

Une famille de gènes du développement : les gènes Pax

La conservation de séquences dans les gènes du développement de la drosophile a fait naître l'idée que des gènes porteurs de séquences similaires pourraient être impliqués dans le développement d'organismes éloignés phylogénétiquement et, en conséquence, que les mécanismes fondamentaux du développement pourraient, en dépit des apparences, être communs à des organismes aussi différents qu'une mouche, une souris ou un nématode [1]. De fait, cette conjecture, nullement évidente à l'origine, s'est trouvée amplement confirmée ; en utilisant comme sonde des séquences d'ADN conservées dans les gènes du développement de la drosophile pour cribler des banques d'ADN génomique d'autres espèces, nombre de

gènes portant ces séquences conservées ont été trouvés chez des organismes d'espèces très éloignées, le cas le mieux documenté étant celui du domaine homéotique présent dans des gènes de tous les eucaryotes étudiés jusqu'ici. Cependant, d'autres gènes du développement ont pu être également découverts en adoptant la même stratégie : c'est le cas d'une famille de huit gènes de la souris, isolés sur la base d'une homologie avec le domaine *paired* [2] trouvé à l'origine dans trois gènes de segmentation de la drosophile dont le gène *paired*. Aussi, ces huit gènes sont-ils appelés gènes *Pax* (pour *paired box*).

Plusieurs propriétés remarquables communes à l'ensemble des gènes de cette famille en font de bons candi-

dates comme gènes du développement. Propriétés structurales d'abord : tous ont en commun un domaine appelé *paired* codant pour un motif protéique de 128 acides aminés toujours présent dans la partie NH₂-terminale de la protéine et capable de se lier à l'ADN ; tous, sauf *Pax-1*, possèdent également une séquence apparentée à un domaine homéo (le domaine homéo de type *paired*), tronquée toutefois dans le cas de *Pax-2*, *Pax-5* et *Pax-8* (figure 1) ; enfin, excepté *Pax-4* et *Pax-6*, ils contiennent, entre les domaines *paired* et homéo, un motif octapeptide conservé. Bien que les rôles respectifs des différents domaines des protéines Pax ne soient pas encore élucidés, il a été montré dans plusieurs cas que celles-ci avaient une

