

Vers une sélection génomique chez les caprins laitiers

C. CARILLIER-JACQUIN¹, H. LARROQUE¹, C. ROBERT-GRANIÉ¹

¹ GenPhySE, Université de Toulouse, INRA, INPT, ENVT, 31320, Castanet-Tolosan, France

Courriel : celine.carillier@inra.fr

L'arrivée récente d'une puce caprine de 53 000 marqueurs permet d'envisager la sélection génomique dans cette espèce. Cette étude présente, d'une part, un état des lieux de la population génotypée et, d'autre part, les premiers résultats d'évaluations génomiques obtenus dans cette espèce. Quelques perspectives d'applications ainsi que quelques pistes d'amélioration de ces résultats sont également évoquées.

La sélection génomique permet le plus souvent de prédire plus précisément la valeur génétique des reproducteurs en utilisant l'information de marqueurs moléculaires répartis sur le génome, les « *Single Nucleotide Polymorphism* » (SNP). Le principe des évaluations génomiques est basé sur l'utilisation d'une population dite « de référence » (population d'animaux disposant à la fois de génotypes et de phénotypes) à partir de laquelle sont estimées les relations entre les allèles aux marqueurs et les phénotypes, ces relations sont ensuite extrapolées aux candidats génotypés dont on ne connaît pas encore les phénotypes. Cette population de référence est généralement constituée de mâles de testage ancêtres (pères, grands-pères...) de la population candidate constituée de jeunes mâles ou femelles dont on souhaite prédire la valeur génétique.

Le progrès génétique annuel est une fonction des intensités de sélection et des intervalles de générations sur les quatre voies de sélection (*i.e.* père-fils, père-fille, mère-fils et mère-fille) ainsi que de la précision des évaluations et de la variance génétique. L'utilisation des puces à SNP permet de capter l'aléa de méiose, qui correspond à l'écart entre la moyenne des valeurs génétiques des parents et la valeur génétique du descendant, dès la naissance d'un individu. La précision des valeurs génétiques des individus dès leur plus jeune âge pouvant ainsi être augmentée, la sélection génomique peut être synonyme de changements dans le schéma de sélection. En effet, selon les contraintes de chaque espèce, elle peut permettre de sélectionner des individus plus jeunes et donc de diminuer les intervalles de génération, ou bien d'augmenter les intensités de sélection en génotypant plus de candidats

que ceux actuellement testés sur descendance. En bovins laitiers, la sélection génomique, déjà mise en place, connaît un franc succès en raison de sa capacité à prédire de manière plus précise la valeur génétique des mâles reproducteurs dès la naissance. Elle a donc permis d'éliminer le testage sur descendance, très long et coûteux pour les entreprises de sélection, diminuant ainsi les intervalles de générations. En revanche, l'économie du testage sur descendance peut permettre d'augmenter les intensités de sélection en génotypant un nombre plus élevé de candidats. La sélection génomique a permis un accroissement du progrès génétique d'au moins 80% (Colleau *et al* 2015).

La sélection génomique est actuellement déployée en bovins laitiers (depuis 2010 pour les principales races laitières, Fritz *et al* 2010), en bovins allaitants (depuis 2015, Tribout *et al* 2015), en ovins laitiers Lacaune (depuis 2015, Astruc *et al* 2016) et envisagée dans un avenir proche pour d'autres espèces d'élevage (volailles, Le Roy *et al* 2014 ; porcs, Tribout 2011, races ovines laitières des Pyrénées). En caprins, la disponibilité depuis 2011 d'une puce SNP contenant un peu plus de 53 000 marqueurs (dite 50k) permet d'envisager l'application de la sélection génomique dans cette espèce.

1 / Structure génétique et génomique de la population caprine

La structure génétique de la population comme le niveau de consanguinité et le déséquilibre de liaison, influence les précisions des valeurs génomiques atten-

dues. En effet, plus la diversité génétique d'une population est élevée, plus la relation marqueur-phénotype sera difficile à construire. Il faudra alors utiliser, quand cela est possible, une population de référence suffisamment grande pour être représentative de la population candidate à sélectionner. Par ailleurs, l'apparement entre la population candidate et la population de référence est également un élément clé de la précision des valeurs génomiques des candidats. En effet, si la population candidate est trop distante de la population de référence, le lien marqueur-phénotype établi dans la population de référence ne correspondra pas à celui existant dans la population candidate.

Après un contrôle qualité des génotypes, destiné à éliminer les marqueurs ou les génotypes complets sources possibles d'erreur, 677 mâles testés sur descendance dont 384 de race Alpine et 293 de race Saanen, 148 jeunes mâles (87 de race Alpine et 61 de race Saanen) nés en 2010 et 2011 et n'ayant pas encore de filles au moment de l'étude, ainsi que 1 985 femelles (1 243 de race Alpine et 742 de race Saanen) pour 46 959 marqueurs ont été conservés. Les femelles, issues de 20 pères différents (11 de race Alpine et 9 de race Saanen), ont été génotypées dans le cadre d'un dispositif de détection de QTL (Quantitative Trait Locus) c'est-à-dire de régions du génome ayant un effet sur les caractères d'intérêt zootechnique. Dans notre étude, les 677 mâles testés sur descendance (nés entre 1993 et 2009, dont la totalité des mâles de testage de races Alpine et Saanen nés à partir de 1998) constituent la population de référence, cette dernière peut inclure ou non, selon le cas testé, les femelles génotypées (nées en 2008 et 2009). La population candidate est constituée des

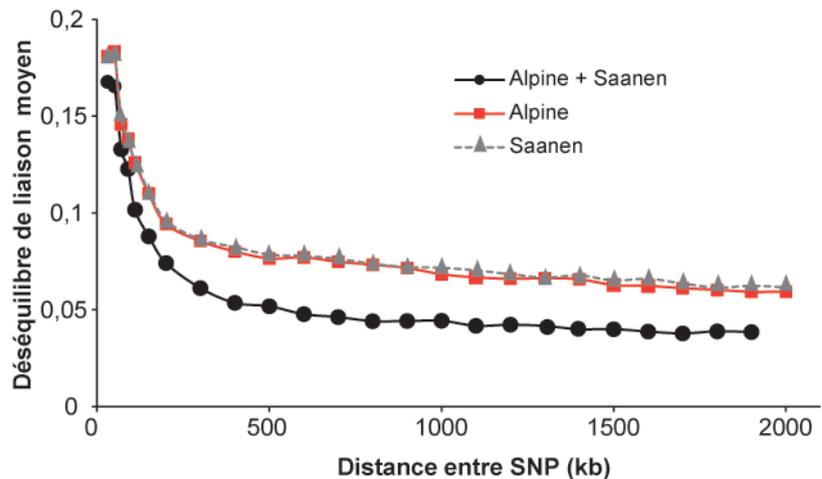
148 jeunes mâles génotypés, nés entre 2010 et 2011.

Étant donné la petite taille des populations génotypées (384 et 293 mâles génotypés et phénotypés respectivement en race Alpine et Saanen), l'intérêt d'utiliser une population multiraciale pour les évaluations génomiques a été étudié. Une attention particulière a été portée à la description de la structure génétique de la population multiraciale, comme à celle des populations « uniraciales » Alpine et Saanen, et de la distance génétique entre ces deux races.

1.1 / Déséquilibre de liaison

Le Déséquilibre de Liaison (DL) qui est une mesure de l'association préférentielle entre marqueurs a été calculé dans notre population de mâles génotypés (figure 1) selon la méthode de Rogers et Huff (2009). Comme attendu, le niveau de déséquilibre de liaison décroît avec la distance entre marqueurs pour se stabiliser à partir de 1000 kb à une valeur minimale de 0,03. Le niveau de DL dans la population multiraciale, comprenant les animaux de races Saanen et Alpine, est plus faible que celui estimé séparément dans chaque race. En effet, le déséquilibre de liaison estimé entre deux SNP adjacents (*i.e.* distance moyenne de 50 kb sur la puce) est d'environ 0,17 pour les populations « uniraciales » Alpine et Saanen et de 0,14 pour la population multiraciale. Cette différence entre le DL obtenu entre les populations multiraciales et « uniraciales » a également été observée en bovins laitiers (Hozé *et al* 2013). Ce niveau de DL estimé dans la population de mâles génotypés est plus faible que celui observé dans la population de femelles (0,22 à 50 kb pour la population multiraciale, résultats non montrés). Cette différence peut s'expliquer par le fort niveau d'apparentement entre les femelles (groupes de 100 demi-sœurs) qui a une influence positive sur le déséquilibre de liaison. Le DL obtenu dans la population caprine française concorde avec celui trouvé dans les populations canadiennes de mêmes

Figure 1. Niveau de déséquilibre de liaison moyen estimé dans la population de mâles génotypés multiraciale (Alpine + Saanen), « uniraciale » Alpine ou « uniraciale » Saanen.



rares (entre 0,14 et 0,15 ; Brito *et al* 2015). Ce niveau de déséquilibre de liaison est légèrement inférieur au niveau estimé en bovins laitiers français (Hozé *et al* 2013), mais proche de celui observé en ovins laitiers (Baloche *et al* 2013, Legarra *et al* 2014a).

1.2 / Variabilité génétique de la population de référence caprine

La population de référence caprine (hormis les femelles génotypées) a été choisie comme étant représentative de la population en sélection (*i.e.* animaux d'élevages adhérents au contrôle laitier). Il est intéressant de vérifier si le niveau de diversité génétique, estimée dans la population de référence choisie est similaire à celui de la population en sélection à l'aide, par exemple, des coefficients de consanguinité et de parenté.

