



## Cartographier le génome du cheval : les avancées

S. Godard, L. Schibler, D. Vaiman, A. Oustry  
M. Nocard, M. Bertaud, E.P. Cribiu, G. Guérin.  
INRA, Centre de Recherches de Jouy,  
Laboratoire de Génétique biochimique et de Cytogénétique,  
78352, Jouy-en-Josas Cedex.

### Résumé

La cartographie du génome prend une importance considérable dans la recherche de maladies génétiques chez l'homme et la compréhension des mécanismes de base chez la souris. Pour les espèces domestiques, elle vise à une amélioration de l'élevage par l'identification de régions, et finalement de gènes, d'intérêt zootechnique. Chez le cheval, le retard accumulé par le faible nombre d'équipes et par une sensibilisation tardive des milieux concernés tend à se combler notamment par une action internationale combinée à de fortes initiatives personnelles. Nous présentons ici les avancées acquises par notre laboratoire sur la carte génétique (microsatellites) et la carte physique constituant des points d'ancrages (gènes) sur les chromosomes. La construction et l'exploitation d'une banque de grands fragments (BACs) visant à intégrer les deux cartes vont encore faciliter la comparaison du génome équin avec celui des autres espèces et permettre un repérage plus aisé des régions susceptibles de présenter un intérêt pour l'élevage.

Mots-clés : cheval, cartographie, gènes, banque de BACs, amélioration génétique.

### Summary

Genome mapping provides information that is particularly important for research on genetic diseases in humans, and for further understanding of basic biological mechanisms in the mouse. In domestic species, knowledge of the genome will make breeding improvements possible by establishing regions and ultimately genes of zootechnic interest. In the horse, work in this area has progressed particularly slowly as only a few research teams are involved and the appreciation of the potential benefits has been slow. This situation, however, is changing, largely due to international co-operative efforts and some dynamic initiatives on the part of a few individuals. In this paper we presented the recent developments from our laboratory with respect to the genetic map (microsatellites) and the physical map, where we were able to establish some anchor points (genes) on the chromosomes. The construction and exploitation of a large fragment library (BACs) was undertaken in the hopes of integrating the two maps. The results will also facilitate the comparison of the horse genome with those of other species. This, in turn, will make it easier to identify regions that are potentially interesting for breeding improvement.

Key-words : horse, mapping, genes, BAC library, genetic improvement.

## LES OBJECTIFS DE LA CARTOGRAPHIE

Depuis longtemps, le progrès génétique des espèces d'élevage a été apporté par les méthodes de génétique quantitative qui ignorent les mécanismes biologiques sous-jacents.

Depuis peu, la génétique moléculaire a mis à notre portée l'explication moléculaire de la variabilité apparente observée, c'est-à-dire les variations moléculaires des gènes, ces fragments d'ADN spécifiant les protéines. Il est maintenant acquis qu'une mutation à l'intérieur d'un gène risque d'entraîner la production d'une protéine variante, différente ou anormale, et générant ainsi l'apparition d'une pathologie ou au contraire d'un caractère bénéfique.

Ces nouvelles options permettent d'envisager l'identification de gènes responsables de maladies chez l'homme, mais aussi de caractères d'intérêt économique chez les animaux domestiques.

Quelle technique peut-on utiliser pour identifier un gène responsable d'un caractère complexe ? Dans de rares cas, un gène connu auparavant, peut être testé en tant que « gène candidat » car il est susceptible d'expliquer les différences phénotypiques observées, en raison de ses spécificités (symptômes similaires dans une autre espèce, localisation tissulaire ou schéma chronologique de son expression chez l'animal). Dans la plupart des cas, cependant, les mécanismes biologiques sous-jacents sont inconnus, et l'on ne dispose pas de gènes candidats convaincants. Pourtant, les techniques dites de **clonage positionnel** permettent d'identifier le gène responsable. Ces techniques consistent à restreindre un intervalle chromosomique jusqu'au moment où un gène unique peut seul s'y trouver ; le gène est trouvé par sa position et non pas, en théorie, par sa fonction.

