

ELABORATION DE LA CARTE GENETIQUE DU CANARD COMMUN

Dans le cadre du programme de recherche « GENEKAN »

**Feve Katia¹, Bounet Marie¹, Vignoles Florence, Leroux Sophie¹, Bardes Suzanne¹,
Vignal Alain¹, Marie-Etancelin Christel²**

¹INRA, UMR444, Laboratoire de Génétique Cellulaire, B.P. 52627, 31326 Castanet-Tolosan

²INRA, UPR631, Station d'Amélioration Génétique des Animaux, B.P. 52627, 31326
Castanet-Tolosan

RESUME

Les travaux sur le génome du canard ont pour cadre une recherche de QTL de qualité des produits, du métabolisme hépatique et du comportement. Afin de réaliser cette détection, l'établissement d'une carte génétique est indispensable. Dans un premier temps, 161 marqueurs microsatellites développés à l'INRA-LGC et 162 publiés dans les banques internationales ont été positionnés « *in-silico* » sur le génome de la poule par similarité de séquence (Feve et al., 2006 ; Marie-Etancelin et al., 2006). Dans un second temps, après sélection de 112 marqueurs microsatellites informatifs dans notre dispositif QTL sur les 323 testés, nous avons génotypé la plus grande famille du croisement Back-cross (Marie-Etancelin et al., 2008) avec ceux-ci. Ce jeu de données a permis, à l'aide du logiciel Crimap, la construction d'une première carte génétique du canard commun (caryotype composé de 40 paires de chromosomes), agrégeant 65 marqueurs en 13 groupes de liaison, eux-mêmes assignés à 9 chromosomes différents. De plus, une comparaison entre les groupes de liaison établis chez le canard et le positionnement de ces marqueurs sur la séquence de la poule a pu être établie : elle n'a pas pour l'instant démontré de différences majeures, avec seulement 4 remaniements intra-chromosomiques possibles. Enfin, le développement de marqueurs complémentaires de type SNP, qui nous permettra de compléter cette carte, est évoqué.

ABSTRACT

Our current work on the duck genome is within the frame of a QTL detection program for traits such as the quality of products, liver metabolism and behaviour. To this end, the construction of a genetic map is essential. A choice of microsatellite markers was made amongst the 161 developed at INRA-LGC and 162 published in international databases. First the markers were positioned on the chicken genome by sequence similarity (Feve et al., 2006 ; Marie-Etancelin et al., 2006), so as to optimise the genome coverage in duck. Then 112 were selected based on informativity in our QTL pedigree, for genotyping our largest back-cross family (Marie-Etancelin et al., 2008). This dataset enabled us to build a first genetic map of the mallard duck (*Anas platyrhynchos*), grouping 65 markers into 13 linkage groups (n=40), by using the Crimap software. Cross-localisation of markers on the chicken genome showed that no major differences could be demonstrated with our dataset, with only 4 possible intra-chromosomal rearrangements identified. Finally, a project for the development of complementary SNP markers to improve our map will be described.

INTRODUCTION

Afin d'augmenter l'efficacité de la sélection pour des caractères difficiles ou coûteux à mesurer (par exemple caractères de qualité des produits ou de résistance à une maladie), il peut être judicieux de disposer de marqueurs génétiques servant à identifier les zones des chromosomes porteuses d'allèles favorables pour ces caractères. Si dans de rares cas, une approche par gène candidat peut fournir un marqueur utilisable, la voie la plus efficace est de réaliser une recherche de QTL (Quantitative Trait Loci) sur l'ensemble du génome. Néanmoins, cette démarche nécessite l'établissement d'une carte génétique.

Une carte génétique correspond à l'ordonnement de marqueurs moléculaires ou phénotypiques le long des chromosomes, l'ordre et la distance entre les marqueurs étant déterminés par la fréquence des recombinaisons (crossing-over) lors de la méiose entre les deux chromosomes d'une paire. La probabilité qu'un tel événement se produise est proportionnelle à la distance qui sépare ces deux marqueurs. Ainsi, la fréquence de recombinaison reflète les distances entre marqueurs, et l'unité de distance est le centiMorgan (cM), sachant qu'1 cM est égal à 1% de recombinaison, soit 1 recombinaison pour 100 méioses analysées.