La moyenne de la consanguinité estimée à partir du pedigree dans la population de référence caprine (2,1%, tableau 1) est proche de celle estimée dans la population de l'évaluation génétique (ou indexation) (2,3%, tableau 1). Cette dernière correspond aux mâles de monte naturelle, y compris les mâles de testage

non retenus comme améliorateurs, nés la même année que les animaux génotypés. Cependant, la population de jeunes mâles candidats génotypés est en moyenne légèrement moins consanguine (2,1 vs 3,1%, tableau 1) que la population totale de mâles nés entre 2010 et 2011. La plus faible consanguinité observée dans la population génotypée peut être expliquée par une récente gestion des accouplements (Colleau et Moureaux 2006, Colleau *et al* 2011,) pour le renouvellement des mâles d'Insémination Artificielle (IA) dans le but de préserver la diversité génétique.

Le coefficient de parenté moyen, estimé entre individus 2 à 2, constaté au sein de la population génotypée est proche de celui estimé dans la population d'indexation (tableau 1). Il est cependant légèrement supérieur pour la population de référence génotypée (1,8% vs 1,2% ; tableau 1) alors que l'inverse est observé pour la population candidate (1,3% vs 1,9% ; tableau 1). En revanche l'apparentement moyen estimé entre la population de référence et la population candidate (1,4%) est le même que celui constaté entre les mâles nés entre 1993 et 2009 et les mâles nés entre 2010 et 2011 de la population d'indexation.

Tableau 1. Coefficients de consanguinité et de parenté moyens (en %) estimé à partir du pedigree dans la population génotypée ainsi que dans la population d'animaux en élevage (population d'indexation).

	Population génotypée		Population d'indexation	
	Consanguinité	Parenté	Consanguinité	Parenté
Mâles nés entre 1993 et 2009¹	2,1	1,8	2,3	1,2
Mâles nés entre 2010 et 2011²	2,1	1,3	3,1	1,9

¹ Mâles correspondant à la population de référence.

² Mâles correspondant à la population candidate.

Les caractéristiques présentées précédemment font état de la diversité génétique observée dans la population multiraciale génotypée contenant les animaux de races Alpine et Saanen. Cependant les moyennes des coefficients de consanguinité et de parenté estimés dans la population génotypée de race Saanen sont plus élevées que celles estimées dans la population génotypée de race Alpine (tableau 2). De même l'apparementement entre les populations candidate et de référence estimé à partir des données de génotypages est plus élevé en race Saanen (2,4%) qu'en race Alpine (1,1%). La structure génétique observée dans la population génotypée Saanen est donc plus propice à la mise en place d'évaluations génomiques que celle en race Alpine, car elle permet d'établir plus facilement les relations marqueurs-phénotypes.

La diversité génétique constatée au sein de la population caprine génotypée est donc proche de celle estimée dans la population d'indexation. Les coefficients de consanguinité et de parenté estimés dans la population caprine, bien qu'inférieurs à ceux estimés dans les populations bovines laitières Holstein, Normande et Montbéliarde (Danchin-Burge 2009), sont comparables à ceux estimés en ovins Lacaune (Danchin-Burge 2011). Le niveau de diversité génétique présent dans la population caprine est donc similaire à celui présent dans la population ovine Lacaune pour laquelle la sélection génomique est déjà mise en place.

1.3 / Distance génétique entre les races Alpine et Saanen

Une analyse en composante principale des génotypes des mâles de race Alpine et Saanen a été réalisée. Les deux races se distinguent nettement sur les deux premiers axes (figure 2) qui expliquent plus de 12% de la variabilité observée, ce qui permet de conclure que ce sont deux races bien distinctes d'un point de vue génomique.

La persistance du déséquilibre de liaison entre les races Alpine et Saanen présentée pour des courtes distances entre SNP (figure 3) est représentative de l'histoire ancienne des populations. En effet pour de courtes distances entre SNP, peu de recombinaisons sont possibles, les segments chromosomiques conservés se sont donc transmis de génération en génération. Pour ces courtes distances le niveau de persistance de DL entre les deux races caprines est plus proche de celui observé entre les races ovines Manech tête rousse et tête noire, relativement proches génétiquement, que de celui observé entre la race ovine Lacaune et la race ovine Manech

Tableau 2. Coefficient de consanguinité moyen (en %) estimé à partir du pedigree dans la population de mâles génotypés et phénotypés et coefficient de parenté moyens (en %) entre la population de référence et la population de candidats estimés séparément dans chaque race à partir des génotypes.

	Race Alpine	Race Saanen
Coefficient de consanguinité	1,8	2,3
Coefficient de parenté	1,1	2,4

Figure 2. Analyse en composantes principales (deux premiers axes) réalisée sur les génotypes des 825 mâles comprenant 471 individus de race Alpine (en noir) et 354 individus de race Saanen (en rouge).

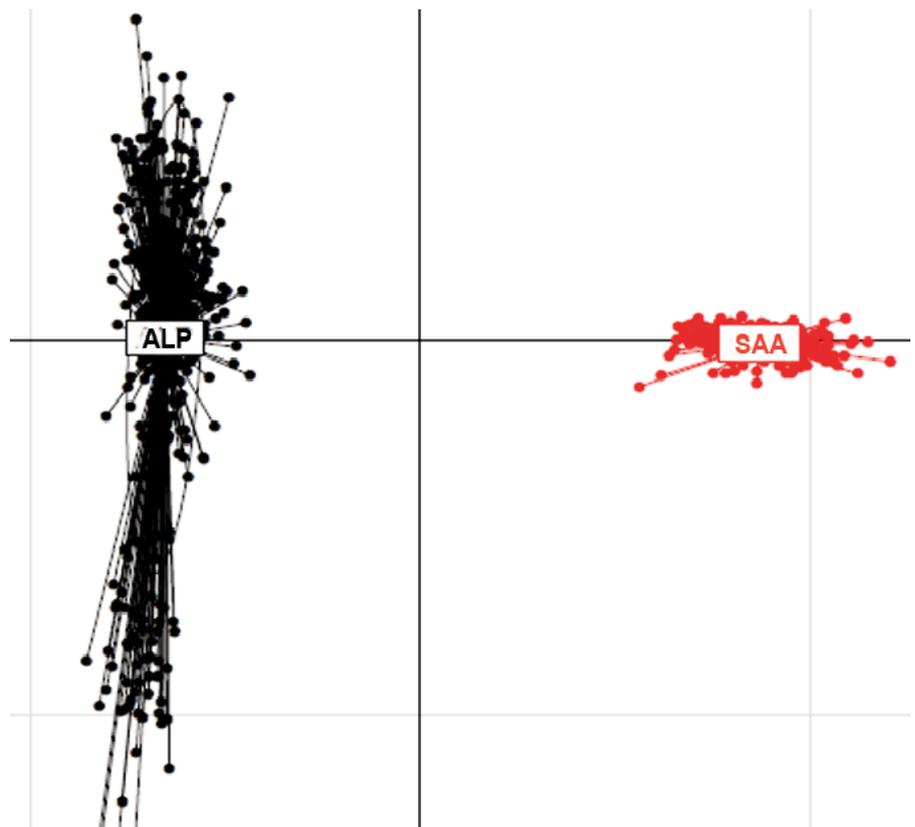
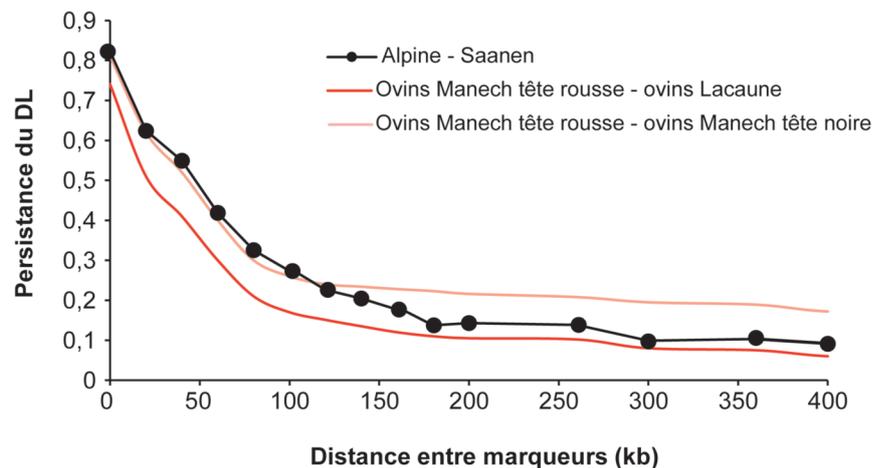


Figure 3. Persistance du déséquilibre de liaison (DL) selon la distance entre SNP (en kb) entre (1) les ovins laitiers de races Manech tête rousse et Manech tête noire, (2) les ovins laitiers de races Manech tête rousse et Lacaune et (3) les caprins de races Alpine et Saanen.



tête rousse, plus éloignées. La persistance de DL pour la distance moyenne entre deux SNP sur la puce (50 kb) est même plus élevée que celle observée entre les deux races Manech. Ces résultats concordent avec une origine commune de l'Alpine et la Saanen à partir d'une même race Suisse dont la version albinos (*i.e.* blanche) a été sélectionnée en race pure en France depuis plus de quarante ans pour former la race Saanen (Babo 2000). Cependant, la persistance du déséquilibre de liaison entre les races Alpine et Saanen est plus faible que celle constatée entre les deux rameaux de la race Lacaune laitière (Guillaume Baloche, INRA GenPhySE, communication personnelle) bien que le temps de divergence entre les races ou rameaux considérés soit similaire.