Un réseau de **marqueurs polymorphes ordonnés**, les **microsatellites**, répartis sur l'ensemble du génome va permettre, par l'analyse de familles, de mettre en évidence une association entre certains de ces marqueurs et le caractère observé (on parle alors de l'établissement d'une **liaison génétique**). Réaliser une carte de tels marqueurs (que l'on appelle **carte génétique**) chez le cheval constitue donc une étape indispensable pour identifier des gènes d'intérêt dans cette espèce.

En effet, chez les espèces d'élevage, la localisation et l'étude de caractères quantitatifs d'intérêts économiques ont pour objectif ultime l'amélioration des schémas de sélection. Chez de nombreuses espèces des cartes génétiques et physiques de moyenne densité sont désormais disponibles (vache, porc, mouton, chèvre) ; ces outils ont d'ores et déjà portés leurs fruits et permis la localisation de certains gènes.

Parmi la vingtaine de gènes déjà localisés dans les espèces domestiques, une mention particulière peut être faite pour le gène **culard bovin** localisé au début de l'année 1996. Par la suite, et de façon indépendante, un gène résultant en un phénotype similaire (augmentation de 200 % de la masse de muscles squelettiques) fut cloné chez la souris. La cartographie comparée et l'analyse de mutations du gène bovin homologue ont révélé qu'il s'agissait du même gène (myostatine). C'est le premier exemple d'un clonage positionnel ayant abouti chez les animaux domestiques (Grobet, 1997). Un simple test génétique permettra alors la détection de cette mutation et l'identification des animaux homozygotes pour ce génotype. Ces tests pourront être mis en place sur les animaux reproducteurs dès la naissance et permettre de réduire les coûts d'entretien (5 ans) des jeunes taurillons en testage.

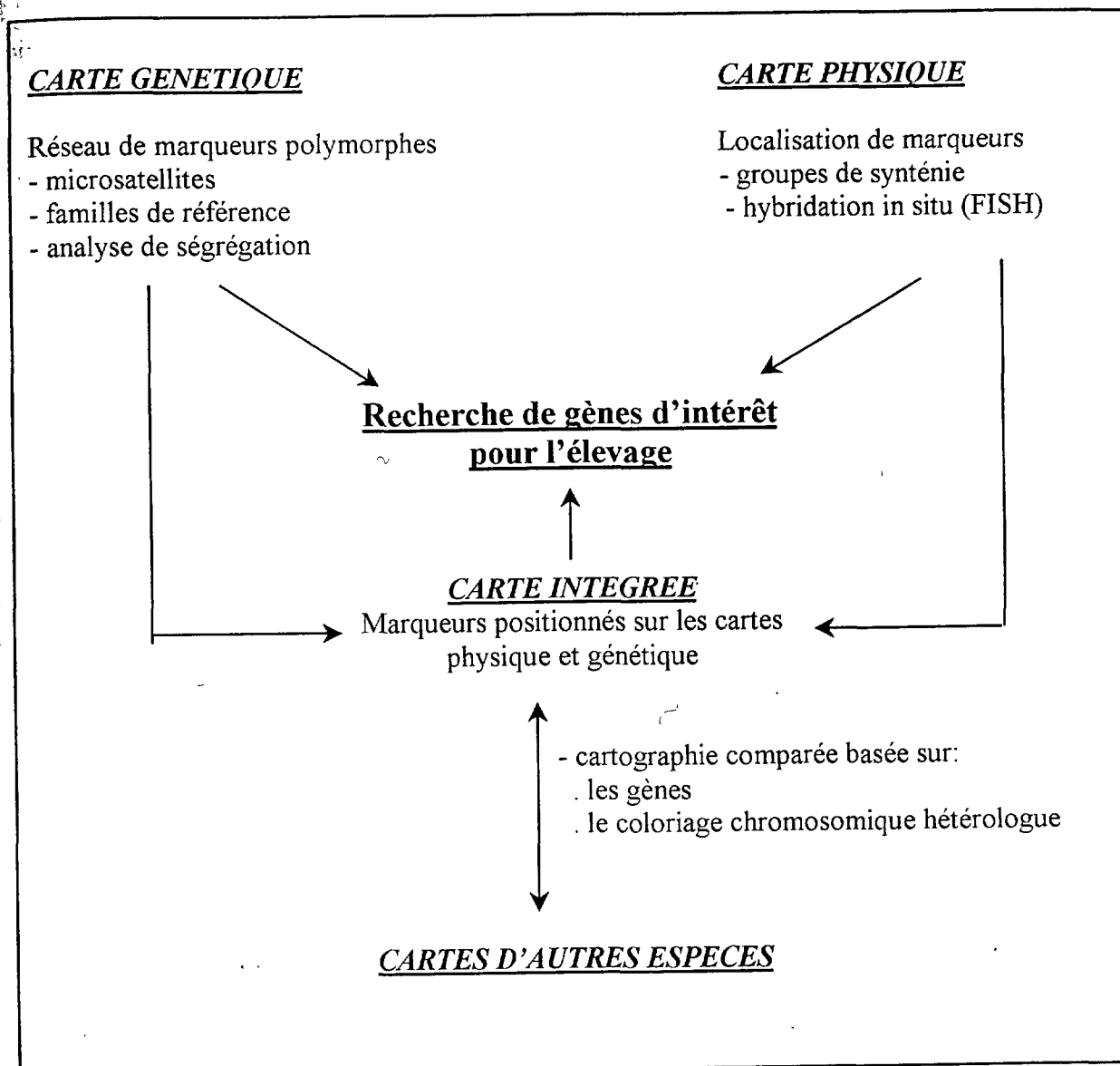
De même chez le cheval, le gène responsable du syndrome d'immunodéficience équine (CID) a été identifié à la fois par clonage positionnel et par l'approche gène candidat par une équipe américaine (Bailey, 1997). Des études ont montré que chez la souris des mutations dans le gène **DNApk** généraient des symptômes identiques à ceux observés chez les chevaux atteints. Des **liaisons génétiques** entre ce gène et des marqueurs équins polymorphes ont permis de le localiser. Les localisations de ce gène étant cohérentes entre la souris, l'homme et le cheval, il semblerait que ce soit le même gène qui soit responsable de la maladie dans l'espèce équine. Un test génétique est désormais disponible pour identifier les animaux porteurs de ce gène dans les élevages à risque.

