

## À la recherche de gènes impliqués dans le développement embryonnaire de la souris : le « piégeage » de gènes

Anne Camus  
Jacqueline Barra  
Charles Babinet

La recherche de gènes essentiels au développement embryonnaire de la souris a été abordée par la mise en œuvre de stratégies variées. Comparée à toutes ces approches, la stratégie de piégeage de gènes présente un important avantage : elle associe l'isolement et la caractérisation moléculaire d'un gène à l'étude de son profil d'expression ainsi qu'à la détermination de sa fonction *in vivo*. Au-delà d'un moyen d'identification des gènes impliqués dans le développement de la souris, cette stratégie ouvre la voie à la constitution de véritables banques de cellules, étiquetées chacune dans un gène différent, base de départ nécessitant ensuite une analyse détaillée gène par gène. Par ailleurs, les informations relatives au profil d'expression des gènes piégés devraient permettre d'élargir l'intérêt du piégeage de gènes à la production d'animaux chimères ou la réalisation de greffes. En outre, le piégeage d'une activité transcriptionnelle particulière pourrait servir à l'expression ectopique de gènes d'intérêt ou à la destruction spécifique d'une structure embryonnaire ou d'un lignage cellulaire donnés.

### ADRESSES

A. Camus : stagiaire postdoctorale. Laboratoire du docteur P. Tam, Children's Medical Research Institute, Locked Bag 23, Wentworthville, NSW 2145, Australie. J. Barra : chargée de recherche à l'Institut Pasteur. C. Babinet : directeur de recherche au Cnrs, directeur de l'Unité de Biologie du Développement de l'Institut Pasteur. Unité de biologie du développement, Ura Cnrs 1960, Institut Pasteur, 25, rue du Docteur-Roux, 75724 Paris Cedex 15, France.

La compréhension des mécanismes moléculaires qui sous-tendent l'organisation et le développement de l'embryon de souris requiert l'identification des gènes impliqués, la caractérisation de leurs profils d'expression, ainsi que la détermination de leurs rôles au cours du développement. Chez la souris, environ mille gènes ont été identifiés, par des méthodes

génétiques ou biochimiques [1], ce qui ne représente qu'une petite fraction des 50 à 100 000 gènes estimés faire partie du génome des mammifères. Il est donc important de développer de nouvelles stratégies permettant l'accès, si possible sans restriction, à l'ensemble des gènes. Au cours des dernières années, la recherche de gènes essentiels au développement embryonnaire de la

souris a été abordée par la mise en œuvre de stratégies variées. La plus répandue a certainement été celle qui a consisté à identifier de nouveaux gènes chez la souris sur la base de leurs similitudes de séquences avec des gènes déjà caractérisés chez d'autres organismes comme la drosophile, le nématode ou le xénope (pour revue, voir [2]). Souvent, la fonction même de ces gènes a été conservée au cours de l'évolution: ainsi, les connaissances acquises chez un organisme peuvent aider à la compréhension des mécanismes contrôlant certains aspects du développement chez un autre (pour revue, voir [3]).

D'autres voies de recherche des gènes impliqués dans des phénomènes de différenciation cellulaire et de développement consistent en la réalisation de banques soustractives d'ADN complémentaires (ADNc) ou encore en des expériences de *differential mRNA display*. Ces techniques permettent, en principe, l'isolement d'ARNm représentés spécifiquement dans un type cellulaire donné, dans un organe particulier, ou à un stade donné du développement embryonnaire. Ces banques peuvent être élaborées à partir de cellules différenciées et de leurs précurseurs comme, par exemple, les cellules souches embryonnaires et les cellules différenciées provenant de corps embryonnaires [4]. Une autre approche de l'étude de la différenciation cellulaire et du développement réside en l'utilisation de banques soustraites à partir de tissus différents, tels que l'ectoderme, le mésoderme, l'endoderme et la ligne primitive composant l'embryon au moment de la gastrulation [5]. Usant au départ de la même procédure, B. Bettenhausen et A. Gossler ont utilisé des séquences spécifiques d'une banque d'ADNc d'embryons de 10,5 jours de développement pour réaliser des hybridations *in situ*, sur des embryons de même stade, et ainsi détecter et isoler les clones donnant des profils restreints à un type cellulaire ou spécifiques d'une région particulière de l'embryon [6]. L'ensemble de ces stratégies permet l'isolement de gènes, connus ou inconnus, mais présente une limitation importante: en effet, ces stratégies ne donnent aucun renseigne-

ment sur la fonction biologique des gènes isolés. Il faut pour cela utiliser une étape supplémentaire d'inactivation de ces gènes chez l'animal.

Comparée à toutes ces approches, la stratégie de piégeage de gènes, développée chez la souris à l'origine par Gossler *et al.* et ultérieurement par plusieurs autres laboratoires [7-10], présente un important avantage, car elle associe l'isolement et la caractérisation moléculaire d'un gène à l'étude de son profil d'expression ainsi qu'à la détermination de sa fonction *in vivo*. Cette stratégie repose sur l'introduction de « pièges » à gènes dans le génome des cellules embryonnaires souches (cellules ES): celles-ci ont la capacité remarquable de coloniser un embryon hôte, y compris sa lignée germinale, ce qui permet le transfert à l'animal d'une modification génétique introduite dans leur génome [11]. Cette stratégie constitue ainsi une extension naturelle de la mutagenèse par insertion mais permet, en outre, de muter spécifiquement et de manière délibérée les régions du génome transcriptionnellement actives.

## Le piégeage de gènes

Le principe général des stratégies de piégeage, qu'il s'agisse de piéger des séquences régulatrices (*enhancer-trap*) ou des gènes (*gene-trap*), repose sur l'introduction dans le génome d'un gène rapporteur dépourvu d'une partie ou de la totalité de ses éléments de contrôle transcriptionnel. L'expression du gène rapporteur dépendra donc des éléments de régulation agissant *in cis* de l'unité de transcription à l'intérieur ou à proximité de laquelle l'insertion aura eu lieu et sera donc le reflet fidèle de l'expression de cette dernière. A la suite de certains événements d'intégration, la « construction-piège » aura pu interrompre un gène endogène et donc se comporter comme un agent mutagène. L'insertion de cette séquence étrangère représente un « hameçon » moléculaire et facilite ainsi le clonage ultérieur du gène muté.

### Les vecteurs de piégeage

De façon générale, tous les vecteurs de piégeage possèdent, d'une part,

un gène rapporteur qui d'ordinaire est le gène bactérien *lacZ* (son activité est facilement décelable et n'est pas affectée par de larges fusions à l'extrémité amino-terminale) et, d'autre part, un gène codant pour une protéine de résistance à un facteur de sélection, généralement la néomycine phospho-transférase. On trouve trois catégories principales de constructions-pièges qui sont présentées schématiquement dans la *figure 1* avec, d'une part, les séquences génomiques minimales requises pour leur transcription et, d'autre part, les produits protéiques synthétisés en conséquence.

On distingue tout d'abord le piège à séquences régulatrices (*figure 1A*) qui met en œuvre un gène rapporteur, portant un codon ATG de début de traduction et un signal de polyadénylation, dirigé par un promoteur minimal (idéalement une boîte TATA et un site de début d'initiation de transcription). Le gène rapporteur ne sera donc transcrit que s'il s'est intégré au voisinage de régions activatrices agissant *in cis*.

