



dépendante de la méthylation de l'ADN (DmRd) [10]. La méthylation d'une séquence promotrice implique généralement une inhibition de la transcription d'un gène. On pourrait donc envisager de réprimer l'expression d'un gène, par méthylation de ces séquences régulatrices, en utilisant comme support signalétique un ARNdb spécifique d'une séquence donnée. ♦

RNA interference in mammalian cells

REMERCIEMENTS

Nous remercions le Professeur W. Filipowicz et le Docteur J.P. Jost pour leurs conseils et la lecture critique de ce manuscrit.

RÉFÉRENCES

- Bernstein E, Denli AM, Hannon GJ. The rest is silence. *RNA* 2001 ; 7 : 1509-21.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998 ; 391 : 806-11.
- Samuel CE. Antiviral actions of interferons. *Clin Microbiol Rev* 2001 ; 14 : 778-809.
- Billy E, Brondani V, Zhang H, Muller U, Filipowicz W. Specific interference with gene expression induced by long double-stranded RNA in mouse embryonal teratocarcinoma cell Lines. *Proc Natl Acad Sci USA* 2001 ; 98 : 14428-33
- Zamore PD, Tuschl T, Sharp PA, Bartel DP. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 2000 ; 101 : 25-33.
- Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* 2001 ; 294 : 853-8.
- Reinhart BJ, Slack FJ, Basson M, et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 2000 ; 403 : 901-6.
- Nishikura K. A short primer on RNAi: RNA-directed RNA polymerase acts as a key catalyst. *Cell* 2001 ; 107 : 415-8.
- Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001 ; 411 : 494-8
- Mette MF, Aufsatz W, Van der Winden J, Matzke MA, Matzke AJ. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J* 2000 ; 19 : 5194-201.

NOUVELLE

« Le piégeage de gènes » : un outil efficace de la génomique fonctionnelle

Sophie Gomez, Patrice Dubreuil, Sophie Lopez

Inserm U.119,
27 boulevard Leï Roure,
13009 Marseille, France.

> Divers stimulus relevant de mécanismes physiologiques fondamentaux comme la survie, la prolifération, la différenciation cellulaire ou le développement embryonnaire, entraînent la modulation de nombreux gènes au sein des cellules concernées. Dans le but d'identifier les gènes-clé de ces phénomènes, des outils permettant leur analyse à grande échelle ont été développés, qui se fondent sur l'isolement de transcrits dont l'expression varie selon que la cellule est ou n'est pas stimulée. Au vu du nombre restreint de gènes dont la fonction est identifiée, des stratégies d'analyse fonctionnelle du génome deviennent nécessaires.

Une de ces approches, « le piégeage de gènes », permet de révéler des locus

dont la transcription est activée lors d'une stimulation donnée, indépendamment de leur régulation post-transcriptionnelle. Cette technique offre l'avantage de coupler l'étude de la modulation d'un transcrit et la conséquence fonctionnelle de cette variation, elle-même directement liée au stimulus appliqué. Ce couplage est le fondement même de la technique du « piégeage de gènes ». Le principe (Figure 1) est fondé sur l'insertion aléatoire, dans le génome cellulaire, d'un vecteur plasmidique ou rétroviral dépourvu de promoteur propre, et contenant un gène révélateur immédiatement précédé d'un site accepteur d'épissage. Lorsque l'insertion survient dans un locus « actif », en présence du

stimulus étudié, un transcrit de fusion composé d'une partie 5' du gène « piégé » et du gène révélateur, s'exprime sous le contrôle du promoteur du gène « piégé ». L'analyse de la modulation de l'expression de ce gène de fusion révèle ainsi celle du gène piégé dont l'identification est accessible par 5'RACE (*rapid amplification of cDNA ends*). Si l'insertion est aléatoire, l'expression de la protéine « révélatrice », elle, sera d'autant plus efficace que le vecteur est inséré en 5' du gène. En effet, la protéine révélatrice peut être traduite à partir de l'ATG du gène révélateur. Pour que celui-ci fonctionne en dehors de l'ATG de l'exon « piégé », les intégrations doivent prendre place avant l'exon codant du gène « piégé »,

