

Précision de l'imputation de génotypages haute densité à partir de puces basse densité pour des individus de race pure et croisés Piétrain

Alban BOUQUET (1), Katia FEVE (2), Juliette RIQUET (2), Catherine LARZUL (2)

(1) IFIP – Institut du Porc, BP 35104, 35651 Le Rheu Cedex, France

(2) INRA, UMR GenPhySE, Chemin de borde rouge, 31328 Castanet-Tolosan Cedex, France

Alban.bouquet@ifip.asso.fr

Précision de l'imputation de génotypages haute densité à partir de puces basse densité pour des individus de race pure et croisés Piétrain

L'objectif de cette étude est d'évaluer la précision d'une méthode d'imputation pour reconstruire les génotypages haute-densité (HD) d'animaux de race pure et croisés issus de verrats Piétrain à partir de puces basse-densité (BD). Les données analysées proviennent du projet ANR Utopige et incluent le génotypage sur le panel Illumina PorcineSNP60 de 792 animaux Piétrain et 733 F1 Piétrain x Large White (LW). Les génotypages HD de 491 individus LW et 219 Piétrain, antérieurs au projet Utopige, ont été inclus dans l'analyse. Pour un échantillon de 100 individus Piétrain et 100 F1, des puces BD ont été créées *in silico* en conservant les génotypes situés tous les 5, 10, 25, 50, 75 et 100 marqueurs de la puce HD. Les génotypes manquants ont été imputés avec le logiciel BEAGLE en considérant l'information HD des autres individus de race pure 1) Piétrain ou 2) Piétrain et LW. Pour les individus Piétrain échantillonnés, 99% des génotypes sont correctement imputés pour les puces BD comportant au moins 10% des marqueurs de la puce HD. Pour les individus croisés, l'imputation s'est révélée peu précise lorsque les génotypages HD des animaux LW ne sont pas inclus dans l'analyse. Lorsque les génotypages HD des deux populations parentales sont pris en compte dans l'analyse, la précision de l'imputation augmente significativement pour les croisés et est comparable à celle des individus Piétrain si la densité des panels BD est suffisante.

Imputation accuracy of high density genotypings using low density panels in purebred and two-way crossbred Pietrain pigs

The aim of this study was to assess the accuracy of imputation techniques to reconstruct high density (HD) genotyping using low density (LD) marker panels for purebred and crossbred pigs sired by Pietrain boars. Analysed data were collected in the Utopige project and included HD genotypings on the Illumina PorcineSNP60 beadchip for 792 purebred Pietrain animals and 733 Pietrain x Large White (LW) crossbreds. Genotypic data of 491 LW and 219 Pietrain older individuals were also taken into account. Low density panels were designed *in silico* from HD genotypic data for a sample of 100 purebred and 100 crossbred animals. To create these panels, 1 genotype was kept every 5, 10, 25, 50, 75 and 100 markers; all other marker data were removed. Discarded genotypes were imputed with the Beagle software using HD genotypic information from other 1) purebred Pietrain and 2) Pietrain and LW individuals. For the sample of purebred Pietrain, 99% of missing genotypes on the LD panel were correctly imputed when at least 10% of marker data were kept on the LD chip. For crossbreds, HD genotypings were imputed with low accuracy when HD genotypings of LW animals were not included in the reference population. When HD genotypings of both parental populations were included in the analysis, the accuracy of imputation significantly increased for crossbreds and was similar to the one achieved for purebred animals when LD panels considered for genotyping were sufficiently dense.