Il existe ensuite les pièges à promoteurs, qui contiennent le gène rapporteur (avec ou sans codon ATG) dépourvu de séquence promotrice (*figure 1B*). Pour qu'ils soient fonctionnels, les pièges à promoteurs doivent s'intégrer dans un exon (dans la bonne orientation et dans la phase de lecture du gène), ou encore dans la partie 5' terminale transcrite mais non traduite du gène si le vecteur possède un codon ATG. Lorsque l'intégration d'un piège à promoteurs survient dans un intron, aucune expression n'est observée puisque le processus d'épissage des ARN entraîne son élimination.

L'avantage du piège à gènes par rapport au piège à promoteurs est qu'il possède, fusionné en 5' terminal du gène rapporteur, un site accepteur d'épissage (*figure 1C*). Des travaux antérieurs avaient montré que des sites donneurs et accepteurs d'origine différente peuvent interagir fonctionnellement [12]. Ainsi, dans les cas où l'insertion s'effectue dans un intron, la présence du site accepteur d'épissage dans la construction permet d'engendrer un transcrit de fusion comprenant une partie de longueur variable du gène endogène suivie du gène rapporteur. Le fonc-

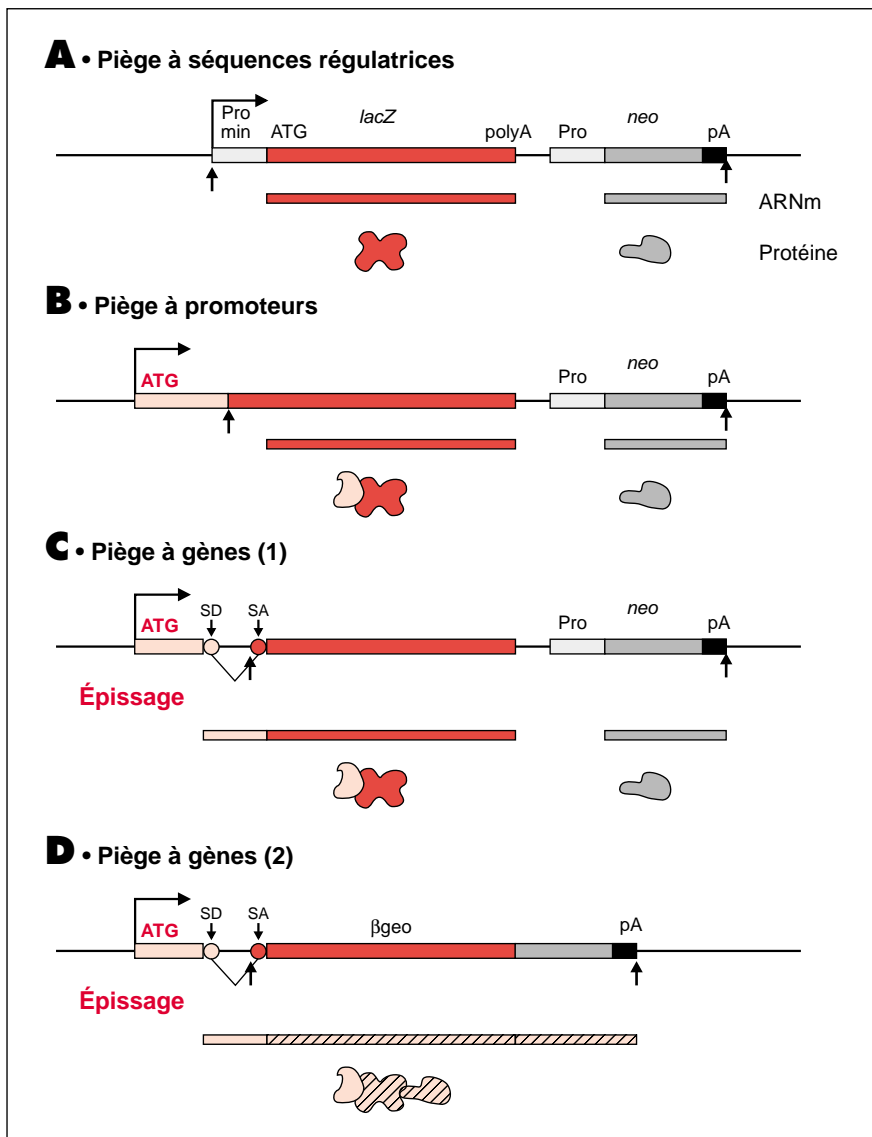


Figure 1. **Présentation des principaux types de construction utilisés.** Les extrémités de chacune des constructions, représentées après intégration dans le génome, sont indiquées par des flèches verticales. Chaque construction comprend un gène rapporteur, la  $\beta$ -galactosidase (rectangles rouges) et un gène de sélection, ici un gène de résistance à la néomycine (neo, rectangles grisés). Dans le cas de la construction A, l'expression du gène rapporteur dépend de la présence, dans les séquences génomiques flanquantes, de séquences régulatrices (schématisées par un rectangle R); dans le cas des constructions B, C et D, l'expression dépend de l'intégration fonctionnelle à l'intérieur d'un gène. Pour le piège à promoteurs (B), cette expression dépend de l'intégration en phase dans une séquence exonique (ou dans les régions non codantes, mais dans ce cas le gène rapporteur doit posséder son propre site ATG de début de la traduction). Pour les pièges à gènes (C et D), l'adjonction d'un site accepteur d'épissage en 5' de la construction permet l'expression du gène rapporteur même quand l'intégration a lieu dans un intron. Dans les constructions A, B et C l'expression de neo dépend d'un promoteur (Pro) actif dans les cellules ES; dans la construction D, au contraire, l'expression de neo qui est fusionné aux séquences codantes LacZ ( $\beta$  geo) dépend du promoteur du gène dans lequel la construction s'est intégrée. Rectangles roses hachurés: exon endogène; rectangles rouges: gène lacZ; Pro: promoteur constitutif; Pro min: promoteur minimal; PolyA: signal de polyadénylation; SA: site accepteur d'épissage; SD: donneur d'épissage.

tionnement des pièges à gènes, comme celui d'ailleurs des pièges à promoteurs, dépend évidemment du sens de leur intégration. S'il n'y a pas de codon ATG dans le gène rapporteur, celui-ci ne sera exprimé que si l'événement d'épissage le place dans la phase de lecture du gène endogène.

L'organisation des premiers vecteurs plaçait le gène de sélection sous le contrôle indépendant d'un promoteur constitutivement actif dans les cellules ES, tel celui des gènes thymidine kinase (*tk*), phosphoglycérate kinase (*pgk-1*) ou  $\beta$ -actine humaine (*hba*) [7, 8, 10]. Cela permettait de sélectionner positivement toutes les cellules ayant intégré la construction introduite mais exigeait un second criblage pour repérer les événements-pièges fonctionnels. Plus récemment, la création d'une construction-piège, composée d'un gène de fusion entre le gène conférant la résistance à un antibiotique et le gène *lacZ*, a rendu possible la sélection directe et exclusive des événements d'insertions fonctionnels, autrement dit, dépendants de l'activité d'un promoteur endogène [8, 13]. Dans ce cas de figure, tout clone cellulaire sélectionné *in vitro* pour sa résistance à l'antibiotique correspondra à une insertion dans un gène exprimé dans ces cellules ES et synthétisera la  $\beta$ -galactosidase (figure 1D).