La structure génétique de la population caprine génotypée est représentative de la population utilisée en indexation. La taille de cette population étant petite et la diversité génétique observée élevée, la structure de la population caprine n'est pas *a priori* optimale pour une mise en place d'évaluations génomiques. Cependant les caractéristiques de la population caprine (DL, consanguinité et parenté) sont similaires à celles observées en ovins laitiers Lacaune, race qui s'est lancée dans la sélection génomique récemment. Enfin, bien qu'ayant une origine commune, les races Alpine et Saanen sont aujourd'hui sélectionnées en race pure et se distinguent clairement l'une de l'autre d'un point de vue génomique.

2 / Les premiers résultats d'évaluations génomiques caprines

La précision des évaluations génomiques a été estimée par validation croisée. Celle-ci consiste à diviser la population de référence qui contient l'ensemble des mâles phénotypés et génotypés (677 mâles) en deux sous-populations : la population d'apprentissage comprenant les 425 mâles nés entre 1993 et 2005, et la population de validation contenant les 252 mâles plus récents nés entre 2006 et 2009. La population d'apprentissage comprenant les phénotypes et les génotypes des animaux, permet d'établir les équations de prédiction faisant le lien entre les allèles aux marqueurs SNP et les phénotypes. Ces équations sont ensuite appliquées à la population de validation qui ne contient que les génotypes des plus jeunes individus. La précision de validation (figure 4) estimée correspond à la corrélation entre les valeurs génomiques prédites (ou GEBV pour « *Genomic Estimated Breeding Values* ») pour les individus de la population de validation et une approxima-

tion de leurs valeurs génétiques vraies par les « *Daughter Yield Deviation* » (DYD). Les DYD sont issues de l'évaluation génétique BLUP (« *Best Linear Unbiased Predictor* ») classique, et correspondent à la performance moyenne des filles de chaque mâle, chacune corrigée des effets du milieu et de la moitié de la valeur génétique de la mère.

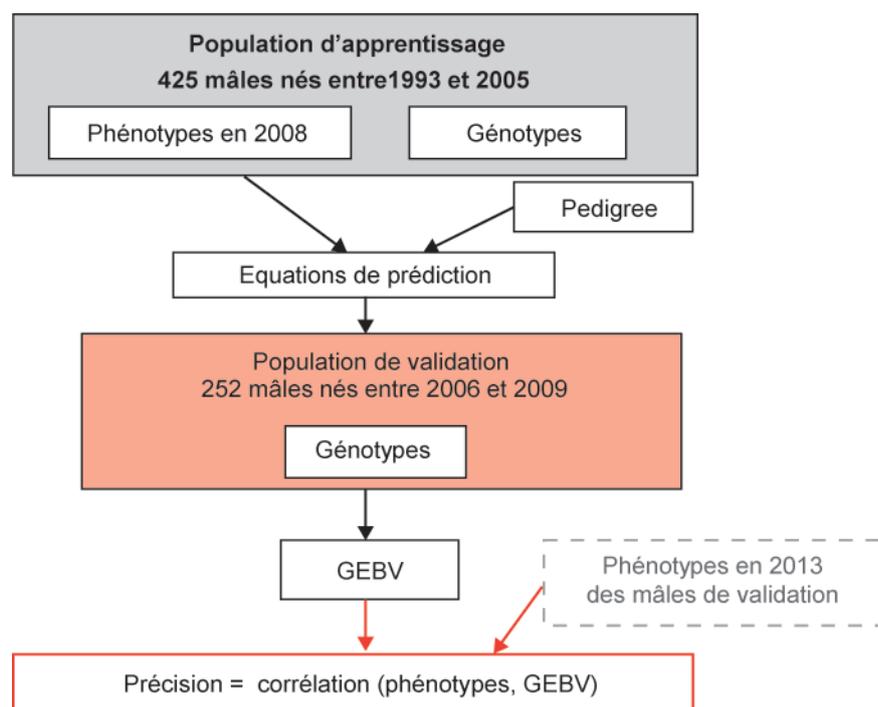
2.1 / Évaluations génomiques basées uniquement sur les mâles génotypés

Les premières évaluations génomiques testées en caprins sont des évaluations basées uniquement sur l'information phénotypique (c'est-à-dire les DYD) et génotypique des seuls mâles génotypés. Elles correspondent aux évaluations mises en place en premier lieu en bovins laitiers. Dans un premier temps, les évaluations génomiques caprines ont été réalisées pour la population multiraciale, considérant la race Alpine et la race Saanen comme une seule et même race. Des évaluations réalisées selon la méthode d'estimation GBLUP (« *Genomic BLUP* ») ont permis d'obtenir les précisions de validation pour les 11 caractères sélectionnés en caprins laitiers. Les 5 caractères de production : les quantités de lait, de matière grasse et de matière protéique et les taux butyreux et protéique, ainsi que le comptage de cellules somatiques sont issus des données collectées lors du contrôle laitier et standardisés pour une durée de 250 jours (Clément *et al* 2014). Les 5 caractères de morphologie mammaire (le profil de

la mamelle, la distance plancher-jarret, la qualité de l'attache arrière, la forme de l'avant pis et l'orientation des trayons) sont notés une fois dans la carrière des chèvres le plus souvent en première lactation lors du pointage. Les précisions de validation obtenues se situent entre 0,32 pour le score de cellules somatiques qui est le caractère le moins héritable ($h^2 = 0,2$) et 0,53 pour le taux butyreux ($h^2 = 0,5$) (Carillier *et al* 2013, Carillier 2015). Les gains de précision constatés avec l'apport de l'information génomique, entre 3,4 et 21,3%, sont inférieurs à ceux estimés en ovins laitiers Lacaune (Astruc *et al* 2012) et en bovins laitiers de race Holstein, Montbéliarde ou Normande (Fritz *et al* 2010). Ces faibles valeurs de précision de validation ainsi que ces faibles gains de précision peuvent être expliqués par la petite taille de la population de référence caprine qui comprend seulement 677 mâles contre plus de 2 000 béliers pour les évaluations génomiques ovines. Les gains observés en caprins sont proches de ceux constatés en race bovine Simmental dont la population de référence est constituée de 229 taureaux (Hozé *et al* 2014a).

Les évaluations génomiques bovines des principales races laitières françaises étant basées sur l'utilisation de méthodes bayésiennes (Boichard *et al* 2012), des méthodes de type Lasso Bayésien ainsi que de type Bayes $C\pi$ ont été testées pour les évaluations génomiques en caprins laitiers (Carillier 2015). Contrairement aux méthodes de type GBLUP, les méthodes bayésiennes permettent d'al-

Figure 4. Schéma de la validation croisée utilisée pour estimer les précisions génomiques dans le cas de la population caprine.



louer une part de variance différente à certains SNP. Ces parts de variance peuvent différer selon le caractère considéré. Les précisions obtenues pour la population de validation des 252 mâles caprins avec ces méthodes sont plus faibles (entre 0,21 pour l'orientation des trayons avec le Lasso Bayésien et 0,43 pour le profil de la mamelle avec le Bayes $C\pi$) que celles obtenues avec le GBLUP. Ces résultats ne concordent pas avec ceux trouvés dans les autres espèces qui sont similaires voire meilleurs avec les méthodes Bayésiennes qu'avec le GBLUP (Duchemin *et al* 2012, Colombani *et al* 2013, Jiménez-Montero *et al* 2013, Hozé *et al* 2014b). Les plus faibles précisions constatées peuvent être expliquées par l'utilisation d'une population multiraciale dans le cas des évaluations caprines. En effet, les méthodes Bayésiennes permettent de sélectionner les SNP ayant un effet fort sur les caractères. Dans le cas multiracial, ces méthodes sélectionnent donc les SNP ayant un effet fort dans les deux races, or il existe des QTL qui ne sont pas communs aux deux races caprines Alpine et Saanen (Palhière *et al* 2012, Maroteau *et al* 2013).

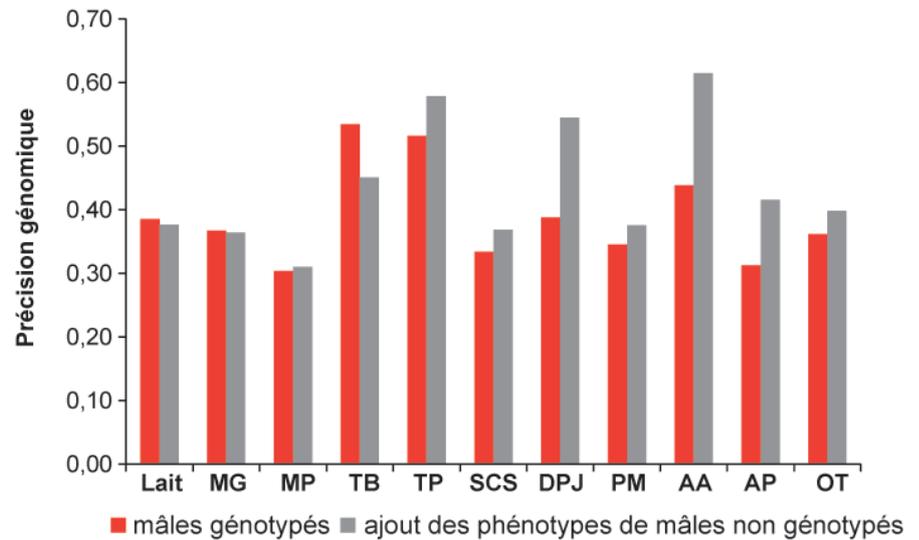
De plus, les évaluations présentées ici sont basées uniquement sur les performances des mâles génotypés. Dans le cas des bovins laitiers (races Holstein, Normande et Montbéliarde), ces évaluations permettent de prendre en compte la quasi-totalité de l'information des performances brutes des femelles puisque la totalité des mâles d'insémination artificielle (IA) sont génotypés et que le taux d'insémination artificielle est de plus de 85%. En caprins laitiers français, le taux d'insémination dans le noyau de sélection est beaucoup moins élevé qu'en bovins laitiers (40%, Larroque *et al* 2014). L'information des performances des femelles issues de boucs de monte naturelle, qui sont en général des fils de boucs d'IA, n'est donc pas prise en compte. L'ajout des phénotypes (DYD) de ces mâles non génotypés dans les évaluations génomiques a donc été testé afin d'améliorer les précisions génomiques.