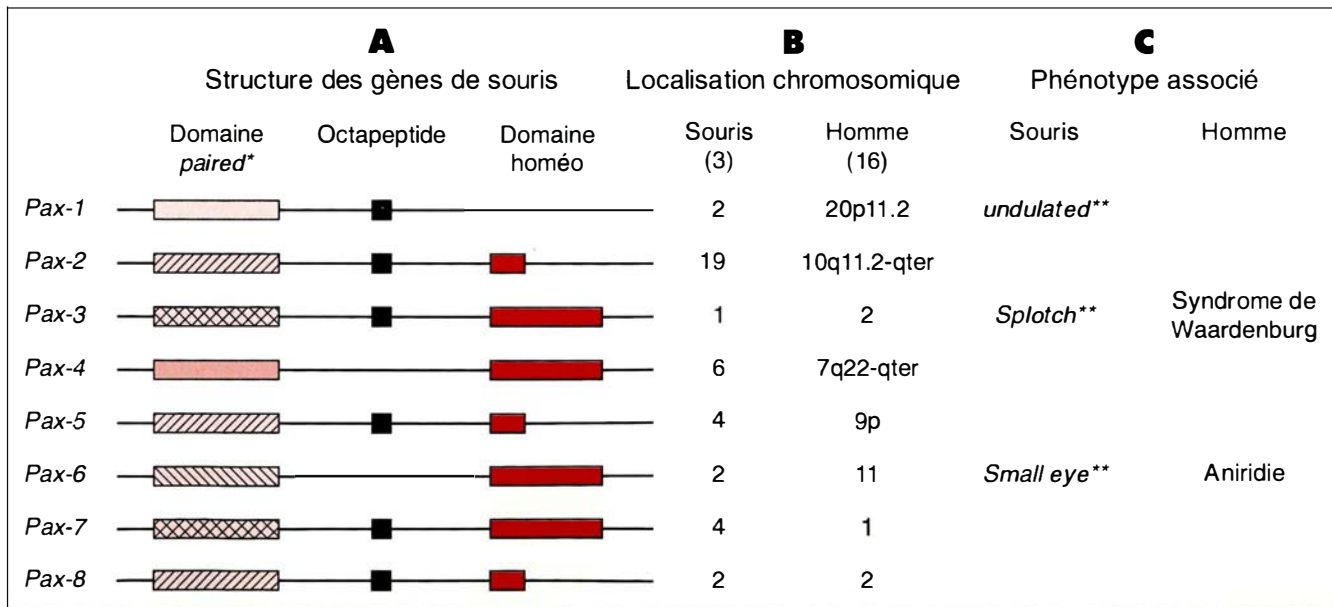


Figure 1. **Les gènes Pax : structure, localisation et phénotypes associés.** * Les différents dessins des domaines *paired* indiquant que chaque gène Pax peut être classé dans des sous-groupes en fonction des homologies relatives et de la présence de tel ou tel acide aminé à des positions particulières dans ces domaines (voir [3]). ** Les mutations connues de Pax-3 et Pax-6 sont létales à l'état homozygote ; pour Pax-1, seule la mutation un^S est létale, mais, comme il s'agit d'une délétion, la létalité pourrait être due à l'inactivation d'autres gènes.

activité de stimulation transcriptionnelle. L'analyse comparée et détaillée de la structure et de la séquence codante de chacun de ces gènes rend très vraisemblable l'idée qu'ils seraient nés de duplications d'un gène ancestral suivies de divergences des gènes ainsi dupliqués [3].

L'examen du profil d'expression des différents gènes *Pax* (Tableau I) est également très suggestif de leur implication dans le développement : (1) tous sauf deux, sont exprimés uniquement pendant l'embryogenèse ; (2) tous les gènes *Pax*, sauf *Pax-1*, sont exprimés au cours de la formation du système nerveux ; (3) enfin, les gènes *Pax*, en dehors du système nerveux (d'origine ectodermique), ont des domaines d'expression remarquables dans les tissus dérivés du mésoderme, notamment les somites et le système excréteur rénal. Cette expression dans des structures segmentées rapprochée du fait que les gènes *Pax* ont été isolés sur la base d'une homologie avec des séquences de gènes de la drosophile impliqués dans la segmentation, renforce l'idée d'une conservation fonctionnelle de certains gènes au cours de l'évolution.

Mutations des gènes *Pax*

Les arguments que nous venons d'indiquer militent en faveur de l'implication des gènes *Pax* dans le développement, mais la génétique a permis d'apporter une preuve directe pour trois d'entre eux. En effet, la localisation génétique de ceux-ci a fait apparaître qu'ils étaient très liés à des mutations connues de la souris présentant des anomalies dans des tissus où ces gènes s'exprimaient ; l'analyse moléculaire a ensuite montré que ces mutants possédaient une mutation dans un gène *Pax*.

Pax-1 et le mutant de souris *undulated* (*un*)

Pax-1 code pour un gène ayant un domaine *paired* mais pas de domaine homéo. L'analyse génétique a permis de le localiser sur le chromosome 2 entre les *locus* *agouti* et $\beta 2$ -microglobuline. Dans cette région se trouve une mutation récessive appelée *undulated* (*un*). Or l'expression de *Pax-1* a lieu dans les régions de l'embryon qui sont à l'origine des

Tableau I			
PROFILS D'EXPRESSION DES DIFFÉRENTS GÈNES <i>Pax</i> CHEZ LA SOURIS			
Gènes	Embryon		Adulte
	période	site*	
<i>Pax-1</i>	j10-17	Mésoderme somitique : sclérotome	Non
<i>Pax-2</i>	j10-18	Rein en développement . Tube nerveux, rhombencéphale, vésicule otique, vésicule optique	Non
<i>Pax-3</i>	j8,5-16	Certaines régions du cerveau, partie dorsale du tube nerveux, cellules de la crête neurale Dermomyotome	Non
<i>Pax-4</i> <i>Pax-5</i>	ND j10-14	ND Mesencéphale, tube nerveux Foie fœtal : lignage lymphocytaires B	ND Rate, ganglions, lymphatiques, sang : lignage lymphocytes B Testicule
<i>Pax-6</i>	j8-18	Cerveau, antérieur et postérieur, hypophyse, épithélium olfactif, œil, tube nerveux (zone ventrale)	Non
<i>Pax-7</i>	j8-17	Groupe de cellules dans tout le cerveau, puis limité au mésencéphale, tube neural (zone dorsale) Somites : dermomyotome, myotome, muscle squelettique	Non
<i>Pax-8</i>	A partir de j11,5	Expression transitoire dans tout le tube nerveux et le myélinocéphale (j11,5 à j12,5) Thyroïde et rein en développement	Thyroïde, rein

