

A. EGGEN

INRA, Laboratoire de Génétique Biochimique et de Cytogénétique, 78352 Jouy-en-Josas cedex

e-mail : Andre.Eggen@jouy.inra.fr

4 - Recherche de gènes associés à des fonctions

Cartographie fine d'un gène et clonage positionnel

Résumé. Le clonage positionnel consiste à identifier un gène d'intérêt à partir de la seule connaissance de sa localisation chromosomique. Cette stratégie est utilisée lorsqu'il n'existe aucun gène candidat évident ou aucun défaut biochimique spécifique. Elle comprend une localisation primaire, faite généralement par une analyse de liaison classique, puis une cartographie fine de la région d'intérêt permettant de réduire progressivement l'intervalle critique contenant le gène, d'identifier un ensemble de clones chevauchant et couvrant la région, de mettre en évidence les séquences codantes contenues dans cette région et enfin d'identifier le gène responsable du caractère étudié. Etant donné qu'une approche de clonage positionnel pure est lourde et longue à mettre en œuvre, elle est souvent associée à une approche de cartographie comparée et de gènes candidats.

Les efforts internationaux de cartographie des génomes des animaux domestiques ont grandement favorisé le développement de cartes de marqueurs de type microsatellite pour les différentes espèces d'animaux de rente (Georges et Andersson 1996). Ainsi, il est aujourd'hui possible, dans la plupart des populations, de constituer un panel de 200-300 marqueurs informatifs et répartis de façon homogène sur l'ensemble du génome, utilisables pour cartographier, par l'analyse de liaison conventionnelle, des caractères d'intérêt agronomique, et tout particulièrement des QTL. Depuis 1993, de nombreux gènes et QTL ont été localisés : *Booroola* chez le mouton (Montgomery *et al* 1993), *polled* chez le bovin (Georges *et al* 1993), *RN* chez le porc (Milan *et al* 1995), *PIS* chez la chèvre (Vaiman *et al* 1996), etc. Une fois que des régions chromosomiques jouant un rôle dans le déterminisme d'un caractère d'intérêt ont été identifiées, il est possible de mettre en place une sélection assistée par marqueurs (SAM) sans connaissance préalable de la nature des gènes impliqués. L'utilisation de marqueurs comme aide à la sélection reste cependant limitée du fait du risque de recombinaison entre l'un des marqueurs utilisés et le gène impliqué. L'identification du gène d'intérêt et de la mutation causale représente donc un avantage certain : la procédure de typage est simplifiée à un seul test génétique, elle est transposable à d'autres populations et permet d'appréhender le déterminisme génétique du caractère. Quatre stratégies générales d'identification de gènes ont été développées en génétique humaine ; elles sont également utilisées dans l'identification de gènes pour les espèces animales (figure 1).

1 / Stratégies d'identification des gènes

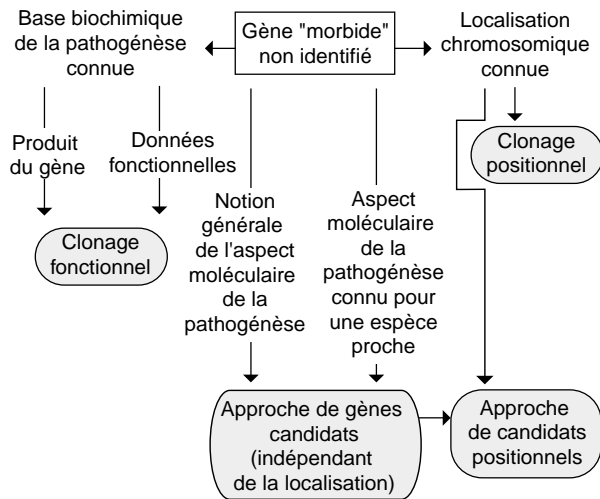
Clonage fonctionnel

Une information concernant la fonction du gène est utilisée pour isoler un clone du gène. Si le produit du gène est connu, la purification partielle de ce produit peut permettre différentes stratégies d'identification du gène sous-jacent. Alternativement, un test fonctionnel peut être utilisé pour le détecter. Chez l'Homme, cette approche a été notamment fructueuse pour l'identification du gène de la phénylalanine hydroxylase, responsable de la phénylcétonurie, et des hémoglobines dans le cas des thalassémies.