INTRODUCTION

A l'instar d'autres espèces d'élevage, la sélection génomique est intégrée progressivement dans les schémas de sélection des porcs. Cette méthode consiste à utiliser l'information du génome dans les modèles d'évaluation génétique (Meuwissen *et al.*, 2001). L'information génomique est révélée par génotypage des individus pour un panel de marqueurs génétiques (appelés SNP). A partir de l'information génomique et phénotypique d'un groupe d'individus dits « de référence », il est alors possible de prédire les valeurs génétiques des candidats à la sélection à l'aune de leur seul génotypage avec une précision plus élevée que le modèle d'évaluation génétique conventionnel. La précision des index prédits dépend fortement de la taille de la population de référence et de la densité en marqueurs de la puce utilisée pour le génotypage. En effet, en utilisant un nombre élevé de marqueurs répartis sur tout le génome, il est plus probable de détecter des SNP fortement associés à un caractère d'intérêt.

A l'heure actuelle, le coût du génotypage haute densité (panels HD $\geq 60\ 000$ SNP) constitue un des principaux freins à la mise en place de la sélection génomique. Une stratégie consiste alors à génotyper les reproducteurs à l'aide de panels HD et les candidats sur des panels basse densité moins chers (panels BD $< 10\ 000$ SNP). En utilisant des méthodes statistiques dites « d'imputation », il est possible d'inférer les génotypes manquants sur la puce BD des candidats à partir de l'information HD des reproducteurs. Les méthodes d'imputation se sont révélées être très précises pour reconstituer des génotypes HD à partir de puces BD lorsqu'appliquées en race pure, par exemple chez les bovins mais aussi chez les porcs (Zhang et Druet, 2010; Cleveland et Hickey, 2013).

Bien que le croisement soit communément utilisé en production porcine, les choix de reproducteurs effectués dans les noyaux de sélection intègrent principalement des données enregistrées sur des individus de race pure. A plus long terme, la sélection génomique pourrait permettre d'intégrer plus aisément l'information d'individus croisés dans les modèles d'évaluation génétique. Les bénéfices de cette stratégie seraient multiples : prise en compte de phénotypes récoltés dans des environnements plus proches des conditions rencontrées dans les élevages de production, modélisation de l'effet d'hétérosis et des effets génétiques sous-jacents (Tribout, 2011). Cependant, les individus croisés ne participent pas au renouvellement de la population de race pure et n'ont donc d'intérêt que par les phénotypes qu'ils apportent. Pour réduire les coûts de génotypage, il peut alors sembler pertinent de les génotyper avec une puce BD et d'imputer leur génotypage HD. Cette stratégie serait économiquement avantageuse si l'imputation du génotypage des croisés peut être réalisée en exploitant les génotypes HD d'individus de race pure des populations parentales. Récemment, Xiang *et al.* (2014) ont montré qu'il est possible d'imputer avec précision une puce 7K à partir d'une puce 5K pour des croisés Landrace x Yorkshire en utilisant les génotypes HD des populations parentales. Toutefois, pour une application en routine, il est nécessaire d'imputer des génotypes HD pour les croisés.

L'objectif de cette étude est donc d'évaluer, à partir de données réelles, la faisabilité et la précision de l'imputation du génotypage d'individus croisés en utilisant les génotypes HD d'individus de race pure des populations parentales.

1. MATERIEL ET METHODES

1.1. Données et génotypages

Les données analysées dans cette étude ont été collectées dans le cadre du projet ANR Utopige. Elles comprennent les données de génotypage de 792 individus de race pure Piétrain inscrits aux Livres Généalogiques Porcins Collectifs (LGPC) – dont 99 pères et 693 descendants - et de 733 individus croisés Piétrain x Large White issus des mêmes pères. Tous ces individus ont été génotypés à l'aide du panel Illumina PorcineSNP60 Beadchip (Illumina, San Diego, CA). Ces données ont été complétées par l'information génomique de 491 individus Large White (LW) et de 219 individus Piétrain de la même population. Ces animaux ont été génotypés en dehors du projet Utopige pour le même panel de marqueurs.

