

Les gènes homéotiques

JA Lepasant

Institut Jacques-Monod, Centre National de la Recherche Scientifique et Université Paris VII,
2, pl Jussieu, 75251 Paris Cedex 05, France

(28^e Réunion de la Société Française pour l'Étude de la Fertilité; Paris, 19–21 octobre 1989)

Résumé — Les gènes homéotiques occupent une position clé dans la hiérarchie des gènes impliqués dans le déroulement des premières étapes du développement embryonnaire chez la drosophile, car ils spécifient l'identité individuelle des segments qui composent le corps de la mouche. La majorité des gènes homéotiques de la drosophile sont regroupés dans deux grands complexes géniques distincts, le complexe *Antennapedia* (ANT-C) et le complexe *Bithorax* (BX-C). Ces gènes sont exprimés dans deux domaines spatiaux partiellement chevauchants, dont la position selon l'axe antéro-postérieur de l'organisme est colinéaire avec leur position respective à l'intérieur de chaque complexe. Cinq gènes de l'ANT-C sont impliqués dans le déterminisme de l'identité de segments composant la tête, et dans celle des premier et deuxième segments thoraciques. Les trois gènes du BX-C contrôlent les identités de la partie postérieure du deuxième segment thoracique, du troisième segment thoracique et des huit segments abdominaux. Les études moléculaires ont révélé que les protéines codées par les gènes homéotiques possèdent en commun un motif structural de 60 acides aminés, le «motif homéotique» ou *homeobox*. La structure en hélice-coude-hélice de ce motif permet aux protéines homéotiques de se fixer au niveau de séquences d'ADN spécifiques dans les régions régulatrices en *cis* des gènes cibles dont ils modulent l'expression en qualité de facteurs de transcription.

Le haut degré de conservation phylogénétique du motif homéotique a rendu possible l'identification et l'isolement de plus de cinquante gènes *Hox* codant pour des protéines porteuses de ce motif chez un grand nombre d'autres organismes eucaryotes dont le nématode, le xénope, la souris et l'homme. Chez la souris et l'homme les gènes *Hox* sont regroupés en tandem au sein de quatre grands complexes géniques situés sur des chromosomes différents. Une colinéarité a été mise en évidence entre la position relative des gènes *Hox* dans ces complexes et la frontière antérieure de leur domaine spatial d'expression le long de l'axe antéro-postérieur du système nerveux central et de la colonne prévertébrale de l'embryon de souris. Les similitudes entre la structure, l'organisation et le patron d'expression des gènes homéotiques de la drosophile et ceux des gènes à boîte homéotique des vertébrés suggèrent qu'un certain nombre de principes fondamentaux d'organisation et de développement de l'embryon ont été conservés au cours de la longue période évolutive écoulée depuis la divergence des lignées phylogénétiques ayant conduit aux vertébrés et aux arthropodes.

homéose / homeobox / drosophile / gènes Hox / développement embryonnaire

Summary — **Homeotic genes.** Homeotic genes occupy a central position in the hierarchy of genes controlling the early steps of embryonic development in *Drosophila*. They are involved in the specification of the individual identity of each segment of the insect's body. Most homeotic genes of *Drosophila* are located in two large gene clusters, the *Antennapedia* complex (ANT-C) and the *Bithorax* complex (BX-C). These genes are expressed in two partially overlapping domains whose positions along the anterior-posterior axis of the organism are colinear with their position within each complex. Five genes in the ANT-C are involved in the specification of the identity of some segments of the head and of that of the first and second thoracic segments. The three genes of the BX-C determine the identities of the posterior compartment of the second and third thoracic segments and of

the eight abdominal segments. Molecular studies have revealed that the proteins coded by the homeotic genes share a 60 amino acid motif, the homeobox, whose helix–turn–helix structure enables them to bind as transcription factors to specific DNA sequences in the cis-acting regulatory regions of their target genes.

The high degree of phylogenetic conservation of the homeobox has made possible the identification and the isolation of more than fifty homeobox genes (Hox genes) in many eukaryotic organisms including nematode, *Xenopus*, mouse and man. In the mouse and human genomes Hox genes are clustered in tandem in four large gene complexes located on separate chromosomes. The relative positions of the genes within the clusters and the positions of the anterior boundaries of their expression domains along the anterior–posterior axis of the central nervous system and prevertebral column of the mouse embryo were found to be colinear. The similarities between the structure, organization and pattern of expression of the homeotic genes of *Drosophila* and of the homeobox genes of vertebrates suggest that some basic principles of embryonic organization and development have been preserved during the long period of evolution that has elapsed since the divergence of the phylogenetic lineages leading to vertebrates and arthropods.

homeosis / homeobox / *Drosophila* / Hox genes / embryonic development

INTRODUCTION

Le terme d'homoéose, qui tend actuellement à être supplanté par celui équivalent d'homéose, a été proposé à la fin du siècle dernier par le biologiste anglais Bateson (1894) pour désigner le remplacement d'une structure différenciée par une structure homologue présente sur une autre partie du corps. Les transformations homéotiques, comme par exemple le remplacement d'un balancier par une aile chez la mouche, sont facilement identifiables au niveau des segments qui composent le corps des arthropodes (Ouweneel, 1976). Elles ont été surtout étudiées chez les insectes et tout particulièrement chez la drosophile en raison des possibilités d'analyse génétiques et moléculaires qu'offre cet organisme (Rubin, 1988). Grâce en particulier aux travaux précurseurs du généticien américain Lewis (1978), il a pu être montré que les gènes dont la mutation entraînent ces transformations contrôlent l'identité des segments qui composent le corps de la mouche. Au cours des cinq dernières années, l'étude des gènes homéotiques s'est considérablement développée en raison principalement de l'utili-

sation des techniques du génie génétique qui a permis d'isoler les gènes homéotiques et d'étudier leur structure et leur expression au cours du développement. L'ensemble des résultats obtenus a confirmé que ces gènes occupent une position clé dans la hiérarchie des gènes contrôlant le développement. L'analyse de la séquence nucléotidique des gènes homéotiques a conduit à la découverte d'un court motif protéique, la «boîte homéotique» ou *homeobox*, présent dans tous ces gènes et responsable de la fixation des protéines homéotiques au niveau de séquences d'ADN spécifiques. En utilisant des sondes dérivées des gènes homéotiques de la drosophile, de nombreux gènes codant pour des protéines à boîte homéotique ont été identifiés dans le génome d'autres organismes incluant la souris et l'homme.

Le but de cette revue est de dégager les points saillants de ces avancées récentes qui introduisent un bouleversement certain dans les concepts concernant les mécanismes fondamentaux de l'ontogenèse et de la morphogénèse, leur origine et leur conservation au cours de l'évolution.

Il en saurait être possible de discuter ici en détail tous les aspects de l'étude des gènes homéotiques. Pour des informations détaillées et un complément de références bibliographiques, il est conseillé de se reporter aux excellentes mises au point publiées récemment sur les gènes homéotiques de la drosophile (Gehring et Hiromi, 1986; Scott et O'Farrell, 1986; Akam, 1987; Peifer *et al*, 1987; Akam *et al*, 1988; Ingham, 1988; Mahaffey et Kaufman, 1988), la boîte homéotique ou *homeobox* (Gehring, 1987; Scott *et al*, 1989) et les gènes à boîte homéotique des autres organismes (Holland et Hogan, 1988; Akam, 1989; De Robertis *et al*, 1989; Duboule et Dollé, 1989; Graham *et al*, 1989; Kappen *et al*, 1989; Wright *et al*, 1989).

RÔLE ET POSITION DES GÈNES HOMÉOTIQUES DANS LA HIÉRARCHIE DES GÈNES RÉGULANT LE DÉVELOPPEMENT

Au travers des études réalisées chez la drosophile, il ressort clairement que les gènes homéotiques sont impliqués dans les processus d'acquisition et de conservation de l'identité individuelle des différents segments qui composent l'embryon, la larve ou l'insecte adulte (Lewis, 1978; Akam *et al*, 1988; Mahaffey et Kaufman, 1988). Ils ne sont pas impliqués, en revanche, dans les processus qui conduisent à l'établissement du plan d'architecture segmentée de l'insecte. En effet la mutation d'un gène homéotique peut entraîner une transformation partielle ou totale du segment dont l'identité est contrôlée par ce gène, mais elle n'entraîne ni la délétion de ce segment ni l'altération du nombre et de la disposition des autres segments. Ce point peut être illustré par les mutations dominantes du gène *Antennapedia* de la drosophile où l'antenne est remplacée par

une patte normalement différenciée et où il est possible d'établir une correspondance directe entre les articles de l'antenne transformée et les articles homologues de la patte de remplacement (Mahaffey et Kaufman, 1988). De même, la délétion des gènes contrôlant l'identité du troisième segment thoracique et des segments abdominaux n'entraîne pas la délétion de ces segments mais leur transformation en segments ayant une identité de deuxième segment thoracique (Lewis, 1978).