#### Mise en œuvre de la stratégie de piégeage

L'introduction du vecteur-piège dans les cellules ES se fait généralement par électroporation de la construction plasmidique linéarisée [14] ou par infection à l'aide de rétrovirus défectueux MuLV (pour *murine leukemia virus*), c'est-à-dire dont les séquences promotrices et *enhancer* du LTR (*long terminal repeat*) ont été délétées [15]. Les événements de piégeage sont repérés *in vitro* et les insertions choisies, selon des critères qui sont développés ultérieurement. Ces mutations peuvent aussi être introduites chez l'animal, *via* l'obtention de chimères, engendrées par l'injection de quelques cellules dans des blastocystes qui sont ensuite implantés chez des souris porteuses. Les premières expériences de piégeage chez la souris ont été publiées

par les groupes de A. Joyner et J. Rosant en 1989 [7]. Deux types de vecteurs étaient utilisés : un piège à séquences régulatrices et un piège à gènes. Si le rendement en événements de piégeage est plus élevé avec le premier, les événements mutationnels y sont aussi plus rares. La lignée Etl1, issue de ce travail et étudiée en détail par la suite [16], en est un bon exemple : la construction piège à séquences régulatrices s'est insérée à une distance de 2,5 kb d'un gène mais sans qu'il en résulte une perturbation de la fonction.

En revanche, dans le cas du piège à gènes, le recrutement « obligé » d'un promoteur de gène endogène entraîne, dans la majorité des cas, la création d'un transcrite hybride composé d'une portion du gène interrompu, ce qui en facilite le clonage. Ainsi, le piège à gènes constitue un puissant mutagène. C'est principalement pour ces deux raisons que l'utilisation de constructions piège à gènes tend à se généraliser aux dépens des pièges à séquences régulatrices.

En résumé, la stratégie de piégeage de gènes dans les cellules ES se dis-

tingue par trois avantages importants.

1. Le gène rapporteur *lacZ* d'*Escherichia coli*, qui code pour la  $\beta$ -galactosidase (dont l'activité est révélée par une coloration bleue en présence d'un substrat chromogène, le X-gal), permet de suivre dans les embryons chimères le profil d'expression du gène dans lequel la construction s'est insérée.

2. L'événement d'insertion a la potentialité d'être mutagène ; il entraîne en effet l'interruption de l'unité de transcription. L'analyse du phénotype des souris porteuses de la mutation à l'état hétérozygote et homozygote permet alors d'obtenir des informations sur la fonction du gène piégé.

3. La formation d'un transcrite de fusion entre la région 5' du messager endogène et le gène rapporteur, dans les cellules dans lesquelles le gène piégé est exprimé, permet le clonage du gène concerné par l'utilisation de la technique de 5' RACE PCR (*rapid amplification of cDNA ends*) [17], ou par la construction de banques d'ADNc mises en route à partir de la séquence du gène rapporteur [18].

## Le choix des stratégies de criblage

Lorsqu'ils ont obtenu des clones de cellules ES modifiés génétiquement, les expérimentateurs effectuent une étape de criblage des événements de piégeage afin de choisir ceux dont ils poursuivront l'analyse. Au cours de ces dix dernières années, des évolutions, à la fois dans la conception des vecteurs de piégeage et dans la stratégie de tri des gènes piégés, se sont dessinées que nous allons examiner. Le but des équipes qui se sont intéressées initialement au piégeage de gènes était de trouver de nouveaux gènes nécessaires au développement de l'embryon de souris et de les caractériser. Aussi ont-elles mis en œuvre un criblage direct par l'analyse *in situ* de l'expression du gène rapporteur afin d'identifier les gènes exprimés et réglés au cours du développement embryonnaire. En pratique, cette expression était analysée dans les embryons chimères au cours de la gastrulation ou encore à des stades plus avancés dans l'organogénèse, soit entre les jours 7,5 et 12,5 du développement [7, 8, 19-21].

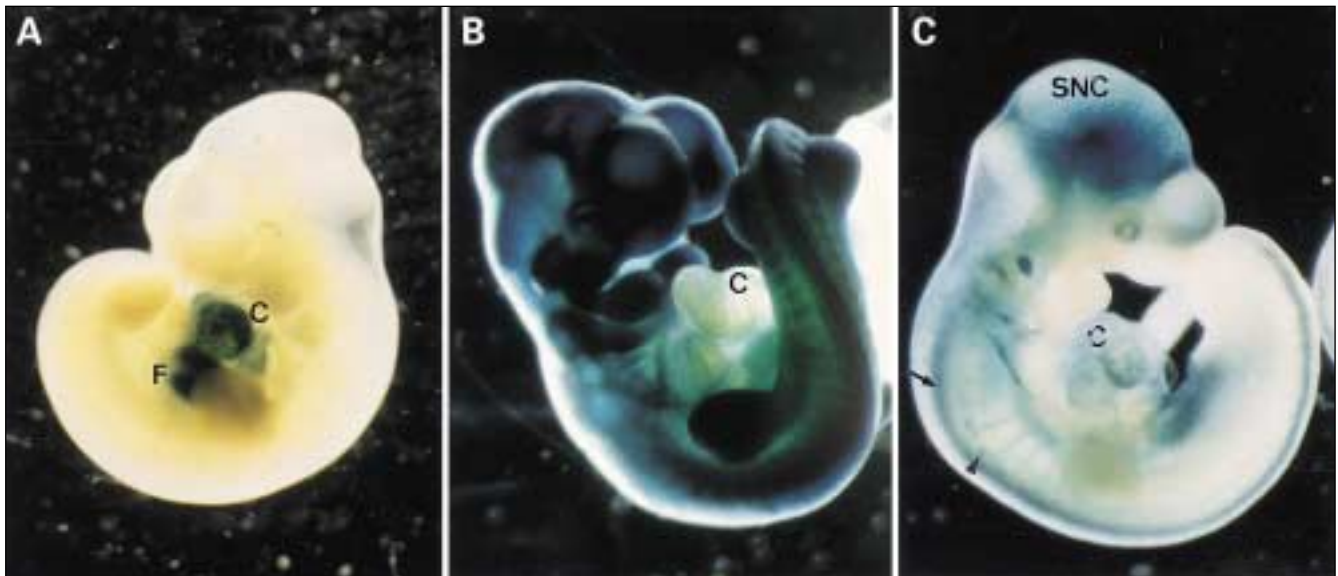


Figure 2. Profil d'expression du gène rapporteur *lacZ* dans des embryons E 10,5 (x 1000) de trois lignées phléal obtenues à partir d'une construction piège codant pour une protéine hybride de résistance à la phléomycine et à la  $\beta$ -galactosidase. **A.** Lignée TK5, l'activité  $\beta$ -galactosidase est observée dans le cœur et le foie, dans des cellules éparées sur l'embryon et dans des cellules hématopoïétiques. Ce profil a été retrouvé chez 45% des clones indépendants obtenus avec cette construction. **B.** Lignée TK14, le gène rapporteur est dans ce cas exprimé de manière très large, presque ubiquitaire. **C.** Lignée TL15, cette lignée présente un marquage plus restreint, spécifique du système nerveux (central et périphérique). C : cœur ; F : foie ; SNC : système nerveux central.

Notons ici que, d'une manière générale, il ne semble pas qu'un tissu particulier ait été exclu du marquage ou très préférentiellement marqué; une exception cependant concerne des observations que nous avons faites avec la construction piège à gènes Phleal [13]. Celle-ci est constituée d'une fusion entre le gène *sh-ble* codant pour une protéine de résistance à la phléomycine et le gène *lacZ*. Nos résultats montrent que 45 % des événements pièges indépendants analysés présentent un profil d'expression très voisin au jour 10,5 du développement (dans le cœur, le foie et la lignée hématopoïétique) (figure 2). La raison de ce biais n'est pas claire mais pourrait être due à une insertion préférentielle du piège à gène dans une région chromosomique particulière.