Clonage positionnel

Le clonage positionnel correspond à l'isolement du gène uniquement à partir de sa localisation sub-chromosomique, sans nécessiter aucune information sur sa fonction biochimique. L'approche générale consiste à essayer de construire des cartes physique et génétique de la région, à définir la localisation sub-chromosomique, puis à identifier les gènes de la région pour mettre en évidence un gène candidat. Le clonage positionnel demeure difficile et devient de moins en moins nécessaire lorsque les informations s'accroissent, permettant une approche positionnelle du gène candidat.

Figure 1. Stratégies d'identification de gène (ex d'une anomalie génétique chez l'Homme, Strachan et Read 1998).



Approche de gènes candidats, indépendante de la position

Le gène candidat d'un caractère peut être suggéré sans que l'on connaisse sa localisation chromosomique. C'est le cas lorsque l'on distingue un phénotype similaire entre deux espèces, ou si la fonction moléculaire suggère que le gène appartient à une famille de gènes connus. De telles approches ont apporté plusieurs résultats en génétique animale par extrapolation de connaissances des génomes humain et murin (Dennis *et al* 1989, Klungland *et al* 1995).

Approche de gènes candidats positionnels

Lorsqu'un caractère a été décrit de façon détaillée, il devient possible d'utiliser des banques de données pour identifier les gènes candidats. De plus en plus de gènes humains et murins ayant été cartographiés dans des régions sub-chromosomiques spécifiques, les approches positionnelles de gènes candidats, via les connaissances de la cartographie comparée, sont souvent privilégiées. C'est par cette approche que le gène responsable de l'hypertrophie musculaire bovine a été identifié (Grobet *et al* 1997).

2 / Le clonage positionnel

Lorsqu'aucun gène candidat évident ou aucun défaut biochimique n'existe, la seule stratégie possible est d'approcher le gène à partir de sa seule position dans le génome. Cette stratégie comprend quatre étapes (figure 2). Elle nécessite toutefois que l'on dispose de suffisamment de familles informatives pour établir une liaison et affiner cette liaison primaire, ainsi que d'un nombre suffisant de marqueurs génétiques informatifs.

Première étape : localisation primaire de la région contenant le gène d'intérêt

Le but de cette étape est l'identification de marqueurs qui ségrègent spécifiquement avec le caractère d'intérêt. Cela permet de définir un intervalle de localisation contenant le gène responsable. Toutefois, une approche par clonage positionnel ne peut pas être tentée de façon sérieuse sans avoir localisé le gène d'intérêt par cartographie génétique ou physique à l'intérieur d'un segment de quelques centimorgans, soit moins de 5 Mb d'ADN. Il convient donc de chercher à réduire progressivement l'intervalle contenant le locus d'intérêt, soit par l'isolement de nouveaux marqueurs de la région, soit par une extension des familles utilisées, ce qui paraît pourtant difficile dans le cadre de programmes de recherche de QTL. Une autre méthode, actuellement en cours d'évaluation, consiste à comparer, dans diverses populations, les segments chromosomiques flanquant l'allèle QTL hypothétique et supposés identiques par descendance. Cette démarche (IBD : Identical-by-descent) a notamment été utilisée pour cartographier une maladie autosomale récessive en race bovine Holstein (syndactylie : Charlier *et al* 1996) et pour réduire considérablement l'intervalle contenant un QTL laitier cartographié dans la région centromérique du chromosome bovin 14 (Riquet *et al* 1999).

Deuxième étape : isolement de la région recouvrant cette région

Les marqueurs qui ségrègent avec le caractère étudié peuvent être utilisés pour identifier des clones

Figure 2. Les différentes étapes du clonage positionnel (Bernot 1999).