La recherche systématique par Nüsslein-Volhard et Weischaus (1980) de mutations affectant les étapes initiales du développement embryonnaire de la drosophile a ouvert la voie à de très nombreux travaux génétiques et moléculaires qui ont permis d'obtenir une image cohérente de la nature et de la hiérarchie des gènes qui sont impliqués dans l'élaboration de l'architecture de cet organisme selon les axes antéro-postérieur et dorso-ventral (fig 1). Il est permis de penser que la majorité de ces gènes ont été identifiés, mais la compréhension de leur rôle et des interactions régulateurs complexes qui prennent place entre eux est probablement encore incomplète. Toutefois il est clairement établi que le rôle primaire de certains gènes maternels et des différentes classes de gènes de segmentation est de définir, au cours des étapes précédant la gastrulation, l'architecture métamérisée de l'embryon et en particulier le nombre de segments et leur polarité selon l'axe antéro-postérieur (Scott et O'Farrell, 1986, Akam, 1987; Ingham, 1988) (fig 1). A partir de la gastrulation, c'est la combinaison de l'expression d'un ou de plusieurs gènes homéotiques dans la limite des domaines spatiaux définis par les gènes de segmentation qui confèrera à chaque segment son identité particulière. Celle-ci peut être déterminée grâce à une série de marqueurs cuticulaires tels que

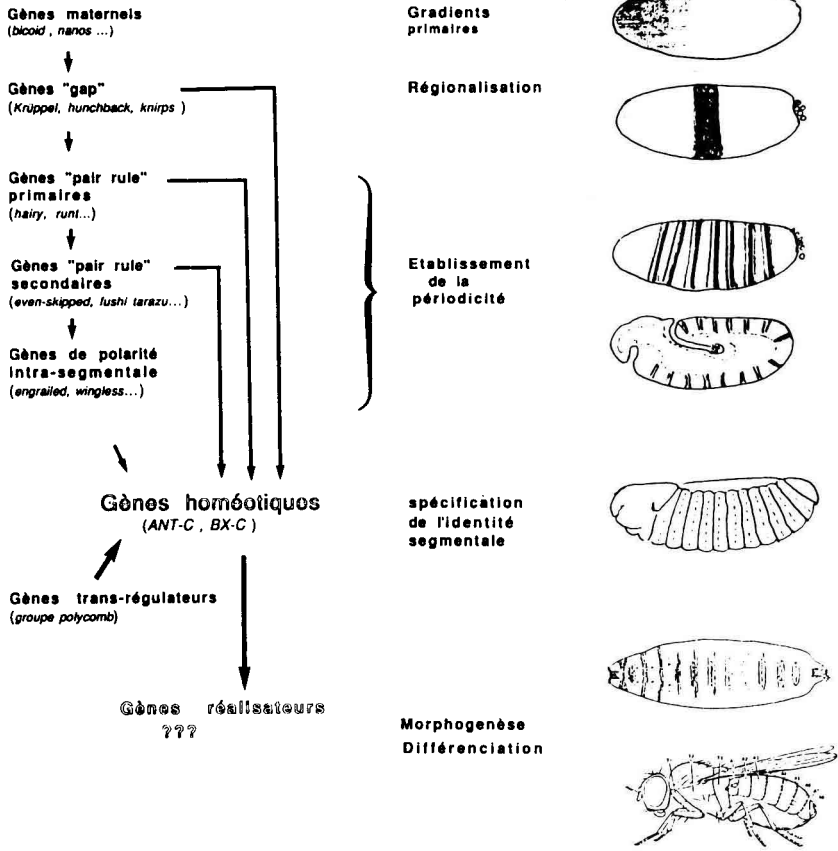


Fig 1. Hiérarchie des gènes impliqués dans le développement de la drosophile. Schéma établi d'après les revues de Akam (1987), Akam *et al* (1988) et Ingham (1988). Les flèches indiquent les interactions régulateurs entre les différentes classes de gènes.

les soies, dont la structure et l'arrangement sont propre à chaque segment (Lewis, 1978).

Situés en quelque sorte au bas de la hiérarchie des gènes de segmentation, les gènes homéotiques sont de fait au sommet d'une autre hiérarchie importante de gènes «réalisateurs» dont l'expression aboutira à la différenciation segmentale

correcte (fig 1). Cette position des gènes homéotiques justifie l'appellation de gènes «sélecteurs» donnée par Garcia-Bellido (1977) dans la mesure où au niveau de chaque segment ils orientent des groupes de cellules dans des voies de développement qui requièrent le choix de sous-programmes particuliers dans l'ensemble de l'information génétique du génome.

ORGANISATION DES GÈNES HOMÉOTIQUES DE LA DROSOPHILE

Les gènes homéotiques de la drosophile sont en majeure partie rassemblés dans deux grands complexes distincts, le com-

plexe *Antennapedia* (ANT-C) et le complexe *Bithorax* (BX-C) situés sur le bras droit du chromosome 3 (fig 2A) (Gehring et Hiromi, 1986; Duncan, 1987; Scott, 1987; Akam *et al*, 1988; Mahaffey et Kaufman, 1988). L'ANT-C localisé dans la région

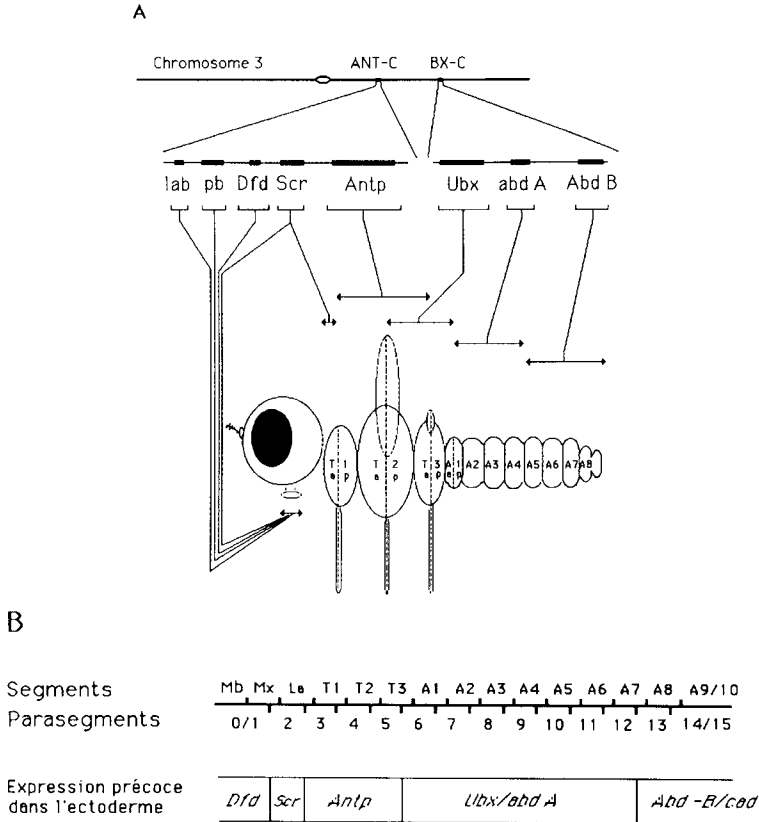


Fig 2. Organisation chromosomique et domaines d'activité des gènes homéotiques des complexes *Antennapedia* (ANT-C) et *Bithorax* (BX-C). **A.** La position des séquences codantes des gènes *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *sex combs reduced* (*Scr*), *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) et *Abdominal-B* (*Abd-B*) sont symbolisées par des rectangles noirs. Les segments dans lesquels l'activité des gènes homéotiques est requise sont délimités par des flèches horizontales. T1, T2, T3 : segments thoraciques. A1-8 : segments abdominaux; a et p indiquent les compartiments antérieurs et postérieurs des segments délimités par les lignes verticales en pointillé. **B.** Correspondance entre segments et parasegments et domaines d'expression des gènes homéotiques au cours du développement embryonnaire précoce. Mb, Mx, La indiquent les segments mandibulaire, maxillaire et labial respectivement, *cad* indique l'activité du gène *caudal* requis pour l'identité des structures postérieures (reproduit de Akam *et al* (1988) avec permission de l'éditeur).

