HAL
open science

Mise au point d'un panel d'assignation de parenté SNP unique pour deux espèces de canards et leur hybride

Sophie Brard-Fudulea, Hervé Chapuis, Sophie Leroux, Romuald Rouger, Alain Vignal, Christian Diot, Marc Teissier

To cite this version:

Sophie Brard-Fudulea, Hervé Chapuis, Sophie Leroux, Romuald Rouger, Alain Vignal, et al.. Mise au point d'un panel d'assignation de parenté SNP unique pour deux espèces de canards et leur hybride. 14. Journées de la Recherche Avicole et Palmipèdes à Foie Gras, Itavi, Mar 2022, Tours, France. pp.628, 10.1016/j.anscip.2022.05.026 . hal-03617695

HAL Id: hal-03617695

https://hal.inrae.fr/hal-03617695

Submitted on 23 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MISE AU POINT D'UN PANEL D'ASSIGNATION DE PARENTE SNP UNIQUE POUR DEUX ESPECES DE CANARDS ET LEUR HYBRIDE

Sophie Brard-Fudulea[1], Hervé Chapuis[2], Sophie Leroux[2], Romuald Rouger[1], Alain Vignal[2], Christian Diot[3], Marc Teissier[2]

[1]SYSAAF, UMR BOA, Centre INRAE Val de Loire, 37380, NOUZILLY
[2] GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326, CASTANET TOLOSAN
[3]PEGASE UMR 1348, Agrocampus Ouest, INRA, 35590 SAINT-GILLES,

sophie.brard-fudulea@inrae.fr

RÉSUMÉ
En espèces avicoles, l'assignation de parenté représente une alternative au suivi du pedigree réalisé via la cage individuelle et l'insémination avec la semence d'un mâle identifié. Compte-tenu des effectifs d'animaux par génération et de la valeur de chaque individu, un enjeu majeur est le coût du panel d'assignation, conditionné par le nombre de marqueurs. Un défi supplémentaire existe chez le canard, où deux espèces sont croisées pour obtenir un hybride : l'assignation du mulard nécessite des marqueurs fonctionnant chez le mulard, ainsi que chez le canard commun et le canard de Barbarie. L'objectif est le développement d'un panel d'assignation de parenté de 96 marqueurs SNP pour l'assignation de canetons Pékin, Barbarie et mulards obtenus sans recours à la cage. Compte-tenu des effectifs importants éclos par génération, il est important que le panel soit de la plus petite taille possible afin de minimiser son coût d'utilisation. A partir de la puce bi-espèce canard 670K SNP, 192 marqueurs ont été présélectionnés. Les critères de choix ont porté sur la faisabilité d'un changement de technologie, et sur la qualité et le polymorphisme des marqueurs. Les 192 marqueurs retenus ont été génotypés sur les parents des animaux des lignées expérimentales à assigner, ainsi que sur des animaux ayant des liens de filiation connus, pour un total de 334 échantillons. Les marqueurs présentant un taux de génotypage supérieur à 95% et aucune incompatibilité mendélienne ont été conservés, puis un tri a été effectué sur la fréquence allélique, et le déséquilibre de liaison. Une liste de 96 marqueurs a été proposée à l'issue de ces filtres. Les taux d'assignation élevés obtenus avec ces 96 marqueurs (autour de 98%) ont permis de valider la qualité du panel. Les 192 marqueurs utilisés étant issus d'une puce développée sur plusieurs lignées commerciales aux orientations diverses, il devrait être possible de proposer des listes de 96 marqueurs opérationnelles pour d'autres lignées.