Si l'on considère l'ensemble des travaux publiés à ce jour, on constate que les résultats obtenus sont généralement concordants et sont à l'image de ceux présentés dans l'étude la plus étendue, celle réalisée par les équipes de A. Joyner et de J. Rossant (Toronto, Canada). Au total, 279 clones indépendants de cellules ES transformées par une construction piège à gènes et exprimant le gène rapporteur *lacZ* ont été criblés. Dans ces clones, les gènes piégés se répartissent de la manière suivante: un tiers présente une expression restreinte à un type tissulaire ou à des structures embryonnaires particulières, un autre tiers des gènes possède une expression ubiquitaire ou touchant un grand nombre de tissus différents et le dernier tiers comprend des gènes pour lesquels aucun marquage n'est observé aux différents stades de développement étudiés [21].

De manière générale, forts de l'idée que l'expression du gène rapporteur reflète l'expression du gène endogène dans lequel l'insertion s'est opérée, les expérimentateurs se sont intéressés aux gènes piégés dont l'analyse de l'expression dans les embryons chimères révélait des profils réglés de façon dynamique au cours du développement. Cette étape leur permettait de sélectionner, avant l'analyse moléculaire, un gène dont l'expression touchait un lignage cellulaire précis, un tissu ou une structure embryonnaire particulière.

Cette démarche a permis, par exemple, à une équipe japonaise s'intéressant plus particulièrement aux gènes impliqués dans la morphogénèse du cerveau, d'isoler *jumonji*, un nouveau gène essentiel pour la formation du tube neural [22].

Ce type de présélection possède ses propres limites. En effet, les informations apportées par les expériences de mutations par recombinaison homologue nous suggèrent que le profil d'expression d'un gène n'est pas nécessairement un bon indicateur de sa fonction [23]. Une étape supplémentaire est donc nécessaire, à savoir l'établissement de lignées de souris afin de repérer les phénotypes des homozygotes avant de décider quel gène cloner. Les études utilisant ce crible requièrent un travail intensif, des moyens importants, particulièrement en animalerie, et demandent généralement un temps très long entre l'expérience de transfection et la caractérisation d'un gène. Les expérimentateurs s'orientent donc actuellement vers d'autres modes de criblage que nous allons décrire.

### Tri après différenciation

Dans ce cas, le crible consiste à analyser l'expression du gène rapporteur dans les cellules ES avant et après différenciation *ex vivo*. Il est établi depuis longtemps que les cellules ES, privées de facteurs inhibant leur différenciation (comme le LIF, *leukemia inhibitory factor*) et cultivées dans certaines conditions, ont la propriété de former des corps embryonnaires composés d'ectoderme et d'endoderme. Ces corps embryonnaires peuvent à leur tour évoluer en un grand nombre de types cellulaires différenciés incluant les lignages hématopoïétiques, cardiaques et nerveux, mais de manière quelque peu chaotique et difficile à reproduire [26, 27]. Cependant, dans les années récentes, des conditions de culture ont commencé à être mises au point permettant d'induire, majoritairement et de manière reproductible, la différenciation des cellules ES vers un lignage particulier [28-33]. L'analyse de l'expression du gène rapporteur au cours de ces cultures permet alors de détecter les gènes piégés, réprimés dans les cellules ES, mais

activés lors de la différenciation. Dans certains cas, les auteurs repèrent simplement des gènes activés ou inhibés après différenciation [34-36]. Une autre approche pourrait consister à utiliser deux vecteurs: un premier marquant spécifiquement un type cellulaire donné à l'aide d'un gène rapporteur dirigé par un promoteur spécifique du tissu choisi (voir par exemple [33] pour le cas des myocytes cardiaques) et un second portant une construction piège comportant un autre gène rapporteur ainsi qu'un gène de sélection dirigé par un promoteur constitutif. Si dans un clone de cellules ES doublement transformées et engagées dans une voie de différenciation, il est possible d'observer la co-expression des deux types de marqueurs, tout indique alors le piégeage de gènes activés au cours du processus de différenciation ou bien de gènes spécifiques du type cellulaire étudié. Une approche tout à fait similaire consiste à effectuer un marquage avec des anticorps spécifiques d'un type cellulaire donné. Il suffit alors de rechercher la co-expression du gène rapporteur contenu dans le piège à gènes avec ces marqueurs pour identifier des gènes exprimés spécifiquement au cours de la différenciation du type cellulaire considéré. Cette approche a été utilisée dans le cas des cellules nerveuses [37], des cellules neuronales et gliales, des chondrocytes ou des myocytes [38]. Dans ces deux expériences, chacun des gènes sélectionnés *in vitro* dans les cellules ES différenciées s'est révélé exprimé *in vivo* dans les tissus prédits.

### Tri après réponse à un inducteur

Un second crible peut consister en la sélection *ex vivo* de cellules ES chez lesquelles l'événement de piégeage implique des gènes inductibles. En effet, il est possible de faire agir une substance ou un facteur de croissance dans le milieu de culture des cellules ES: l'examen de l'activité du gène rapporteur, avant et après ajout de l'inducteur, permet alors d'identifier des gènes inductibles ou répressibles directement par cet inducteur ou, indirectement, en aval d'une cascade de transmission du signal cellulaire. Plusieurs travaux, faisant sou-

vent appel à des systèmes expérimentaux de sélection ingénieux, utilisent cette approche. L'acide rétinoïque a été utilisé pour sélectionner des gènes dont l'activité transcriptionnelle est influencée positivement ou négativement par cet agent dans les cellules ES [39]. Après ce criblage, les expérimentateurs montrent que 95 % des gènes piégés ont un profil d'expression restreint *in vivo* et réglé (entre les jours 8,5 et 11,5 de développement) contre les 30 % observés ordinairement sans cette induction. De plus, aucun clone ne conduit à l'expression constitutive du gène rapporteur alors qu'un pourcentage relativement élevé (environ 30 %) est observé lors d'un crible direct de gènes actifs dans les cellules ES [8, 21].

Il faut mentionner également ici des expériences effectuées dans d'autres types cellulaires que les cellules ES mais qui pourraient être appliquées à ces dernières; en effet, elles illustrent l'utilisation de nouveaux types de vecteurs de piégeage comportant des systèmes particuliers de sélection, destinés à identifier des gènes induits ou réprimés dans une séquence donnée d'événements cellulaires. Ainsi, on peut citer la mise au point d'un système de sélection conditionnel permettant de piéger les gènes impliqués dans le processus apoptotique dans une lignée de cellules hématopoïétiques cultivées en l'absence d'interleukine-3 [40]. De même, la stratégie de piégeage a été mise en œuvre pour isoler les gènes réglés par les glucocorticoïdes dans des cellules hypophysaires [41], par les lipopolysaccharides dans les cellules du lignage lymphoïde B [42] ou par MyoD dans une lignée fibroblastique [43].

