque segment. Et les frontières limitant l'expression de ces gènes sont calquées sur les frontières des segments. Quel peut donc être alors l'équivalent d'un segment d'insecte chez un mammifère ? C'est évidemment à la division en somites qu'on a d'abord pensé. Mais la segmentation du mésoblaste et la formation des somites se font d'une manière très différente de la segmentation des insectes [18]. En outre, l'expression des gènes homéotiques est limitée par des frontières qui ne correspondent nullement à celles des somites.

Récemment, une autre hypothèse a été proposée, beaucoup plus probable celle-là. Pendant le développement embryonnaire des vertébrés,

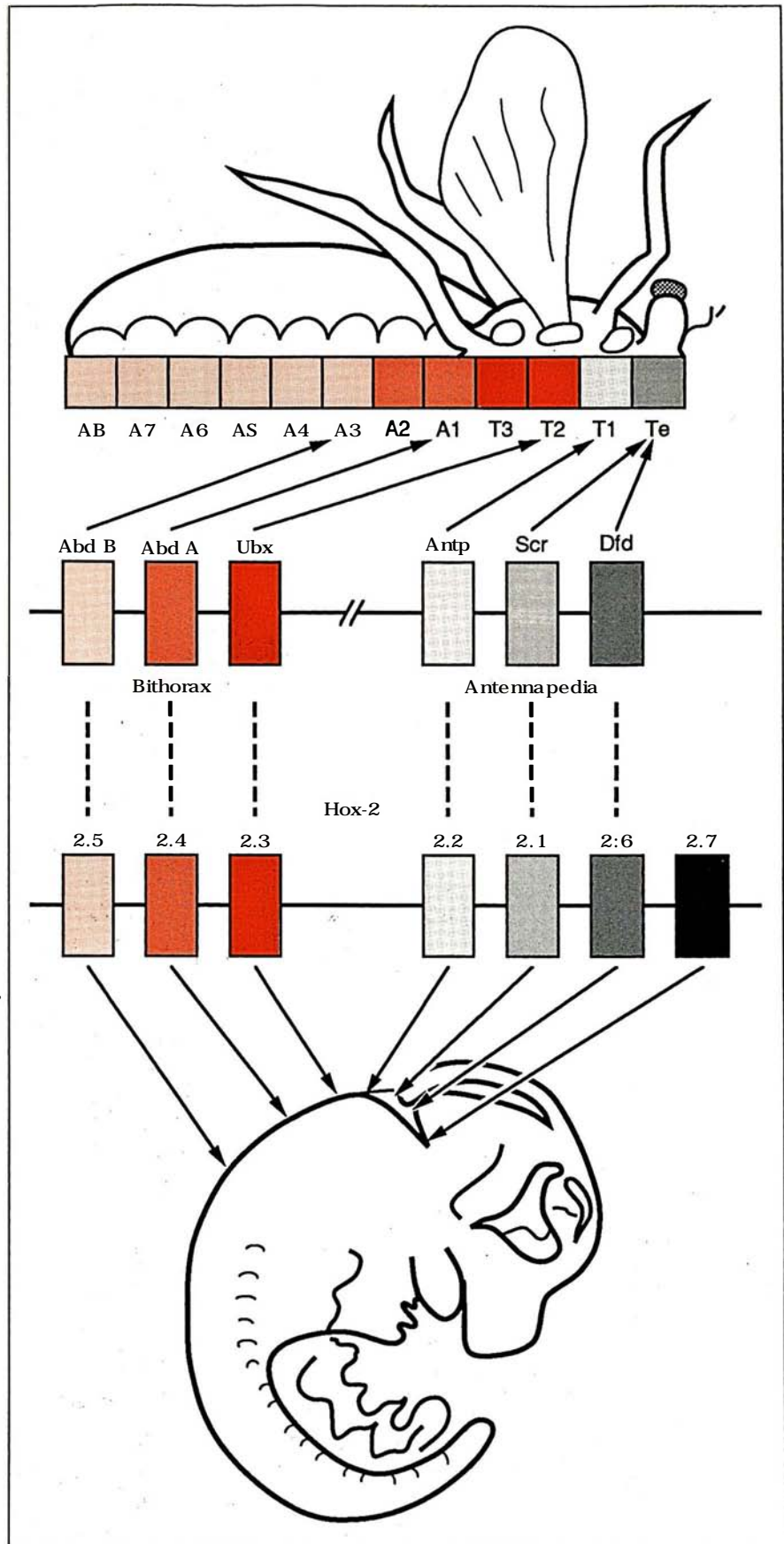


Figure 1. **Représentation schématique des relations entre groupes de gènes homéotiques et leurs domaines d'expression chez la drosophile et la souris.** Au milieu, la comparaison des séquences de drosophile et de souris permet d'aligner verticalement les groupes de gènes qui se correspondent : complexes bithorax et antennapedia chez la drosophile, complexe Hox-2 chez la souris. De part et d'autre sont schématisés : en haut, le corps segmenté d'une larve de drosophile ainsi qu'une demi-mouche adulte ; en bas, un embryon de souris au jour 12. Les gènes homéotiques sont représentés par des rectangles : Abd = abdomen ; Ubx = ultrabithorax ; Antp = antennapedia ; Dfd = deformed ; Scr = sex comb reduced. Les flèches représentent, dans les deux cas, les limites antérieures des domaines d'action des gènes homéotiques correspondants. (D'après Graham et al. [16].)