ABSTRACT
Design of a SNP parentage assignment panel for two different species of duck and their hybrid
In avian species, parentage assignment is one of the key to get the pedigree of animals breed without cages nor insemination. To allow its use in a large number of individuals (several hundred each generation in each line), the SNP marker panel should be as cheap as possible (96 SNP). In ducks, another need is the possibility of assigning two different species, the common duck and the Muscovy duck, and also hybrids from these two species: the mules duck. The objective of this study was to develop a small SNP chip panel efficient for both species and their hybrid. First, 192 SNP from the 670K SNP chip were selected, based on the feasibility of moving from one genotyping technology to a new one, on their quality, and on their polymorphism. The 192 SNP were genotyped on 334 samples, including the parents of experimental lines for which a parentage assignment was needed, and samples already genotyped on the 670K SNP chip for validation purpose. Among the 192 SNP, the 96 best SNP were chosen based on several quality filters: call rate < 0.95, zero mendelian incompatibility, high minor allelic frequency, and low linkage disequilibrium. Progenies of the experimental lines were successfully assigned with the 96 selected SNP. As the 192 SNP used here come from a high-density SNP chip developed using several commercial breeds, specific panels of 96 SNP should be informative for parentage assignment of other breeds than the two experimental lines targeted in this study.

INTRODUCTION

La sélection du canard nécessite actuellement d'élever les reproducteurs en cages individuelles, afin de permettre l'enregistrement du pedigree qui est nécessaire pour la réalisation des évaluations génétiques. Ce mode d'élevage ne correspond cependant plus aux attentes de la société, et l'interdiction ou la restriction de l'usage de la cage est actuellement envisagée à l'échelle européenne. Pour le suivi du pedigree, une alternative à la cage est l'assignation de parenté, qui consiste à génotyper parents et descendants sur un panel restreint de marqueurs, et à utiliser l'information génomique pour reconstituer les liens de parenté. Chez le canard, un panel de microsatellites existe (Chapuis et al. 2010), mais son taux de succès est insuffisant pour une utilisation en routine. Depuis plusieurs années, ce sont les marqueurs de type Single Nucleotide Polymorphism (SNP) qui sont préférés pour ce type d'analyses, avec des taux d'assignation élevés atteignables avec un nombre restreint de marqueurs (Tortereau et al. 2017). Il est important que le panel soit de petite taille afin de limiter son coût. Chez le canard, une contrainte majeure est le besoin d'un outil unique pour les deux genres (*Anas platyrhynchos* et *Cairina moschata*) et leur hybride. Ces canards pékin, barbarie et mulards font l'objet d'un programme de recherche INRAE sur le comportement et l'efficacité alimentaire, dans lequel l'enregistrement du pedigree n'a pu se réaliser dans les conditions habituelles du fait de la pandémie de Covid-19. L'objectif de travaux présentés ici est de produire un panel SNP de petite taille en technologie KASPar à partir de la puce haute-densité 670K Axiom, afin de permettre l'assignation des animaux de ce dispositif expérimental. La technologie KASPar a été retenue en raison de son faible coût et de sa flexibilité : il est possible d'une analyse à l'autre de modifier la composition du panel si cela est nécessaire, avec pour seul coût additionnel celui de la commande des amorces des éventuels nouveaux marqueurs. Le passage de la technologie Axiom à la technologie KASPar nécessite cependant d'utiliser des amorces plus longues (50 pb au lieu de 35 pb), ce qui peut amener à écarter certains marqueurs pour lesquels du polymorphisme serait présent dans les amorces allongées. Le coût d'un génotypage étant établi par multiple de 96 marqueurs utilisés, les travaux présentés ci-après testeront 192 SNP afin d'en retenir 96 de bonne qualité pour l'objectif d'assignation de parenté. A terme, l'objectif est que les 192 marqueurs puissent également être testés et utilisés par les sélectionneurs français sur des lignées commerciales, et par les gestionnaires de races locales de canards.