de fait en 5' du gène. Si les intégrations ont lieu après cet exon, c'est l'ATG de l'exon codant qui permettra la traduction de la protéine chimérique et, dans ce cas, la traduction de la protéine du gène révélateur ne se produirait que dans environ 30 % des cas, et la protéine chimérique pourrait être moins stable que la protéine sauvage. Quant à l'utilisation de rétrovirus, dont les LTR (*long terminal repeat*) contiennent des régions riches en CG, fréquentes dans les parties 5' des gènes, elle augmente potentiellement le nombre d'intégrations dans les parties 5' des gènes. Il faut signaler le développement récent d'un vecteur lentiviral, dont l'utilité est surtout de pouvoir « piéger » des gènes impliqués dans l'engagement de cellules souches quiescentes vers différents lignages cellulaires [2].

Les différents vecteurs utilisés varient essentiellement par la nature des gènes révélateurs. Ceux-ci codent pour des activités enzymatiques particulièrement sensibles comme le gène β -gal codant pour la β -galactosidase ou le gène *geo* codant pour un produit de fusion entre l'activité β -galactosidase et la néomycine-phosphotransférase conférant la résistance au G418. Cette sensibilité permet l'identification de transcrits faiblement exprimés et faiblement modulés, mais dont les produits peuvent avoir cependant des fonctions majeures, et leur variation d'expression entraîne un changement de comportement cellulaire très important. Récemment, l'utilisation comme gène révélateur du gène codant pour la GFP (*green fluorescent protein*) [1] permet un criblage direct du niveau

d'expression des gènes piégés, proportionnel à l'intensité de fluorescence, visualisable par cytométrie de flux selon une cinétique en temps réel. Un autre vecteur, dans lequel un domaine transmembranaire de type II est fusionné en N-terminal du gène *géo*, permet de « piéger » des locus codant pour un large éventail de molécules transmembranaires comme des molécules d'adhérence ou des récepteurs membranaires [3].

L'application de cette technique à des cellules ES a permis de mettre en évidence des gènes critiques pour le développement embryonnaire non seulement lors d'étapes précoces, mais également tardives [4]. L'intégration des vecteurs de « piégeage de gènes » conduit à une mutation insertionnelle du gène dans lequel l'intégration a eu lieu, et l'analyse phénotypique des embryons mutants obtenus révèle directement le rôle des produits des gènes piégés au cours du développement. De plus, l'expression spatio-temporelle des gènes piégés est directement visualisable, indépendamment de leur identification, par le suivi de l'expression du gène révélateur. Un exemple bien répertorié est celui du gène *Shrm*, un déterminant critique de l'architecture cellulaire nécessaire à la neurulation dont la fonction requiert sa liaison à la F-actine [5].

Dans les cellules plus spécialisées, le « piégeage de gènes » permet d'identifier des gènes liés à des événements de différenciation ou d'activation cellulaire. Citons ceux qui sont induits lors de réponses cellulaires