Pour réaliser une carte génétique à haute densité, il est nécessaire de disposer d'un grand nombre de marqueurs informatifs, c'est à dire des marqueurs polymorphes permettant de différencier les deux chromosomes d'une paire en un point précis du génome afin d'identifier le segment du chromosome provenant du père de celui provenant de la mère. Cette différence de la forme du marqueur –dénommé allèle- se trouve au niveau de la séquence de l'ADN du marqueur et l'on utilisera généralement soit une variation du nombre de répétitions d'un motif nucléotidique court (marqueurs microsatellites), soit plus récemment le changement d'une base nucléotidique (marqueurs SNP: Single Nucleotide Polymorphism).

Depuis le début des années 1990, la stratégie type pour développer une carte génétique est (i) le développement de marqueurs microsatellites par criblage de banque d'ADN génomique (Genet et al., 2003) ou à partir des séquences disponibles dans les bases de données (Feve et al., 2006 ; Marie-Etancelin et al., 2006), (ii) la constitution de familles dans lesquelles vont être analysées les liaisons entre marqueurs par étude de la transmission des allèles des parents vers les descendants (Marie-Etancelin et al., 2008) et (iii) le génotypage des individus, c'est-à-dire la détermination des allèles portés par chaque individu pour chacun des marqueurs. Pour plus de détails, voir Boichard et al. (1998).

Les données sur le génome du canard sont rares et seule une carte génétique a été publiée à ce jour par une équipe Chinoise (Huang et al., 2006). Notre objectif est donc de combler cette lacune en développant une carte qui permettra le positionnement de nos marqueurs, mais surtout l'identification et la localisation de QTL pour des caractères d'intérêt pour le canard en France, comme la qualité des produits issus du gavage ou la résistance au portage de maladies (programme ANR « GENEKAN »).

1. MATERIELS ET METHODES

1.1. Le dispositif animal

Une des 7 familles du dispositif animal back-cross (BC) du programme ANR « GENEKAN » a été utilisée (Marie-Etancelin et al., 2008) pour la constitution de cette carte (cf. figure 1). Cette famille est composée au total de 75 individus, dont 62 femelles BC descendantes d'un seul père. L'ADN de ces 75 individus a été extrait à partir d'échantillons de sang.

1.2 Sélection des marqueurs

Au début des travaux, nous disposions d'une part de 161 marqueurs microsatellites développés au laboratoire INRA de Génétique Cellulaire (LGC) (Genet et al., 2003 ; Feve et al., 2006), et d'autre part, d'un plus grand nombre de marqueurs dont les séquences avaient été récemment rendues disponibles dans les bases de données internationales de séquences nucléotidiques.

Afin d'optimiser le choix des marqueurs le long du génome du canard, nous avons tiré parti de la grande similarité entre les caryotypes de canard et de poule (Fillon et al., 2007). Ainsi, toutes les séquences nucléotidiques de marqueurs de canard ont été localisées par similarité de séquence sur le génome de la poule, permettant de prédire leurs positions chez le canard. De cette façon, un total de 323 marqueurs microsatellites a été retenu (cf. figure 2).

1.3 Génotypage et cartographie

Les génotypages ont été principalement effectués à l'aide des différents outils disponibles sur la plateforme génomique de Toulouse : les robots pipetteurs ont permis la préparation des réactions d'amplification en plaques (format 384 puits) et le mélange des marqueurs amplifiés (cf. figure 3). Puis la séparation des différents produits d'amplifications a été effectuée sur des séquenceurs ABI3100 et ABI3730.

L'analyse des électrophorégrammes a été réalisée à l'aide du logiciel Genemapper V4.0, puis les données de génotypages ont été intégrées dans notre base de

données (GEMMA, LIMS développé au laboratoire), afin de vérifier leurs cohérences par un test de compatibilité mendélienne (analyse de la transmission des allèles parentaux à la descendance). La construction des groupes de liaisons a été effectuée à l'aide du logiciel Crimap (version 2.4).