1. MATERIELS ET METHODES

1.1. Choix de 192 marqueurs pour design en technologie KASPar

La puce Axiom bi-espèce 670K SNP compte 343 950 SNP dans la librairie du canard barbarie, et 331 241 SNP dans la librairie du canard pékin. (Thébault et al. 2019) Au total, 1305 SNP sont partagés entre ces deux librairies et sont *a priori* susceptibles de fonctionner chez le mulard. Cependant, il avait été constaté dans le programme CanArray qu'en dehors de l'intersection de ces deux listes, des marqueurs de la librairie canard pékin fonctionnaient chez le canard barbarie, et inversement. La totalité des marqueurs de la puce ont donc été considérés comme candidats pour le choix des 192 marqueurs. Afin de retenir 192 marqueurs pour le design en technologie KASPar, les amorces des SNP de la puce 600K, d'une longueur de 35 paires de bases (pb) ont été allongées à 50pb en alignant les séquences des marqueurs sur les génomes de référence des deux espèces, et en éliminant les marqueurs présentant un risque de polymorphisme dans les amorces compte-tenu des polymorphismes détectés lors de la conception de la puce 670K. Seuls les marqueurs présentant exactement les mêmes séquences d'amorces et le même polymorphisme attendu dans les deux espèces ont été retenus. Le logiciel Axiom Analysis Suite (AxAS) a été utilisé pour produire les fichiers de génotypage, en ne conservant que les marqueurs de statut PolyHighResolution (taux de génotypage élevé et présence des trois génotypes possibles AA, AB et BB).

Les SNP répondant à ces exigences ont ensuite été filtrés en utilisant des échantillons de bonne qualité (moins de 5% de données manquantes) : 139 canards pékin, 79 canards barbarie, et 39 mulards (incluant 125 trios père-mère-descendant), analysés avec le logiciel Plink v2.0 (Bender et al. 2007). Les filtres suivant ont été appliqués : les marqueurs devaient avoir un call rate d'au moins 95%, aucune incompatibilité mendélienne, une fréquence de l'allèle minoritaire (MAF) d'au moins 0,10 chez le canard pékin et le canard de Barbarie, ne pas être fixés chez les mulards et être peut redondants (déséquilibre de liaison inférieur à 0,25). Les marqueurs retenus ont été soumis au laboratoire INRAE Gentyane (Clermont-Ferrand) pour design des amorces.

1.2. Sélection de 96 marqueurs pour le panel AsParCan

Afin de choisir parmi les 192 SNP retenus les 96 marqueurs qui seront utilisés pour l'assignation dans le cadre du programme DuckoDac, des échantillons ont été génotypés au laboratoire INRAE Gentyane. Il s'agissait pour les lignées expérimentales INRAE des parents des canetons à assigner (134 canards pékins et 128 canards barbarie), et d'échantillons issus du projet CanArray (44 canards pékins, 15 canards barbaries, et 13 mulards). Ces deux types d'échantillons ont été utilisés afin d'une part de sélectionner les marqueurs

les plus polymorphes dans les lignées d'intérêt, et d'autre part de valider le changement de technologie de l'Axiom vers le KASPar via des échantillons déjà génotypés sur la puce 670K. Pour réaliser les analyses, les outils suivants ont été utilisés : SNP Genotyping Analysis pour visualiser le clustering et produire les fichiers de génotypages, et Plink v2.0 pour le calcul des critères de qualité sur les échantillons et les SNP.

Les échantillons présentant beaucoup de données manquantes (plus de 50%) ont été éliminés avant d'appliquer les filtres de qualité aux SNP. Les critères suivants ont été calculés sur le jeu de données : taux de génotypage (CR), fréquence de l'allèle minoritaire intra-lignée (MAF), équilibre d'Hardy-Weinberg intra lignée (HWE), taux d'incompatibilités mendéliennes entre les échantillons CanArray entre lesquels il y avait des liens de filiation connus (11 trios et 4 duos). Trois échantillon génotypés deux fois ont été utilisés pour regarder la cohérence des deux résultats de génotypage. Les résultats du génotypage 670K en Axiom et 192 SNP en KASPar ont été comparés pour les échantillons génotypés dans les deux projets. La capacité des SNP retenus à assigner les animaux de pedigree connu a été testée en utilisant le logiciel APIS.