2.2 / Ajout de phénotypes de mâles non génotypés dans l'évaluation génomique

Les DYD de l'ensemble des mâles ayant plus de 10 filles soit 54 412 mâles (31 024 de race Alpine, 20 522 de race Saanen et 2 866 mâles croisés) nés entre 1980 et 2005 ont été ajoutés aux performances des mâles génotypés (Carillier 2015) dans le cadre d'une évaluation génomique GBLUP. Les précisions génomiques obtenues pour les 252 mâles de validation, entre 0,31 et 0,58 (figure 5), sont en général améliorées : de 3% pour

Figure 5. Précisions génomiques (corrélations entre les « Daughter Yield Deviation » (DYD) et les « Genomic Estimated Breeding Values » (GEBV)) des 252 mâles de validation obtenues à l'aide d'évaluations génomiques de type « Genomic Best Linear Unbiased Predictor » (GBLUP) basées sur les phénotypes des mâles (DYD) dans le cas d'une population d'apprentissage constituée uniquement de mâles génotypés (en rouge) et avec l'ajout de phénotypes de mâles non génotypés (en gris).

MG = Matière Grasse ; MP = Matière Protéique ; TB = Taux Butyreux ; TP = Taux Protéique ; SCS = Score de Cellules Somatiques ; DPJ = Distance Placher-Jarret ; PM = Profil de la Mamelle ; AA = Attache Arrière ; AP = Avant Pis ; OT = Orientation des Trayons.



la quantité de matière protéique à 47% pour la qualité de l'attache arrière par rapport aux précisions génomiques uniquement basées sur les mâles génotypés, exceptées pour les quantités de lait et de matière protéique et le taux butyreux. Cette augmentation des précisions génomiques avec l'ajout de phénotypes des mâles de monte naturelle est plus conséquente pour les caractères de morphologie mammaire. Pour les caractères de production laitière, les performances des femelles sont enregistrées depuis plus longtemps avec une information répétée à chaque lactation. L'information contenue dans les DYD des mâles génotypés est donc plus précise pour les caractères de production que celle contenue dans les DYD des autres caractères. Cela peut ainsi expliquer le fait que l'ajout des DYD des mâles de monte naturelle ait moins d'effet sur les caractères de production que sur les caractères de morphologie ou du score de cellules somatiques.

En revanche dans cette approche, les précisions non génomiques, c'est-à-dire uniquement basées sur l'information du pedigree et des phénotypes, sont plus élevées qu'avec une approche basée uniquement sur les animaux génotypés. Le gain relatif de précision avec l'ajout des génotypes est par conséquent moins important : entre 2 et 6% pour les caractères de morphologie. Ce gain de précision est négatif pour tous les autres

caractères car les précisions génomiques obtenues sont inférieures aux précisions non génomiques de 10% à 3%.

De nombreux résultats dans la littérature montrent que les précisions obtenues à l'aide d'évaluations génomiques basées directement sur les performances des femelles surpassent nettement les précisions obtenues à l'aide d'évaluations basées sur les phénotypes des mâles (Aguilar *et al* 2010, Vitezica *et al* 2010, Legarra *et al* 2014b, Li *et al* 2014). Ce type d'évaluations génomiques n'a pas pu être testé dès le début de notre étude. En effet, les logiciels permettant de faire face à la complexité des calculs dans cette approche n'étaient alors pas encore disponibles.

2.3 / Évaluations génomiques « single step »

a) Intérêt des évaluations multiraciales

Les précisions génomiques obtenues avec des évaluations basées sur les performances brutes des femelles, aussi appelées évaluations « single step », ont donc été examinées dans le cas de la population caprine. Afin d'évaluer l'intérêt d'évaluations génomiques multiraciales ou « uniraciales », les précisions génomiques obtenues avec trois modèles différents ont été comparées (Carillier *et al* 2014). Le premier modèle testé correspond à une évaluation multiraciale

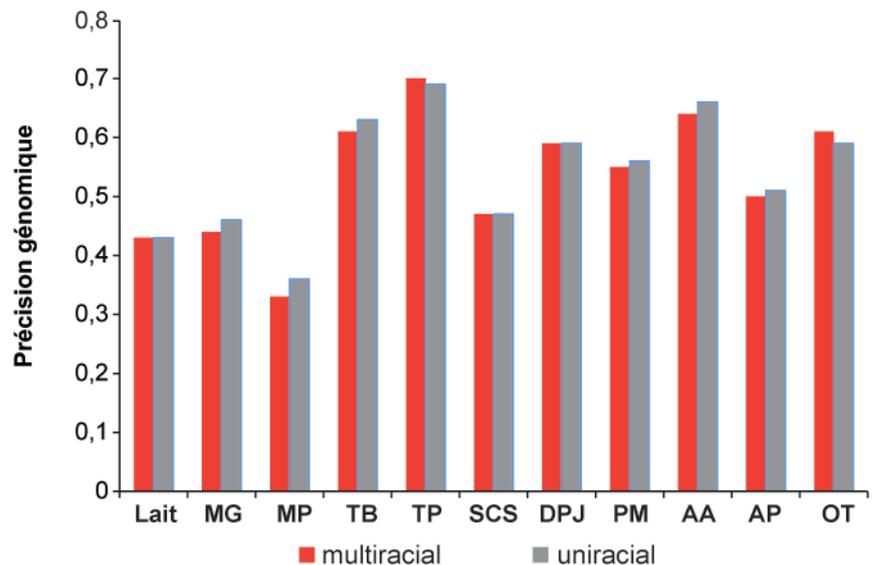
comme décrite dans le paragraphe précédent, les races Alpine et Saanen étant considérées comme une seule et même race ayant les mêmes paramètres génétiques. Le deuxième modèle correspond à une évaluation bi-caractère pour laquelle le premier caractère considéré correspond à un caractère donné (par exemple le taux butyreux) en race Alpine et le deuxième caractère, corrélé génétiquement au premier, correspond au même caractère (le taux butyreux dans cet exemple) en race Saanen. La corrélation génétique estimée à l'aide de ce modèle entre deux caractères analogues varie de 0,45 pour la quantité de lait à 0,76 pour la distance plancher-jarret. Le dernier modèle utilisé correspond à deux évaluations « uniraiales » : une évaluation basée uniquement sur les informations des animaux de race Alpine (pedigree, phénotypes et génotypes) et une autre uniquement avec les animaux de race Saanen.

Les précisions de validation obtenues avec le modèle multiracial ont été estimées entre 0,47 pour le score de cellules somatiques et 0,70 pour le taux protéique (Carillier *et al* 2014, Carillier 2015). Ces précisions sont nettement supérieures à celles estimées à partir des évaluations basées sur les phénotypes des mâles. En revanche, le gain de précision apporté par l'information génomique, entre + 1,7% et + 25,3% est similaire à celui observé avec les évaluations basées sur les phénotypes des mâles génotypés. Les précisions génomiques sont cependant légèrement inférieures aux précisions non génomiques pour les caractères taux butyreux et taux protéique d'environ 3%. Cette diminution de précision, bien que non significative, pourrait être expliquée par la petite taille de la population de validation disponible en validation croisée (425 mâles) qui dans le cas de ces caractères ne serait pas assez importante pour prédire de façon suffisamment précise la population candidate.

Le niveau de précision génomique obtenu avec le modèle multiracial est similaire à ceux obtenus avec les modèles bi-caractère et « uniraial ». Très peu de différences de précisions sont observées quels que soient les animaux à partir desquels sont estimés ces précisions (race Alpine, race Saanen, races Alpine et Saanen). La figure 6 présente les précisions génomiques des mâles de validation en modèles multiracial ou « uniraial ». Pour l'ensemble des caractères, les précisions sont globalement très proches. Bien que la taille de la population soit augmentée en considérant un modèle multiracial, cette étude ne montre pas d'intérêt des évaluations multiraciales. Ces résultats concordent avec ceux de la littérature qui ne montrent pas de différence entre les précisions des évaluations

Figure 6. Précisions génomiques des 252 mâles de validation obtenues à l'aide d'évaluations génomiques de type « Genomic Best Linear Unbiased Predictor » (GBLUP) basées sur les performances des femelles dans le cas d'évaluations multiraciales (en rouge) et « uniraiales » (en gris).

MG = Matière Grasse ; MP = Matière Protéique ; TB = Taux Butyreux ; TP = Taux Protéique ; SCS = Score de Cellules Somatiques ; DPJ = Distance Placher-Jarret ; PM = Profil de la Mamelle ; AA = Attache Arrière ; AP = Avant Pis ; OT = Orientation des Trayons.



multiraciales et celles des évaluations « uniraiales » lorsque les races utilisées sont peu liées d'un point de vue génétique (Hayes *et al* 2009; Karoui *et al* 2012). Il ne semble donc pas y avoir d'intérêt à utiliser un modèle multiracial pour augmenter les précisions génomiques des caprins. Cependant, le modèle multiracial qui ne consiste qu'en une seule évaluation, peut être intéressant pour son temps de calcul réduit.

b) Apport de l'information génomique des femelles

Le génotypage de 1 985 femelles, d'un dispositif de détection de QTL, nous a permis d'étudier l'intérêt de l'ajout des génotypages de femelles pour l'augmentation des précisions génomiques. Ces femelles étant nées entre 2008 et 2009, le dispositif de validation croisée présenté précédemment n'est pas adapté. En effet, dans un contexte d'application de la sélection génomique en temps réel, les évaluations génomiques utilisent l'information des animaux les plus jeunes. Or les mâles de validation utilisés dans les études précédentes sont nés entre 2006 et 2009, ils sont donc en partie plus vieux que les femelles génotypées. Le design de validation croisée revu est maintenant composé d'une population d'apprentissage de 677 mâles phénotypés et génotypés nés entre 1993 et 2009 comprenant ou non les 1985 femelles génotypées, et d'une population de

validation de 148 jeunes mâles génotypés nés entre 2010 et 2011 dont les filles ont des performances depuis janvier 2015.