84AB contient cinq gènes qui contrôlent l'identité d'une partie des segments composant la tête et les segments du thorax. Ce complexe inclut également un certain nombre de gènes tels que les gènes *fushi tarazu* (*ftz*) et *bicoid* (*bcd*), qui interviennent dans la mise en place du patron initial de segmentation ainsi que les gènes *zerknüllt* (*zen*) 1 et 2 impliqués dans le développement des tissus dorsaux de l'embryon (Scott, 1985). Le BX-C situé dans la région 89E comprend trois gènes impliqués dans l'établissement de l'identité du deuxième et du troisième segments thoracique ainsi que celle des huit segments abdominaux (Duncan, 1987; Mahaffey et Kaufman, 1988) (fig 2A).

Il est à noter que ces deux complexes ne regroupent pas tous les gènes homéotiques et n'incluent pas en particulier les gènes *trans*-régulateurs de l'activité des gènes des complexes ANT-C et BX-C, qui peuvent eux-mêmes être assimilés à des gènes homéotiques.

DOMAINES D'ACTION DES GÈNES HOMÉOTIQUES

Les gènes homéotiques des deux complexes, ANT-C et BX-C, ont été initialement identifiés et localisés génétiquement grâce à l'isolement d'un grand nombre de mutations. L'analyse de l'effet phénotypique de ces mutations a permis de définir les limites spatiales du domaine dans lequel l'activité du gène considéré est requise pour une différenciation normale et une identité des tissus correspondant à leur situation segmentale (Lewis, 1978). Deux grandes classes de mutations ont été examinées, des mutations «perte de fonction» récessives conduisant à une réduction plus ou moins importante ou à une élimination totale du produit du gène et des mutations «gain de fonction» domi-

nantes dont on peut considérer maintenant qu'elles correspondent à une expression ectopique du gène homéotique affecté. Ces études ont été facilitées par la présence des marqueurs cuticulaires caractéristiques de chaque segment aussi bien au niveau de la larve que de l'adulte. L'interprétation des résultats a pu être limitée par le fait que les transformations peuvent être partielles ou incomplètes en particulier dans le cas d'allèles «faibles» correspondant vraisemblablement à des mutations n'abolissant pas complètement l'activité ou la synthèse du produit du gène étudié. Une seconde limitation provient du fait de la létalité de certaines mutations qui empêche d'obtenir des adultes. Cette difficulté peut être tournée par l'observation des transformations homéotiques au niveau de la cuticule des embryons dans la mesure où les mêmes principes de base régissent l'architecture et la détermination de l'identité des segments de la larve et de l'adulte (Lewis, 1978; Duncan, 1987, Mahaffey et Kaufman, 1988).

La seconde approche de la définition des domaines d'action des gènes homéotiques repose sur la localisation des produits de l'expression des gènes, soit au niveau ARN par hybridation *in situ* sur des coupes à l'aide de sondes nucléiques marquées de manière appropriée, soit au niveau protéine à l'aide d'anticorps spécifiques dirigés contre les produits d'expression des ADNc clonés dans *E coli* (Akam, 1987).

Des nombreuses études réalisées jusqu'à présent il est permis de conclure à une très bonne concordance entre les domaines d'action des gènes homéotiques déduits de l'analyse des phénotypes embryonnaires larvaires ou adultes des mutations et ceux déduits de l'analyse des domaines d'expression de ces gènes au niveau ARN et au niveau protéine. L'interprétation des résultats a été rendue plus

simple depuis la mise en évidence par Martinez-Arias et Lawrence (1985) du fait que la subdivision métamérique initiale de l'embryon, visualisée en particulier par l'expression des gènes de polarité segmentale tel qu'*engrailed*, est parasegmentale et décalée d'un demi-segment par rapport aux segments définis sur la base des critères morphologiques (fig 2B). Un segment morphologique donné est en fait constitué de la réunion de la partie postérieure d'un parasegment et de la partie antérieure du parasegment suivant. Les limites spatiales antérieures et postérieures des domaines d'expression des gènes homéotiques qui sont observées par exemple dans le cas des gènes *Antp* et *Ubx* reflètent clairement cette origine parasegmentale mixte des segments (fig 2) (Akam, 1987; Mahaffey et Kaufman, 1988; Scott, 1988; Irish *et al*, 1989).

Domaines d'action des gènes de l'ANT-C

Les quatre premiers gènes de l'ANT-C sont requis pour l'identité des structures de la tête, mais ils ne représentent pas les seuls gènes impliqués dans cette morphogénèse complexe dont le détail demeure encore imprécis. Une difficulté provient en particulier de la fusion des segments mandibulaire, maxillaire et labial avec les autres structures constitutives de la tête.

Labial (lab), le gène le plus proximal du complexe (Mahaffey et Kaufman, 1988; Mlodzik *et al*, 1988; Diederich *et al*, 1989) est également exprimé le plus antérieurement dans l'embryon et son fonctionnement apparaît requis pour une différenciation d'au moins quatre segments distincts contribuant à la formation de la tête.

Le gène *proboscipedia (pb)*, dont la structure moléculaire a été analysée récemment (Pultz *et al*, 1988), détermine

l'identité segmentale des structures labiales et maxillaires de l'insecte adulte. L'inactivation complète du gène conduit à des transformations des pièces buccales en pattes et une altération de la morphologie des palpes maxillaires. En revanche ce gène ne semble pas avoir de rôle particulier au cours du développement embryonnaire et de la vie larvaire (Mahaffey et Kaufman, 1988).

Le gène *deformed (Dfd)* est requis pour une différenciation correcte de certaines structures dérivant des lobes mandibulaire et maxillaire ainsi que celle de certains éléments cuticulaires du squelette de la tête (Regulski *et al*, 1987; Mahaffey et Kaufman, 1988).

Le gène *sex combs reduced (Scr)* (Mahaffey et Kaufman, 1988; Le Motte *et al*, 1989) contrôle l'identité des structures de la tête dérivant du lobe labial et de l'identité du compartiment antérieur du premier segment thoracique. La fonction de *Scr* est certainement complexe, car elle s'exerce à la transition entre la tête et la région thoracique. Ceci pourrait expliquer que contrairement aux autres gènes homéotiques, la perte de fonction du gène *Scr* a des effets dans deux directions opposées selon l'axe antéro-postérieur. Au niveau du lobe labial, elle conduit à une transformation en structures du type de celles dérivant du lobe maxillaire plus antérieur. En revanche, au niveau du premier segment thoracique, elle conduit à une nouvelle identité correspondant à celle du deuxième segment thoracique avec, en particulier, comme l'indique le nom du gène, une diminution du nombre des soies du peigne sexuel, qui est une structure caractéristique de la première patte chez le mâle adulte.

Le gène *Antennapedia (Antp)*, premier gène identifié du complexe auquel il a donné son nom, est requis pour l'identité des segments thoraciques. Il a fait l'objet d'une étude très détaillée de sa structure

complexe et de son patron d'expression au cours du développement (Gehring et Hiromi, 1986; Schneuwly *et al*, 1986; Birmingham et Scott, 1988; Mahaffey et Kaufman, 1988). Contrairement à ce que pourrait conduire à penser le phénotype des mutations dominantes, ce gène n'est pas exprimé dans les segments de tête, mais dans les segments T1 postérieur, T2 et T3 antérieur, qui correspondent aux parasegments PS4 et PS5. Dans ces parasegments, son rôle normal serait de réprimer le développement de structures de tête et de structures T1 antérieures. Le phénotype de remplacement de l'antenne par une patte observé dans le cas des mutations dominantes est explicable par une expression spatiale inappropriée du gène *Antp* se trouvant placé dans ce cas sous le contrôle d'un promoteur dont le fonctionnement est non restreint spatialement. Ceci a été confirmé récemment par Schneuwly *et al* (1987) qui ont placé le gène *Antp* sous le contrôle du promoteur du gène *hsp 70* de *D melanogaster*. Pour une raison qui demeure inexplicée, les cellules du disque imaginal de l'antenne sont particulièrement sensibles à l'expression ectopique du gène *Antp*.