2. RESULTATS ET DISCUSSION

2.1 Sélection des marqueurs et génotypage

Trois cent vingt trois marqueurs microsatellites ont été sélectionnés, principalement en fonction de la position d'une séquence similaire sur le génome de la poule. Pour réaliser une carte, le plus grand nombre de marqueurs possible doit être hétérozygote chez un même parent, afin d'obtenir l'information de transmission de l'un ou l'autre des 2 allèles dans la descendance (marqueur informatif). Les 323 marqueurs ont donc été testés sur les 7 mâles F1 du dispositif « GENECA ». Seuls 112 se sont révélés informatifs pour au moins 3 des 7 mâles : 36 sur 161 marqueurs développés au LGC étaient informatifs (28.5%) alors que 75 sur 162 marqueurs développés à partir des bases de données étaient informatifs (45.8%). Cette meilleure informativité des marqueurs développés à partir des bases de données résulte de contraintes plus fortes sur le choix des séquences (seuls les marqueurs possédant un motif répété de grande taille ont été retenus). Parmi les 112 marqueurs informatifs, 105 présentent des génotypes sûrs (i.e. avec profil interprétable et sans erreur de compatibilité mendélienne) permettant de les retenir pour la suite des premières analyses.

La famille retenue pour réaliser cette première carte est celle dont le mâle F1 présentait le plus grand nombre de descendants et, qui plus est, disposait du plus grand nombre de marqueurs informatifs.

2.2 Prédiction de la localisation des marqueurs

Pour 92 de ces 105 marqueurs, une prédiction fiable ($e\text{-value} > 10^{-15}$) de la localisation sur le génome du canard peut être déduite de la position d'une séquence similaire sur le génome de la poule (cf. tableau 1).

2.3 Construction des groupes de liaison

En première approche, l'analyse des taux de recombinaison entre marqueurs 2 à 2 a permis la construction de 13 groupes de liaisons dans lesquels se répartissent 65 marqueurs. Parmi les 40 marqueurs restants, 30 ne présentent pas suffisamment de méioses informatives dans cette famille pour être localisés et 10 marqueurs ont un profil particulier, laissant supposer qu'ils sont localisés sur le chromosome Z ou porteur d'un allèle nul.

En tenant compte des prédictions de localisation des marqueurs sur la séquence de la poule préalablement décrite, des groupes de liaison correspondant à un même chromosome ont été regroupés pour une analyse multipoints. Au final, les 13 groupes de

liaison ont pu être assignés à 9 chromosomes différents (cf. figure 4).

En premier lieu, 4 marqueurs sans séquence similaire chez la poule, et donc sans position prédite chez le canard, ont été localisés sur les groupes de liaisons 1 (CAM93 et APT21) et 3 (AM1033 et CAM022) par l'analyse de liaison. Par ailleurs, un seul marqueur (APH24) n'a pas été localisé dans le groupe de liaison prédit. L'approche de la sélection de marqueurs par comparaison de séquence avec la poule qui avait été démontrée chez la caille (Kayang et al., 2006), reste efficace chez le canard, malgré une distance phylogénétique plus grande.

L'ordre des 65 marqueurs répartis en 13 groupes de liaison de la carte du canard a été comparé à celui des séquences similaires sur la séquence de la poule (cf. figure 4). Pour les 13 groupes de liaison construits, seuls 4 remaniements intra-chromosomiques potentiels sont détectés, sur les groupes de liaison 1, 2, 3 et 4.

On notera une répartition des marqueurs globalement homogène sur l'ensemble des grands chromosomes : seul le bas du chromosome 2 et le haut du chromosome 5 sont pour l'instant mal couverts.