## LES OUTILS DE LA CARTOGRAPHIE

La détection de liaisons entre gènes et marqueurs, requiert l'utilisation de **cartes génétiques**, comme celles mentionnées précédemment (figure 1). La construction de telles cartes nécessite la production de **marqueurs polymorphes** qui ne sont pas conservés entre différentes espèces. L'étape suivante, c'est-à-dire le clonage du gène impliqué, nécessite une carte de marqueurs d'une forte densité. De telles cartes sont difficilement réalisables chez les espèces domestiques. En revanche, pour deux espèces de mammifères, l'homme et la souris, un tel effort a été mené à terme. Alors que les cartes génétiques ne sont pas transposables entre différentes espèces, les séquences codantes, gènes ou EST (expressed sequence tag : séquence d'ADN codant

dont la fonction est inconnue) sont largement conservées ; les gènes servent ainsi de "passerelles" entre les différentes cartes. Il est indispensable aujourd'hui de pouvoir déduire à partir des localisations de gènes murins et humains, la localisation de ces mêmes gènes dans l'espèce étudiée. A cet effet, la **cartographie comparée** doit être développée.

**Figure I** : La cartographie du génome



La carte comparée consiste à repérer les segments chromosomiques équivalents entre deux espèces. Le « painting hétérologue » (ou coloriage chromosomique hétérologue) (Raudsepp, 1996) est le premier travail qui a posé les bases de la cartographie comparée entre l'homme et le cheval. Cela consiste à hybrider des chromosomes humains triés un à un sur les chromosomes équins pour repérer les segments conservés entre les deux espèces. La position précise d'un gène donné est alors recherchée à l'intérieur d'un segment qui le contient.

La cartographie comparée ne peut se construire que sur la base d'une **carte physique** du cheval. Cette carte a pour objectif d'associer les groupes de liaison aux chromosomes (ancrage) puis de mieux définir la position relative des marqueurs d'un groupe de liaison sur les chromosomes (orientation). La **cartographie physique** consiste donc à **localiser directement des marqueurs** (qu'ils soient polymorphes ou non) sur les chromosomes. Enfin, son enrichissement en gènes permet de préciser plus finement la position des gènes conservés et de les comparer entre espèces.