Domaines d'action des gènes du BX-C

L'analyse génétique initiale reposant presque exclusivement sur des mutations dues à des réarrangements chromosomiques plus ou moins complexes, des insertions d'éléments transposables ou des petites délétions, avait donné à penser que le complexe *Bithorax* était composé d'au moins huit gènes homéotiques (Lewis, 1978). L'isolement de mutations ponctuelles a permis de montrer l'existence de seulement trois gènes capables de coder pour des protéines (Mahaffey et Kaufman, 1988).

Le gène *Ultrabithorax (Ubx)* est requis pour la détermination de l'identité des tissus s'étendant des segments T2 postérieur à A1 antérieur et correspondant aux parasegments PS5 et PS6. L'absence de tout produit du gène *Ubx* est létale à la fin de l'embryogenèse et entraîne une transformation complète des parasegments PS5 et PS6 en parasegment PS4, soit une identité de segment T1 postérieur et T2 antérieur (Lewis, 1978). Toutefois, quatre classes de mutations conduisant à des transformations partielles affectant un seul parasegment comme *anterobithorax (abx)* (PS5) et *bithoraxoid (bxd)* (PS6) ou un demi-parasegment comme *bithorax (bx)* (PS5) et *postbithorax (pbx)* (PS6) ont des effets visibles surtout chez l'adulte et conduisent à des transformations en structures plus antérieures à celles des segments affectés (Lewis, 1978; Duncan, 1987). Ces mutations qui correspondent à des réarrangements chromosomiques plus ou moins importants perturbent probablement des régions de régulation en *cis* de l'activité du gène *Ubx* dans les différents segments. Ceci est en particulier le cas des mutations *bxd* et *pbx* qui affectent la région transcrite *bxd* située immédiatement en amont du gène *Ubx* et dont l'altération entraîne une réduction de l'expression de ce gène (Peifer *et al*, 1987).

Le gène *abdominal-A (abd-A)* détermine l'identité des segments abdominaux A1 postérieur à A4 inclus et le gène *abdominal-B (Abd-B)* est requis pour l'identité des autres segments abdominaux. Les mutations perte de fonction de *abd-A* conduisent à une transformation de type parasegmentale des segments abdominaux A2 à A4 en segments ayant une identité de T3 postérieur et d'abdominal A1 antérieur. L'inactivation du gène *Abd-B* transforme les segments A5 à A8 en segments ayant une identité de segment abdominal A4 (Peifer *et al*, 1987; Mahaffey et Kaufman, 1988; Celniker *et al*, 1989).

De manière remarquable on peut noter une correspondance entre l'ordre des gènes au sein des deux complexes et leur ordre d'expression segmentale selon l'axe antéro-postérieur de l'insecte (fig 2A). La signification fonctionnelle, si elle existe, de cet arrangement demeure inconnue, mais cette disposition pourrait refléter, comme il sera discuté ci-dessous, l'organisation initiale d'un unique complexe homéotique (HOM-C) ancestral dont on retrouve la trace chez le coléoptère *Tribolium cataneum* (Beeman, 1987) considéré comme un insecte plus primitif que la drosophile (voir Akam *et al*, 1988, pour une hypothèse intéressante sur l'évolution de la lignée myriapodes-insectes en relation avec l'origine des gènes homéotiques).

STRUCTURE MOLÉCULAIRE DES GÈNES HOMÉOTIQUES ET DE LEURS PRODUITS. LE MOTIF HOMÉOTIQUE OU HOMEBOX

Les régions chromosomiques correspondant aux deux complexes ANT-C et BX-C ont été clonées par marche sur le chromosome sur une étendue de plus de 360 et 300 kilobases respectivement (Duncan, 1987; Mahaffey et Kaufman, 1988, et références citées). Ceci a permis d'isoler les ADNc correspondant aux transcrits codés par ces régions génomiques et d'établir l'organisation génomique des gènes et de leurs ARNm. Ce matériel cloné a permis également d'analyser en détail la nature des lésions mutationnelles induites dans ces régions au cours des études génétiques. L'ensemble de ces travaux fait ressortir au moins trois points importants.

Le premier à trait à la complexité de l'organisation et de l'expression des gènes homéotiques. Bien que cette situation ne soit unique ni chez la drosophile ni chez les autres organismes eucaryotes, il appa-

raît que les gènes homéotiques sont en majorité de très grande taille, les gènes *Antp* et *Ubx* s'étendant par exemple sur plus de 100 et 80 kilobases (Schneuwly *et al*, 1986; Peifer *et al*, 1987; Bermingham et Scott, 1988). Les ARNm et les protéines codées étant de taille moyenne, la grande longueur des gènes homéotiques est attribuable à la présence d'introns dont la taille atteint plusieurs dizaines de kilobases. L'analyse de la structure des ADNc révèle également la complexité de l'expression des gènes homéotiques : un même gène peut coder pour une famille de transcrits résultant d'un épissage alternatif des introns ou de l'utilisation de plusieurs promoteurs sans que ces deux possibilités soient exclusives (Schneuwly *et al*, 1986; Bermingham et Scott, 1988; O'Connor *et al*, 1988; Kornfeld *et al*, 1989). Dans la mesure où les ADNc isolés jusqu'à présent correspondent vraisemblablement aux transcrits les plus abondants, cette analyse moléculaire de l'expression des gènes homéotiques ne peut pas encore être considérée avec certitude comme exhaustive dans le cas de tous les gènes. La multiplicité des ARNm implique qu'un même gène homéotique peut coder pour plusieurs produits polypeptidiques (Bermingham et Scott, 1988; Kornfeld *et al*, 1989) et la caractérisation biochimique de ces protéines demeure encore très incomplète. Il est clair que ces possibilités de générer de manière régulée plusieurs produits distincts pourraient avoir des implications fonctionnelles différentes pour un même gène homéotique dans différents segments.

Par ailleurs, l'analyse moléculaire des mutations résultant de réarrangements chromosomiques, de délétions ou d'insertion d'éléments transposables divers fait ressortir la grande étendue des régions de régulation en *cis* des gènes homéotiques qui peuvent s'espacer sur plusieurs di-

zaines de kilobases et se situer à l'intérieur, en amont ou en aval des séquences codantes (Peifer *et al*, 1987). La localisation précise et la détermination du rôle exact de ces régions dans la modulation de l'activité des gènes homéotiques demeurent des aspects fondamentaux de l'étude des gènes homéotiques. Il est probable que sont inclus dans ces séquences des éléments de régulation de type «augmenteur» (*enhancer*) de transcription, dont la présence a été mise en évidence au voisinage d'un grand nombre de gènes viraux et cellulaires (Atchison, 1988; Dynan, 1989, et les références citées pour des revues sur ces éléments de régulation en *cis*). Il a été suggéré que c'est l'action combinée et polarisée de ces éléments de régulation en *cis* qui détermine dans les cellules ou groupes de cellules individuels le niveau d'expression d'un gène homéotique dans les différents segments ou son activité est requise (Peifer *et al*, 1987). De plus, certaines de ces régions, telle que la région *bithoraxoid* (*bxo*) codent pour des transcrits ne contenant pas de cadre ouvert de lecture et dont la fonction exacte demeure non comprise (Peifer *et al*, 1987; Mahaffey et Kaufman, 1988).

Le résultat le plus important de l'analyse moléculaire a été certainement de montrer que les protéines homéotiques sont localisées dans le noyau des cellules et possèdent toutes un motif protéique conservé d'une soixantaine d'acides aminés, qui a été appelé *homeobox* ou «motif homéotique» (McGinnis *et al*, 1984a, b; Scott et Weiner, 1984). Sur la base de comparaisons de séquences protéiques, une analogie structurale a été très rapidement proposée par les groupes de Gehring et Scott entre le motif homéotique et le motif hélice-coude-hélice présent dans un certain nombre de protéines très bien caractérisées ayant une fonction de répresseur de l'activité génique chez les

bactéries ou les bactériophages (Laughon et Scott, 1984; McGinnis *et al*, 1984b). Cette hypothèse a été confirmée récemment par Otting *et al* (1988) qui ont déterminé par résonance magnétique nucléaire la structure tridimensionnelle en solution d'un peptide correspondant au motif homéotique du gène *Antp*.