2. RESULTATS ET DISCUSSION

2.1. Choix de 192 marqueurs pour design en technologie KASPar

L'allongement des amorces des marqueurs de la puce 670K pour passer de 35 à 50 pb, sans polymorphisme dans les amorces, a permis de conserver 229 138 SNP dans la librairie barbarie, et 198 091 SNP dans la librairie pékin. L'intersection de ces deux listes comptait 399 marqueurs ayant exactement les mêmes séquences d'amorces et le même polymorphisme attendu dans les deux espèces. Ces 399 marqueurs ont été considérés comme candidats pour la constitution du panel d'assignation de parenté. Trois de ces marqueurs n'avaient pas le statut PolyHighResolution après analyse dans AxAS et ont été écartés. Seuls les échantillons présentant moins de 5% de données manquantes ont été utilisés pour filtrer les SNP sur leur qualité : après application de ce filtre, il restait 139 canards pékin, 79 canards barbarie, et 39 canards mulards.

Quarante-huit marqueurs présentant plus de 5% de données manquantes ont été éliminés. Après application du filtre sur les incompatibilités mendéliennes, il restait 336 SNP. Les filtres sur la MAF chez les pékins, barbarie et mulards a conduit à retenir 232 SNP. Le filtre sur le DL a permis d'éliminer 12 marqueurs présentant un DL supérieur à 0,25 avec d'autres marqueurs. Ce dernier filtre a permis de produire une liste de 192 marqueurs, ainsi qu'une liste complémentaire de 8 marqueurs dans le cas où pour des raisons techniques un ou plusieurs SNP de la liste principale ne pourrait être mis au point.

2.2. Sélection de 96 marqueurs pour le panel AsParCan

A l'issue du génotypage sur 192 SNP, des CR plus faibles ont été observés chez les barbarie INRAE que chez les pékin INRAE (tableau 1). Cela concerne 57 SNP, qui ont des CR compris entre 0.25 et 0.58. Ces marqueurs présentent par ailleurs des CR proches de 1 dans les échantillons barbarie CanArray. Notre hypothèse est qu'il existe du polymorphisme dans les amorces de ces marqueurs, non détectés dans CanArray (i) du fait du faible nombre d'animaux expérimentaux INRAE génotypés dans ce précédent projet (et non-séquencés), et potentiellement (ii) du fait que globalement moins de lignées barbaries que de lignées pékin ont été analysées dans le programme CanArray. Ce résultat souligne l'importance lors d'un changement de technologie de génotypage d'une part de confronter les résultats obtenus avec les deux outils, et d'autre part de prévoir pour la phase de design plus de marqueurs que l'objectif visé, afin de s'assurer que le nombre de marqueurs analysables *in fine* sera suffisant.

Six marqueurs présentant une incompatibilité mendélienne ou plus ont été écartés. Parmi les SNP restant, aucun n'avait une *P-value* inférieure à 10^{-5} pour le test d'Hardy-Weinberg.

Les SNP les moins polymorphes dans les deux lignées expérimentales INRAE (MAF < 0.15) ont été éliminés. Il restait après ce filtre 111 marqueurs. Afin de ne conserver que les SNP présentant les meilleurs taux de génotypage chez les mulards, les marqueurs ayant un CR < 0.95 chez les mulards ont été supprimés, réduisant la liste à 104 marqueurs. Huit marqueurs ont finalement été supprimés sur la qualité de leur clustering.

La figure 1 présente la distribution de la MAF des marqueurs chez les parents du dispositif pékin INRAE d'une part, et chez les parents barbarie INRAE d'autre part. Toutes les classes de MAF entre 0.15 et 0.5 sont représentées. La figure 2 permet d'observer que les marqueurs sont répartis sur 22 des 28 chromosomes et groupes de liaison du canard Commun.