Différents critères ont été pris en compte pour évaluer la qualité des données de génotypage. En particulier, nous avons écarté les données 1) des marqueurs ayant plus de 10% de génotypes manquants dans la population, 2) des individus ayant un taux d'hybridation $< 90\%$ (i.e. plus de 10% de génotypes manquants pour l'individu), 3) des marqueurs avec une fréquence de l'allèle mineur $< 1\%$ et 4) des individus dont les génotypes sont incompatibles avec celui de leur père. La structure familiale du dispositif expérimental a été utilisée pour déduire si les erreurs de génotypage sont attribuables au père ou au descendant, chaque père possédant plusieurs descendants de race pure et croisés. Ainsi, si plus de 2 couples père-descendant présentaient une incompatibilité mendélienne au même marqueur, le génotype du père a été invalidé à ce marqueur. Inversement, le génotype au marqueur a été invalidé pour le descendant si le conflit ne concernait qu'un seul couple père-descendant. Enfin, les marqueurs pour lesquels la répartition des génotypes observés déviait fortement des proportions attendues sous l'hypothèse d'équilibre de Hardy Weinberg ont été éliminés (statistique de test du $\chi^2 > 300$). Dans cette étude, seules les données des marqueurs situés sur le chromosome 1 ont été prises en compte (7221 SNP). Ainsi, après élimination des données aberrantes, nous avons conservé l'information génotypique de 5397 marqueurs pour 491 individus LW, 988 Piétrain et 685 croisés Piétrain x LW.

1.2. Création des panels *in silico* et scénarios d'imputation

Pour évaluer la précision de l'imputation, des panels BD ont été créés *in silico* par échantillonnage des marqueurs contenus sur la puce HD pour un groupe de 100 descendants de race pure et de 100 descendants croisés. Ces 200 individus ont été choisis aléatoirement en veillant toutefois à ne pas sélectionner de pleins-frères.

Différentes densités en marqueurs ont été considérées pour créer les panels BD. Ces panels ont été générés en conservant les informations génotypiques situées tous les 5, 10, 25, 50, 75 et 100 marqueurs (noté respectivement panels 1:5, 1:10, 1:25, 1:50, 1:75, 1:100), c'est-à-dire en ne retenant que 20%, 10%, 4%, 2%, 1,3% et 1% des marqueurs de la puce HD. Ces densités correspondent à l'utilisation de puces BD comprenant l'équivalent d'environ 10K, 5K, 2,5K, 1K, 750 et 450 SNP respectivement, répartis sur tout le génome.

Les génotypes manquants sur les puces BD des 200 individus échantillonnés ont été imputés pour le chromosome 1 à l'aide du logiciel Beagle (Browning et Browning, 2009).

Trois scénarios ont été envisagés pour réaliser l'imputation et diffèrent par la composition de la population de référence des individus génotypés sur la puce HD.

Dans le premier scénario (S1), seuls les individus de race pure Piétrain génotypés dans le cadre du projet Utopige ont été inclus dans la population de référence (hormis les individus échantillonnés), soit 669 individus. Ce scénario doit permettre d'évaluer la précision de l'imputation pour les individus de race pure et croisés en ayant à disposition les génotypes HD d'une seule population parentale comprenant des individus fortement apparentés aux animaux inférés (pères, pleins-frères, demi-frères).

Dans le second scénario (S2), les 491 individus de race pure LW (non directement apparentés aux animaux Utopige) ont été ajoutés à la population de référence du scénario S1. Dans ce scénario, nous évaluerons le gain de précision obtenu par la prise en compte des génotypes HD de la deuxième population parentale.

Dans le troisième scénario (S3), la population de référence comprend les 491 individus LW et 219 Piétrain génotypés en dehors du projet Utopige. Ce scénario a été étudié pour mesurer l'effet de taille de la population de référence et de son apparentement avec les candidats sur la précision des génotypes imputés.