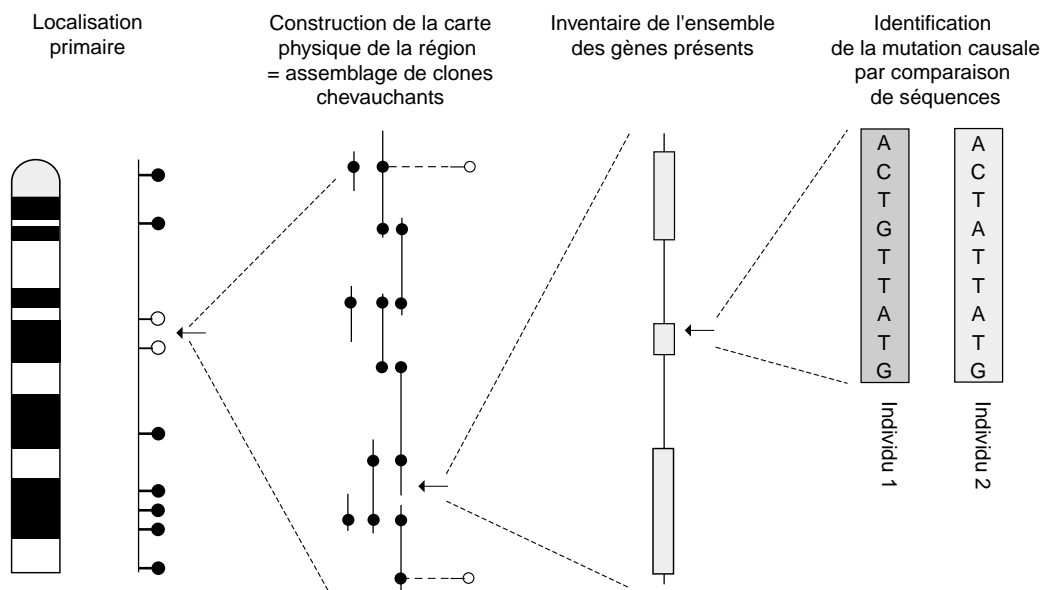


Tableau 1. Méthodes communément utilisées pour identifier des gènes dans l'ADN cloné (d'après Strachan et Read)

Méthode	Commentaires
Zoo-Blot	Un clone d'ADN est hybridé dans des conditions de stringence réduite avec un Southern blot d'échantillons d'ADN génomique de différentes espèces animales, un zoo-blot.
Identification d'îlots CpG	De nombreux gènes de vertébrés contiennent des îlots CpG, séquences hypométhylées riches en GC, présentant habituellement plusieurs sites de restriction pour des enzymes coupant rarement.
Hybridation avec ARNm/ADNc	Une clone d'ADN génomique peut être hybridé avec un Northern blot d'ARNm obtenu à partir de différentes lignées cellulaires en culture ou avec des banques d'ADNc appropriées.
Piège à exons (Exon trapping)	Il s'agit d'un test d'épissage artificiel d'ARN. Il repose sur l'observation que la grande majorité des gènes de mammifères contiennent plusieurs exons qui doivent être épissés au niveau de l'ARN.
Sélection d'ADNc	Ces techniques reposent sur la purification répétée d'un sous-groupe de clones d'ADN génomique qui s'hybrident avec une population donnée d'ADNc.
Analyse informatique de séquence d'ADN	Recherche d'homologie avec une séquence d'ADN connue. Prédiction d'exons à l'aide de programmes informatiques identifiant des exons possibles dans des séquences nucléotidiques.
Expression différentielle à partir du YAC	Mises en évidence de différences d'expression entre un clone levure et un clone contenant un chromosome artificiel de levure. Ces différences peuvent être attribuées à des séquences codantes du chromosome artificiel provenant de l'espèce utilisée pour construire le YAC.

d'ADN recouvrant la région critique. Il s'agit en effet de construire une trame facilitant la construction d'une carte physique et comprenant une série de fragments d'ADN clonés (principalement YAC et BACs) qui, collectivement, contiennent la totalité de

la séquence d'intérêt. La série de clones doit donc contenir des inserts chevauchants formant un ensemble complet de clones contigus (contig). La marche sur le chromosome (uni- ou bi-directionnelle) permet d'établir des contigs à partir d'un ou de plusieurs points de départ fixés : ces points de départ peuvent être les marqueurs utilisés dans l'analyse de liaison, des gènes déjà localisés dans cette région ou même des gènes hypothétiquement présents dans la région critique et identifiés par le biais de la cartographie comparée. L'obtention de tels clones permet alors de disposer d'un matériel facile à utiliser pour étudier les gènes présents dans la région et pour identifier le gène responsable.

Troisième étape : inventaire de tous les gènes présents dans la région

Le plus souvent, l'identification du gène d'intérêt passe par la recherche de tous les gènes présents dans la région. La tâche est d'autant plus difficile que les gènes sont généralement petits par rapport à la taille de la région considérée, qu'ils sont fragmentés en exons et en introns et, enfin, qu'ils sont disséminés au hasard, séparés par de vastes régions non codantes. Cette étape reste limitante et délicate car il n'existe pas de méthode simple et directe pour reconnaître à coup sûr une séquence codante. Les différentes techniques utilisées sont : le criblage de banques d'ADN complémentaire, la sélection d'ADN complémentaire, le Zoo-Blot, l'identification d'îlots CpG, le piège à exons, l'analyse de séquence, l'expression différentielle à partir de clones contenant un chromosome artificiel de levure (YAC) (tableau 1).