L'inclusion des femelles dans la population d'apprentissage permet d'augmenter les précisions génomiques de 2% à 14% pour la grande majorité des caractères étudiés, exceptés la quantité de lait, la matière et le taux protéiques (Carillier 2015). De telles augmentations de précisions, non négligeables, avec l'ajout de génotypes de femelles n'ont pas été observées en bovins laitiers de race Holstein (Wiggans *et al* 2011, Dasseville *et al* 2012, Pryce *et al* 2012), mais dans des populations avicoles (Lourenco *et al* 2015). Ceci peut s'expliquer d'une part par la taille de la population d'apprentissage caprine, nettement plus petite qu'en bovins, et d'autre part, par l'existence de traitement préférentiel de certaines femelles bovines conduisant à un phénotypage biaisé de ces femelles, ce qui n'est pas observé dans le cas des troupeaux caprins (Dasseville *et al* 2012).

Bien que plus importantes qu'en bovins laitiers, les augmentations des précisions avec l'ajout des génotypes femelles restent faibles. L'utilisation de femelles issues d'un dispositif de détection de QTL (*i.e.* filles de 20 pères) n'est probablement pas optimale, les mères des jeunes boucs pourraient être de bien meilleures candidates au génotypage. Cependant, ces femelles étant filles de

boucs d'insémination artificielle eux-mêmes génotypés, leur génotypage propre n'apporterait que peu d'information en termes de diversité génétique, nécessaire pour prédire les jeunes animaux. Une étude en bovins laitiers Holstein, a d'ailleurs montré qu'utiliser les génotypes de femelles choisies au hasard permet d'obtenir les mêmes précisions génomiques que d'utiliser les meilleures femelles (Jiménez-Montero *et al* 2010). Chez les caprins, de nombreuses femelles (44% du pedigree) sont issues de pères inconnus, du fait du peu d'inséminations artificielles réalisées en pratique et d'une conduite de monte en lots. L'apparement sur pedigree de ces femelles avec les jeunes boucs candidats à prédire est donc nul. Il pourrait donc être intéressant de génotyper certaines de ces femelles afin de révéler leur parenté avec les candidats, d'apporter une plus grande diversité génétique et ainsi améliorer les précisions génomiques des candidats.

Les évaluations génomiques de type « *single step* » (basées sur les performances brutes des animaux) permettent d'obtenir de meilleures précisions de validation que celles obtenues à l'aide d'évaluations génomiques basées sur les performances des mâles. L'ajout des génotypes des femelles du dispositif QTL permet d'améliorer modestement les précisions des évaluations génomiques. Dans l'objectif d'augmenter davantage ces précisions, le génotypage d'autres femelles pourrait être envisagé. En revanche, il n'est pas possible de conclure à l'avantage des évaluations multiraciales comparées aux évaluations par race.

Au vu de ces premiers résultats encourageants, la mise en place d'un modèle d'évaluation génomique « *single step* » multiraciale pour les caprins laitiers français est envisageable. Il a l'avantage d'être simple à implémenter et rapide en temps de calcul.

3 / Perspectives pour une application de la sélection génomique

3.1 / Augmenter la taille de la population génotypée

La taille de la population de référence, c'est-à-dire le nombre d'animaux génotypés, est présentée dans de nombreuses études comme un élément clé de la précision des évaluations génomiques. Or nous avons vu que la taille de la population caprine est à l'heure actuelle limitée (825 mâles génotypés), bien que la stratégie de génotypage adoptée ait consisté à génotyper l'ensemble des mâles des séries de testage depuis les années 2000

ainsi que l'ensemble des mâles d'insémination artificielle nés entre 1993 et 2000. Le nombre de mâles génotypés ne pourra donc être augmenté que par le génotypage des nouveaux mâles candidats à la sélection (soit actuellement environ 70 par an). Une première stratégie d'augmentation de la taille de la population de référence caprine a consisté à mutualiser les génotypes des 2 principales races pour constituer une population de référence multiraciale, cela n'a pas permis d'augmenter les précisions génomiques.

La deuxième stratégie envisagée a été l'inclusion de génotypes de femelles, avec la prise en compte des génotypes des femelles d'un dispositif de détection de QTL. Les premiers résultats sont encourageants. Il pourrait donc être envisagé de génotyper d'autres femelles afin d'améliorer de façon plus conséquente les précisions des évaluations génomiques. La structure génétique des populations de référence et candidate ayant un impact fort sur les précisions génomiques, le choix des femelles à génotyper est un élément sensible. Afin de maximiser l'apparement entre la population candidate et la population de référence, il pourrait être envisagé de génotyper l'ensemble des mères à boucs jeunes candidats. Cependant, génotyper l'ensemble des mères à boucs aurait un coût financier important et n'apporterait que peu de diversité génétique au sein de la population de référence (car elles sont pour la plupart issues de mâles génotypés). Une solution pourrait être de génotyper les mères à boucs qui sont le moins apparentées aux mâles déjà génotypés. Par ailleurs, la proportion de femelles issues de pères inconnus étant élevée, il pourrait être intéressant de génotyper certaines de ces femelles apparentées aux candidats comme des demi-sœurs ou des cousines du côté maternel. Enfin, un outil d'assignation de parenté (Barbotte *et al* 2012) est actuellement en cours de développement (OPA-Action Innovante, Isabelle Palhière, INRA). Son utilisation pour les femelles issues de pères inconnus qui ne seraient pas génotypées à l'aide de la puce de 50K pourrait permettre de révéler la parenté de ces animaux et donc d'améliorer la précision de leurs valeurs génétiques.

L'augmentation de la taille de population de référence n'est pas uniquement à réfléchir au sein de la seule population caprine française. En effet, la constitution d'une population de référence internationale en race Holstein bovine, multipliant la taille de la population par 4 a permis d'augmenter les précisions génomiques de 8 à 12% (Lund *et al* 2011). Les races Saanen et Alpine caprine bénéficient de programmes de sélection dans de nombreux pays en Europe et en Amérique du Nord. Des collaborations

pourraient ainsi être envisagées avec le Canada, dont 403 femelles de race Alpine et 318 de race Saanen sont actuellement génotypées avec la puce 50k d'Illumina (Brito *et al* 2015), les États-Unis (Kennedy *et al* 1982, Luo *et al* 1997), l'Italie (Frattini *et al* 2014), le Brésil dont la majeure partie des animaux sont à l'origine importés de France (Araujo *et al* 2006). Cependant, le bénéfice de la constitution d'une population de référence internationale dépend de la proximité génétique entre les populations des différents pays, qu'il faudra étudier. Un tel projet, bien qu'assurément bénéfique pour les précisions génomiques des animaux français, nécessite de mettre en place un organisme comme Interbull en bovins laitiers afin de mettre en commun et standardiser non seulement les génotypes mais aussi les données phénotypiques des animaux dont les mesures peuvent être différentes entre pays.

3.2 / Utilisation de différents types de génotypes

L'évolution rapide des technologies ainsi que des coûts de génotypage pousse à s'interroger sur la technologie de génotypage à utiliser pour les futurs candidats. En effet, étant donné le prix élevé (environ 150€ par animal, tout compris hors prélèvement) de la puce SNP50k caprine jusqu'en 2014, il était envisagé de développer une puce basse densité, moins chère, qui aurait pu servir au génotypage des femelles. La baisse récente du prix du génotypage 50k (environ 70€ par animal, tout compris) permet d'envisager d'autres stratégies notamment le développement de puce SNP haute densité ou l'exploitation des données de séquences. Ce type d'outils, en utilisant un nombre élevé de marqueurs répartis sur tout le génome, permettrait de détecter des SNP fortement associés à un caractère d'intérêt, en augmentant le déséquilibre de liaison entre les SNP et les QTL. Cependant, le gain de précision génomique obtenu avec l'augmentation du nombre de marqueurs n'est pas observé dans toutes les études (Hozé *et al* 2014b).

Les données de séquences complètes des 20 boucs du dispositif de détection de QTL pourraient être utilisées afin d'identifier les gènes ou mutations causales impliqués dans la régulation des caractères d'intérêt. Ces informations pourraient être utilisées pour prédire plus précisément les valeurs génomiques. Les récentes études ne permettent cependant pas de conclure unanimement à l'intérêt de l'utilisation des données de séquences dans les évaluations génomiques (Hayes *et al* 2014; Pérez-Enciso *et al* 2015). En effet, il n'apparaît pas toujours intéressant d'utiliser l'ensemble de ces données de séquence pour construire la matrice de parenté génomique

comme c'est le cas pour les évaluations génomiques classiques. En revanche, l'utilisation des régions d'intérêt détectées plus précisément grâce aux données de séquence pourraient être intégrées dans un modèle d'évaluation génomique à déterminer afin d'améliorer les précisions génomiques.