Les vérifications complémentaires réalisées sur les 96 SNP sélectionnés ont donné les résultats suivants : sur les 55 échantillons génotypés à la fois en technologie Axiom et KASPar, le taux de concordance entre les deux technologies était de 0.997. Les génotypages de trois animaux analysés deux fois sur le panel KASPar ont été comparés et ont montré 100% de cohérence dans leurs résultats. Onze trios père-mère-descendant déjà validés sur puce 670K étaient présents dans le jeu de donnée KASPar : les descendants ont pu être assignés sans erreur à leurs parents.

Des places disponibles sur les plaques d'échantillons ont permis de génotyper également sur les 192 SNP des canards de races locales : 34 canards de Duclair, et 10 canards de Rouen. Les MAF moyennes calculées dans ces deux races sont respectivement de

0.26 et 0.29. Les probabilités d'identité estimées sont très faibles ($5{,}5.10^{-49}$ et $2{,}6*10^{-45}$), et les probabilités d'exclusion sont supérieures à 0.99. Ces résultats préliminaires permettent de supposer que les 192 SNP, bien que choisis pour l'assignation de deux lignées expérimentales, devraient pouvoir être également valorisés chez des races locales.

CONCLUSION

Les travaux présentés ici ont permis, à partir d'une puce Axiom haute-densité 670K, de produire un panel d'assignation de parenté de 96 marqueurs en technologie KASPar fonctionnant chez le canard pékin, le canard de barbarie, et le mulard. Le critère ayant le plus contraint le choix des SNP était leur pertinence simultanée pour les deux genres de canard et leur hybride. Les filtres appliqués ensuite sur la qualité de ces marqueurs ont permis de réduire la liste de SNP candidats aux 96 marqueurs nécessaires pour maximiser le taux de succès de l'assignation par rapport au coût de l'outil de génotypage. Le panel a été optimisé pour les deux lignées expérimentales INRAE, mais la prise en compte des génotypages de lignées commerciales devrait permettre son utilisation pour les lignées des sélectionneurs français. Les 1ers résultats de génotypage obtenus sur des individus de races locales sont prometteurs, bien que ces races n'aient pas été prises en compte dans la phase de filtre des marqueurs. La 1ère utilisation fructueuse de cet outil a été réalisée par Chapuis et al. 2022.

REMERCIEMENTS

Les auteurs remercient le Département Génétique Animale INRAE pour son soutien au programme ASPARCAN dont est issu ce panel. Merci également au sélectionneur Grimaud Frères pour l'autorisation d'exploitation de leurs résultats de génotypage issus de CanArray, et à Jocelyn Marguerie (Collectif pour la Sauvegarde des Races Avicoles Normandes - CSRAN) pour la collecte des échantillons de races locales.

REFERENCES BIBLIOGRAPHIQUES

Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. 2007. American Journal of Human Genetics, 81

Chapuis H., Faugeras R., Rossignol M.N., Feve K., Genestout L., Chantry-Darmon C., Guemene D., 2010. 9èmes Journ. Rech. Palmipèdes à Foie Gras, Bordeaux, France, 1-4.

Chapuis H., Hazard A., Brard-Fudulea S., Martin X., Amadio B., Barrieu J., Mallet J.M., Morganx P., Garrigues-Cardinet G., Gilbert H. 2022. 14èmes Journ. Rech. Palmipèdes à Foie Gras, Tours, France

Thébault N., Riquet J., Diot C., Brard-Fudulea S., Guémené D., Alletru B., Cornil M., Bouleau P.,Blanchet M., le Mignon G., Demeure O., Vignal A. 2019. 13èmes Journ. Rech. Palmipèdes à Foie Gras, Tours, France

Tortereau, F., Moreno, C.R., Tosser-Klopp, G., Servin, B., Raoul, J., 2017. BMC Genet. 18, 50