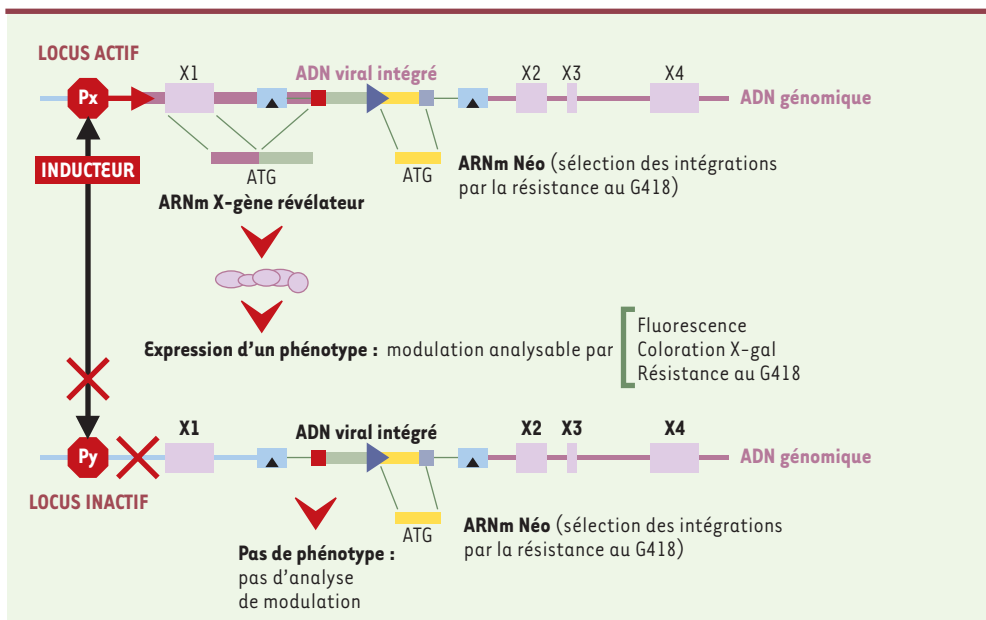


Figure 1. Fonctionnement du piégeage de gènes. Le vecteur de piégeage de gènes présenté sur ce schéma possède un gène de sélection d'intégration dans le génome : le gène *neo* (jaune) sous la dépendance du promoteur pGK (flèche bleue). Les cellules résistantes au G418 représentent les cellules ayant intégré le vecteur. La sélection des cellules ayant intégré le vecteur dans un locus actif se fait grâce à trois principales caractéristiques du vecteur : (1) la présence d'une délétion (triangle noir) dans les séquences régulatrices du vecteur (bleu turquoise) empêche l'activité promotrice du vecteur ; et (2) la transcription du gène piégé est sous le contrôle de son propre promoteur ; (3) il existe un site accepteur d'épissage (rouge) en amont de la séquence du gène révélateur qui permettra, après intégration dans un locus actif pour un inducteur donné, la production d'un produit de fusion (rose-vert) entre le gène piégé (rose) et le gène révélateur (vert). Le gène révélateur peut être d'origine différente : le gène β gal, le gène *géo*, le gène codant pour la protéine fluorescente GFP ou autres. Le produit de ce transcrit permet la sélection des cellules d'intérêt par coloration β gal, résistance au G418 ou fluorescence.



spécifiques à des cytokines comme le HGF (*hepatocyte growth factor*) [1], le TGF- β (*transforming growth factor*) [6], le GM-CSF (*granulocyte-macrophage colony-stimulating factor*) [7] ou à des facteurs apoptotiques [8] ou encore liés à l'activation des lymphocytes B par les lipo-polysaccharides [9] et ont pu être révélés grâce à cette approche. Il est important de remarquer que, dans chacune de ces études, plusieurs nouveaux gènes ont été découverts à partir d'une séquence courte identifiée par 5' RACE, complétée ensuite par la détection d'une séquence codante complète présente dans les banques de données. D'autres ont révélé de nouvelles fonctions pour des gènes déjà identifiés, ou de nouvelles combinaisons de domaines protéiques [10] du produit de gènes connus. C'est le cas par exemple de la sémaphorine 6, de l'EPH4 [11], de la neuropiline [12]. Certains intermédiaires dans des voies de transduction pourtant largement explorées ont aussi été mis en évidence. Ce dernier exemple est illustré par l'analyse du phénotype des embryons résultant de l'implantation de cellules dont le gène piégé est la molécule LD6, un nouveau membre d'une famille de récepteur de type LDL. Ces animaux présentent les mêmes défauts de développement embryonnaire que ceux qui sont observés pour des mutants de différents gènes *Wnt* démontrant l'implication de LD6 dans de nouveaux aspects de la signalisation par *Wnt* [13].

L'ensemble de ces résultats montre la puissance de cet outil dans l'analyse fonctionnelle du génome à grande échelle. Plusieurs milliers de gènes sont accessibles par cette technique expérimentale, au vu du grand nombre d'intégrations aléatoires révélées. L'analyse des gènes qui en résulte reste encore loin d'une saturation de la stratégie. Bien que les données de l'ensemble du génome de différents organismes, dont le génome humain, soient disponibles dans les banques de données, les informations fonctionnelles ne concernent

pas plus de 15 % des 35 000 gènes humains répertoriés ou potentiels. Les informations que nous offre cette analyse génomique, associées à l'analyse des produits des gènes « piégés », augmenteront le nombre de gènes auxquels les informations fonctionnelles pourront être assignées.