1.3. Mesures de la précision de l'imputation

Deux critères ont permis d'évaluer la précision de l'imputation des génotypes HD pour chaque individu échantillonné :

- le taux de concordance entre génotypes observés et imputés, qui correspond au pourcentage de génotypes correctement imputés parmi les marqueurs soumis à l'imputation.
- La corrélation entre les génotypes observés et imputés

Toutes les analyses ont été réalisées de façon indépendante en considérant successivement chaque densité de puces BD et chaque scénario. Les critères de précision de l'imputation ont été moyennés pour l'ensemble des individus échantillonnés pour chaque densité de puces BD et chaque scénario.

2. RESULTATS

2.1. Imputation à partir des génotypes HD d'une seule population parentale

La précision d'imputation varie fortement entre individus Piétrain et croisés dans le scénario S1 où l'imputation est réalisée à partir des génotypes HD de la seule population de race pure Piétrain (Figure 1). Le taux de concordance entre SNP imputés et observés est supérieur à 99% pour les puces BD avec les densités en marqueurs les plus élevées (10K (1:5), 5K (1:10)) et le coefficient de corrélation est supérieur à 0,98. Pour les puces de plus faible densité, la précision de l'imputation diminue progressivement : 96% des SNP sont correctement imputés si l'on utilise la puce BD de densité 2,5K (1:25) et 83% des SNP sont correctement imputés à partir de la puce comprenant 450 SNP (densité 1:100).

La précision de l'imputation est beaucoup plus faible pour les individus croisés que ceux de race pure (Figure 1). Le taux de concordance est de 89% dans le scénario le plus favorable (puce 10K (densité 1:5)) et diminue à 62% dans le scénario le moins favorable (puce 450 SNP (densité 1:100)). La même évolution est observée pour les coefficients de corrélation entre génotypes observés et imputés.

2.2. Imputation avec les génotypes HD des deux populations parentales

Lorsque les génotypes HD des deux populations parentales sont intégrés dans la population de référence, la précision de l'imputation reste élevée pour les individus de race pure pour les puces BD les plus denses (Figure 2).

Dans ce scénario, la précision de l'imputation est comparable entre les individus de race pure et croisés lorsque les puces BD ont une densité suffisante (10K (1:5), 5K (1:10)). Ainsi, le logiciel utilisé permet de reconstruire avec une grande précision des génotypes HD à partir de puces BD pour les individus croisés en exploitant les génotypes HD des deux populations parentales.

En utilisant les puces BD de plus faible densité, le taux de concordance et plus particulièrement le coefficient de corrélation entre génotypes observés et imputés sont plus faibles à la fois pour les individus purs et croisés.