### Constructions-pièges pour la sélection de gènes codant pour un type particulier de protéines

Toutes les approches que nous avons décrites jusqu'à présent étaient fondées sur l'intégration du piège à gènes dans une unité de transcription, quel que soit par ailleurs le type de protéine codée par cette dernière. Une alternative intéressante consisterait à utiliser des constructions-pièges conçues pour piéger spécifiquement des gènes codant pour un type particulier de protéines. Un bon exemple

de cette approche originale est fourni par l'étude de Skarnes *et al.* [47]. Ceux-ci cherchaient à identifier des gènes codant pour des protéines sécrétées ou transmembranaires, dont un nombre croissant d'études ont montré ces dernières années l'importance dans le développement précoce de l'embryon de souris (voir par exemple dans [48-50]). Ils ont tout d'abord démontré que les vecteurs pièges classiques mais porteurs en 5' d'une séquence signal (SS)

entraînent une inactivation de la  $\beta$ -galactosidase (probablement liée à une orientation non fonctionnelle de celle-ci vers la lumière du réticulum endoplasmique) et que, par conséquent, les gènes piégés contenant cette séquence sont exclus des cribles classiques (figure 3). En revanche, l'adjonction dans le vecteur piège d'une séquence transmembranaire (TM) permet de restaurer l'activité de la  $\beta$ -galactosidase qui est cette fois ancrée dans la membrane du réticu-

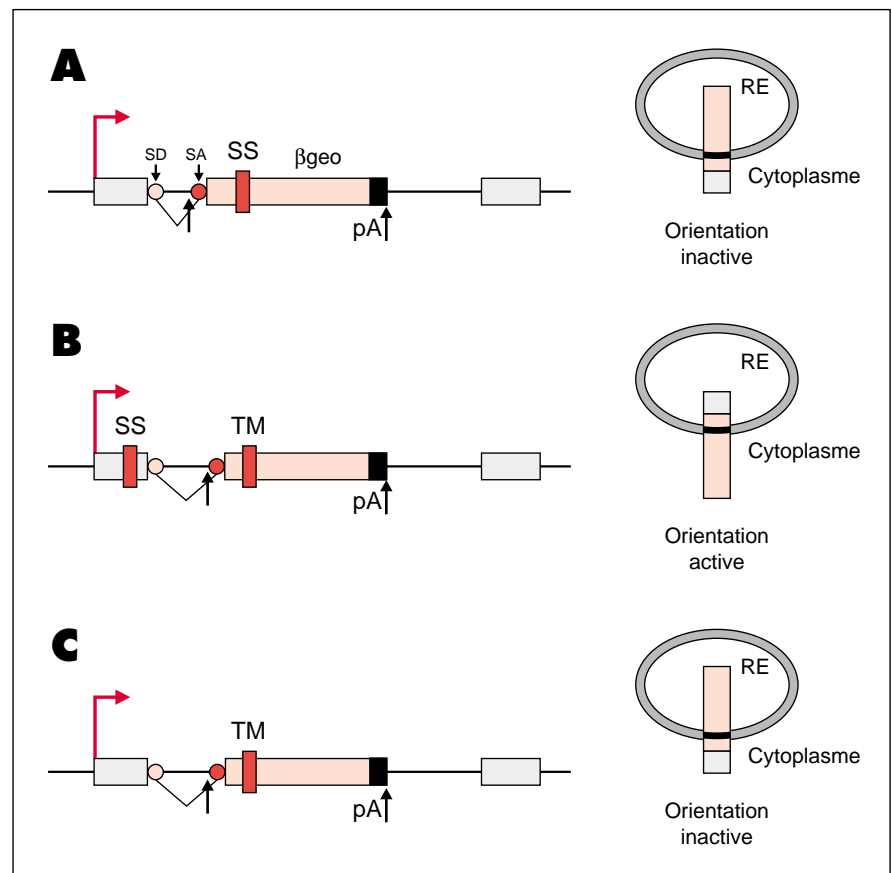


Figure 3. **Vecteurs réalisés dans le but de piéger spécifiquement des gènes codant pour des protéines sécrétées ou ancrées dans la membrane.** **A.** Construction classique porteuse en 5' d'une séquence signal (SS) qui est responsable de la localisation de la  $\beta$ -galactosidase qui est fusionnée au gène néo ( $\beta$ geo) dans la lumière du réticulum endoplasmique (RE). Dans cette situation, l'enzyme est inactive. **B.** Construction utilisée pour piéger les protéines excrétées ou membranaires. L'ajout d'un segment (TM) transmembranaire à une construction classique permet, si cette construction est insérée en aval d'une séquence signal, de repérer l'activité  $\beta$ -galactosidase puisque, dans ce cas, la protéine est ancrée dans la membrane du réticulum endoplasmique mais orientée vers le cytoplasme: ainsi, l'enzyme conserve son activité. **C.** Lorsque le gène piégé ne contient pas de séquence signal, la séquence transmembranaire oriente la  $\beta$ -galactosidase dans la lumière du réticulum endoplasmique la rendant inactive. Les flèches verticales indiquent les limites de la construction. Rectangles roses: construction-piège; rectangles gris: exon endogène; pA: site de polyadénylation; SA: site accepteur d'épissage; SD: site donneur d'épissage. (D'après [47].)

lum, mais orientée vers le cytoplasme (figure 3). Sur la base de ces observations, Skarnes *et al.* ont utilisé un vecteur-piège porteur en 5' d'une séquence TM (figure 3B): ainsi devraient être piégés et spécifiquement identifiés les gènes porteurs d'une séquence SS et donc codant pour des protéines sécrétées ou constituées en partie d'un domaine transmembranaire. De fait, l'analyse de cinq sur six des insertions obtenues a montré qu'elles avaient eu lieu dans ce type de gènes [47].

### Le piégeage généralisé et la création de cellules ES « étiquetées »

Dans les expériences de piégeage que nous avons décrites jusqu'ici, seuls étaient examinés les gènes exprimés dans les cellules ES; cela vaut évidemment lorsque le piège à gènes comprend une fusion du gène de sélection et du gène rapporteur. Dans les pièges dans lesquels les deux gènes sont indépendants, tous les événements d'intégration sont

sélectionnés et seule une minorité de ces événements correspondra à l'intégration dans un gène actif, cette intégration étant identifiée par la coloration bleue en présence de X-gal. Pour les événements n'entraînant pas l'expression du gène rapporteur, la situation est plus problématique, car la majorité de ces événements correspond à une intégration en dehors d'un gène. Aussi, pour tourner la difficulté, Yoshida *et al.* (Iharaki, Japon) et Niwa *et al.* (Kumamoto, Japon) ont construit des vecteurs pièges contenant en 5' le gène *lacZ*, précédé d'un site accepteur d'épissage, et dans lesquels la cassette de résistance à la néomycine est dépourvue de séquences de polyadénylation. En conséquence, le piège à gènes doit utiliser des séquences de polyadénylation endogènes pour que la résistance à la néomycine s'exprime [24, 25]. Cette « astuce » doit donc permettre d'enrichir en événements de piégeage pour lesquels l'intégration du « piège » a eu lieu dans des unités de

transcription, que celles-ci soient exprimées ou non dans les cellules ES. Les résultats obtenus dans ces deux études sont intéressants mais il restait cependant à analyser un plus grand nombre d'insertions afin de vérifier que cette approche permet de piéger des gènes de manière efficace, qu'ils soient ou non exprimés dans les cellules ES. De fait, ce scénario a été validé très récemment dans une étude [63] utilisant des vecteurs bâtis sur le principe que nous venons de décrire (figure 4). De plus, dans cette étude, une fois les clones ES sélectionnés, un système de séquençage direct des produits de RACE-PCR a été mis au point (voir aussi [46]). Deux mille clones ont été analysés de cette manière correspondant à autant de gènes « étiquetés ». L'analyse comparative, à l'aide des banques de données, des EST (*expressed sequence tags*) ainsi obtenus a montré que 18 % correspondent à des gènes déjà connus, 10 % à des EST humains ou murins, 10 % à des séquences de type B1, B2 ou élé-