CONCLUSION

Aujourd'hui cette première carte du canard se compose de 65 marqueurs informatifs pour cette famille de père : le travail se poursuit donc avec le génotypage des 6 autres familles de pères, ceci afin d'une part de localiser les 30 marqueurs microsatellites complémentaires et d'atteindre l'objectif fixé d'une centaine de microsatellites informatifs et d'autre part de réaliser une première analyse de détection de QTL sur ce dispositif complet. Néanmoins, vu le faible taux de polymorphisme des marqueurs microsatellites (moins de 50%), il sera long et difficile d'obtenir une carte dense en marqueurs de ce type. Par contre, avec les nouvelles technologies, il est maintenant possible d'envisager le développement rapide de marqueurs SNP. Nous avons donc prévu d'enrichir notre carte en identifiant un très grand nombre de SNP par séquençage d'une fraction du génome des individus F1 issus du croisement des 2 lignées du dispositif QTL (projet AIP Bioressources 2008 « Pyresavi », Pitel et al., JRA 2009).

REMERCIEMENTS

Les auteurs remercient l'ANR Genanimal, AGENAVI/CIFOG et l'InterRégion Midi-Pyrénées Aquitaine pour le financement du programme. Les génotypages ont été effectués sur la plateforme génomique de la Génopole de Toulouse Midi-Pyrénées. <http://genomique.genotoul.fr/>

REFERENCES BIBLIOGRAPHIQUES

- Boichard D., Le Roy P., Levéziel H., Elsen J.M., 1998, INRA Prod. Anim., 11(1), 67-80.
- Feve K., Genet C., Cardinet G., Vignoles F., Vignal A., Marie-Etancelin C., 2006, JRPFG p13-16.
- Fillon V., Vignoles M., Crooijmans R.P.M.A., Groenen M.A.M., Zoorob R., Vignal A., 2007, Animal Genetics, 38, 303-307.
- Genet C., Vignal A., Larzul C., 2003, 5e JRA, p 395-398.
- Huang Y., Zhao Y., Haley C.S., Hu S., Hao J., WU C., Li N., 2006, Genetics, 173 : 287-296.
- Kayang B., Fillon V., Inoue-Murayama M., Miwa M., Leroux S., Fève K., Monvoisin J.L., Pitel F., Vignoles M., Mouilhayrat C., Beaumont C., Ito S., Minvielle F., Vignal A., 2006, BMC Genomics, 7: 101.
- Marie-Etancelin C., Feve K., Genet C., Cardinet G., Vignoles F., Pitel F., Vignal A. 2006. Symposium COA (Taiwan) and INRA (France).
- Marie-Etancelin C., Baéza E., Basso B., Bastianelli D., Bernadet M.D., Brun J.M., Davail S., Dubos F., Fernandez X., Guéméné D., Guy G., Manse H., Mialon M.M., Larzul C., 2008. 8ème JRPFG, 30 et 31 octobre 2008, Arcachon, France.
- Pitel F., Vignal A., Leroux S., Feve K., Tircazes A., Morisson M., Gourichon D., Minvielle F., Leterrier C., Arnould C., Dubos F., Bernadet M.D., Marie-Etancelin C., Basso B., Herault F., Lecerf F., Besnard J., Calenge F., Beaumont C., Klopp C., Diot C., 2008, 8ème JRA, St Malo, 25 et 26 mars 2009.