* Pour une comparaison plus détaillée des profils d'expression, en particulier dans le tube nerveux, voir [2] et [10].

ND : non déterminé.

deux tissus, les disques intervertébraux et le sternum, où l'on observe des anomalies chez les individus homozygotes *un*. Le clonage puis le séquençage du gène *Pax-1* chez ces mutants a permis d'identifier une mutation ponctuelle entraînant un changement de sérine en glycine, et ceci à une position extrêmement conservée dans le domaine *paired* [4].

Pax-1 est donc selon toute vraisemblance le gène dont la mutation est à l'origine du phénotype observé chez les animaux *undulated*. La vraisemblance devient certitude si l'on ajoute que deux autres allèles mutés sur le *locus undulated* présentent également une mutation du gène *Pax-1* ; pour l'un des allèles mutés, *undulated exten-*

sive (*un^{ex}*), *Pax-1* est délété de sa partie COOH-terminale et, pour l'autre, *undulated-short tail* (*un^s*), il est complètement délété ([5] et Rudi Balling, communication personnelle). L'étude moléculaire du produit de *Pax-1* a été effectuée et indique qu'il s'agit d'une protéine capable de se lier de manière spécifique à l'ADN et d'activer la transcription à partir d'un promoteur minimal, porteur des séquences de reconnaissance. Enfin, observation particulièrement intéressante, le produit du gène muté *un*, non seulement a perdu une grande partie de son affinité pour la séquence cible d'ADN, mais il a acquis une nouvelle spécificité de liaison. On pourrait donc s'attendre à ce que cette muta-

tion se comporte comme une mutation gain de fonction et entraîne des effets à l'état hétérozygote. Bien que les individus hétérozygotes *unl* / + aient une allure générale normale, il semble cependant que l'on puisse observer de subtiles modifications phénotypiques.

Le mutant de souris *Pax-3* - *Splotch* et le syndrome humain de Waardenburg

Pax-3 est exprimé très tôt lors de la formation du tube neural dans la zone ventriculaire dorsale, siège d'une intense activité mitotique. Il est également exprimé dans différentes régions du cerveau ainsi que dans certaines cellules qui dérivent de la crête neurale [6]. Il a d'autre part été montré que *Pax-3* code pour une protéine capable de se lier, comme la protéine *paired*, à des séquences en amont du gène de segmentation de drosophile *evenskipped* (appelées « e 5 ») et que aussi bien le domaine *paired* que le domaine homéo contribuaient à l'efficacité de la liaison. Comme dans le cas de *Pax-1*, la localisation génétique de *Pax-3* le place tout près d'un locus, *Splotch* (*Sp*) où se trouvent plusieurs mutations entraînant chez l'homozygote un phénotype évocateur d'un dysfonctionnement du gène *Pax-3* : en effet, les tissus atteints chez les homozygotes correspondent aux tissus dans lesquels s'exprime *Pax-3* : malformation de l'axe neural avec exencéphalie, méningocèle et *spina bifida*. En outre, tous les hétérozygotes pour les différents variants alléliques *Sp* présentent des anomalies du pelage avec des tâches blanches (*white spotting*) qui pourraient avoir pour origine un défaut de migration des mélanocytes, dont on sait qu'ils sont dérivés de la crête neurale. Le clonage du gène *Pax-3* à partir d'homozygotes *Splotch* 2H (*Sp*^{2H}, l'un des variants alléliques de *splotch*) ainsi que de l'ADNc correspondant a permis d'identifier une délétion entraînant la formation d'une protéine privée de toute sa partie COOH-terminale à partir de l'homéodomaine [7]. Ajoutons que chez le variant *Splotch retarded* (*Sp*), on observe une délétion complète du gène *Pax-3*. L'ensemble de ces données montre que *Pax-3* est impliqué