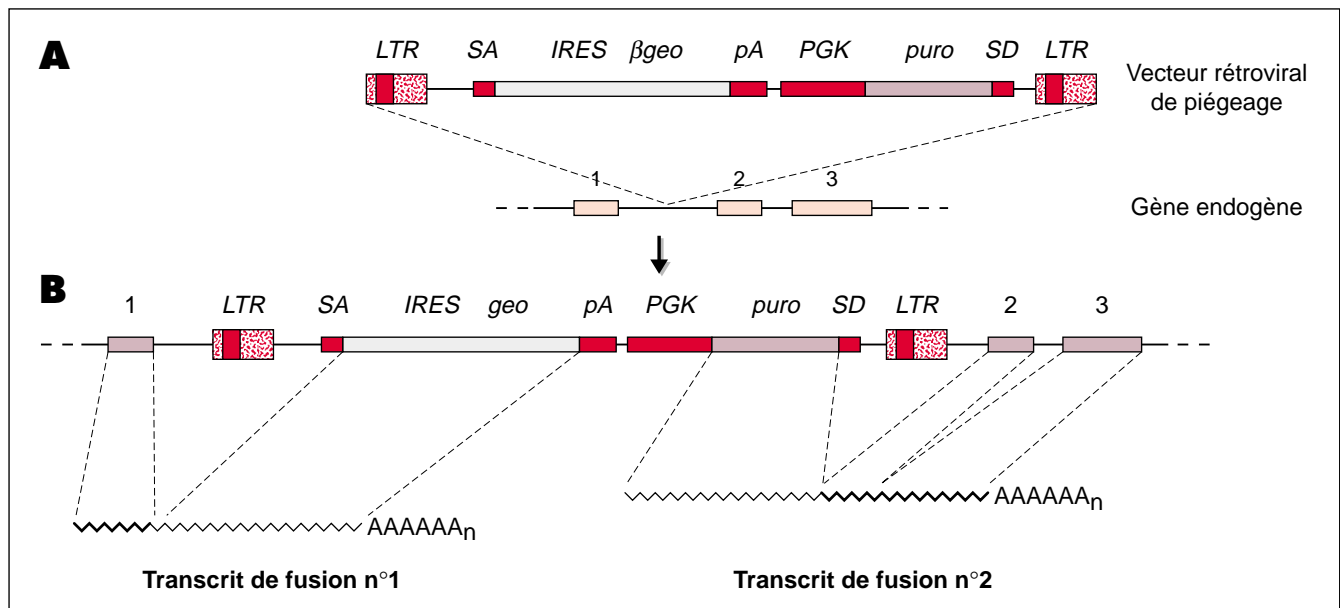


Figure 4. Structure d'un vecteur permettant le piégeage de gènes dans les cellules ES, indépendamment de leur expression. A. Le vecteur utilisé est de type rétroviral (LTR: long terminal repeat) et contient, de 5' en 3', un accepteur d'épissage (SA), une séquence IRES (internal ribosomal entry site), une séquence codant pour une protéine de fusion  $\beta$ -galactosidase/protéine de résistance à la néomycine ( $\beta$ geo) suivie d'un site de polyadénylation (pA), le promoteur de la phosphoglycérate kinase (PGK) suivi des séquences codant pour une protéine de résistance à la puromycine (puro) et, enfin, un site donneur d'épissage (SD). B. Exemple d'intégration du vecteur rétroviral dans le premier intron d'un gène contenant trois exons. Deux transcrits de fusion sont attendus: (1) l'un contenant les séquences 5' du gène piégé et celles de  $\beta$ geo; (2) l'autre contenant les séquences puro et les séquences 3' du gène piégé, y compris le site de polyadénylation. Ainsi, que le gène piégé soit exprimé ou non dans les cellules ES, ce type d'intégration sera sélectionné en présence de puromycine. (D'après [63].)

ments transposables et, enfin, 61 % ne correspondent à aucune séquence contenue dans les banques de données.

## Conclusion

Depuis sa description initiale en 1989 [7], la stratégie de piégeage de gènes chez la souris s'est considérablement développée. De nouveaux vecteurs ont été décrits, des méthodes de sélection pour un type particulier de gènes ont été mises au point et différentes approches pour le criblage de gènes ont été éprouvées. Au total, les prédictions initiales sur la possibilité d'effectuer un très large « balayage » du génome de la souris et, plus spécifiquement, d'identifier de nouveaux gènes du développement ont été largement vérifiées. La majorité, sinon la totalité, des gènes de la souris paraissent susceptibles d'être « étiquetés » par un gène rapporteur et, la plupart du temps, ce dernier semble reproduire fidèlement l'expression du gène piégé. Les résultats globaux obtenus dans des expé-

riences dans lesquelles un nombre (relativement) élevé d'insertions ont été étudiées [8, 21, 44, 45, 51], rendent bien compte de ces potentialités : ainsi, parmi 279 insertions dans des gènes exprimés dans les cellules ES, au moins 30 % concernent des gènes dont l'expression est réglée au cours du développement embryonnaire [21]. Par ailleurs, comme nous venons de le voir, dans une étude extensive concernant plus de 2 000 insertions, le séquençage des EST correspondants a montré que près des deux tiers correspondaient à des séquences non répertoriées dans les banques de données [63]. Enfin, le transfert, *in vivo*, de 42 insertions dans des gènes exprimés dans les cellules ES a montré que près de la moitié d'entre elles entraînait une létalité à l'état homozygote, indiquant le caractère essentiel de ces gènes pour le développement de l'embryon [8, 18]. L'importance de la stratégie de piégeage pour isoler ce type de gène est confirmée par l'examen des résultats obtenus par différents groupes qui ont étudié *in vivo* les gènes qu'ils

avaient piégés (Tableau I). Cinquante pour cent des gènes mutés présentent un phénotype létal embryonnaire.

Nous nous sommes intéressés dans cet article à la mise en œuvre de la stratégie de piégeage de gènes comme moyen d'identification des gènes impliqués dans le développement de la souris. Cependant, il faut noter, au-delà de ce but spécifique, que cette stratégie ouvre la voie à la constitution de véritables « banques » de cellules ES, « étiquetées » chacune dans un gène différent. Du fait des progrès réalisés dans la mise en œuvre des techniques de 5' et 3' RACE-PCR d'une part, et dans les techniques de séquençage, d'autre part [46], ce sont autant d'EST qui pourront être identifiés et caractérisés. Par ailleurs, les informations relatives au profil d'expression des gènes piégés devraient permettre d'élargir l'intérêt du piégeage de gènes à l'utilisation d'une expression particulière du gène rapporteur *lacZ*, dans des études de lignage cellulaire, pour la production d'animaux chimeres ou la réalisation de greffes [52,

Tableau I  
GÈNES ISOLÉS PAR PIÉGEAGE DANS LES CELLULES ES ET ANALYSÉS CHEZ L'ANIMAL

Lignée	Type de construction	Gène et localisation	Phénotype	Références
EI11	piège à séquences régulatrices	nouveau, <i>et 11</i> , famille <i>brahma</i>	viable	[16]
GT4-1	piège à gènes	nouveau, chr* 16	mort à la naissance	[10]
GT4-2	piège à gènes	nouveau, chr 4	retard croissance ou létal	[10]
Fug-1	piège à promoteurs (rétrovirus)	homologue de <i>RNA 1 (S. cerevisiae)</i> , chr 15	létal à jour 6,5	[58]
Rosa 5	piège à gènes (rétrovirus)	<i>TEF-1</i> , chr 15	létal à jour 11,5	[18]
Rosa 3-7	piège à gènes (rétrovirus)	<i>BTF3</i>	létal à jour 6,5	[59]
GT cordon bleu	piège à gènes	nouveau, <i>cob1</i> , chr 11	viable	[60]
Jumonji	piège à gènes	nouveau, <i>jmj</i>	létal à jour 15,5	[22]
J3A3	piège à gènes (rétrovirus)	<i>eck</i> , chr 4	viable	[61]
I.114	piège à gènes	nouveau	viable	[39]
I.163	piège à gènes	nouveau	viable	[39]
I. 193	piège à gènes	nouveau	viable	[39]
R. 140	piège à gènes	nouveau	létal à jour 12,5	[39]
ST 534	piège à gènes ( <i>secretory trap</i> )	<i>lar</i> , chr 7	viable	[47]
ST 531	piège à gènes ( <i>secretory trap</i> )	<i>PTP κ</i> , chr 10	viable	[47]
PGT1.8	piège à gènes	<i>αE caténine</i>	létal au stade blastocyste	[62]

\* chr: chromosome.