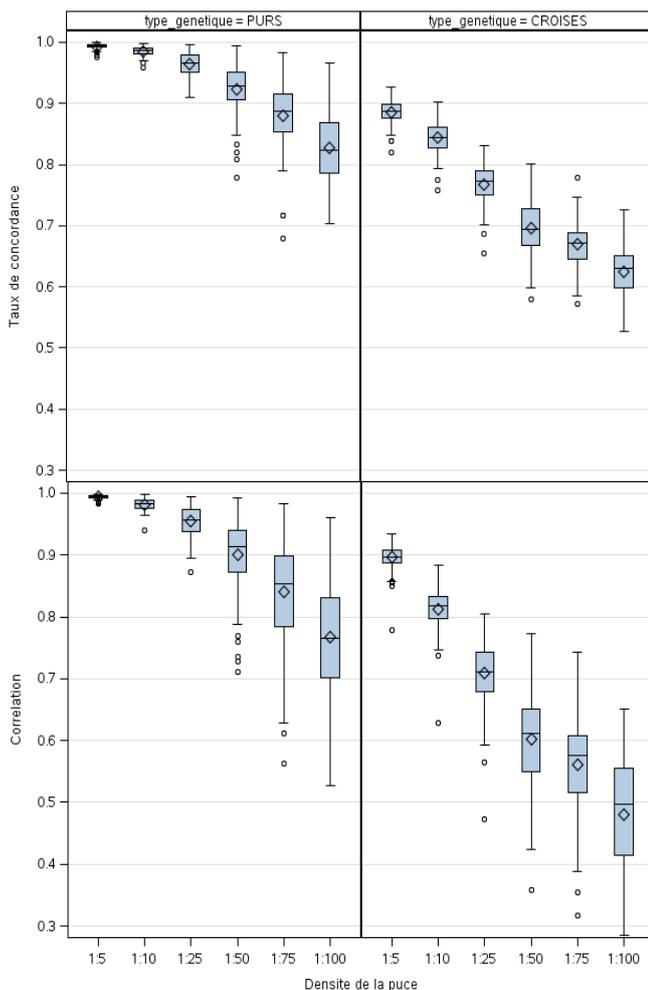


Figure 1 – Taux de concordance et corrélation entre génotypes observés et imputés pour les individus de race pure et croisés lorsque l'imputation est réalisée à partir des seuls génotypes HD d'individus de race pure Piétrain (scénario S1).

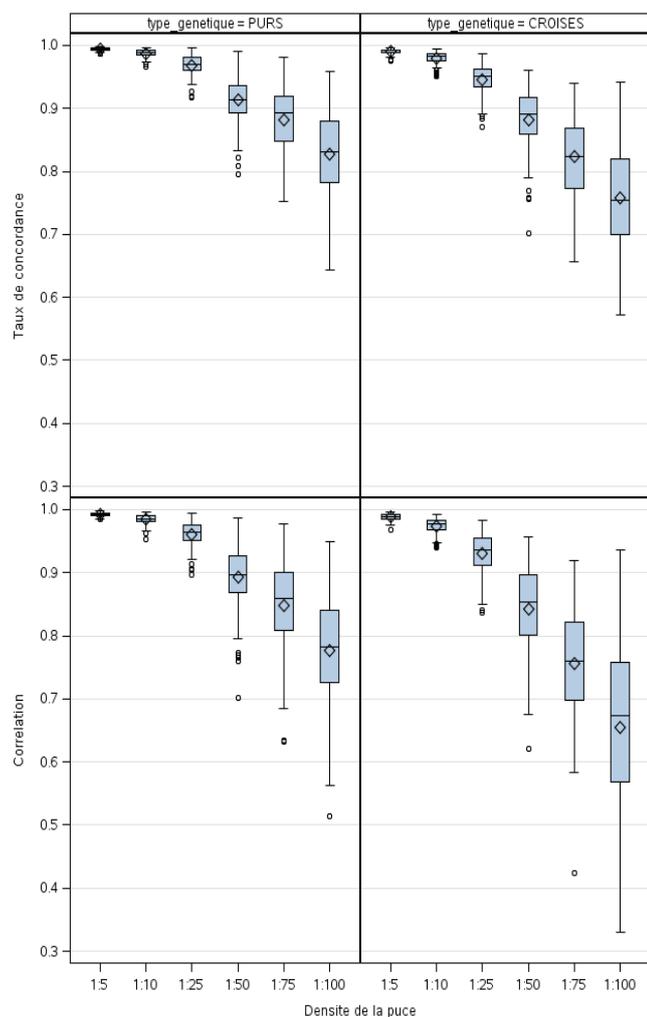


Figure 2 - Taux de concordance et corrélation entre génotypes observés et imputés pour les individus de race pure et croisés dans le scénario S2

2.3. Effet de la taille de la population de référence et de son apparentement avec les candidats

Dans le dernier scénario (S3), l'imputation a été réalisée en exploitant les génotypes HD des deux populations parentales, mais en considérant des individus de référence Piétrain moins apparentés aux candidats à imputer. Les taux de concordance et les coefficients de corrélation présentent des évolutions similaires quelle que soit la densité de la puce BD. Ainsi, seuls les coefficients de corrélation de la présente étude sont présentés dans le Tableau 1.