de manière déterminante dans certains aspects du développement du tube neural. Cette conclusion est encore renforcée par l'étude génétique et moléculaire du syndrome humain de Waardenburg. Ce syndrome est transmis comme un caractère autosomique dominant et se manifeste par une surdité associée à des anomalies de la pigmentation dont l'origine pourrait être due à une mutation affectant des populations cellulaires dérivées de la crête neurale. Or il se trouve que *Pax-3* est localisé sur le chromosome 1 de la souris dans une région à laquelle correspond un fragment synténique qu'on trouve sur le bras long du chromosome 2 humain. Il était donc tentant de rechercher la présence dans cette région d'un gène humain de type *Pax* : de plus, un candidat se présentait puisque l'un des gènes humains de type *paired* connus, HuP2, contient un domaine *paired* strictement identique à celui de *Pax-3* et constitue probablement l'homologue humain de *Pax-3*. De fait, la cartographie a montré que HuP2 était dans une région identique à celle où avait été localisé le déterminant génétique du syndrome de Waardenburg et, d'autre part, l'analyse de plusieurs familles atteintes a mis en évidence l'existence, chez les individus porteurs, de mutations dans le gène HuP2 [8, 9].

Pax-6, le mutant de souris *Small-eye* et l'aniridie humaine

Le gène codant pour *Pax-6* est exprimé pendant l'embryogenèse dans certaines parties du cerveau, dans le tube neural et dans l'œil [10]. Il a été localisé sur le chromosome 2 de la souris au voisinage d'un locus, appelé *small eye* et auquel correspondent plusieurs mutations alléliques, qui ont un caractère semi-dominant : en particulier, les hétérozygotes ont l'œil plus petit que le type sauvage. A l'état homozygote, on constate l'absence totale d'œil ainsi que de cavité nasale. Pour trois mutations alléliques du locus *Sey*, l'analyse moléculaire a montré l'existence de mutations dans le gène *Pax-6*, entraînant soit l'absence totale du produit du gène (*Sey*^H) soit la formation d'un produit tronqué (*Sey*, *Sey*^{Nes}) [11]. Il apparaît donc clairement que le locus

Sey code pour la protéine *Pax-6* et les phénotypes résultant des mutations de ce locus, sont une illustration supplémentaire du rôle des gènes *Pax* dans le développement embryonnaire. En outre, comme dans le cas de *Pax-3*, il existe une maladie génétique humaine, l'aniridie, dont les signes cliniques rappellent le phénotype des souris *Sey* et qui est transmise comme un caractère autosomique dominant ; or la localisation génétique, puis le clonage de l'ADNc du gène candidat à l'origine de cette maladie ont montré que, chez deux malades appartenant à des familles différentes, le gène *Pax-6* était soit délété soit tronqué [12].

Un point mérite d'être souligné qui caractérise aussi bien les souris mutantes que les individus malades dans les trois cas que nous avons examinés (*Pax-1*, *Pax-3*, *Pax-6*) : dans tous ces cas, les mutations se manifestent à l'état hétérozygote et ont donc un effet semi-dominant. Cela pourrait s'expliquer par une mutation de type gain de fonction (voir ci-dessus le cas de *un*) ; alternativement, il pourrait s'agir d'un effet de dosage génique, deux allèles actifs du gène sauvage étant nécessaires à un développement normal. Seule l'élucidation complète du fonctionnement des gènes *Pax* permettra de trancher entre les deux hypothèses.