Il est à souligner que chez la drosophile ce motif n'est pas présent exclusivement dans les protéines codées par les gènes homéotiques, mais se retrouve, compte tenu d'un certain nombre de remplacements d'acides aminés qui préservent la possibilité d'une configuration hélice-coude-hélice, dans plus de vingt autres protéines de la drosophile (Scott *et al*, 1989). La plupart de ces protéines sont codées par des gènes impliqués dans les premières étapes du développement, comme les gènes contrôlant le patron de segmentation de l'embryon. De même, ainsi qu'il sera discuté ci-dessous, un motif analogue est retrouvé dans plus de cinquante autres protéines chez des organismes eucaryotes très variés des annélides à l'homme.

MODÈLES ET MÉCANISMES D'ACTION DES GÈNES HOMÉOTIQUES

Les modèles actuels rendant compte du contrôle de l'identité segmentale des gènes homéotiques sont fondés sur les travaux initiaux de Lewis (1978). Celui-ci envisageait, dans le cas du complexe *Bithorax*, la dérépression progressive de l'activité d'un nombre de plus en plus grand de gènes, depuis le segment T2 jusqu'au segment A8. Dans ce modèle, c'est la présence combinée dans une cellule des produits de plusieurs gènes qui définit son identité, celle-ci étant déterminée de manière identique au niveau de toutes les cellules d'un même segment (fig 3). Compte

tenu des résultats accumulés récemment, ce modèle de code combinatoire doit être très certainement révisé en faveur d'un modèle mosaïque (fig 3 et Akam *et al*, 1988). Dans ce nouveau modèle, chaque cellule d'un segment n'exprimerait à un niveau élevé qu'une seule protéine homéotique à la fois. Dans le cas où plus d'un gène homéotique est exprimé dans un segment, l'ensemble des cellules du segment forme une mosaïque de cellules exprimant des protéines différentes. Dans ce modèle, la proportion des cellules exprimant l'un ou l'autre gène peut également varier au cours du développement ou bien d'un segment au suivant. Ce modèle s'accorde avec le nombre relativement restreint des gènes homéotiques, dont trois seulement sont requis pour spécifier l'identité de deux segments thoraciques et de huit segments abdominaux. Par ailleurs, les observations concernant la distribution des protéines indiquent clairement une ré-

partition mosaïque de leur expression au niveau de l'ensemble des cellules d'un même segment, certaines cellules pouvant exprimer séquentiellement et de manière exclusive deux protéines homéotiques différentes (Peifer *et al*, 1987; Akam *et al*, 1988).

La présence du motif homéotique dans les protéines codées par les gènes homéotiques et leur localisation nucléaire sont de très forts arguments en faveur du fait qu'elles interviennent directement dans la modulation de l'expression d'autres gènes par l'intermédiaire de leur fixation au niveau de séquences spécifiques (Desplan *et al*, 1988; Scott *et al*, 1989, Treisman *et al*, 1989). La nature et le nombre de gènes cibles ou de gènes réalisateurs dont l'activité serait ainsi modulée demeurent actuellement complètement inconnus. En dehors de la compréhension intime des mécanismes qui régulent spatialement et temporellement l'activité des gènes homéotiques, l'identification de ces gènes constituerait une des étapes importantes pour comprendre le mécanisme d'action des gènes homéotiques.

L'étude du mécanisme d'action moléculaire des gènes homéotiques et des gènes à motif homéotique a déjà été largement entamée. Il a pu être démontré que le motif homéotique permet une fixation *in vitro* au niveau de séquences d'ADN spécifiques (Beachy *et al*, 1988; Desplan *et al*, 1988; Scott *et al*, 1989; Treisman *et al*, 1989) et qu'il se retrouve dans des protéines ayant une fonction avérée de facteur de transcription (voir Levine et Hoey, 1988; Scott *et al*, 1989, pour des revues). Il est maintenant couramment accepté que la présence du motif homéotique dans une protéine signe sa fonction de régulation de l'activité génique au niveau de la transcription (Struhl, 1989). Il convient de souligner toutefois que jusqu'à présent ceci n'a pu être démontré directement et sans ambi-

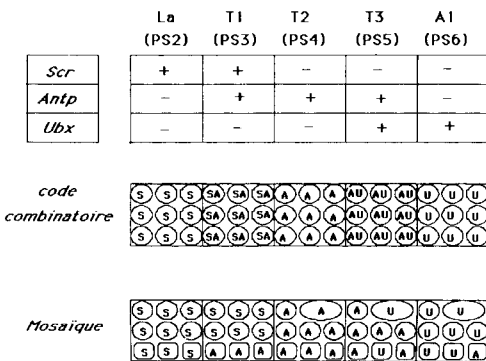


Fig 3. Modèles d'expression des gènes homéotiques. Voir les légendes des figures 1 et 2 pour la signification des abréviations et le texte pour la discussion des deux modèles fondés sur les travaux de Lewis (1978) (code combinatoire) et les travaux récents (mosaïque). Les cercles, les ovales et les rectangles symbolisent les cellules des différents segments et les lettres inscrites la présence de la protéine codée par le gène homéotique correspondant (reproduit de Akam *et al* (1988) avec permission de l'éditeur).

guité que pour un nombre relativement limité de protéines homéotiques. Une approche originale et fructueuse de ce problème est d'étudier les capacités d'activation de la transcription des produits des gènes à motif homéotique par leur expression dans des cellules de drosophile en culture (Jaynes et O'Farrell, 1988; Krasnow *et al*, 1989; Winslow *et al*, 1989), des cellules de mammifères en culture (Thali *et al*, 1988) ou dans la levure (Samson *et al*, 1989).

RÉGULATION DE L'EXPRESSION DES GÈNES HOMÉOTIQUES

Le processus de segmentation résulte de la superposition de deux patrons, le patron périodique des segments dont la mise en place dépend de l'activité d'un certain nombre de gènes maternels et de gènes de segmentation et le patron séquentiel des identités segmentales sous la dépendance des gènes homéotiques (fig 1).

La distribution spatiale de l'activité des gènes homéotiques est complexe et évolue de manière dynamique (Irish *et al*, 1989) au cours du développement. Sans faire d'hypothèses sur les mécanismes moléculaires mis en jeu, on peut considérer qu'elle est régulée à trois niveaux distincts.

Le premier niveau est la définition des domaines spatiaux d'activation des gènes homéotiques au début du développement embryonnaire qui conduira à leur expression dans des limites parasegmentales précises. Les gènes de segmentation de la classe *gap* et certains gènes de la classe *pair rule*, comme *ftz* sont directement impliqués dans la définition précoce des domaines d'activité des gènes des complexes ANT-C et BX-C (Ingham et Martinez-Arias, 1986). Ces mêmes gènes déterminent aussi probablement l'acti-

tion initiale directe de la transcription des gènes homéotiques (Struhl et White, 1985; Scott et O'Farrell, 1986; Akam, 1987; Scott et Carroll, 1987; Carroll *et al*, 1988; Harding et Levine, 1988; Irish *et al*, 1989).

Un deuxième niveau représentant un mécanisme de régulation commun à tous les gènes homéotiques est déterminé par l'activité de gènes «*trans-régulateurs*» tels que *Polycomb (Pc)* (Ingham, 1985). Ces gènes au nombre d'environ cinquante ne jouent probablement pas de rôle dans le déterminisme primaire des domaines d'expression et dans l'activation de la transcription des gènes homéotiques. Ils apparaissent, en revanche, requis pour stabiliser les patrons appropriés d'expression mis en place par les gènes de segmentation. Leur action serait essentiellement celle de régulateurs négatifs de l'expression des gènes des ANT-C et BX-C. Les mutations négatives de ces gènes altèrent profondément la distribution spatiale des transcrits des gènes homéotiques. Elles conduisent à des transformations homéotiques totales, comme par exemple la transformation de l'identité de tous les segments de l'embryon en une identité de segment A8 chez les mutants défectifs du gène *Pc*. La mise en évidence récente de la liaison du produit du gène *Pc* au voisinage de la majorité des gènes des complexes ANT-C et BX-C et des autres gènes *trans-régulateurs* apporte un argument direct en faveur de l'interaction de *Pc* avec les gènes homéotiques (Zink et Paro, 1989).