Figure 1. Dispositif expérimental utilisé pour la construction de la carte

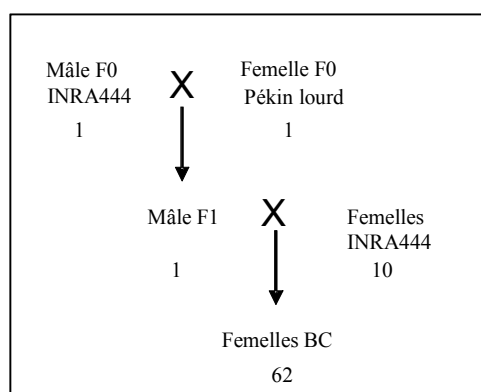
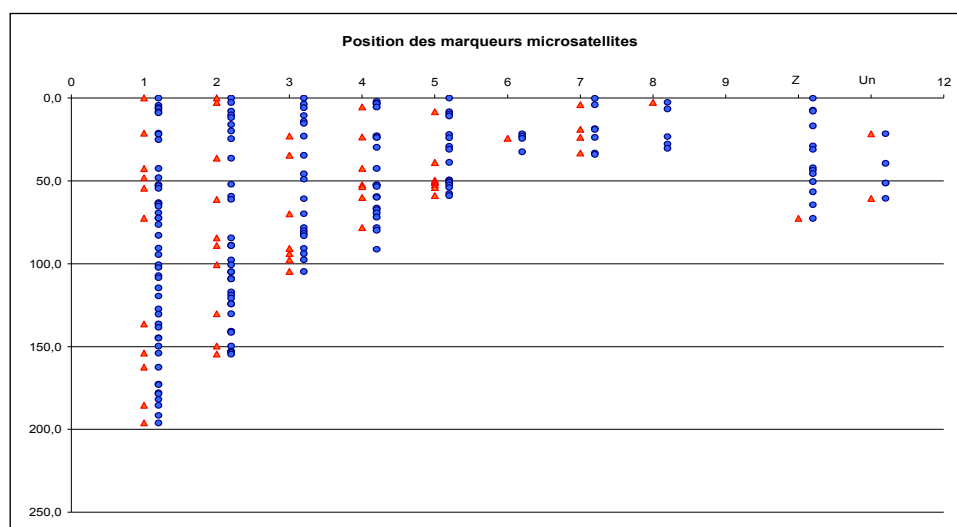


Figure 2 : Répartition des microsatellites canard sélectionnés sur la séquence de la poule (seuls les 10 plus grands chromosomes sont représentés)



Légende : les ronds bleus représentent les 323 marqueurs testés et les triangles rouges les marqueurs informatifs pour au moins 3 des 7 mâles F1 et donc retenus pour le génotypage.

Figure 3 : Electrophorégramme d'un mélange de 8 marqueurs microsatellites, pour 1 individu. Les marqueurs sont mélangés en fonction de leur taille et de leur fluorophore associé.

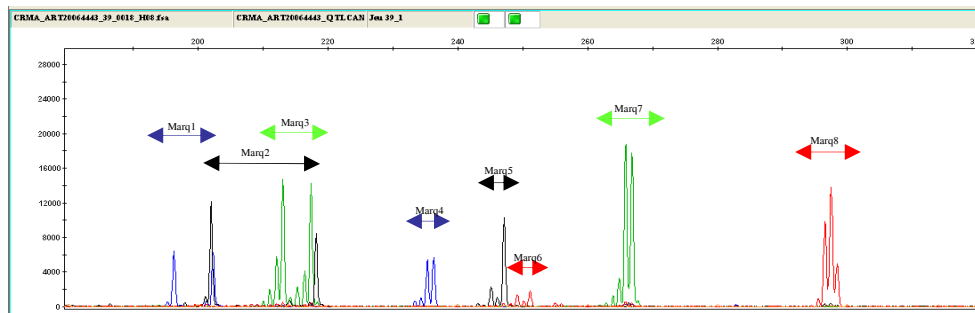


Tableau 1. Répartition des 112 marqueurs informatifs canard sur la séquence des chromosomes de la poule.

GGA01 = 19	GGA09 = 3	GGA19 = 3
GGA02 = 13	GGA10 = 1	GGA20 = 3
GGA03 = 9	GGA11 = 1	GGA21 = 2
GGA04 = 8	GGA12 = 3	GGA23 = 1
GGA05 = 8	GGA13 = 3	GGA24 = 1
GGA06 = 2	GGA15 = 1	GGA26 = 1
GGA07 = 8	GGA17 = 2	GGA Z = 2
GGA08 = 2	GGA18 = 2	Unk = 13

Figure 4 : Comparaison entre la localisation des marqueurs sur les groupes de liaisons (GL) du canard (en centiMorgan) et la position « virtuelle » sur les chromosomes (GGA) de la poule (en MégaBases).