La construction d'une telle carte repose sur deux techniques. La FISH (hybridation in situ en fluorescence) permet de visualiser la position d'une sonde marquée de grande taille sur des chromosomes métaphysiques observés au microscope et précise ainsi la position cytogénétique d'un marqueur.

Les **hybrides somatiques**, résultent de la fusion entre les cellules du cheval et des cellules d'une autre espèce (hamster ou souris) ; chaque hybride ne conserve qu'un ou quelques chromosomes du cheval. L'analyse statistique de la répartition des marqueurs sur les hybrides, permet de constituer des **groupes de synténies** (regroupements de marqueurs sur un chromosome donné), sans connaissance de leur positions relatives.

Ce travail de recherche, par la masse de travail qu'il représente et les nombreuses techniques utilisées ne peut être le fruit que de laboratoires disposant de beaucoup de moyens (type Généthon) ou d'une coopération internationale entre différents laboratoires.

L'axe de notre travail à l'INRA a été de développer des marqueurs qui puissent permettre de réaliser directement une carte intégrée. Ce parti pris permet de s'affranchir d'une recherche de marqueurs aléatoires pour développer une approche ciblée en s'appuyant sur la cartographie comparée. Cette orientation demande un investissement plus important au départ mais est plus intéressante à long terme puisque ce type de marqueurs permet alors d'étoffer à la fois la carte génétique et la carte physique.

## LES AVANCEES DE LA CARTE GENETIQUE

La carte génétique n'a pu se développer que lorsque le nombre de marqueurs polymorphes microsatellites existant a été suffisamment important pour permettre de réaliser une analyse de liaison.

Il y a trois ans, au début de ce travail, la carte du cheval était embryonnaire. La carte génétique n'était composée que de 5 groupes de liaison comprenant 23 marqueurs. Seuls une vingtaine de marqueurs génétiques polymorphes de type microsatellites étaient publiés.

En 1998, plus de 200 marqueurs microsatellites ont été mis à jour par différentes équipes ; l'INRA de Jouy-en-Josas en a produit 64 ce qui fait de nous une des équipes qui a le plus investi dans la production de marqueurs. Commencé sur des familles de référence françaises, le typage de ces microsatellites a mis à jour 4 nouveaux groupes de liaison (Godard, 1997).

Aujourd'hui la carte génétique du cheval comprend 54 marqueurs répartis sur 16 groupes de liaison, pour la plupart assignés à des chromosomes. Le travail en synergie de différentes équipes et leur complémentarité a abouti à la mise en commun d'ADN de familles internationales et au regroupement des données de typage de l'ensemble des microsatellites. L'analyse des données dans notre laboratoire doit permettre la réalisation d'une carte génétique de 200 marqueurs dans le courant de l'année 1998.

## LES AVANCEES DE LA CARTE PHYSIQUE

Jusqu'en 1996 seuls 5 gènes étaient localisés physiquement. Aujourd'hui 92 sondes sont localisées par FISH, dont 63 dans notre laboratoire de l'INRA (tableau 1).

Parmi ces sondes, 45 sont des microsatellites issus de cosmides (25) ou de phages (20) (Godard, 1997, Breen, 1997). Cette recherche aléatoire de cosmides avait déjà permis de placer un marqueur sur 14 des 32 chromosomes équins ( $2N = 64$ ).

La mise en place d'un nouvel outil, une banque de grands fragments ordonnés (BACs, bacterial artificial chromosome) a permis de progresser rapidement. La recherche de séquences d'ADN dans la banque par PCR est rapide et illimitée. Jusqu'à maintenant, la recherche au hasard d'un marqueur amenait une localisation aléatoire ; la banque de BACs permet maintenant d'établir des localisations ciblées par régions chromosomiques.