Le dernier niveau d'interactions régulateurs concerne celles qui prennent place entre gènes homéotiques eux-mêmes. Il a été montré que le produit du gène *Ubx* réprime l'expression du gène *Antp* dans les parasegments PS5 et PS6 et vraisemblablement dans les parasegments postérieurs. La délétion de *Ubx* entraîne la dérégulation de *Antp* dans les PS5 et PS6 qui accompagne le glissement d'identité de

ces parasegments vers l'identité PS4, comme décrit ci-dessus. De même le produit du gène *abd-A* réprime l'expression du gène *Ubx* dans les segments postérieurs au parasegment PS6 (Peifer *et al*, 1987). Ces interactions sont confirmées par la démonstration qu'une des protéines codées par le gène *Ubx* se fixe *in vitro* près du promoteur du gène *Antp* (Beachy *et al*, 1988).

LES GÈNES À MOTIF HOMÉOTIQUE DES AUTRES ORGANISMES EUCARYOTES

La découverte du motif homéotique dans les protéines codées par les gènes homéotiques de la drosophile et la mise en évidence de sa très grande conservation de séquence entre gènes différents a conduit très rapidement à rechercher la présence d'un motif analogue chez les autres organismes eucaryotes en utilisant des sondes correspondant à la région codant pour le motif homéotique du gène *Antp*. Actuellement, plus de cinquante gènes contenant un motif homéotique ont été caractérisés chez une grande variété d'organismes eucaryotes incluant le nématode, l'abeille, l'oursin, et plusieurs vertébrés dont la lamproie, le poisson zèbre, le xénope, la souris et l'homme (Bürghin *et al*, 1989; Dressler, 1989; Duboule et Dollé, 1989; Kappen *et al*, 1989; Scott *et al*, 1989).

Ces gènes font actuellement l'objet de très nombreuses études concernant leur structure, leur organisation génomique et leur expression au cours du développement. Un certain nombre de points importants émergent clairement de l'ensemble des données déjà disponibles.

Mis à part un très petit nombre d'entre eux, l'ensemble des motifs homéotiques connus, soit environ soixante-dix, peut être classé en neuf familles distinctes. Ces familles ont en commun un certain nombre

de remplacements d'acides aminés, si on les compare à la séquence du motif homéotique du gène *Antp* considérée comme séquence de référence (Graham *et al*, 1989; Scott *et al*, 1989). Six de ces familles regroupent des gènes présentant une identité partielle de séquence significative avec les gènes homéotiques de drosophile *lab*, *pb*, *Dfd*, *Scr*, *Antp* et *Abd-B* (Graham *et al*, 1989; Duboule et Dollé, 1989; Kappen *et al*, 1989; Scott *et al*, 1989). Les séquences des motifs homéotiques des gènes *Antp*, *Ubx* et *abd-A* étant très proches, un certain nombre d'entre eux sont rassemblés dans un groupe apparenté sans qu'il soit possible à ce stade de distinguer des sous-familles (Graham *et al*, 1989; Scott *et al*, 1989) (fig 4). Le regroupement en familles distinctes est renforcé par la comparaison des séquences en dehors de la région du motif homéotique. Ce type de comparaison montre de plus des identités de séquence très significatives entre les gènes *Hox 1.6*, *Hox 2.1* et *Hox 2.6* de souris et les gènes *lab*, *Scr* et *Dfd* de drosophile respectivement (Baron *et al*, 1987; Graham *et al*, 1988; Graham *et al*, 1989). Par ailleurs, un gène homologue du gène *caudal* impliqué dans le déterminisme des structures postérieures de l'embryon de drosophile a été identifié chez la souris (Duprey *et al*, 1988).

Le regroupement en familles prend en outre une signification particulière dans le cas des gènes de souris et des gènes humains. Leur localisation génomique montre en effet qu'ils sont rassemblés au sein de quatre grands complexes *Hox* situés sur des chromosomes différents (Kappen *et al*, 1989). L'alignement des complexes *Hox* fait ressortir que les différentes familles sont disposées dans le même ordre dans les quatre complexes et que cette disposition est la même chez l'homme et chez la souris. L'intérêt de cette colinéarité est considérablement renforcé par le fait qu'il est facile d'aligner les deux complexes de

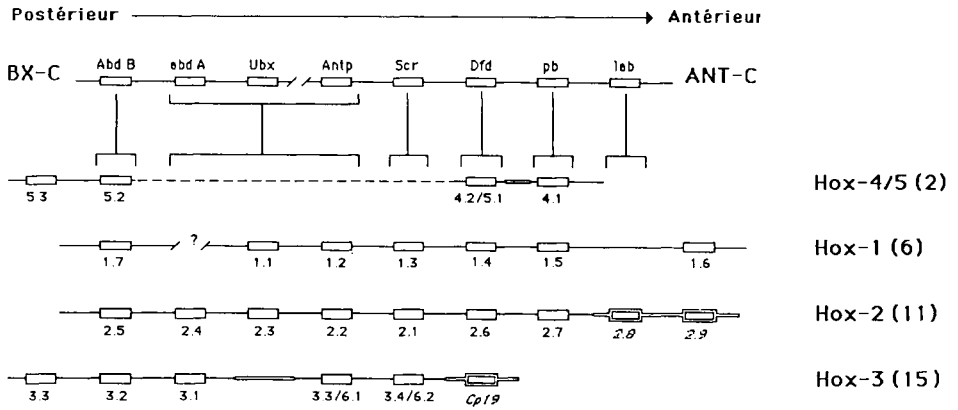


Fig 4. Comparaison de l'organisation des gènes des complexes ANT-C et BX-C de la drosophile et des gènes des complexes *Hox* de la souris et de l'homme. Le schéma inclut les informations compilées par Duboule et Dollé (1989), Graham *et al* (1989), et Kappen *et al* (1989) pour les gènes de souris et les gènes humains. Les gènes de souris et leurs séquences génomiques flanquantes sont indiqués par des rectangles et des lignes simples, les gènes humains par des rectangles et des lignes doubles. Pour les gènes identifiés chez l'homme et chez la souris, le nom du gène est indiqué en caractères romains. Les gènes humains dont l'homologue n'est pas encore connu chez la souris sont indiqués en italiques. Les chiffres entre parenthèses indiquent la localisation chromosomique des gènes murins.

gènes homéotiques (HOM-C) de la drosophile ANT-C et BX-C avec les quatre complexes *Hox* de façon à faire coïncider les différents gènes homéotiques de drosophile avec les familles *Hox* qui leur sont reliées par la séquence du motif homéotique (Akam, 1989; Duboule et Dollé, 1989; Graham *et al*, 1989).

L'ensemble de ces faits tend à indiquer que les complexes HOM-C de la drosophile et les complexes *Hox* de l'homme et de la souris dérivent d'un complexe ancestral unique de gènes à motif homéotique déjà présent chez l'hypothétique organisme à partir duquel ont dérivé des lignées phylogénétiques ayant conduit aux vertébrés et aux insectes. Les données de séquence sont encore trop incomplètes pour qu'il soit actuellement possible d'identifier parmi les gènes actuels les membres du complexe ancestral et de retracer une

histoire évolutive probable des complexes géniques qui en sont dérivés. On peut toutefois proposer deux phases dans la génération des complexes *Hox* (Kappen *et al*, 1989). En effet les identités de séquence plus grandes entre membres d'une même famille qu'entre les différents gènes d'un même complexe indiquent que les quatre complexes *Hox* actuels sont le résultat d'une duplication partielle ou totale d'un complexe *Hox* primordial. Ces duplications auraient pris place avant la divergence des poissons, des amphibiens et des mammifères (Kappen *et al*, 1989). Le complexe vertébré primordial aurait pu lui-même être formé par un processus de multiplication des gènes au sein du complexe ancestral qui aurait conduit à l'individualisation des différentes familles de gènes à motif homéotique conservées chez tous les vertébrés (Kappen *et al*, 1989). Séparément, un

processus de multiplication analogue aurait pu conduire au complexe HOM-C primordial des insectes qui se serait ensuite séparé dans le cas de la drosophile en deux sous-complexes à la suite d'un remaniement chromosomique (Akam *et al*, 1988). Un fort argument en faveur de ce dernier point a été fourni par Beeman (1987) et Beeman *et al* (1989) qui ont démontré l'existence chez le coléoptère *Tribolium cataneum* d'un grand complexe unique rassemblant les gènes homologues des gènes des ANT-C et BX-C de la drosophile.

