

# Utilisation d'une puce très basse densité (1 100 SNP) pour la sélection génomique chez 3 races porcines françaises

Céline CARILLIER-JACQUIN (1), Alban BOUQUET (2), Yann LABRUNE (1), Pauline BRENAUT (2), Juliette RIQUET (1),  
Catherine LARZUL (1)

(1) GenPhySE, Université de Toulouse, INRAE, INPT, ENVT, 31320, Castanet-Tolosan, France

(2) IFIP-Institut du Porc, 35651 Le Rheu Cedex, France

[celine.carillier-jacquin@inrae.fr](mailto:celine.carillier-jacquin@inrae.fr)

## Design of a very low density chip (1 100 SNP) for genomic selection in 3 French pig breeds

To reduce genotyping costs for genomic selection, a Low-Density SNP (LD) chip, designed in 2016, is now used routinely. This panel is composed of approximately 1 100 equidistant SNPs. The relevance of this chip has been studied in French populations of the Landrace, Large White and Pietrain pig breeds. The quality of imputation was estimated by the correlation between actual and imputed genotypes and error rates. The impact of imputation on the genomic evaluations was estimated by the correlation between the genomic values obtained for the candidates with imputed genotypes, and those obtained with the high-density genotypes. Average error rates of imputation estimated on all the chromosomes were 0.03, 0.11 and 0.14 for Landrace, Large White and Pietrain, respectively. The estimated correlations between actual and imputed genotypes were relatively high at 0.93, 0.92 and 0.88 for Landrace, Large White and Pietrain populations, respectively. Correlations between genomic breeding values predicted with high-density genomic data or imputed genomic data from the LD SNP panel ranged from 0.89-0.97 for Large White and Landrace populations for reproductive traits. They were higher than those obtained for the Pietrain population (0.80 and 0.97 for production traits, respectively). In conclusion, despite the limited number of SNPs on the low-density panel used in this study, the imputation accuracy is sufficient to use the imputed genotypes in the genomic evaluations. In practice, genotyping candidates with a LD chip is a solution for selecting future breeding pigs at lower cost.

## INTRODUCTION

Le coût actuel des puces SNP constitue une limite au développement de la sélection génomique dans les populations porcines (Badke *et al.*, 2014 ; Wellmann *et al.*, 2013). Afin de réduire les coûts de génotypage, une puce SNP basse densité (LD), conçue en 2016, est utilisée en routine. Ce panel LD d'environ 1100 SNP a été optimisé pour maximiser la précision d'imputation dans la population de porcs Landrace français. Dans la présente étude, nous avons proposé d'étudier l'impact de l'utilisation de ce panel pour deux autres grandes races de porc françaises, le Large White et le Piétrain.

## 1. MATERIEL ET METHODES

### 1.1. Données

Le jeu de données est composé de 3 724 Landrace, 1 061 Large White et 984 Piétrain (projet Utopige) génotypés avec la puce porcine Illumina SNP60 ou la puce GeneSeek 80k. Le panel basse densité d'environ 1 100 marqueurs contient des SNP sélectionnés pour le contrôle de filiation et un ensemble de SNP additionnels situés de manière équidistante les uns des autres pour optimiser l'imputation.

La population d'animaux génotypés a été divisée en deux sous-ensembles : 1) une population d'apprentissage dont l'ensemble des SNP sur puces moyenne densité sont connus (Large White

et Landrace nés avant 2016, 96 pères Piétrain ainsi qu'un maximum de 9 fils/père Piétrain choisis au hasard) et 2) une population de validation (Large White et Landrace nés en 2016 et 96 fils Piétrain non inclus dans la population d'apprentissage) dont les génotypes connus ne sont que ceux des SNP présents sur la puce basse densité.

### 1.2. Imputation

L'imputation a été réalisée séparément dans chaque race. Le logiciel FImpute (Sargolzaei *et al.*, 2011) a été utilisé pour imputer les génotypes sur la base d'informations du pedigree et des SNP connus. La précision d'imputation a été évaluée à l'aide des corrélations de Pearson entre génotypes imputés et les génotypes réels, par animal et par SNP, et le taux d'erreur d'imputation estimé comme le rapport entre la somme des erreurs d'imputation pour le SNP considéré et le nombre d'animaux génotypés.

### 1.3. Prédiction des valeurs génomiques

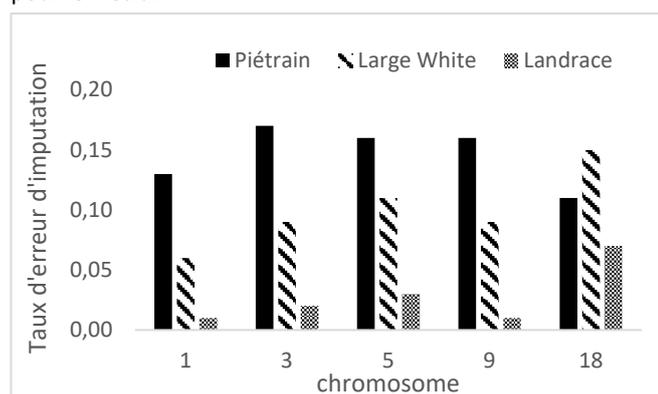
Des évaluations génomiques ont été menées séparément dans chaque race en utilisant la méthode BLUP génomique en une étape (Legarra *et al.*, 2009) avec le programme blupf90 (Misztal *et al.*, 2002) pour les principaux caractères sélectionnés dans chacune des trois races en utilisant les paramètres génétiques des évaluations de routine. L'impact de l'utilisation du panel LD pour l'évaluation génomique a été évalué par la corrélation

entre les valeurs génomiques estimées avec les génotypes imputés et celles estimées avec les génotypes moyenne densité pour la population de validation.