A



B



C



D

beaucoup de gènes *HOM* sont exprimés dans le système nerveux central, moelle épinière et cerveau postérieur notamment (figure 2). A certains stades du développement, le cerveau postérieur comprend une série de sept ou huit renflements, appelés rhombomères ou neuromères, séparés par des constrictions. On a montré, chez le poulet, que renflements et constrictions reflètent une organisation répétitive du tissu nerveux le long du tube neural. Cette segmentation correspond à la série de racines des nerfs moteurs crâniens et des ganglions sensoriels crâniens, donc à la série des arcs branchiaux qu'ils innervent [19]. A cette organisation anatomique semble bien répondre une organisation embryologique. Il apparaît, en effet, que toute une série de gènes constituant un complexe *HOM* est exprimée le long du cerveau postérieur et de la moelle épinière. Comme chez la drosophile, on trouve une corrélation entre position d'un gène sur le chromosome et domaine d'expression le long du système nerveux : plus un gène est placé loin du début du complexe, plus la région où il est exprimé se situe en arrière. En outre, le domaine d'expression de chaque gène est limité en avant par une frontière qui coïncide exactement avec la frontière d'un rhombomère [20]. Mieux, quand on passe d'un gène au gène adjacent sur le chromosome, la frontière antérieure du domaine où s'exprime ce gène saute en arrière de deux rhombomères ; comme chez la drosophile où les frontières qui limitent en avant les régions exprimant des gènes *HOM* adjacents sont, bien souvent, séparés par un intervalle de deux segments.

Dans les années 1930, la biochimie avait montré l'unité de composition qui règne dans le monde vivant. Dans les années 1950-1960, la biologie moléculaire en a montré l'unité de structure et de fonctionnement. Aujourd'hui, cette même biologie moléculaire est en train de mettre en évidence l'unité de certains des principes qui sous-tendent le développement des organismes multicellulaires. Qui eût dit, il y a seulement dix ans, que la construction d'un être humain pourrait, par certains aspects, ressembler à celle d'une mouche ! ■

Summary

The unity of life

In every field of science, boring and exciting periods alternate. During the years 70, molecular biology was suddenly provided with new, very powerful tools for the analysis of nucleic acids. All laboratories in the world started to isolate genes, to sequence them and find out their regulation. Hence, to begin with, a series of rather tedious results. After a few years, however, the situation has changed. There is today a burst of new results which, by many aspects, drastically change current ideas. Especially in the fields of evolution, of embryonic development, of the cell cycle and its aberrations represented by oncogenes. And all the results appear to converge. So that, in order to analyze certain phenomena in mammals, in which genetics is so difficult, it becomes often possible to use data obtained in *Drosophila*, in which genetics is so advanced. Examples are given.

◀ **Figure 2. Embryons de souris chimeres exprimant la β -galactosidase d'*Escherichia coli* sous commande d'un promoteur Hox 3.1.** Le gène de structure de la β -galactosidase d'*E. coli* a été introduit, par recombinaison homologue (voir *m/s*, article de V. Lemarchandel et X. Montagutelli, n° 1, vol. 6, p. 18), à la place du gène *Hox-3.1* dans l'un des chromosomes de cellules embryonnaires ES. L'expression de la β -galactosidase est alors commandée par le promoteur *Hox-3.1*. Chacune de ces cellules ES transformées a alors été injectée dans un blastocyste. Après réimplantation, les embryons chimériques formés ont été examinés et la production de β -galactosidase révélée par le 4CI-5Br-indolyl- β -galactoside (X-Gal) qui donne une coloration bleue. Les photos représentent de tels embryons chimériques de 10,5 jours après fécondation, c'est-à-dire avant localisation précise de l'expression du gène en fonction de sa localisation dans le groupe des gènes *Hox-3*. **A et B** : deux embryons avec une activité β -galactosidase (en bleu) au niveau du tube neural dans la moitié postérieure. **C** : région caudale d'un autre embryon chimère. **D** : région dorsale postérieure d'un autre embryon. Les blastocystes (3,5 jours après fécondation) ayant reçu une cellule ES transformée ont été introduits dans des femelles pseudo-gravides à un temps considéré comme étant 2,5 jours après fécondation pour la mère porteuse. Les embryons ont été récoltés après implantation et l'activité β -galactosidase mise en évidence par la réaction X-gal. (Expériences inédites d'Hervé Le Mouellic, Yvan Lallemand et Philippe Brûlet.)

m/s n° 3 vol. 6, mars 90

TIRÉS A PART

F. Jacob.