Tableau 1. Taux de génotypage (CR) et Fréquence de l'Allèle Minoritaire (MAF) observés sur les 192 SNP génotypés sur les lignées expérimentales INRAE

	Taux de génotypage par SNP		**Fréquence de l'allèle minoritaire**	
	Pékin INRAE	Barbarie INRAE	Pékin INRAE	Barbarie INRAE
Minimum	0.84	0.25	0.02	0.05
Médiane	1.00	0.99	0.34	0.32
Maximum	1.00	1.00	0.50	0.50

Figure 1. Distribution de la MAF des 96 marqueurs du panel d'assignation dans les lignées expérimentales INRAE pékin (à gauche) et barbarie (à droite)

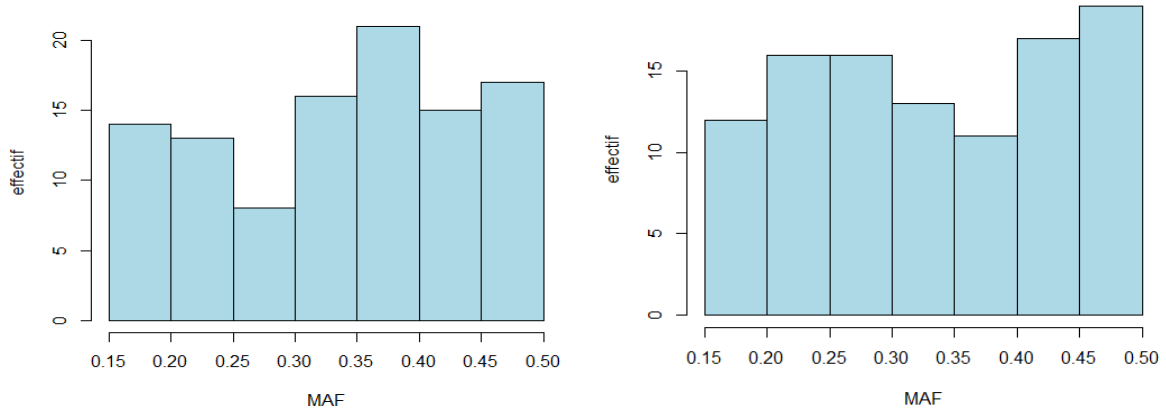
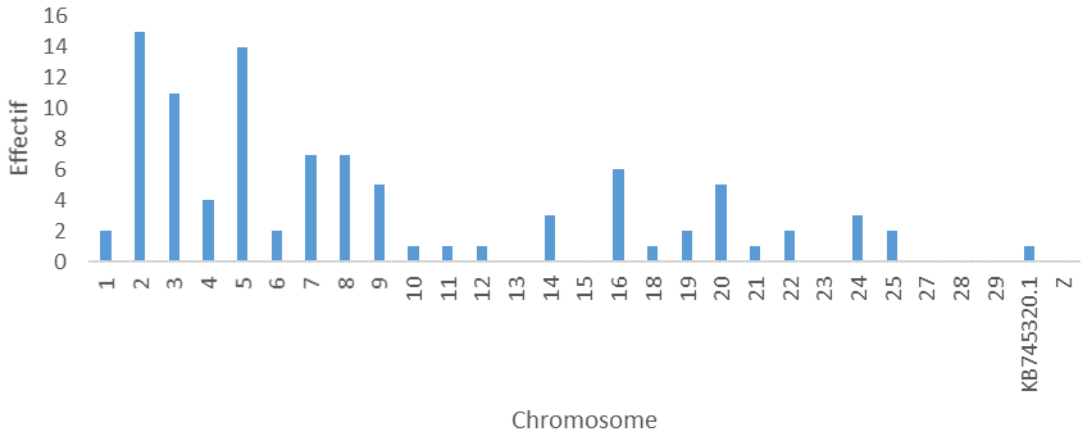
Figure 2. Répartition des 96 SNP du panel d'assignation entre les chromosomes