Tableau 1 – Corrélations entre génotypes imputés et observés pour les individus de race pure et croisés dans le scénario S3 et écart avec le scénario S2

Densité de la puce BD	RACE PURE		CROISES	
	Moyenne	Ecart / scénario S2	Moyenne	Ecart / scénario S2
1:5	0,98	-0,01	0,98	-0,01
1:10	0,96	-0,02	0,96	-0,01
1:25	0,90	-0,06	0,88	-0,05
1:50	0,80	-0,09	0,76	-0,08
1:75	0,73	-0,12	0,67	-0,09
1:100	0,63	-0,15	0,57	-0,08

La précision de l'imputation est élevée et équivalente pour les individus de race pure et croisés lorsque le génotypage est réalisée à partir de la puce BD la plus dense (10K (1:5)). La corrélation entre génotypes observés et imputés diminue ensuite plus rapidement que dans le scénario S2, lorsque l'imputation est réalisée à partir des puces BD de plus faible densité. Cette diminution est particulièrement importante pour les individus de race pure compte tenu du plus faible apparentement entre les individus de la population de référence et les animaux Piétrain inférés dans ce scénario S3. Toutefois, l'imputation est toujours en moyenne plus précise pour les individus de race pure que pour les individus croisés dans ce scénario.

3. DISCUSSION

Les résultats de cette étude confirment que les techniques d'imputation permettent de reconstruire des génotypes HD à partir de puces BD avec une précision élevée pour des individus de race pure. A partir de données réelles, nous avons montré qu'il était également possible d'utiliser ces techniques pour reconstruire les génotypes HD d'individus croisés avec une précision comparable à celle des individus de race pure à la condition que 1) des génotypes HD d'individus des deux populations parentales soient pris en compte dans l'analyse et 2) que la densité de la puce BD soit suffisante.

3.1. Imputation des génotypes d'individus croisés

La méthode d'imputation utilisée semble applicable pour réaliser l'imputation de génotypes d'individus croisés, si tant est que la population de référence comprenne des génotypes HD des deux populations parentales. Cette étude confirme donc les résultats récents présentés par Xiang *et al.* (2014) chez le porc pour réaliser l'imputation d'une puce 5K à une puce 7K. Les résultats de la présente étude montrent que l'imputation peut être utilisée avec succès pour imputer des génotypes HD (équivalent 60K) à partir de puces 5K. Par ailleurs, la prise en compte d'individus LW dans la population de référence n'a pas affecté la précision de l'imputation des génotypes des individus Piétrain de race pure. En pratique, ce résultat est intéressant parce qu'il suggère que l'imputation des individus de race pure et croisés peut être réalisée de façon concomitante sans affecter la précision des génotypes imputés.

Comme la plupart des méthodes d'imputation, le logiciel Beagle (Browning et Browning, 2009) utilise le déséquilibre de liaison existant entre marqueurs génétiques pour inférer les génotypes manquants sur la puce BD. Le déséquilibre de liaison correspond à l'association statistique observée entre marqueurs liés sur un même chromosome. Il est inhérent à chaque population et dépend notamment de son historique (évolution démographique, sélection). Pour réaliser l'imputation, le logiciel Beagle utilise une méthode itérative composée de deux étapes. Tout d'abord, les haplotypes parentaux les plus probables sont reconstruits à l'aide d'un modèle statistique utilisant des chaînes de Markov cachées. Ensuite, un algorithme de classification est utilisé pour déterminer les haplotypes paternel et maternel les plus probables à partir de combinaisons d'haplotypes présents dans la population de référence. Le génotype à chaque marqueur est alors déduit comme la somme des allèles prédits sur chaque haplotype parental.

Cette méthode a été conçue pour des populations de très grande taille dans lesquelles les deux parents peuvent être considérés comme non apparentés pour garantir une efficacité maximale à l'algorithme de classification. En pratique, ce logiciel s'est avéré très précis lorsqu'appliqué à des populations d'animaux sélectionnées (Calus *et al.*, 2014).