L'utilisation simultanée de l'ensemble de ces technologies, actuellement à l'étude dans le cadre du projet INCoMINGS (méta-programme SELGEN), nécessite de s'intéresser aux méthodes d'imputation (*i.e.* estimer les génotypes manquants) des données les moins denses ainsi qu'au choix des animaux à génotyper pour chacune des technologies envisagées. Comme en bovins laitiers, il pourrait être proposé de génotyper les candidats à la sélection ainsi que certaines femelles avec une puce 50k puis une fois leur descendance connue, de génotyper à nouveau les meilleurs mâles (futurs pères de candidats) sur une puce haute densité. Malgré une évolution du coût du séquençage qui a très faible densité est comparable à celui d'une puce haute densité en bovins, il reste aujourd'hui encore relativement élevé en caprin. Le choix des quelques animaux à séquencer dans cette espèce est donc à raisonner.

3.3 / Prise en compte des principaux gènes majeurs

L'information des puces à SNP ne permet pas toujours de capter l'information parfois complexe des gènes majeurs (lorsqu'ils sont multi-alléliques par exemple). C'est le cas du gène majeur de la caséine α_{s1} en caprins laitiers, pour lequel les variants sont pour la plupart des additions ou délétions de bases. Cependant, pour ce gène une autre information moléculaire est disponible, puisque son génotypage est réalisé en routine depuis 1998 (Barbieri *et al* 1995) et utilisé pour la présélection des animaux candidats au testage. Il serait intéressant de pouvoir intégrer, en plus de l'information génomique par les SNP, l'effet de ces gènes majeurs dans les modèles d'évaluations génomiques afin d'améliorer la prédiction des valeurs génomiques des animaux.

Le gène de la caséine α_{s1} est connu pour son effet majeur sur le taux protéique dont il explique à lui seul 40% de la variabilité génétique (Barbieri *et al* 1995 ; Carillier *et al* 2016). Le génotype complet du gène de la caséine α_{s1} obtenu à l'aide des technologies PCR (réaction en chaîne de la polymérase) et RFLP (polymorphisme des fragments de restriction) est actuellement disponible pour un peu plus de 6 000 animaux (2 946 femelles et 3 327 mâles). Nous l'avons utilisé afin d'étudier l'intérêt de la prise en compte des gènes majeurs dans les évaluations génomiques. Du fait du multi-

allélisme de ce gène ainsi que du caractère monomorphe de certains allèles dans notre population, les seuls variants pouvant être captés par des SNP ont été éliminés lors du contrôle qualité réalisé sur les marqueurs de la puce 50K. La disponibilité du génotype au locus de la caséine α_{s1} pour une faible proportion (moins de 1%) des femelles phénotypées nous a poussés à étudier dans un premier temps la prise en compte de ce génotype dans les évaluations génomiques basées uniquement sur les phénotypes des mâles génotypés (DYD). Nous avons pu montrer que dans ce cas, l'inclusion du génotype au locus de la caséine α_{s1} permettait d'augmenter les précisions génomiques de 19% pour le taux protéique. Cependant, les précisions des évaluations génomiques basées sur les performances brutes des femelles (sans prise en compte d'un effet du génotype au locus de la caséine α_{s1}) sont meilleures que celles obtenues avec les évaluations basées sur les phénotypes des mâles. Il était nécessaire de trouver une méthode permettant de prendre en compte l'information manquante au locus de la caséine α_{s1} pour la majorité des femelles. La première méthode envisagée a consisté à prédire le génotype pour toutes les femelles non génotypées à l'aide d'algorithmes itératifs (Fernando *et al* 1993, Vitezica *et al* 2005). La deuxième solution envisagée était basée sur la méthode du « *gene content* » (Gengler *et al* 2007, Legarra et Vitezica, 2015) permettant d'évaluer simultanément, par modèle multicaractère, la valeur génétique pour le taux protéique et pour le « *gene content* », c'est-à-dire le nombre de chaque allèle dans le génotype des individus. Il a été montré que seule cette deuxième solution permettait des gains de précisions conséquents entre 4 et 14% pour le taux protéique dépendant de la population considérée (Alpine ou Saanen (Carillier *et al* 2016).

Outre le gène majeur de la caséine α_{s1} , le gène DGAT1 ayant un effet très fort sur le taux butyreux pourrait également être étudié. En effet, deux mutations récemment identifiées pour ce gène en caprin (mutations R396W et R251L (Maroteau 2014)) ont permis de développer un outil de génotypage. Cependant, l'intégration de l'effet du gène DGAT1 dans les évaluations génomiques semble plus complexe dans la mesure où certains auteurs montrent qu'il existe un effet de dominance d'un des allèles du gène pour certains caractères (Kuehn *et al* 2007, Koopaei *et al* 2012, Molee *et al* 2012) chez les bovins laitiers Holstein. Des effets de dominance semblent également avoir été trouvés pour le gène de la caséine α_{s1} dans des populations caprines étrangères (Dagnachew *et al* 2011, Vázquez-Flores *et al* 2012). Il existe des modèles d'évaluations génomiques, développés récemment, permettant de pren-

dre en compte des effets de dominance (Vitezica *et al* 2013). Cependant, ces mêmes auteurs rapportent que cette prise en compte n'augmente que faiblement les précisions génomiques.

En dehors de ces informations de gènes majeurs, il pourrait être envisagé, à l'instar de la méthode de sélection par marqueurs génomiques mise en place en bovins laitiers (Guillaume *et al* 2011), d'utiliser l'ensemble des QTL détectés en caprins (Maroteau 2014). Cependant, de même que pour le génotype au locus de la caséine α_{s1} , il sera nécessaire de prédire les allèles portés au QTL pour la totalité des femelles non génotypées. L'approche de type « *gene content* » permet de prédire (bien que de façon imprécise) l'effet d'un QTL ou d'un gène majeur pour l'ensemble des femelles. Cette méthode semble cependant difficilement applicable pour des centaines de QTL car la prédiction n'est possible que QTL par QTL. Des études et modèles complémentaires sont donc nécessaires et en cours d'investigation.

3.4 / Évolution des schémas de sélection caprins avec la sélection génomique

L'augmentation des précisions n'est pas le seul levier existant pour augmenter le gain génétique annuel. La sélection génomique pourrait permettre par exemple de réduire les intervalles de génération de la voie père, plutôt longs à l'heure actuelle (environ 4 ans pour l'intervalle père-fille et 5,5 pour l'intervalle père-fils), en utilisant les mâles en IA dès leur maturité sexuelle. La période d'évaluation de leur capacité sexuelle et de la qualité de leur semence étant relativement longue et incompressible (entre 7 et 8 mois), les intervalles de génération père-fille et père-fils pourraient être réduits à 1 an et 7 mois. Cependant, le choix politique de l'organisme de sélection caprin (Capgènes) pourrait s'orienter vers la conservation d'une précision élevée pour les mâles utilisés comme pères à boucs ce qui ne permettrait pas une diminution de l'intervalle de génération père-fils.

De plus, d'un point de vue économique, la sélection génomique ne permettra *a priori* pas d'économiser les coûts du testage comme en bovins. En effet, contrairement au cas des bovins laitiers, la gestion actuelle des boucs d'IA les élimine une fois le nombre de doses d'IA suffisant obtenu. La mise en place de la sélection génomique pourrait cependant modifier cette gestion. Il semblerait donc qu'hormis un gain de précision et une diminution modeste des intervalles de génération, l'accélération du progrès génétique en caprin passera essentiellement par l'intensité de sélection.

tion, c'est-à-dire le nombre de mâles candidats à génotyper. Le choix de l'intensité de sélection devra tenir compte du coût de génotypage relativement élevé, selon les différentes stratégies de génotypage citées plus haut, ainsi que du besoin des éleveurs en doses d'IA. En effet, aucune économie sur le testage ne pouvant être réalisée, l'augmentation de l'intensité de sélection nécessitera donc une stratégie économique de la part de l'entreprise de sélection qui doit étudier différents scénarios. Capgènes étudie actuellement différents scénarios pour son futur schéma génomique.

Par ailleurs, la mise en place de la sélection génomique pour les caractères déjà en sélection pourrait également permettre de sélectionner de nouveaux caractères tels que la robustesse ou des caractères d'adaptation. Ces nouveaux caractères étant corrélés faiblement ou négativement avec les caractères de production (Isabelle Palhière et Rachel Rupp, INRA GenPhySE, communication personnelle), leur sélection engendrerait à l'heure actuelle une diminution du progrès génétique pour les caractères principaux. Cependant, cette diminution du progrès serait contrebalancée par l'augmentation du progrès génétique engendré

par la sélection génomique. Les nouveaux caractères qui intéressent la filière caprine sont entre autres les caractères de fonction sexuelle et de morphologie des mâles (projet CASDAR Maxi'mâle) ainsi que les caractères de robustesse c'est-à-dire la longévité fonctionnelle et la persistance laitière (projet CASDAR RUSTIC).

Conclusion

Dans un contexte très marqué par la mise en place de la sélection génomique chez les bovins laitiers, l'arrivée très récente de la puce 50k caprine (2011) a poussé la filière à s'interroger sur l'intérêt de la sélection génomique dans cette espèce. Cette étude présente une synthèse des premiers résultats d'évaluations génomiques, pour cette espèce, dans le monde. Elle a permis de caractériser la structure génétique de la population de référence et d'identifier ses atouts et faiblesses afin d'aider au choix des futures stratégies de génotypage. Les résultats obtenus montrent que la sélection génomique peut être envisagée en caprins. En effet, les précisions des évaluations génomiques basées sur les performances des femelles sont supérieures aux précisions

obtenues à l'aide d'évaluations génétiques classiques. Il est toutefois possibles d'aller plus loin et d'améliorer ces précisions en incluant l'effet de gènes majeurs ou en génotypant des femelles.