Le réseau des gènes *Pax*

Les gènes *Pax* de la souris, nous l'avons vu, ont en commun d'être exprimés dans le tube neural (sauf *Pax-1*) et dans les structures du mésoderme segmenté. Pour tous, l'expression a lieu très tôt durant le développement et dans la plupart des cas, elle est limitée à la période embryonnaire. L'examen détaillé du profil d'expression tel qu'il a été réalisé pour certains gènes *Pax* est encore plus suggestif de l'implication de ces gènes dans le développement et éventuellement de leurs rôles complémentaires dans la formation de certaines structures. Par exemple, si l'on considère le tube neural, l'expression des différents gènes *Pax* qui a lieu sur toute sa longueur a un profil bien particulier à la fois dans le temps — certains de ces gènes étant exprimés très tôt au moment de l'intense

activité mitotique dans le neuro-épithélium — et dans l'espace puis- que l'on observe également une expression différentielle de différents gènes *Pax* le long de l'axe dorso-ventral du tube neural [2].

De même l'expression de *Pax-3* dans certaines cellules dérivées de la crête neurale [6] et les défauts observés dans ces mêmes cellules chez les mutants de souris *Spotch* [7] ou chez les malades atteints du syndrome de Waardenburg [8, 9] suggèrent très fortement un rôle de *Pax-3* dans la fonction de ces cellules. A cet égard, il est intéressant de noter que Moase et Trasler [17] ont mis en évidence, chez le mutant *Spotch*, un retard de migration des cellules de la crête neurale. Ils ont en outre mis à jour chez ce mutant des altérations de la synthèse de la protéine d'adhésion N-CAM qui est, comme *Pax-3*, exprimée fortement dans le tube neural et qui pourrait être impliquée dans le processus de migration cellulaire à partir de la crête neurale. Il est donc tentant de supposer que l'un des rôles de *Pax-3* serait de contrôler l'expression de N-CAM [17]. L'exemple de l'organogenèse rénale est également très suggestif : la formation de cet organe implique des relations d'induction entre le canal néphritique et le mésenchyme néphrogénique ; or, *Pax-8* et *Pax-2*, deux gènes exprimés spécifiquement dans le rein embryonnaire et qui ont par ailleurs une grande similarité structurale, sont exprimés l'un uniquement dans le tissu répondeur (*Pax-8*) et l'autre à la fois dans les tissus inducteurs et répondeurs (*Pax-2*) [13].

Les propriétés régulatrices des protéines Pax

La structure des protéines Pax avec leur domaine *paired*, parfois associé à un domaine homéo en fait des protéines capables de liaison spécifique avec l'ADN et susceptibles d'agir comme des facteurs de transcription [4, 6, 14, 15]. Tout récemment, une preuve directe en a été apportée pour *Pax-5* et *Pax-8* [15]. Le cas de *Pax-5* est particulier puisque son identification comme facteur de transcription est passée par le détour de l'oursin et de gènes codant pour les histones. En effet, chez cet animal, un facteur

de transcription, TSAP, avait été trouvé, contrôlant l'expression des gènes d'histone H2A-2 et H2B-2. En examinant par la suite différentes lignées cellulaires murines et humaines dans des expériences de retard sur gel avec les séquences de reconnaissance de TSAP, une protéine ayant une activité de liaison à l'ADN et homologue de TSAP, appelée BSAP, a été trouvée dans une lignée de cellules B humaine. Le clonage puis le séquençage de l'ADNc de BSAP a révélé qu'il codait pour une protéine identique à la protéine *Pax-5* [3, 14] et possédait donc un domaine homéo, une séquence octamère et une séquence de 23 acides aminés homologue de la partie NH₂-terminale du domaine homéo de type *paired*, à l'instar de *Pax-2* et *Pax-8*. Le profil d'expression de *Pax-5* est, lui aussi, très intéressant : chez l'embryon ainsi que chez le fœtus, il est exprimé entre les jours 9,5 et 15 dans des régions particulières du cerveau et le long du tube neural et, chez l'adulte, dans les organes lymphoïdes, uniquement dans les lignages des cellules B, et dans le testicule [14]. Enfin, un premier pas vient d'être fait dans l'identification de gènes cibles des protéines *Pax*, puisqu'il a été montré que *Pax-5* contrôlait l'expression du gène codant pour CD19, protéine spécifique des lymphocytes B [14], et que *Pax-8* contrôlait l'expression de deux gènes exclusivement exprimés dans la thyroïde [15].

