

Assignation de parenté par marqueurs SNP chez les bovins français

Parentage assignment with SNP markers in French cattle

BARBOTTE L. (1), GENESTOUT L. (1), FRITZ S. (2), CHANTRY DARMON C. (1), BOICHARD D. (1, 3)

(1) LABOGENA, Domaine de Vilvert, 78350 Jouy-en-Josas, France

(2) UNCEIA, Domaine de Vilvert 78350 Jouy-en-Josas, France

(3) INRA, UMR 1313GABI, 78350 Jouy-en-Josas, France

INTRODUCTION

Les marqueurs génétiques peuvent contribuer à une meilleure connaissance des parentés de deux façons: a) si les parentés sont a priori connues, les marqueurs peuvent les confirmer, en détectant les parentés fausses. C'est l'utilisation actuellement la plus fréquente chez les ruminants. b) si les parentés sont inconnues comme en aquaculture, les marqueurs peuvent aider à les déterminer, et on parle alors d'assignation.

Ces 15 dernières années, les marqueurs de choix étaient les microsatellites, du fait de leur fort polymorphisme (environ 10 allèles par marqueur) mais dont l'inconvénient est l'impossibilité d'une complète automatisation de la chaîne d'analyse. On tend actuellement à utiliser les SNP (ou mutations ponctuelles), moins informatifs (en général 2 allèles) mais dont le génotypage est complètement automatisable et multiplexable, permettant leur analyse en grand nombre simultanément.

Cette étude vise à établir les caractéristiques du jeu de SNP minimum (coût d'analyse le plus bas possible) permettant une assignation totale dans l'ensemble des populations françaises. On se place dans des conditions défavorables : les mères sont non typées, les parents potentiels sont fortement apparentés et les pères ne sont pas tous typés.

1. METHODES D'ASSIGNATION

L'assignation repose sur la recherche de compatibilité entre un individu et ses parents potentiels. Non seulement les parents incompatibles sont exclus mais les parents compatibles sont classés en fonction de leur vraisemblance, calculée à partir des probabilités de transmission mendélienne des marqueurs. L'assignation est parfaite si un seul parent, ou un seul couple de parents, est identifié sans ambiguïté, car significativement le plus vraisemblable. Nous avons développé un logiciel d'assignation par maximum de vraisemblance qui accepte des erreurs de typages.

2. EFFET DES FREQUENCES ALLELIQUES

Un SNP est caractérisé par la fréquence de son allèle le plus rare dans une population cible (MAF pour Minor Allele Frequency). L'effet de la MAF est étudié par simulation d'une population de 1000 animaux issus de 50 pères et 100 mères fortement apparentés, les individus étant génotypés pour 25 à 125 SNP. Le tableau 1 montre que la puissance d'assignation augmente avec la MAF et avec le nombre de marqueurs.

Tableau 1 : Puissance d'assignation au père (la mère étant non typée), selon la MAF et le nombre de SNP (%)

	25	50	75	100	125
0,1	9,3	67,8	92,7	99,1	99,7
0,2	26,4	94,6	99,5	100,0	99,9
0,3	41,6	98,9	99,9	100,0	100,0
0,4	44,7	99,5	100,0	100,0	100,0

3. EFFET D'ERREURS DE GENOTYPAGE

Toute technique de génotypage produit une proportion de typages erronés (variable selon les technologies) qu'il faut prendre en compte dans l'assignation. Basé sur des simulations similaires au tableau 1, le tableau 2 montre qu'un taux d'erreur raisonnable ne remet pas en cause la capacité

d'assignation mais augmente légèrement le nombre de marqueurs nécessaires.

Tableau 2 : Puissance d'assignation au père (la mère étant non typée), selon le taux d'erreur de typage et le nombre de SNP de MAF égale à 0,3.

	25	50	75	100	125
0,00	43,3	99,5	100,0	100,0	100,0
0,01	47,6	98,9	99,9	100,0	100,0
0,02	37,0	97,0	99,8	100,0	100,0
0,03	32,1	96,1	99,5	99,9	100,0

Le nombre minimal de SNP utilisés doit être majoré par rapport au nombre théorique pour prendre en compte un taux de typages manquants de 2% par individu, ce qui peut atteindre 6% pour un triplet père-mère-produit.

4. SELECTION D'UN JEU DE SNP POUR LE BOVIN

Cent marqueurs ont été ajoutés au panel bovin initial de 100 SNP établi par l'ISAG (International Society for Animal Genetics). Les critères de choix ont été les suivants : présence sur les puces Illumina; excellente qualité technique (taux d'erreur très faible et fréquence de typage manquant <1%) avec ces outils ; MAF moyennes élevées sur 8 races (moyenne=0,41) et confirmées sur 14 autres, à partir des typages réalisés à LABOGENA ; répartition sur tous les chromosomes avec au moins 10 cM entre marqueurs.

5. VALIDATION SUR DONNEES REELLES

Un lot de 704 génotypages dans 19 races pour ces 200 marqueurs a été extrait. Avec 200 marqueurs, 292 produits ont été assignés sans erreur à leur père (parmi 211 pères possibles). Les 201 autres produits, dont le père certifié n'était pas dans l'échantillon, n'ont pas obtenu de résultat d'assignation, indiquant l'absence de faux positif. La puissance d'assignation réelle a été étudiée en diminuant progressivement le nombre de marqueurs utilisés. Le tableau 3 montre que le vrai père est trouvé à partir de 75 marqueurs lorsque le père est typé (puissance), mais qu'il existe un risque important d'assignation fautive lorsque le typage du père n'est pas présent et que l'on dispose de moins de 150 SNP (spécificité).

Tableau 3 : Puissance et spécificité(*) d'assignation selon le nombre de SNP (493 produits dont 292 avec père typé et 201 avec père absent de l'échantillon)

Nb SNP	200	175	150	125	100	75	50
Puissance	100	100	99	99	99	99	91
Spécificité	100	100	93	78	57	32	13

(*) spécificité=1 – fréquence d'assignation à un père incorrect

6. PERSPECTIVES

Des travaux analogues sont conduits en ovins et le passage aux SNP peut être envisagé prochainement pour le contrôle de filiation pour les trois espèces de ruminants. Un petit nombre de SNP peut également être ajouté pour le contrôle du sexe et des caractères d'intérêt. Différentes plateformes sont possibles suivant le nombre de SNP à génotyper, 100-200 pour l'assignation et le contrôle de parenté seulement, ou plusieurs milliers pour la sélection génomique.

Nous remercions les responsables des différents programmes de génotypage sur puces dont les données ont été utilisées