Remerciements

Nous remercions les financeurs des projets Genovicap et Phenoflail : l'ANR, Apis-Gène, CASDAR, FranceAgriMer, France Génétique Élevage et le Ministère de l'Agriculture ainsi que les financeurs du projet européen 3SR qui ont permis d'obtenir l'ensemble des génotypes et l'entreprise de sélection Capgènes pour la mise à disposition des données. Nous remercions également la région Midi-Pyrénées et l'INRA (méta-programme SELGEN) qui ont permis de financer les travaux de thèse menés sur ce sujet. Cette étude n'aurait pas été possible sans le consortium international caprin (« *International Goat Genome Consortium* ») (IGGC) : www.goatgenome.org qui a permis de développer la puce caprine « *goat SNP50 BeadChip* ». Enfin nous remercions Ignacy Misztal (University of Georgia, Athens) de nous avoir permis d'utiliser l'ensemble des logiciels d'évaluations génomiques.

Références

- Aguilar I., Misztal I., Johnson D.L., Legarra A., Tsuruta S., Lawlor T.J., 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.*, 93, 743-752.
- Araujo A.M., Guimaraes S.E.F., Machado T.M.M., Lopes P.S., Pereira C.S., Da Silva C.S., Rodrigues M.T., Columbiano V.S., Da Fonseca C.G., 2006. Genetic diversity between herds of Alpine and Saanen dairy goats and the naturalized Brazilian Moxoto breed. *Genet. Mol. Biol.*, 29, 67-74.
- Astruc J.M., Baloche G., Larroque H., Beltran De Heredia I., Labatut J., Lagriffoul G., Moreno C., Robert-Granié C., Boscher M.Y., Chantry Darmond C., Aguerre X., Boulenc P., Fregeat G., Giral-Viala B., Guibert P., Panis P., Soulas C., Legarra A., Barillet F., 2012. La sélection génomique des ovins laitiers en France : stratégies, premiers résultats des évaluations génomiques et perspectives. *Renc. Rech. Rum.*, 19, 81-84.
- Astruc J.M., Baloche G., Buisson D., Labatut J., Lagriffoul G., Larroque H., Robert-Granié C., Legarra A., Barillet F., 2016. La sélection génomique des ovins laitiers en France. In : *Brebis laitières en France : 50 ans de recherche et de développement*. Barillet F., Hassoun P., Astruc J.M., Lagriffoul G., Morin E. (Eds). Dossier, INRA Prod. Anim., 29, 41-56.
- Babo D., 2000. Races ovines et caprines françaises. France Agricole Editions. 310p.
- Baloche G., Legarra A., Sallé G., Larroque H., Astruc J.M., Robert-Granié C., Barillet F., 2013. Assessment of accuracy of genomic prediction for French Lacaune dairy sheep. *J. Dairy Sci.*, 97, 1107-1116.
- Barbieri M.E., Manfredi E., Elsen J.M., Ricordeau G., Bouillon J., Grosclaude F., Mahé M.F., Bibé B., 1995. Effects of the α_{s1} -casein locus on dairy performances and genetic parameters of Alpine goats. *Genet. Select. Evol.*, 27, 437-450.
- Barbotte L., Genestout L., Fritz S., Chantry Darmond C., Boichard D., 2012. Assignment de parenté par marqueurs SNP chez les bovins français. *Renc. Rech. Rum.*, 19, 92.
- Boichard D., Guillaume F., Baur A., Croiseau P., Rossignol M.N., Boscher M.Y., Druet T., Genestout L., Colleau J.J., Journaux L., Ducrocq V., Fritz S., 2012. Genomic selection in French dairy cattle. *Anim. Prod. Sci.*, 52, 115-120.
- Brito L.F., Jafarikia M., Grossi D.A., Kijas J.W., Porto-Neto L.R., Ventura R.V., Salgorzai M., Schenkel F.S., 2015. Characterization of linkage disequilibrium, consistency of gametic phase and admixture in Australian and Canadian goats. *BMC Genet.*, 16, 67.
- Carillier C., 2015. Etude de la prédiction génomique chez la caprins : Intérêt et limites de la sélection génomique dans le cadre d'une population multiraciale à faible effectif. Thèse de doctorat, Ecole doctorale SEVAB, Toulouse. http://oatao.univ-toulouse.fr/14466/1/carillier_jacquin.pdf
- Carillier C., Larroque H., Robert-Granié C., 2013. A first step toward genomic selection in multi-breed French dairy goat population. *J. Dairy Sci.*, 96, 7294-7305.
- Carillier C., Larroque H., Robert-Granié C., 2014. Comparison of join versus purebred genomic evaluation in the French multi-breed dairy goat population. *Genet. Select. Evol.*, 46, 67.
- Carillier C., Larroque H., Robert-Granié C., 2016. Including α_{s1} casein gene information in genomic evaluations of French dairy goats. *Genet. Select. Evol.*, 48, 54.
- Clément V., Palhière I., Larroque H., 2014. Caractères de production laitière, de comptage de cellules somatiques et de morphologie. Compte rendu 00 14 202 041, Collection Résultats, Institut de l'Élevage.
- Colleau J.J., Moureaux S., 2006. Gestion optimisée de la parenté et de la consanguinité. INRA Prod. Anim., 19, 3-14.
- Colleau J.J., Clément V., Martin P., Palhière I., 2011. Optimized diffusion of buck semen for saving genetic variability in selected dairy goat populations. *BMC Genet.*, 12, 25.
- Colleau J.J., Fritz S., Guillaume F., Baur A., Dupassieux D., Boscher M.Y., Journaux L., Eggen A., Boichard D., 2015. Simulation des potentialités de la sélection génomique chez les bovins laitiers, INRA Prod. Anim., 28, 251-258.
- Colombani C., Legarra A., Fritz S., Guillaume F., Croiseau P., Ducrocq V., Robert-Granié C., 2013. Application of Bayesian least absolute shrinkage and selection operator (LASSO) and BayesC π methods for genomic selection in