Quatrième étape : identification du gène muté

Parmi les gènes présents dans la région critique, l'étape ultime est l'identification du gène d'intérêt. Ce gène doit être identifié par la présence de mutations chez les individus porteurs du caractère étudié et par une absence de mutation chez les individus non porteurs. Cette étape représente un défi majeur car il est essentiel, une fois qu'une variation de séquence a été détectée, de pouvoir distinguer un simple polymorphisme d'une mutation causale.

Conclusion

Bien que le clonage positionnel reste une approche lourde et longue à mettre en œuvre, les découvertes encourageantes de ces dernières années, à la fois en génétique humaine et en génétique animale, montrent que cette approche associée à une approche de gènes candidats apporte de réels résultats qui permettront, par l'identification de gènes d'intérêt économique, de mettre au service de la profession de nouveaux outils de sélection.

Références

- Bernot A., 1999. L'analyse des génomes : cartographie, séquençage, identification des gènes (2e édition). Nathan Université, Paris, France.
- Charlier C., Farnir F., Berzi P., Vanmanshoven P., Brouwers B., Georges M., 1996. IBD mapping of recessive traits in livestock : application to map the bovine syndctyly locus to chromosome 15. *Genome Research*, 6, 580-589.
- Dennis J.A., Healy P.J., Beaudet A.L., O'Brien W.E., 1989. Molecular definition of bovine argininosuccinate synthetase deficiency. *Proceedings of the National Academy of Sciences of the USA*, 86, 7947-7951.
- Georges M., Andersson L., 1996. Livestock genomics comes of age. *Genome Research*, 6, 907-921.
- Georges M., Drinkwater R., King T., Mishra A., Moore S.S., Nielsen D., Sargeant L.S., Sorensen A., Steele M.R., Zhao X., Womack J.E., Hetzel J., 1993. Microsatellite mapping of a gene affecting horn development in *Bos taurus*. *Nature Genetics*, 4, 206-210.

- Grobet L., Martin L.J., Poncelet D., Pirottin D., Brouwers B., Riquet J., Schoeberlein A., Dunner S., Ménéssier F., Massabanda J., Fries R., Hanset R., Georges M., 1997. A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. *Nature Genetics*, 17, 71-74.
- Klungland H., Vage D.I., Gomez-Raya L., Adalsteinsson S., Lien S., 1995. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mammalian Genome*, 6, 636-639.
- Milan D., Le Roy P., Woloszyn N., Caritez J., Elsen J., Sellier P., Gellin J. 1995. The RN locus for meat quality maps to pig chromosome 15. *Genetics Selection Evolution*, 27, 195-199.
- Montgomery G.W., Crawford A.M., Penty J.M., Dodds K.G. Ede A.J., Henry H.M., Pierson C.A., Lord E.A., Galloway S.M., Schmack A.E., Sise J.A., Swarbrick P.A., Hanrahan V., Buchanan F.C., Hill D.F., 1993. The ovine Booroola fecundity gene (FecB) is linked to markers from a region of human chromosome 4q. *Nature Genetics*, 4, 410-414.
- Riquet J., Coppieters W., Cambisano N., Arranz J.J., Berzi P., Davis S., Grisart B., Farnir F., Karim L., Mni M., Simon P., Taylor J.F., Vanmanshoven P., Wagenaar D., Womack J.E., Georges M., 1999. Fine-mapping of quantitative trait loci by identity by descent in outbred populations : application to milk production in dairy cattle. *Proceedings of the National Academy of Sciences of the USA*, 96, 9252-9257.
- Strachan T., Read A.P., 1998. *Génétique Moléculaire Humaine* (1ère édition). Médecine-Sciences Flammarion, Paris, France.
- Vaiman D., Koutita O., Oustry A., Elsen J.M., Manfredi E., Fellous M., Cribru E.P., 1996. Genetic mapping of the autosomal region involved in XX sex-reversal and horn development in goats. *Mammalian Genome*, 7, 133-137.