53]. De plus, à l'image des systèmes mis en œuvre chez la drosophile, organisme chez lequel les stratégies de piégeage ont permis de grandes avancées dans l'étude génétique [54-56], le piégeage d'une activité transcriptionnelle particulière pourrait servir également chez la souris à l'expression ectopique de gènes d'intérêt ou à la destruction spécifique d'une structure embryonnaire ou d'un lignage cellulaire donné [57].

Il ne faut cependant pas perdre de vue que le but final de la stratégie de piégeage est d'accéder à la fonction des gènes piégés; de ce point de vue, la disposition de banques de cellules ES « étiquetées » constitue certes une base de départ remarquable (profil d'expression et version mutée du gène), mais l'élucidation de la fonction suppose ensuite une analyse détaillée (phénotype lié à la mutation d'insertion, éventuellement production d'autres types de mutations, réseaux de régulation en amont et en aval du gène étudié, mécanismes d'action de la protéine produite, etc.) et cette analyse ne peut évidemment être effectuée que gène par gène, chacun d'entre eux ayant ses spécificités propres ■

## RÉFÉRENCES

- Copeland NG, Jenkins NA, Gilbert DJ, *et al.* A genetic linkage map of the mouse: current applications and future prospects. *Science* 1993; 262: 57-66.
- Kessel M, Gruss P. Murine developmental control genes. *Science* 1990; 249: 374-9.
- Davidson EH. Molecular biology of embryonic development: how far have we come in the last ten years? *BioEssays* 1994; 16: 603-15.
- Poirier F, Chan CT, Timmons PM, *et al.* The murine H19 gene is activated during embryonic stem cell differentiation *in vitro* and at the time of implantation in the developing embryo. *Development* 1991; 113: 1105-14.
- Harrison SM, Dunwoodie SL, Arkell R, Lehrach H, Beddington RS. Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo. *Development* 1995; 121: 2479-89.
- Bettenhausen B, Gossler A. Efficient isolation of novel mouse genes differentially expressed in early postimplantation embryos. *Genomics* 1995; 28: 436-41.
- Gossler A, Joyner AL, Rossant J, Skarnes WC. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* 1989; 244: 463-5.
- Friedrich G, Soriano P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* 1991; 5: 1513-23.
- Von Melchner H, De Gregori JV, Rayburn H, *et al.* Selective disruption of genes expressed in totipotent embryonic stem cells. *Genes Dev* 1992; 6: 919-27.
- Skarnes WC, Auerbach BA, Joyner AL. A gene trap approach in mouse embryonic stem cells: the *LacZ* reporter is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev* 1992; 6: 903-18.
- Babinet C. Les cellules souches embryonnaires de souris: une voie privilégiée de transformation génétique à l'échelle de l'animal. *Med Sci* 1992; 3: 268-75.
- Chu G, Sharp PA. A gene chimaera of SV40 and mouse  $\beta$ -globin is transcribed and properly spliced. *Nature* 1981; 289: 378-82.
- Camus A, Kress C, Babinet C, Barra J. Unexpected behavior of a gene trap vector comprising a fusion between the *Sh ble* and *lacZ* genes. *Mol Reprod Dev* 1996; 45: 255-63.
- Potter H, Weir L, Leder P. Enhancer-dependent expression of human  $\kappa$  immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. *Proc Natl Acad Sci USA* 1984; 81: 7161-5.
- Brenner DG, Lin CS, Cohen SN. Analysis of mammalian cell genetic regulation *in situ* by using retrovirus-derived « portable exons » carrying the *Escherichia coli lacZ* gene. *Proc Natl Acad Sci USA* 1989; 86: 5517-21.
- Soininen R, Schoor M, Henseling U, *et al.* The mouse *Enhancer trap locus 1 (Etl-1)*: a novel mammalian gene related to *Drosophila* and yeast transcriptional regulator genes. *Mech Dev* 1992; 39: 111-23.
- Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 1988; 85: 8998-9002.
- Chen Z, Friedrich GA, Soriano P. Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. *Genes Dev* 1994; 8: 2293-301.
- Korn R, Schoor M, Neuhaus H, *et al.* Enhancer trap integrations in mouse embryonic stem cells give rise to staining patterns in chimaeric embryos with a high frequency and detect endogenous genes. *Mech Dev* 1992; 39: 95-109.
- Schuster-Gossler K, Zachgo J, Soininen R, *et al.* Gene trap integrations in genes active in mouse embryonic stem cells efficiently detect developmentally regulated gene expression. *Transgenics* 1994; 1: 281-91.
- Wurst W, Rossant J, Prideaux V, *et al.* A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice. *Genetics* 1995; 139: 889-99.
- Takeuchi T, Yamazaki Y, Katoh-Fukui Y, *et al.* Gene trap capture of a novel mouse gene, *jumonji*, required for neural tube formation. *Genes Dev* 1995; 9: 1211-22.
- Erickson HP. Gene knockouts of *c-src*, transforming growth factor beta-1, and tenascin suggest superfluous, nonfunctional expression of proteins. *J Cell Biol* 1993; 120: 1079-81.
- Niwa H, Araki K, Kimura S, *et al.* An efficient gene-trap method using poly A trap vectors and characterization of gene-trap events. *J Biochem (Tokyo)* 1993; 113: 343-9.
- Yoshida M, Yagi T, Furuta Y, *et al.* A new strategy of gene trapping in ES cells using 3' RACE. *Transgenic Res* 1995; 4: 277-87.
- Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R. The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* 1985; 87: 27-45.
- Robertson EJ. Embryo derived stem cell lines. In: Robertson EJ, ed. *Teratocarcinoma and embryonic stem cells. A practical approach.* Oxford: IRL Press, 1987: 71-112.
- Strübing C, Ahnert-Hilger G, Shan J, *et al.* Differentiation of pluripotent embryonic stem cells into the neuronal lineage *in vitro* gives rise to mature inhibitory and excitatory neurons. *Mech Dev* 1995; 53: 275-87.
- Pascoe WS, Kemler R, Wood SA. Genes and functions: trapping and targeting in embryonic stem cells. *Biochim Biophys Acta* 1992; 1114: 202-9.
- Snodgrass HR, Schmitt RM, Bruyns E. Embryonic stem cells and *in vitro* hematopoiesis. *J Cell Biochem* 1992; 49: 225-30.
- Keller GM. *In vitro* differentiation of embryonic stem cells. *Curr Opin Cell Biol* 1995; 7: 862-9.
- Baker RK, Lyons GE. Embryonic stem cells and *in vitro* muscle development. *Curr Top Dev Biol* 1996; 33: 263-79.
- Loubet F, Fiszman MY. Les cellules souches embryonnaires: un modèle cellulaire pour l'étude de la différenciation cardiaque normale et pathologique. *Med Sci* 1998; 14: 1072-6.
- Johnson MD, Mahon KA. Detection of genes that are differentially expressed during mouse embryogenesis by genetic trapping strategies. *Am J Obstet Gynecol* 1993; 169: 683-9.
- Scherer CA, Chen J, Nachabeh A, Hopkins N, Ruley HE. Transcriptional specificity of the pluripotent embryonic stem cell. *Cell Growth Diff* 1996; 7: 1393-401.
- Rijkers T, Rütther U. Sequence and expression pattern of an evolutionarily conserved transcript identified by gene trapping. *Biochim Biophys Acta* 1996; 1307: 294-300.
- Shirai M, Miyashita A, Ishii N, *et al.* A gene trap strategy for identifying the genes expressed in the embryonic nervous system. *Zool Sci* 1996; 13: 277-83.
- Baker RK, Haendel MA, Swanson BJ, *et al.* *In vitro* preselection of gene-trapped embryonic stem cell clones for characterizing novel developmentally regulated genes in the mouse. *Dev Biol* 1997; 185: 201-4.