Tableau 1 · Etat de la cartographie équine

CHROMOSOMES	NOMBRE DE MARQUEURS DONT GENES ( )			
	CARTE EN 1994		CARTE FIN 1997	
	Bibliographie	Bibliographie	INRA	TOTAL
ECA1	1	5 (0)	5(2)	10 (2)
ECA2	1(1)	5 (1)	0	5(1)
ECA3	0	0	4(2)	4(2)
ECA4	0	2(0)	3	5 (0)
ECA5	0	0	0	0
ECA6	0	1(0)	1(1)	2(1)
ECA7	0	0	2(2)	2(2)
ECA8	0	0	1(1)	1(1)
ECA9	0	4 (1)	3	7(1)
ECA10	2(2)	4(2)	2(1)	6(3)
ECA11	0	0	3	3(0)
ECA12	0	1(1)	1	2(1)
ECA13	1(1)	2(1)	1	3(1)
ECA14	0	0	0	0
ECA15	0	3	6(3)	9(3)
ECA16	0	1	4(3)	5(3)
ECA17	0	0	1	1(0)
ECA18	0	0	1	1(0)
ECA19	0	0	2	2(0)
ECA20	1(1)	1(1)	4(3)	5(4)
ECA21	0	0	5(3)	5(3)
ECA22	0	0	3(1)	3(1)
ECA23	0	0	2(1)	2(1)
ECA24	0	0	3(2)	3(2)
ECA25	0	0	0	0
ECA26	0	0	1(1)	1(1)
ECA27	0	0	1(1)	1(1)
ECA28	0	0	1(1)	1(1)
ECA29	0	0	0	0
ECA30	0	0	0	0
ECA31	0	0	0	0
ECAX	0	0	2(1)	2(1)
ECA Y	0	0	1	1(0)
<b>TOTAL:</b>	<b>6(5)</b>	<b>29(7)</b>	<b>63(29)</b>	<b>92(36)</b>

ECA = Equus Caballus

La banque de grands fragments a permis de localiser 10 autres microsatellites déjà associés à des groupes de liaison. La localisation physique de microsatellites issus de cosmides et de BACs a permis l'ancrage et/ou l'orientation de groupes de liaison (Godard, 1998). De même, ces localisations, ont rendu possible ou ont confirmé l'assignation de près de la moitié des 32 groupes de synténies établis par une équipe américaine à partir de 156 microsatellites (Shiue, 1998).

En s'appuyant sur les données de cartographie comparée, nous avons choisi des gènes répartis sur tous les chromosomes humains dans le but de localiser physiquement au moins un gène sur chaque chromosome du cheval. Cet objectif est en passe d'être atteint, car avec 25 nouveaux gènes localisés, les deux tiers des chromosomes du cheval contiennent maintenant un gène identifié (tableau 1).

## VERS UNE CARTOGRAPHIE INTEGREE... EN S'APPUYANT SUR LA CARTOGRAPHIE COMPAREE

Deux approches complémentaires nous ont amené à identifier des gènes ou EST, comme marqueurs susceptibles de développer une carte intégrée.

Des données bibliographiques croissantes rapportaient l'existence de microsatellites associés à des gènes (introns ou partie 3' non codante) dans 0.1 à 10% des cas. Le criblage d'une banque d'ADNc (banque de séquences codantes) de cerveau fœtal équin pour des microsatellites a permis l'identification de 14 nouveaux marqueurs microsatellites associés à des EST chevaux. Ces microsatellites seront localisés à la fois sur la carte génétique et par FISH. L'existence d'homologies fortes entre les EST équins et des EST humains, permet généralement de déduire leur localisation cytogénétique équine de la localisation humaine.

Le nombre croissant d'informations de séquences dans les bases de données permet de trouver, à partir d'amorces de gènes humains ou d'autres espèces, le gène équin en l'absence de séquence équine préalable. La banque de BACs du cheval a été ainsi criblée pour près d'une centaine de gènes ou EST. 25 nouveaux gènes équins et 3 EST ont ainsi été clonés et localisés cytogénétiquement. Ces nouvelles localisations physiques de gènes augmentent de façon considérable les localisations de gènes équins. Certains sont des loci qui permettent l'étude de caractères pouvant directement intéresser les éleveurs (Agouti gène de coloration, par exemple). Cet outil rend aussi possible la recherche de gènes connus "à la demande".

Ces localisations de gènes issus de BACs présentent d'autant plus d'intérêt qu'elles peuvent être "intégrées" à la carte génétique. En effet, de part la grande taille de leurs inserts, les BACs permettent d'obtenir un microsatellite associé au gène, donc un polymorphisme, qui lui assure un ancrage génétique.