Ainsi, outre son application spécifique à la recherche de gènes exprimés de façon variable, cette technique est complémentaire de celles qu'utilise l'analyse du génome. Notamment, la révélation des intégrations ayant pris place dans la région 5' des gènes peut permettre de mettre en évidence l'expression d'exons spécifiques d'un tissu donné, ou d'exons alternatifs spécifiques à la fonction des isoformes résultantes [14]. Les ADNc correspondants peuvent alors compléter la constitution de « bio-puces » spécifiques. De plus, les modulations observées sont indépendantes de régulations post-transcriptionnelles des ARN qui, pour certaines, pourraient ne pas être mises en évidence lors de ces recherches par l'utilisation des « bio-puces » [15]. Une étude menée sur la recherche de gènes impliqués dans les phénomènes d'apoptose par privation en facteur de croissance a démontré par l'approche du « piégeage de gènes », que les gènes *YB1* et *Rad 50* sont sur-exprimés alors qu'aucune modulation n'a été observée par l'utilisation de puces à ADN.

Enfin, la mise en place d'un crible répondant à la recherche de variations de gènes par l'utilisation du piégeage de gènes représente peu d'investissement en terme de matériel expérimental. ♦

Gene trap : how to identify functional novel genes

RÉFÉRENCES

1. Medico E, Gambarotta G, Gentile A, Comoglio PM, Soriano P. A gene trap vector system for identifying transcriptionally responsive genes. *Nat Biotechnol* 2001 ; 19 : 579-82.
2. Lai Z, Han I, Park H, Brady RO. Design of an HIV-1 lentiviral-based gene-trap vector to detect developmentally regulated genes in mammalian cells. *Proc Natl Acad Sci USA* 2002 ; 99 : 3651-6.
3. Skarnes, W. C. Gene trapping methods for the identification and functional analysis of cell surface proteins in mice. *Meth Enzymol* 2000 ; 328 : 592-615.
4. 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.
5. Hildebrand JD, Soriano P. Shroom, a PDZ domain-containing actin-binding protein, is required for neural tube morphogenesis in mice. *Cell* 1999 ; 99 : 485-97.
6. Akiyama N, Matsuo Y, Sai H, Noda M, Kizaka-Kondoh S. Identification of a series of transforming growth factor beta-responsive genes by retrovirus-mediated gene trap screening. *Mol Cell Biol* 2000 ; 20 : 3266-73.
7. Baghdoyan S, Dubreuil P, Eberle F, Gomez S. Capture of cytokine-responsive genes (NACA and RBM3) using a gene trap approach. *Blood* 2000 ; 95 : 3750-7.
8. 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.
9. Kerr WG, Heller M, Herzenberg LA. Analysis of lipopolysaccharide-response genes in B-lineage cells demonstrates that they can have differentiation stage-restricted expression and contain SH2 domains. *Proc Natl Acad Sci USA* 1996 ; 93 : 3947-52.
10. Mitchell KJ, Pinson KI, Kelly OG, et al. Functional analysis of secreted and transmembrane proteins

critical to mouse development. *Nat Genet* 2001 ; 28 : 241-9.

11. Leighton PA, Mitchell KJ, Goodrich LV, et al. Defining brain wiring patterns and mechanisms through gene trapping in mice. *Nature* 2001 ; 410 : 174-9.

12. Tordjman R, Ortega N,

Coulombel L, Plouet J, Romeo PH, Lemarchandel V. Neuropilin-1 is expressed on bone marrow stromal cells: a novel interaction with hematopoietic cells? *Blood* 1999 ; 94 : 2301-9.

13. Pinson KI, Brennan J, Monkley S, Avery BJ,

Skarnes WC. An LDL-receptor-related protein mediates Wnt signalling in mice. *Nature* 2000 ; 407 : 535-8.

14. Burgess RW, Skarnes WC, Sanes JR. Agrin isoforms with distinct amino termini: differential expression, localization,

and function. *J Cell Biol* 2000 ; 151 : 41-52.

15. Wempe F, Yang JY, Hammann J, Melchner H. Gene trapping identifies transiently induced survival genes during programmed cell death. *Genome Biol* 2001 ; 2 : 1-10.