Les résultats décrits ci-dessus concernant l'organisation génomique et la structure des protéines à motif homéotique donnent du poids à l'idée d'un lien évolutif entre les gènes homéotiques de la drosophile et les gènes à motif homéotique des vertébrés. Il est permis de s'interroger sur la possibilité d'une conservation de la fonction de ces gènes et de proposer par analogie un rôle dans la détermination de la polarité axiale de l'embryon de vertébré (De Robertis *et al*, 1989; Ruiz i Altaba et Melton, 1989; Wright *et al*, 1989). L'étude des domaines d'expression de ces gènes par hybridation *in situ* et par détection des protéines à l'aide d'anticorps spécifiques dans les embryons de xénope et de souris fait ressortir une situation complexe et dynamique au cours du développement (Holland et Hogan, 1988; Wright *et al*, 1989). Néanmoins, il est possible de mettre en évidence des limites antérieures nettes de domaines d'expression des gènes *Hox* au niveau du système nerveux central et de la colonne prévertébrale de l'embryon de souris (Duboule et Dollé, 1989; Graham *et al*, 1989). Ce point est particulièrement bien illustré dans le cas du cerveau postérieur de l'embryon de souris à 9,5 j de développement (Murphy *et al*, 1989; Wilkinson *et al*, 1989). Une corrélation nette entre la position antérieure de ces do-

maines d'expression selon l'axe rostro-caudal de l'embryon et la position distale-proximale des gènes au sein des complexes *Hox* peut être ainsi établie.

Deux conclusions provisoires se dégagent des travaux déjà réalisés. En premier lieu, il est permis de penser que le maintien pendant plusieurs centaines de millions d'années de l'organisation génomique regroupée des gènes *Hox* et de leur position relative fixe au sein des complexes a une importance fonctionnelle primordiale dans l'expression spatiale coordonnée des gènes à boîte homéotique des vertébrés.

Par ailleurs et sans préjuger de leur rôle exacte, il convient de souligner que la présence du motif homéotique dans les gènes *Hox* implique qu'ils codent pour des facteurs de transcription. On peut donc supposer qu'ils interviennent, comme les gènes homéotiques de la drosophile, dans la sélection de sous-programmes de développement au cours de l'ontogenèse.

CONCLUSION

L'étude des gènes homéotiques a sans conteste apporté un renouvellement et un début d'unification des concepts concernant les mécanismes moléculaires impliqués dans l'embryogenèse et la morphogénèse chez des organismes aussi éloignés que la souris et la drosophile. Il ne semble pas en effet douteux que l'on assiste à la découverte de mécanismes fondamentaux, dont l'origine est certainement très ancienne.

Les recherches futures dans le domaine nouveau qui s'est ainsi largement ouvert depuis quelques années devraient s'orienter vers un approfondissement de l'analyse du mécanisme d'action de ces gènes clés du développement, passant en particulier

par l'identification de leurs gènes cibles aussi bien chez la souris que chez les autres organismes eucaryotes.

La biologie moléculaire et l'application des méthodes de la transgénose permettent d'aborder ces questions avec des organismes modèles où les approches de la génétique classique sont lourdes ou impossibles à mettre en œuvre comme chez la souris ou le xénope. Il demeure néanmoins que les études effectuées chez les organismes comme le nématode et la drosophile, où les approches moléculaires peuvent être potentialisées remarquablement par la génétique, vont conserver un impact déterminant.

REMERCIEMENTS

Je remercie Mmes Brunello, Genotte et M Scharzmann pour toute l'aide qu'ils m'ont apportée dans la recherche des références bibliographiques et la préparation du manuscrit et des figures de cette revue.

RÉFÉRENCES

- Akam M (1987) The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* 101, 1-22
- Akam M (1989) *Hox* and HOM: homologous gene clusters in insects and vertebrates. *Cell* 57, 347-349
- Akam ME, Dawson I, Tear G (1988) Homeotic genes and the control of segment diversity. *Development* 104 (suppl), 123-133
- Atchison ML (1988) Enhancers: mechanisms of action and cell specificity. *Annu Rev Cell Biol* 4, 127-153
- Baron A, Featherstone MS, Hill RE, Hall A, Galliot B, Duboule D (1987) *Hox-1.6*: a mouse homeobox containing gene member of the *Hox-1* complex *EMBO J* 6, 2977-2986
- Bateson W (1894) Materials for the study of variation. Cambridge University Press, MacMillan, London, New York
- Beachy PA, Krasnow MA, Gavis ER, Hogness DS (1988) An *Ultrabithorax* protein binds sequences near its own and the *Antennapedia* P1 promoters. *Cell* 55, 1069-1081
- Beeman RW (1987) A homeotic gene cluster in the red flour beetle. *Nature* 327, 247-249
- Beeman RW, Stuart JJ, Hass MS, Denell RE (1989) Genetic analysis of the homeotic gene complex (HOM-C) in the beetle *Tribolium castaneum*. *Dev Biol* 133, 196-206
- Birmingham JR, Scott MP (1988) Developmentally regulated alternative splicing of transcripts from the *Drosophila* homeotic gene *Antennapedia* can produce four different proteins. *EMBO J* 7, 3211-3222
- Bürghin T, Finney M, Coulson A, Ruvkun G (1989) *Caenorhabditis elegans* has scores of homeobox-containing genes. *Nature* 341, 239-243
- Carroll SB, Di Nardo S, O'Farrell PH, White RAH, Scott MP (1988) Temporal and spatial relationships between segmentation and homeotic gene expression in *Drosophila* embryos: distributions of the *fushi tarazu*, *engrailed*, *sex combs reduced*, *Antennapedia* and *Ultrabithorax* proteins. *Genes Dev* 2, 350-360
- Celniker SE, Keelan DJ, Lewis EB (1989) The molecular genetics of the bithorax complex of *Drosophila*: characterization of the products of the *Abdominal-B* domain. *Genes Dev* 3, 1424-1436
- De Robertis EM, Oliver G, Wright CVE (1989) Determination of axial polarity in the vertebrate embryo: homeodomain proteins and homeogenetic induction. *Cell* 57, 189-191
- Desplan C, Theis J, O'Farrell PH (1988) The sequence specificity of homeodomain-DNA interaction. *Cell* 54, 1081-1090
- Diederich RJ, Merrill VKL, Pultz MA, Kaufman TC (1989) Isolation, structure and expression of *labial*, a homeotic gene of the *Antennapedia* complex involved in *Drosophila* head development. *Genes Dev* 3, 399-414
- Dressler GR (1989) An update on the vertebrate homeobox. *Trends Genet* 5, 129-130
- Duboule D, Dollé P (1989) The structural and functional organization of the murine *Hox* gene family resembles that of *Drosophila* homeotic genes. *EMBO J* 8, 1497-1505
- Duncan I (1987) The bithorax complex. *Annu Rev Genet* 21, 533-544
- Duprey P, Chowdhury K, Dressler GR, Balling R, Simon D, Guenet JL, Gruss P (1988) A