Dans cette étude, la précision de l'imputation diminue fortement pour des densités inférieures à 2,5K (1:25). Cela signifie en pratique que la densité en marqueurs n'est plus suffisante pour capter le déséquilibre de liaison de la population Piétrain. Cependant, dans cette étude, il faut souligner que les marqueurs de la puce BD ont été choisis au hasard parmi les marqueurs valides de la puce HD. Des stratégies visant à sélectionner les marqueurs à inclure sur le panel BD pour capter le maximum de déséquilibre de liaison dans la population pourraient permettre de réduire le nombre de marqueurs nécessaires sans compromettre la précision de l'imputation (Badke *et al.*, 2013). Enfin, la persistance du déséquilibre de liaison pouvant être différente entre chromosomes, il est envisagé d'estimer par la suite la précision de l'imputation pour l'ensemble des chromosomes. Les travaux de Xiang *et al.* (2014) ont toutefois montré que la précision de l'imputation estimée pour le chromosome 1 se situe dans la moyenne des autres chromosomes pour les populations Landrace et Yorkshire danoises.

3.2. Importance de l'apparement entre population de référence et candidats à l'imputation

Le scénario S1 dans lequel tous les individus du protocole Utopige étaient inclus dans la population de référence est très favorable pour obtenir une précision élevée de l'imputation. En effet, dans ce scénario, tous les candidats à l'imputation ont leur père, des demi-frères voire un plein frère génotypés. Dans la réalité, l'apparement entre la population de référence et les candidats à l'imputation peut être plus faible.

Dans le scénario S3, l'imputation du génotypage des individus purs s'est révélée presque aussi précise que dans le scénario S2 lorsque la densité de la puce BD est suffisante (10K et 5K), et ce malgré la taille modérée de la population de référence et le plus faible apparement entre individus de référence et candidats à l'imputation. Pour les puces moins denses, la précision de l'imputation est fortement diminuée, notamment pour les individus de race pure. Cette baisse de précision est liée à la diminution de l'apparement entre la population de référence et les candidats à l'imputation, mais probablement aussi à la réduction par trois de la taille de la population de référence par rapport au scénario S2. Dans le scénario S3, la taille de la population de référence ne permet certainement plus de décrire avec précision la diversité des haplotypes présents dans la population. Dans les situations où l'apparement entre la population de référence et les candidats à l'imputation est faible, il semble que d'autres méthodes d'imputation utilisant de façon explicite les relations de parenté entre individus permettent d'imputer les génotypes avec une plus grande précision, notamment lorsque la taille de la population de référence est modérée (Ventura *et al.*, 2014).

3.3. Perspectives d'application dans les schémas de sélection

Les puces BD commercialisées actuellement présentent une densité proche du scénario 1:5, soit environ 10K. Leur densité semble donc suffisante pour réaliser l'imputation des génotypes HD d'individus de race pure et croisés. Dans la pratique, il semblerait donc envisageable de réduire encore la densité des puces BD sans diminuer la précision de l'imputation. Les résultats présentés dans cette étude sont valables pour un schéma de croisement à deux voies. Ce type de croisement se rencontre fréquemment pour la production de truies Large White x Landrace. Ces résultats n'augurent toutefois rien de la précision de l'imputation attendue pour des individus issus du croisement de trois lignées. En effet, dans le cas d'un schéma de croisement à deux voies, le logiciel Beagle reconstruit les haplotypes paternel et maternel en utilisant l'information du déséquilibre de liaison existant dans les populations parentales. Si la mère de l'individu est elle-même croisée, il n'est pas sûr que le logiciel puisse identifier avec précision l'haplotype maternel le plus probable.