De plus ces gènes, tous choisis d'après leur localisation chez l'homme de façon ciblée permettent de préciser les homologies entre les génomes homme/cheval définies par le coloriage chromosomique hétérologue. La présence de remaniements importants à l'intérieur des segments conservés a déjà été décelée (ECA15, ECA20), comme cela a été montré de la même façon chez d'autres espèces.

Ces mêmes gènes sont ou seront localisés dans d'autres espèces (chèvre, vache, porc) ce qui ouvre des perspectives intéressantes en terme de cartographie comparée.

## CONCLUSION

En conclusion, il y a deux ans seulement 4 gènes étaient localisés chez le cheval et il existait 5 groupes de liaison dont la localisation était pour la plupart inconnue. Aujourd'hui, les données de cartographie progressent rapidement. 25 gènes issus de BACs viennent récemment d'être positionnés dans notre laboratoire et 16 groupes de liaison existent dont douze sont localisés. De plus, 200 marqueurs microsatellites sont disponibles au niveau international et doivent être localisés sur la carte génétique en 1998. Nos perspectives actuelles concernent la recherche des microsatellites associés aux gènes déjà localisés, afin d'enrichir la carte intégrée.

Pour les autres espèces, les moyens mis en œuvre ont permis le développement d'une carte génétique, puis d'une carte physique pour ancrer la première. Les cartes ont alors été enrichies en séquences codantes pour bénéficier des données de cartographie comparée.

Chez le cheval, nous disposons de moyens plus limités et l'établissement de cartes du génome équin exploitables ne pourra être que le fruit de collaborations internationales. Le décalage en ce domaine nous fait cependant bénéficier des stratégies et des avancées techniques développées par les autres espèces et doit nous permettre de combler ce retard. Un travail en synergie et une focalisation des énergies devrait aboutir directement à la réalisation d'une carte intégrée riche en gènes qui permettra l'analyse des caractères génétiques et de développer des diagnostics précoces pour identifier les animaux les plus performants et, de façon plus globale, améliorer l'élevage équin.

## REMERCIEMENTS

Ce travail a bénéficié en partie de crédits du Service des haras, des courses et de l'équitation. La bourse de thèse de S. Godard est cofinancée par l'INRA et LABOGENA (Laboratoire d'analyses génétiques pour les espèces animales). Nous remercions également le Laboratoire de Physiologie de la reproduction de Nouzilly (E. Palmer) de nous avoir donné accès à des embryons équins.

## BIBLIOGRAPHIE

- Bailey E., Reid R. C., Skow L. C., Mathiason K., Lear T. L., McGuire T.C., 1997. Linkage of the gene for equine combined immunodeficiency disease to microsatellite markers HTG8 and HTG4; synteny and FISH mapping to ECA9. *Animal Genetics*, 28, 268-273.
- Breen M., Lindgren G., Binns M. M., Norman J., Irvin Z., Bell K., Sandberg K., Ellegren H., 1997. Genetical and physical assignments of equine microsatellites - First integration of anchored markers in horse genome mapping. *Mammalian Genome*, 8, 267-273.
- Grobet L., Martin L.J.R., 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 myosatin causes the double-muscled phenotype in cattle. *Nature genetics*, 17, 71-74.
- Godard S., Vaiman D., Oustry A., Nocard M., Bertaud M., Guzylack S., Cribiu E.P., Guérin G., 1997. Characterisation, genetic and physical mapping of 36 horse plasmid and cosmid-derived microsatellites. *Mammalian Genome*, 8, 745-750.
- Godard S., Schibler L., Oustry A., Cribiu E.P., Guérin G., 1998. Construction of a horse BAC library and cytogenetical assignation of 22 type I and type II markers. En préparation.
- Raudsepp T., Fronicke L., Scherthan H., Gustavsson, I., Chowdhary B.P., 1996. Zoo-FISH delineates conserved chromosomal segments in horse and man. *Chromosome research*, 4, 218-225.
- Shiue Y., Bickel L. A., Caetano A. R., Clark R., Eggleston M. L., Delvallz A., Michelmore R., Bailey E., Guérin G., Godard S., Mickelson J. R., Valberg S. J., Murray J. D., Bowling A.T., 1998. A synteny map of the horse genome comprised of 22 microsatellite and RAPD markers. Soumis pour publication.