NOUVELLE

La relaxine : une ancienne hormone trouve enfin ses récepteurs

Nicolas Chartrel, Hubert Vaudry

Inserm U.413, Institut Fédératif de Recherches Multidisciplinaires sur les Peptides n° 23, Université de Rouen, 76821 Mont-Saint-Aignan Cedex, France.

montré que la protéine joue un rôle crucial dans l'induction de la phase de déцидualisation [6]. La relaxine, qui favorise aussi l'expression du facteur de croissance endothélial

> L'histoire de la relaxine commence il y a 75 ans lorsque Frederick Hisaw, en étudiant les activités hormonales du corps jaune du follicule ovarien, identifie deux principes actifs, la « corporine », qui induit des transformations de type progestatif dans l'utérus, et la « relaxine » qui provoque un relâchement de la symphyse pubienne [1]. Ce n'est que 50 ans plus tard que la structure de la molécule responsable de l'activité myorelaxante a pu être élucidée [2]. La relaxine est une hormone polypeptidique de structure hétérodimérique appartenant à la superfamille de l'insuline (Figure 1) qui comprend également les *insulin-like growth factor-I et -II* (IGF-I et -II) et le *relaxin-like factor* (RLF ou INSL3). Chez la femme, la relaxine est essentiellement produite par le corps jaune alors que chez l'homme, la principale source de relaxine réside dans la prostate [3, 4]. Chez de nombreuses espèces de mammifères comme le porc, le rat ou le cobaye, la relaxine joue un rôle essentiel dans la croissance et le remodelage des organes génitaux pendant la gestation et plus particulièrement au moment de la parturition, en provoquant l'élongation des ligaments pelviens et le relâchement

du col de l'utérus [5]. De plus, chez le rat, la relaxine inhibe les contractions spontanées ou induites par l'ocytocine du myomètre utérin. Cependant, chez la femme, la relaxine n'exerce aucune activité sur la relaxation du col de l'utérus pendant l'accouchement. En revanche, dans l'espèce humaine, un pic de relaxine circulante est observé en début de grossesse, au moment de la nidification de l'embryon, et il a été

dans l'endomètre utérin, est directement impliquée dans le développement de nouveaux vaisseaux sanguins indispensables à la croissance de l'embryon [7]. Par ailleurs, la relaxine, qui est produite par de nombreux tissus autres que ceux constituant les organes génitaux, n'est pas uniquement une hormone de la grossesse : elle est notamment responsable de la néoangiogenèse après un infarctus ou pendant

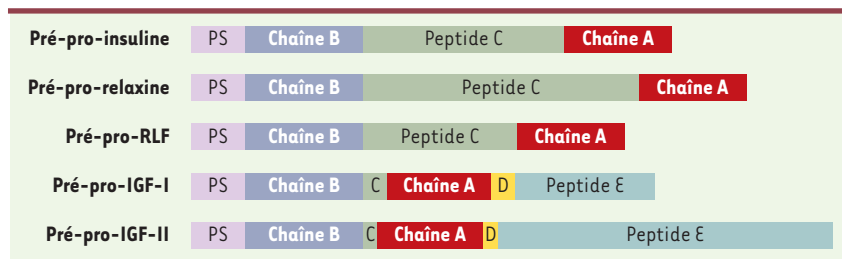


Figure 1. Structure des précurseurs des polypeptides de la superfamille de l'insuline. Toutes les prohormones de cette famille possèdent un peptide C de connexion qui favorise la formation de deux ponts disulfures entre les chaînes A et B de l'hormone mature. Dans la proinsuline, la prorelaxine et le pro-RLF, les chaînes A et B sont respectivement localisées dans les régions C- et N-terminales du précurseur et sont séparées par un peptide C de grande taille qui est excisé au cours de la maturation de l'hormone active. En revanche, pour les pro-IGF-I et -II, le peptide C est de petite taille et il reste présent dans l'hormone mature. Les pro-IGF-I et -II contiennent également dans leur région C-terminale deux autres fragments nommés peptides D et E, le peptide E étant excisé au cours de la maturation de l'hormone. PS : peptide signal.