A l'heure actuelle, les modèles d'évaluation génomique ne permettent pas l'intégration en routine des informations issues d'individus croisés. Cette thématique de recherche fait cependant l'objet de développements méthodologiques importants (Zeng *et al.*, 2013; Christensen *et al.*, 2014). L'intérêt d'inclure ces nouveaux phénotypes d'individus croisés dépendra donc de la précision des modèles d'évaluation génomique pour estimer les indices de sélection des individus purs. Les bénéfices envisagés de l'utilisation de tels modèles sont multiples : modélisation de l'effet d'hétérosis et des effets génétiques non additifs, prise en compte d'effets d'environnement plus proches des conditions rencontrées dans les élevages de production, mutualisation des coûts de génotypage parce qu'une population de référence d'individus croisés peut être utilisée pour l'évaluation génomique de plusieurs populations parentales (Tribout, 2011). A cette fin, l'utilisation de panels BD pour génotyper les individus croisés pourrait être effectivement une stratégie pertinente pour réduire les coûts de génotypage.

CONCLUSION

Les résultats de cette étude suggèrent que l'imputation peut être utilisée pour reconstruire avec précision des génotypes HD à partir de puces BD pour des porcs de race pure et issus du croisement de deux populations. Toutefois, pour obtenir une précision d'imputation élevée, il est nécessaire que la population de référence intègre des génotypes HD des deux populations parentales et que la densité de la puce BD soit suffisante. Ces résultats ouvrent de réelles perspectives pour une mise en application de la sélection génomique utilisant l'information d'individus croisés à moindre coût.

REMERCIEMENTS

Cette étude a bénéficié du soutien financier de l'ANR (programme Utopige ANR10-GENOM BTV015) et de Bioporc. Les auteurs remercient les entreprises de sélection ADN, Gène+ et Nucléus pour la mise à disposition des données de génotypes d'individus Large White et Piétrain.

REFERENCES BIBLIOGRAPHIQUES

- Badke Y.M., Bates R.O., Ernst C.W., Schwab C., Fix J., Van Tassel C.P., Steibel J.P., 2013. Methods of tagSNP selection and other variables affecting imputation accuracy in swine. *BMC Genet.*, 14, 8.
- Browning B.L., Browning S.R., 2009. A Unified Approach to Genotype Imputation and Haplotype-Phase Inference for Large Data Sets of Trios and Unrelated Individuals. *Am. J. Hum. Genet.*, 84, 210-223.
- Calus M.P.L., Bouwman A.C., Hickey J.M., Veerkamp R.F., Mulder H.A., 2014. Evaluation of measures of correctness of genotype imputation in the context of genomic prediction: a review of livestock applications. *Animal*, 8, 1743-1753.
- Christensen O.F., Madsen P., Nielsen B., Su G., 2014. Genomic evaluation of both purebred and crossbred performances. *Genet. Sel. Evol.*, 46, 23.
- Cleveland M.A., Hickey J.M., 2013. Practical implementation of cost-effective genomic selection in commercial pig breeding using imputation. *J. Anim. Sci.*, 91, 3583-3592.
- Meuwissen T.H., Hayes B.J., Goddard M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157, 1819-1829.
- Tribout T., 2011. Perspectives d'application de la sélection génomique dans les schémas d'amélioration génétique porcins. *INRA Prod. Anim.*, 24, 369-376.
- Ventura R.V., Lu D., Schenkel F.S., Wang Z., Li C., Miller S.P., 2014. Impact of reference population on accuracy of imputation from 6K to 50K single nucleotide polymorphism chips in purebred and crossbred beef cattle. *J. Anim. Sci.*, 92, 1433-1444.
- Xiang T., Christensen O.F., Legarra A.C., Ostersen T., 2014. Imputation of genotypes in Danish two-way crossbred pigs using low density panels. In: American Society of Animal Science (Eds), *Proceeding of the 10th World Congress of Genetics Applied to Livestock Production*, Vancouver, Canada.
- Zeng J., Toosi A., Fernando R.L., Dekkers J.C., Garrick D.J., 2013. Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action. *Genet. Sel. Evol.*, 45, 11.
- Zhang Z., Druet T., 2010. Marker imputation with low-density marker panels in Dutch Holstein cattle. *J. Dairy Sci.*, 93, 5487-5494.