- mouse gene homologous to the *Drosophila* gene *caudal* is expressed in epithelial cells from the embryonic intestine. *Genes Dev* 2, 1647-1654
- Dynan WS (1989) Modularity in promoters and enhancers. *Cell* 58, 1-4
- Garcia-Bellido A (1977) Homeotic and atavic mutations in insects. *Am Zool* 17, 613-629
- Gehring WJ (1987) Homeoboxes in the study of development. *Science* 236, 1245-1252
- Gehring WJ, Hiromi Y (1986) Homeotic genes and the homeobox. *Annu Rev Genet* 20, 147-173
- Graham A, Papalopulu N, Krumlauf R (1989) The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* 57, 367-378
- Graham A, Papalopulu N, Lorimer J, McVey JH, Tuddenham EGD, Krumlauf R (1988) Characterization of a murine homeobox gene, *Hox-2-6*, related to the *Drosophila Deformed* gene. *Genes Dev* 2, 1424-1438
- Harding K, Levine M (1988) Gap genes define the limits of *Antennapedia* and *Bithorax* gene expression during early development in *Drosophila*. *EMBO J* 7, 205-214
- Holland PWH, Hogan BLM (1988) Expression of homeobox genes during mouse development: a review. *Genes Dev* 2, 773-782
- Ingham PW (1985) The regulation of the bithorax complex. *Trends Genet* 1, 112-116
- Ingham PW (1988) The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* 335, 25-34
- Ingham PW, Martinez-Arias A (1986) The correct activation of *Antennapedia* and *Bithorax* complex genes requires the *fushi tarazu* gene. *Nature* 324, 592-597
- Irish VF, Martinez-Arias A, Akam M (1989) Spatial regulation of the *Antennapedia* and *Ultrabithorax* homeotic genes during *Drosophila* early development. *EMBO J* 8, 1527-1537
- Jaynes JB, O'Farrell PH (1988) Activation and repression of transcription by homeodomain-containing proteins that bind a common site. *Nature* 336, 744-749
- Kappen C, Schugart K, Ruddle FH (1989) Two steps in the evolution of *Antennapedia*-class vertebrate homeobox genes. *Proc Natl Acad Sci USA* 86, 5459-5463
- Kornfeld K, Saint RB, Beachy PA, Harte PJ, Peattie DA, Hogness DS (1989) Structure and expression of a family of Ultrabithorax mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. *Genes Dev* 3, 243-258
- Krasnow MA, Saffman EE, Kornfeld K, Hogness DS (1989) Transcriptional activation and repression by *Ultrabithorax* proteins in cultured *Drosophila* cells. *Cell* 57, 1031-1043
- Laughon A, Scott MP (1984) Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. *Nature* 310, 25-31
- Le Motte PK, Kuroiwa A, Fessler LI, Gehring WJ (1989) The homeotic gene *sex combs reduced* of *Drosophila*: gene structure and embryonic expression. *EMBO J* 8, 219-227
- Levine M, Hoey T (1988) Homeobox proteins as sequence-specific transcription factors. *Cell* 55, 537-540
- Lewis EB (1978) A gene complex controlling segmentation in *Drosophila*. *Nature* 276, 565-570
- Mahaffey JW, Kaufman TC (1988) The homeotic genes of the *Antennapedia* complex and the bithorax complex of *Drosophila*. In: *Developmental Genetics of Higher Organisms, a Primer in Developmental Biology* (Malacinski GM, ed) Macmillan, New York, London, 329-360
- Martinez-Arias A, Lawrence PA (1985) Parasegments and compartments in the *Drosophila* embryo. *Nature* 313, 639-642
- McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ (1984a) A conserved DNA sequence in homeotic genes of the *Drosophila Antennapedia* and *Bithorax* complexes. *Nature* 308, 428-433
- McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring W (1984b) A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell* 37, 403-408
- Mlodzik M, Fjose A, Gehring WJ (1988) Molecular structure and spatial expression of a homeobox gene from the *labial* region of the *Antennapedia* complex. *EMBO J* 7, 2569-2578
- Murphy P, Davidson DR, Hill RE (1989) Segment-specific expression of a homeobox containing gene in the mouse hindbrain. *Nature* 341, 156-159
- Nüsslein-Volhard C, Weischaus E (1980) Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287, 795-801
- O'Connor MB, Bianari R, Perkins LA, Bender W (1988) Alternative RNA products from the *U-*

- trabithorax* domain of the *Bithorax* complex. *EMBO J* 7, 435-445
- Otting G, Quian YQ, Müller M, Affolter M, Gehring WJ, Wüthrich K (1988) Secondary structure determination for the *Antennapedia* homeodomain by nuclear magnetic resonance and evidence for a helix–turn–helix motif. *EMBO J* 7, 4305-4309
- Ouweneel WJ (1976) Developmental genetics of homeosis. *Adv Genet* 18, 179-248
- Peifer M, Karch F, Bender W (1987) The *Bithorax* complex: control of segmental identity. *Genes Dev* 1, 891-898
- Pultz MA, Diederich RJ, Cribbs DL, Kaufman TC (1988) The *proboscipedia* locus of the *Antennapedia* complex: a molecular and genetic analysis. *Genes Dev* 2, 901-920
- Regulski M, McGinnis N, Chadwick R, McGinnis W (1987) Developmental and molecular analysis of *Deformed*, a homeotic gene controlling *Drosophila* head development. *EMBO J* 6, 767-777
- Rubin GM (1988) *Drosophila melanogaster* as an experimental organism. *Science* 240, 1453-1459
- Ruiz i Altaba A, Melton DA (1989) Interaction between peptide growth factors and homeobox genes in the establishment of antero-posterior polarity in frog embryos. *Nature* 341, 33-38
- Samson ML, Jackson-Grusby L, Brent R (1989) Gene activation and DNA binding by *Drosophila Ubx* and *abd-A* proteins. *Cell* 57, 1045-1052
- Schneuwly S, Klemenz R, Gehring WJ (1987) Redesigning the body plan of *Drosophila* by ectopic expression of the homeotic gene *Antennapedia*. *Nature* 325, 816-818
- Schneuwly S, Kuroiwa A, Baumgartner P, Gehring WJ (1986) Structural organization and sequence of the homeotic gene *Antennapedia* of *Drosophila melanogaster*. *EMBO J* 5, 733-739
- Scott MP (1985) Molecules and puzzles from the *Antennapedia* homeotic gene complex of *Drosophila*. *Trends Genet* 1, 74-80
- Scott MP (1987) Complex loci of *Drosophila*. *Annu Rev Biochem* 56, 195-227
- Scott MP (1988) The molecular biology of pattern formation in the early embryonic development of *Drosophila*. In: *Developmental Biology. A Comprehensive Synthesis*. Vol 5. *The Molecular Biology of Cell Determination and Cell Differentiation*. (Browder LW, ed) 151-185
- Scott MP, Carroll B (1987) The segmentation and homeotic gene network in early *Drosophila* development. *Cell* 51, 689-698
- Scott MP, O'Farrell PH (1986) Spatial programming of gene expression in early *Drosophila* embryogenesis. *Annu Rev Cell Biol* 2, 49-80
- Scott MP, Weiner AJ (1984) Structural relationships among genes that control development: sequence homology between the *Antennapedia*, *Ultrabithorax* and *fushi tarazu* loci of *Drosophila*. *Proc Natl Acad Sci USA* 81, 4115-4119
- Scott MP, Tamkun JW, Hartzell GW III (1989) The structure and function of the homeodomain. *Biochim Biophys Acta* 989, 25-48
- Struhl K (1989) Helix–turn–helix, zinc-finger and leucine-zippers motifs for eukaryotic transcriptional regulatory proteins. *Trends Biochem Sci* 14, 137-140
- Struhl G, White RA (1985) Regulation of the *Ultrabithorax* gene of *Drosophila* by other *Bithorax* complex genes. *Cell* 43, 507-519
- Thali M, Müller MM, DeLorenzi M, Matthias P, Bienz M (1988) *Drosophila* homeotic genes encode transcriptional activators similar to mammalian OTF-2. *Nature* 336, 598-601
- Treisman J, Gönczy P, Vashihtha M, Harris E, Desplan C (1989) A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* 59, 553-562
- Wilkinson DG, Bhatt S, Cook M, Boncinelli E, Krumlauf R (1989) Segmental expression of *Hox-2* homeobox containing gene in the developing mouse hindbrain. *Nature* 341, 405-409
- Winslow GM, Hayashi S, Krasnow M, Hogness DS, Scott MP (1989) Transcriptional activation by the *Antennapedia* and *fushi tarazu* proteins in cultured *Drosophila* cells. *Cell* 57, 1017-1030
- Wright CVE, Cho KWY, Oliver G, De Robertis EM (1989) Vertebrate homeodomain proteins: families of region specific transcription factors. *Trends Biochem Sci* 14, 52-56
- Zink B, Paro R (1989) *In vivo* binding pattern of a transregulator of homeotic genes in *Drosophila melanogaster*. *Nature* 337, 468-471