## RÉFÉRENCES

39. Forrester LM, Nagy A, Sam M, *et al.* An induction gene trap screen in embryonic stem cells: Identification of genes that respond to retinoic acid *in vitro*. *Proc Natl Acad Sci USA* 1996; 93: 1677-82.
40. Russ AP, Friedel C, Ballas K, *et al.* Identification of genes induced by factor deprivation in hematopoietic cells undergoing apoptosis using gene-trap mutagenesis and site-specific recombination. *Proc Natl Acad Sci USA* 1996; 93: 15279-84.
41. Harrison RW, Miller C. Functional identification of genes up- and down-regulated by glucocorticoids in AtT-20 pituitary cells using an enhancer trap. *Endocrinology* 1996; 137: 2758-65.
42. Kerr WG, Heller M, Herzenberg LA. Analysis of lipopolysaccharide-response genes in B-lineage cells demonstrates that they have differentiation stage-restricted expression and contain SH2 domain. *Proc Natl Acad Sci USA* 1996; 93: 3947-52.
43. Gogos JA, Lowry W, Karayiorgou M. Selection for retroviral insertions into regulated genes. *J Virol* 1997; 71: 1644-50.
44. Holzschu D, Lapiere L, Neubaum D, Mark WH. A molecular strategy designed for the rapid screening of gene traps based on sequence identity and gene expression pattern in mice. *Transgenic Res* 1997; 6: 97-106.
45. Hicks GC, Shi ERG, Li XM, *et al.* Functional genomics in mice by tagged sequence mutagenesis. *Nat Genet* 1997; 16: 338-44.
46. Townley DJ, Avery BJ, Rosen B, Skarnes WC. Rapid sequence analysis of gene trap integrations to generate a resource of insertional mutations in mice. *Genet Res* 1997; 7: 293-8.
47. Skarnes WC, Moss JE, Hurlley SM, Beddington RSP. Capturing genes encoding cell surface and secreted proteins in mouse embryonic stem cells. *Proc Natl Acad Sci USA* 1995; 92: 6592-6.
48. Yamaguchi TP, Rossant J. Fibroblast growth factors in mammalian development. *Curr Opin Genet Dev* 1995; 5: 485-91.
49. Wall NA, Hogan BLM. TGF- $\beta$  related genes in development. *Curr Opin Genet Dev* 1994; 4: 517-22.
50. Parr BAN McMahon AP. *Wnt* genes and vertebrate development. *Curr Opin Genet Dev* 1994; 4: 523-8.
51. Chowdhury K, Bonaldo P, Torres M, Stoykova A, Gruss P. Evidence for the stochastic integration of gene trap vectors into the mouse germline. *Nucleic Acids Res* 1997; 25: 1531-6.
52. Beddington RS, Morgernstern J, Land H, Hogan A. An *in situ* transgenic enzyme marker for the midgestation mouse embryo and the visualization of inner cell mass clones during early organogenesis. *Development* 1989; 106: 37-46.
53. Wilson V, Manson L, Skarnes WC, Beddington RS. The *T* gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. *Development* 1995; 121: 877-86.
54. Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 1993; 118: 401-5.
55. Sentry JW, Goodwin SF, Milligan CD, *et al.* Reverse genetics of *Drosophila* brain structure and function. *Prog Neurobiol* 1994; 42: 299-308.
56. Ferveur JF, Storkuhl KF, Stocker RF, Greenspan RJ. Genetic feminization of brain structures and changed sexual orientation in male *Drosophila*. *Science* 1995; 267: 902-5.
57. Kaiser K. Second generation enhancer traps. *Curr Biol* 1993; 3: 560-2.
58. DeGregori J, Russ A, von Melchner H, *et al.* A murine homolog of the yeast *RNAI* gene is required for postimplantation development. *Genes Dev* 1994; 8: 265-76.
59. Deng JM, Behringer RR. An insertional mutation in the *BTF3* transcription factor gene leads to an early postimplantation lethality in mice. *Transgenic Res* 1995; 4: 264-9.
60. Gasca S, Hill DP, Klingensmith J, Rossant J. Characterisation of a gene trap insertion into a novel gene, *cordon-bleu*, expressed in axial structures of the gastrulating mouse embryo. *Dev Genet* 1995; 17: 141-54.
61. Chen J, Nachabah A, Scherer C, *et al.* Germ-line inactivation of the murine Eck receptor kinase by gene trap retroviral insertion. *Oncogene* 1996; 12: 979-88.
62. Torres M, Stoykova A, Huber O, *et al.* An  $\alpha$ -E-catenin gene trap mutation defines its function in preimplantation development. *Proc Natl Acad Sci USA* 1997; 94: 901-6.
63. Zambrowicz BP, Friedrich GA, Buxton EC, *et al.* Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature* 1998; 392: 608-11.

## SOCIÉTÉ DE BIOLOGIE

### Adressage des protéines dans les cellules endothéliales

20 janvier 1999  
16 heures

Institut des Cordeliers  
Amphithéâtre Bilski-Pasquier  
15-21, rue de l'École-de-Médecine  
75006 Paris, France

Renseignements  
Secrétariat de la Société de Biologie  
Collège de France  
3, rue d'Ulm, 75231 Paris Cedex 05, France  
Tél./Fax : 01 44 27 13 40

## Summary

### Gene trap strategies for the identification of developmentally important murine genes

The availability of mouse embryonic stem (ES) cells, which have the remarkable ability to colonize the germline of an host embryo has opened the opportunity to transfer genetic modifications, selected in cultured cells, to the living animal. In this paper, is described a strategy of gene identification which uses particular constructs called trapping vectors basically made with a reporter gene, generally  $\beta$ -galactosidase devoid of regulatory sequences, and with a cassette containing a selection marker. After transfer into ES cells, the reporter gene may be expressed only when integrated in the correct orientation into a transcription unit.  $\beta$ -galactosidase expression can then be monitored not only in ES cells but also in chimeric mice, thereby uncovering the pattern of expression *in vivo* of the « trapped » endogenous gene. Thus the trapping strategy has several virtues: (1) the reporter gene reveals the expression profile of the gene in which the construct has been integrated; (2) the construct, which interrupts the transcription unit, behaves as a mutagen; (3) the synthesis of a fusion transcript made of endogenous and reporter gene sequences, facilitates the cloning of the cDNA of the trapped gene. Due to these virtues, this strategy has become a privileged way to identify genes, whose expression is developmentally regulated and therefore likely candidates to be important in development. More generally, this strategy could serve to tag any endogenous gene, and therefore provide a way of creating whole genome EST libraries.

## TIRÉS À PART

C. Babinet.