Ce faisceau remarquable d'observations établit la famille des gènes *Pax* comme une nouvelle famille de gènes du développement possédant des motifs conservés dans l'évolution et dont l'étude devrait permettre d'aborder, au niveau moléculaire, certains aspects du développement embryonnaire. Le champ qui s'ouvre ainsi est à l'évidence très vaste ; il va falloir comprendre comment les gènes *Pax* se situent dans le réseau de régulations et d'activations géniques qui permet le développement des structures dans lesquelles ils s'expriment comme, par exemple, le tube neural ou le rein. Une première preuve du rôle régulateur des protéines Pax a été apportée, mais un long chemin reste à parcourir pour comprendre la fonction

moléculaire des différents domaines des protéines Pax (attachement à l'ADN, transactivation, coopération éventuelle avec d'autres protéines régulatrices, etc.) et identifier les gènes dont elles modulent l'expression, concourant ainsi à la morphogenèse des organes dans lesquelles elles s'expriment ■

Charles Babinet

Unité de génétique des mammifères, Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, France.

RÉFÉRENCES

1. Lepesant JA. 1992. Le paradigme de la drosophile dans l'étude du développement. *médecine/sciences* 1992 ; 8 : 650-2.
2. Gruss P, Walther C. Pax in development. *Cell* 1992 ; 69 : 719-22.
3. Walther C, Guénet JL, Simon D, et al. Pax : A murine multigene family of paired box-containing genes. *Genomics* 1991 ; 11 : 424-34.
4. Chalepakis G, Fritsch R, Fickenscher H, Deutsch U, Goulding M, Gruss P. The molecular basis of the *undulated/Pax-1* mutation. *Cell* 1991 ; 66 : 873-84.
5. Balling R, Lau CF, Dietrich S, Wallin J, Gruss P. Development of the skeletal system. Postimplantation development in the mouse. Ciba Foundation Symposium 165 Chickester : Wiley, 1992 : 132-43.
6. Goulding MD, Chalepakis G, Deutsch U, Erselius JR, Gruss P. *Pax-3*, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J* ; 1991 ; 10 : 1135-47.
7. Epstein DJ, Vekemans M, Gros P. *Spotch* (*Sp^{2H}*), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of *Pax-3*. *Cell* 1991 ; 67 : 767-74.
8. Baldwin CT, Hoth CF, Amos JA, da Silva, EO, Milunsky A. An exonic mutation in the *HuP2* paired domain gene causes Waardenburg's syndrome. *Nature* 1992 ; 355 : 637-8.

TIRÉS A PART

C. Babinet.

RÉFÉRENCES

9. Tassabehji M, Read AP, Newton VE, *et al.* Waardenburg's syndrome patients have mutations in the human homologue of the *Pax-3* paired box gene. *Nature* 1992 ; 355 : 635-6.
10. Walther C, Gruss P. *Pax-6*, a murine paired box gene, is expressed in the developing CNS. *Development* 1991 ; 113 : 1435-49.
11. Hill RE, Favor J, Hogan BLM, *et al.* Mouse Small eye results from mutations in a paired-like homeobox-containing gene. *Nature* 1991 ; 354 : 522-5.
12. Ton CCT, Hirvonen IL, Miwa H, *et al.* Positional cloning and characterization of a paired box - and homeobox - containing gene from the aniridia region. *Cell* 1991 ; 67 : 1059-74.
13. Plachov D, Chowdhury K, Walther C, Simon D, Guénct JL, Gruss P. *Pax-8* a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development* 1990 ; 110 : 643-51.
14. Adams B, Dörfler P, Aguzzi A, *et al.* Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes Dev* 1992 ; 6 : 1589-607.
15. Zannini M, Francis-Lang H, Plachov D, Di Lauro. Pax-8, a paired domain-containing protein, binds to a sequence overlapping the recognition site of a homeo-domain and activates transcription from two thyroid-specific promoters. *Mol Cell Biol* 1992 ; 12 : 4230-41.
16. Pilz A, Povey S, Gruss P, Abbott C. Mapping of the human homologues of the murine paired box containing genes. *Genomics* 1993 (sous presse).
17. Moase CE, Trasler DG. N-CAM alterations in splotch neural tube defect mouse embryos. *Development* 1991 ; 113 : 1049-58.