## 2. RESULTATS ET DISCUSSION

### 2.1. Qualité d'imputation

Les taux d'erreur d'imputation moyens par chromosome sont de 0,03 pour le Landrace, 0,11 pour le Large White et de 0,14 pour le Piétrain. Ces taux d'erreur sont variables selon le chromosome considéré (Figure 1), allant de 0,01 à 0,07 pour le Landrace, de 0,06 à 0,15 pour le Large White et de 0,11 à 0,17 pour le Piétrain.



**Figure 1** –Taux d'erreur d'imputation pour les chromosomes 1, 3, 5, 9 et 18 pour les trois races : Landrace, Large White et Piétrain.

Les meilleurs résultats ont été obtenus pour la population Landrace, ce qui était attendu puisque le panel LD a été optimisé pour cette race et que le nombre d'animaux génotypés sur un panel moyenne-densité était beaucoup plus élevé. Des tendances similaires ont été obtenues pour les corrélations entre les génotypes réels et imputés estimées par SNP et par animal (entre 0,88 pour les Piétrains et 0,93 pour les Landrace). La précision d'imputation obtenue dans cette étude est similaire à celles rapportées dans d'autres études ayant pris en compte des panels de SNP de très faible densité chez des porcs (Cleveland et Hickey, 2013 ; Huang *et al.*, 2012).

### 2.2. Evaluation génomique

Le tableau 1 présente les corrélations entre les valeurs génomiques prédites à l'aide de génotypes imputés et celles prédites avec les génotypes complets pour les populations de validation Landrace et Large White. Ces corrélations sont comprises entre 0,90 en Large White et 0,97 en Landrace pour le nombre de porcelets nés vivants. Ces résultats sont meilleurs que ceux obtenus pour la population Piétrain (de 0,80 pour l'épaisseur de muscle à 0,97 pour taux de muscle pièce).

**Tableau 1** – Corrélations entre les valeurs génétiques obtenues après imputation et celles obtenues avec les génotypes réels pour les animaux Large White and Landrace

	Landrace	Large White
Porcelets nés vivants	0,97	0,90
Porcelets sevrés par la truie	0,97	0,92
Poids moyen des porcelets à la naissance	0,94	0,91

Ces corrélations sont élevées compte tenu de la qualité de l'imputation trouvée dans cette étude. En effet, elles sont similaires à celles obtenues par Badke *et al.* (2014) pour lesquelles l'imputation était de meilleure qualité.

## CONCLUSION

Le panel basse densité utilisé pour le génotypage des candidats dans les populations françaises a été conçu pour maximiser la qualité d'imputation en Landrace. La qualité d'imputation obtenue pour les populations Large White et Piétrain est moins élevée que pour la population Landrace, mais similaire à celle d'autres études prenant en compte la même densité de SNP. Malgré la petite taille du panel LD utilisé dans cette étude, la précision d'imputation obtenue est suffisante pour utiliser les génotypes imputés dans les évaluations génomiques.

Ce panel semble être une bonne solution pour le génotypage des candidats et des femelles à moindre coût non seulement chez le Landrace mais également chez le Large White et le Piétrain.

## REFERENCES BIBLIOGRAPHIQUES

- Badke Y.M., Bates R.O., Ernst C.W., Fix J., Steibel J.P., 2014. Accuracy of estimation of genomic breeding values in pigs using low-density genotypes and imputation. *G3 Genes Genomes Genet.* 4, 623–631.
- 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.
- Huang Y., Hickey J.M., Cleveland M.A., Maltecca C., 2012. Assessment of alternative genotyping strategies to maximize imputation accuracy at minimal cost. *Genet. Sel. Evol.* 44, 25.
- Legarra A., Aguilar I., Misztal I., 2009. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92, 4656–4663.
- Misztal I., Tsuruta S., Lourenco D., Aguilar I., Legarra A., Vitezica Z., 2014. Manual for BLUPF90 family of programs. University of Georgia, Athens, USA
- Sargolzaei M., Chesnais J.P., Schenkel F.S., 2011. Fimpute - An efficient imputation algorithm for dairy cattle populations. *J. Dairy Sci.* 94: E-Suppl. 1, 421.
- Wellmann R., Preuß S., Tholen E., Heinkel J., Wimmers K., Bennewitz J., 2013. Genomic selection using low density marker panels with application to a sire line in pigs. *Genet. Sel. Evol.* 45, 28.