- French Holstein and Montbéliarde breeds. *J. Dairy Sci.*, 96, 575-591.
- Dagnachew B.S., Thaller G., Lien S., Ådnøy T., 2011. Casein SNP in Norwegian goats: additive and dominance effects on milk composition and quality. *Genet. Select. Evol.*, 24, 43.
- Danchin-Burge C., 2009. Estimation de la variabilité génétique de 19 races bovines à partir de leur généalogie. *Compte-rendu n°000972125* Institut de l'élevage Collection Résultats.
- Danchin-Burge C., 2011. Bilan de variabilité génétique de 9 races de petits ruminants laitiers et à toison. *Compte-rendu n°001172004* Institut de l'élevage Collection Résultats.
- Dassonneville R., Baur A., Fritz S., Boichard D., Ducrocq V., 2012. Inclusion of cow records in genomic evaluations and impact on bias due to preferential treatment. *Genet. Select. Evol.*, 44, 1-8.
- Duchemin S.I., Colombani C., Legarra A., Baloche G., Larroque H., Astruc J.M., Barillet F., Robert-Granié C., Manfredi E., 2012. Genomic selection in the French Lacaune dairy sheep breed. *J. Dairy Sci.*, 95, 2723-2733.
- Fernando R.L., Stricker C., Elston R.C., 1993. An efficient algorithm to compute the posterior genotypic distribution for every member of a pedigree without loops. *Theor. Appl. Genet.*, 87, 89-93.
- Fratini S., Nicoloso L., Coizet B., Chessa S., Rapetti L., Pagnacco G., Crepaldi P., 2014. The unusual genetic trend of α_{S1} -casein in Alpine and Saanen breeds. *J. Dairy Sci.*, 97, 7975-7979.
- Fritz S., Guillaume F., Croiseau P., Baur A., Hoze C., Dassonneville R., Boscher M.Y., Journeaux L., Boichard D., Ducrocq V., 2010. Mise en place de la sélection génomique dans les trois principales races françaises de bovins laitiers. *Renc. Rech. Rum.*, 17, 455-458.
- Gengler N., Mayeres P., Szydlowski M., 2007. A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose Belgian Blue cattle. *Anim. Int. J. Anim. Biosci.*, 1, 21-28.
- Guillaume F., Boichard D., Ducrocq V., Fritz S., 2011. Utilisation de la sélection génomique chez les bovins laitiers. In : Numéro spécial, Amélioration génétique, Mulsant P., Bodin L., Coudurier B., Deretz S., Le Roy P., Quillet E., Perez J.M. (Eds). *INRA Prod. Anim.*, 24, 363-368.
- Hayes B.J., Bowman P.J., Chamberlain A.C., Verbyla K., Goddard M.E., 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Select. Evol.*, 41, 1-9.
- Hayes B.J., Macleod I., Daetwyler H.D., Bowman P.J., Chamberlain A.J., Vander Jagt C.J., Capitan A., Pausch H., Stothard P., Liao X., Schrooten C., Mullaart E., Fries R., Gulbrandsen B., Lund M.S., Boichard D., Veerkamp R.F., VanTassel C.P., Gredler B., Druet T., Bagnato A., Vilkkijä J., de Koning D.J., Santus E., Goddard M.E., 2014. Genomic prediction from whole genome sequence in livestock: the 1000 bull genomes project. 10th World Congr. Genet. Appl. Livest. Prod., Vancouver, Canada.
- Hozé C., Fouilloux M.N., Venot E., Guillaume F., Dassonneville R., Fritz S., Ducrocq V., Phocas F., Boichard D., Croiseau P., 2013. High-density marker imputation accuracy in sixteen French cattle breeds. *Genet. Select. Evol.*, 45, 33.
- Hozé C., Fritz S., Baur A., Phocas F., Boichard D., Ducrocq V., Croiseau P., 2014a. Mise en place d'évaluations génomiques dans les races bovines laitières à effectifs limités en France. *Renc. Rech. Rum.*, 21, 237-240.
- Hozé C., Fritz S., Phocas F., Boichard D., Ducrocq V., Croiseau P., 2014b. Efficiency of multi-breed genomic selection for dairy cattle breeds with different sizes of reference population. *J. Dairy Sci.*, 97, 3918-3929.
- Jiménez-Montero J.A., Gonzalez-Recio O., Alenda R., 2010. Genotyping strategies for genomic selection in dairy cattle. *Animal*, 6, 1216-1224.
- Jiménez-Montero J.A., González-Recio O., Alenda R., 2013. Comparison of methods for the implementation of genome-assisted evaluation of Spanish dairy cattle. *J. Dairy Sci.*, 96, 625-634.
- Karoui S., Carabano M.J., Diaz C., Legarra A., 2012. Joint evaluation of French dairy cattle breeds using multiple-trait models. *Genet. Select. Evol.*, 44, 1-10.
- Kennedy B.W., Finley C.M., Bradford G.E., 1982. Phenotypic and genetic relationships between reproduction and milk production in dairy goats. *J. Dairy Sci.*, 65, 2373-2383.
- Koopaei H.K., Abadi M.R.M., Mahyari S.A., Koshkoiyeh A.E., Tarang A.R., Potki P., 2012. Effect of DGAT1 variants on milk composition traits in Iranian Holstein cattle population. *Anim. Sci. Pap. Rep.*, 30, 231-239.
- Kuehn C., Edel C., Weikard R., Thaller G., 2007. Dominance and parent-of-origin effects of coding and non-coding alleles at the acylCoA-diacylglycerol-acyltransferase (DGAT1) gene on milk production traits in German Holstein cows. *BMC Genet.*, 8, 62.
- Larroque H., Barillet F., Baloche G., Astruc J.M., Buisson D., Shumbusho F., Clément V., Lagriffoul G., Palhière I., Rupp R., Carillier C., Robert-Granié C., Legarra A., 2014. Toward genomic breeding programs in French dairy sheep and goats, 10th World Congr. Genet. Appl. Livest. Prod., Vancouver, Canada.
- Legarra A., Vitezica Z., 2015. Genetic evaluation with major genes and polygenic inheritance when some animals are not genotyped using gene content multiple-trait BLUP. *Genet. Sel. Evol.*, 47, 89.
- Legarra A., Baloche G., Barillet F., Astruc J.M., Soulas C., Aguerre X., Arrese F., Mintegi L., Lasarte M., Maeztu F., Beltrán de Heredia I., Ugarte E., 2014a. Within- and across-breed genomic predictions and genomic relationships for Western Pyrenees dairy sheep breeds Latxa, Manech, and Basco-Béarnaise. *J. Dairy Sci.*, 97, 3200-3212.
- Legarra A., Christensen O.F., Aguilar I., Misztal I., 2014b. Single Step, a general approach for genomic selection. *Livest. Sci.*, 166, 54-65.
- Le Roy P., Chapuis H., Guémené D., 2014. Sélection génomique : quelles perspectives pour les filières avicoles ? *INRA Prod. Anim.*, 27, 331-336.
- Li X., Wang S., Huang J., Li L., Zhang Q., Ding X., 2014. Improving the accuracy of genomic prediction in Chinese Holstein cattle by using one-step blending. *Genet. Select. Evol.*, 46, 1-5.
- Lourenco D.A.L., Fragomeni B.O., Tsuruta S., Aguilar I., Zumbach B., Hawken R.J., Legarra A., Misztal I., 2015. Accuracy of estimated breeding values with genomic information on males, females, or both: an example on broiler chicken. *Genet. Select. Evol.*, 47, 56.
- Lund M.S., De Ross S.P.W., De Vries A.G., Druet T., Ducrocq V., Fritz S., Guillaume F., Gulbrandsen B., Liu Z., Reents R., 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. *Genet. Select. Evol.*, 43, 1-28.
- Luo M.F., Wiggans G.R., Hubbard S.M., 1997. Variance component estimation and multitrait genetic evaluation for type traits of dairy goats. *J. Dairy Sci.*, 80, 594-600.
- Maroteau C., 2014. Cartographie fine de QTL pour des caractères d'intérêt pour la filière caprine. Thèse de doctorat, Université de Toulouse.
- Maroteau C., Palhière I., Larroque H., Clément V., Tosser-Klopp G., Rupp R., 2013. QTL detection for traits of interest for the dairy goat industry. 64th Ann. Meet. Eur. Fed. Anim. Sci., Nantes, France.
- Molee A., Duanghaklang N., Na-Lampang P., 2012. Effects of acyl-CoA:diacylglycerol acyl transferase 1 (DGAT1) gene on milk production traits in crossbred Holstein dairy cattle. *Trop. Anim. Health Prod.*, 44, 751-755.
- Palhière I., Maroteau C., Larroque H., Clément V., Tosser-Klopp G., Rupp R., 2012. QTL detection for milk production traits, fatty acids, udder morphology and milking speed in Alpine and Saanen goats. XIth Int. Conf. Goats, Gran Canaria, Espagne.
- Pérez-Enciso M., Rincón J.C., Legarra A., 2015. Sequence- vs. chip-assisted genomic selection: accurate biological information is advised. *Genet. Select. Evol.*, 47, 1-14.
- Pryce J., Hayes B.J., Goddard M.E., 2012. Genotyping dairy females can improve the reliability of genomic selection for young bulls and heifers and provide farmers with new management tools. 38th Int. Commit. Anim. Record., Cork, Irlande.
- Rogers A.R., Huff C., 2009. Linkage disequilibrium between loci with unknown phase. *Genetics*, 182, 839-844.
- Tribout T., 2011. Perspectives d'application de la sélection génomique dans les schémas d'amélioration génétique porcins. In: Numéro spécial, Amélioration génétique. Mulsant P., Bodin L., Coudurier B., Deretz S., Le Roy P., Quillet E., Perez J.M. (Eds). *INRA Prod. Anim.*, 24, 369-376.
- Tribout T., Barbat M., Saintilan R., Fouilloux M.N., Venot E., Phocas F., 2015. Mise en place d'évaluations génomiques nationales dans les populations bovines allaitantes françaises. *Renc. Rech. Rum.*, 22, 103-106.
- Vázquez-Flores F., Montaldo H.H., Torres-Vázquez J.A., Alonso-Morales R.A., Gayosso-Juárez H., 2012. Additive and dominance

effects of the α_{s1} -casein locus on milk yield and composition traits in dairy goats. *J. Dairy Res.*, 79, 367-374.

Vitezica Z.G., Elsen J.M., Rupp R., Diaz C., 2005. Using genotype probabilities in survival analysis: a scrapie case. *Genet. Select. Evol.*, 37, 403-415.

Vitezica Z.G., Aguilar I., Legarra A., 2010. One-step vs. multi-step methods for genomic prediction in presence of selection. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Allemagne.

Vitezica Z.G., Varona L., Legarra A., 2013. On the additive and dominant variance and cova-

riance of individuals within the genomic selection scope. *Genetics*, 195, 1223-1230.

Wiggans G.R., Cooper T.A., VanRaden P.M., Cole J.B., 2011. Technical note: Adjustment of traditional cow evaluations to improve accuracy of genomic predictions. *J. Dairy Sci.*, 94, 6188-6193.

Résumé

La sélection génomique, qui a révolutionné la sélection génétique des bovins laitiers notamment, est désormais envisagée dans d'autres filières animales. Chez les caprins laitiers français, le gain de précision attendu des valeurs génomiques était un des questionnements de la filière en raison de la petite taille de la population de référence disponible (825 mâles et 1945 femelles génotypés sur une puce SNP 50K). Le but de cette étude est de tester différentes techniques d'évaluation génomique afin d'obtenir les évaluations génomiques les plus précises possibles. Une étude de la structure génétique de la population de référence caprine constituée d'animaux de races Saanen et Alpine, a révélé de faibles niveaux de déséquilibre de liaison (0,17 entre deux SNP consécutifs), de consanguinité et de parenté au sein de la population, ce qui n'est pas favorable à une bonne précision des évaluations génomiques. Les méthodes d'évaluations génomiques (GBLUP ou Bayésiennes), basées sur des performances pré-corrigées n'ont pas permis une amélioration significative des précisions des évaluations génomiques pour les caractères évalués en routine (caractères de production, de morphologie et comptages de cellules somatiques). Cependant les évaluations génomiques basées sur les performances propres des femelles ont permis d'obtenir des précisions supérieures à celles obtenues sur ascendance. La sélection génomique est donc envisageable chez les caprins laitiers français. Ces précisions peuvent également être légèrement augmentées par l'inclusion de gènes majeurs tels que celui de la caséine α_{s1} .

Abstract

Toward genomic selection in dairy goats

Genomic selection, which is revolutionizing genetic selection in dairy cattle is now considered in the breeding of other animals. In French dairy goats, gain in accuracy using genomic selection has been questioned due to the small size of the reference population (825 males and 1 945 females genotyped). The aim of this study was to investigate how to reach adequate genomic evaluation accuracy in the French dairy goat population. The study of a reference population structure (Alpine and Saanen breeds) showed that the level of linkage disequilibrium (0.17 between two consecutive SNP), inbreeding and the relationship between the reference and candidate population were not ideal to maximize genomic evaluation accuracy. Two steps genomic evaluation (GBLUP, Bayesian) based on performances corrected for fixed effects did not improve genetic evaluation accuracy compared to classical evaluations for milk production traits, udder type traits and somatic cells scores. When using genomic evaluation directly based on female performances (single step), accuracy of genomic evaluation is higher than the accuracy level obtained from the ascendance. Genomic selection is feasible in French dairy goats. Accuracies could be slightly improved integrating a major gene such as α_{s1} casein.

CARILLIER-JACQUIN C., LARROQUE H., ROBERT-GRANIÉ C., 2017. Vers une sélection génomique chez les caprins laitiers. *INRA Prod. Anim.*, 30, 19-30.

