

Etude de la capacité immunitaire chez le porc Large White : recherche de marqueurs génétiques liés à l'expression des gènes dans le sang

Tatiana MAROILLEY (1,), Maria BALLESTER (1), Gaétan LEMMONIER (1), Marie-José MERCAT (2), Yvon BILLON (3),
Marco MOROLDO (1), Claire ROGEL-GAILLARD (1), Jordi ESTELLE (1)

(1) INRA-AgroParisTech, UMR1313 Génétique Animale et Biologie Intégrative, Jouy-en-Josas, France

(2) IFIP-BIOPORC, Pôle génétique, Le Rheu, France

(3) INRA, UE967 GENESI, Le Magneraud, France

tatiana.maroilley@jouy.inra.fr

Analysis of the immune capacity in Large White pigs: search for genetic markers linked to the variation of gene expression in the blood

Combining animal production performances and health traits is a main issue in livestock farming. Blood is emerging as a source of biological information linked to the health and physiological state of each individual. Furthermore, it is an interesting surrogate tissue for animal phenotyping because its sampling is almost non-invasive and reproducible. We have already shown that several immune parameters are heritable and that the blood transcriptome is informative to quantify some of them. In this study, our aim was to identify genetic markers associated with gene expression variations in the blood, based on an expression genome-wide association study (eGWAS) using 243 Large White pigs at 60 days of age. Each pig was genotyped with the Illumina iSelect 60K Chip and the blood transcriptome analysis was performed by using a custom gene expression 8X60K microarray (Agilent Technologies). 4,591 significant associations between one SNP and one probe were found, including 62% *cis*-associations (the associated SNP is in a window of 2Mb surrounding the probe). Among the 3,419 probes with at least one associated SNP (eQTL), 70% had a *cis*-eQTL and 89% could be assigned to a porcine gene. The eQTLs were found to be linked with the expression of one to 51 genes. The genomic regions surrounding the eQTLs correspond to candidate regions to further study the genetic architecture of immunity trait variations in pigs, and contribute to paving the way towards a qualification of the individual immune capacity for translational research in precision pig farming.

INTRODUCTION

En élevage porcin, les schémas de sélection cherchent à produire des animaux plus robustes afin, notamment, de réduire l'usage des antibiotiques. Nous avons engagé des recherches qui visent à qualifier la compétence immunitaire des animaux et étudions l'architecture génétique de paramètres immuns dans des populations d'animaux cliniquement sains. Nous avons déjà montré que de nombreux paramètres immunitaires sont héréditaires (Flori *et al.* 2011) et que le transcriptome du sang est informatif pour quantifier certains d'entre eux (Mach *et al.* 2013). Notre objectif est maintenant d'identifier des marqueurs génétiques associés à la variation d'expression des gènes dans le sang.

1. MATERIEL ET METHODES

1.1. Dispositif expérimental

Les animaux ont été produits dans l'unité INRA expérimentale GENESI (Le Magneraud, 17700, Surgères) dans le cadre du projet ANR Sus-Flora, reposant sur un partenariat pluridisciplinaire entre 6 unités INRA et l'association BIOPORC.

Un des objectifs de ce projet est de caractériser les paramètres génétiques de la réponse immunitaire chez le porc Large White. Le sang de 243 porcelets a été prélevé 60 jours après vaccination contre *Mycoplasma Hyopneumoniae*.

1.2. Génotypage des animaux

L'ADN a été extrait à partir d'échantillons de sang et génotypé (Illumina iSelect 60K). Après contrôle qualité, un animal a été exclu et 44 281 SNP ont passé nos critères de qualité : génotype du marqueur disponible pour plus de 95 % des individus, fréquence de l'allèle mineur supérieure à 5 %, proportion allélique respectant l'équilibre d'Hardy-Weinberg (FDR > 0,2).

1.3. Analyses du transcriptome

Le profil d'expression des gènes dans le sang pour chaque animal a été étudié par des puces à ADN 8X60K Agilent Technologies. Le bruit de fond a été corrigé et les niveaux d'expression ont été normalisés entre les puces par la méthode quantile (package R LIMMA, Smyth *et al.*, 2007). L'annotation de la puce à ADN utilisée a été mise à jour en alignant les séquences des sondes sur la séquence de référence du génome porcin (v10.2) avec le logiciel TopHat (v2.0.14).

Les 44 326 sondes alignées à un unique locus ont été ré-annotées en combinant les données des bases Ensembl et NCBI permettant l'assignation de 87 % d'entre elles à un gène porcine. Ainsi, 13 123 gènes ont été inclus dans l'analyse, avec en moyenne trois sondes par gène.

1.4. Analyses eGWAS

Nous avons adapté et parallélisé les fonctions du package R GenABEL (Aulchenko *et al.*, 2007), dont celle implémentant l'algorithme « Family-based Score Test for Association » (Chen et Abecasis, 2007). Ce test d'association est basé sur un modèle polygénique mixte, prenant en compte la structure familiale génomique de la population, le sexe et le lot. Les valeurs de p (« p -values ») obtenues ont été corrigées (taux global de faux positifs (FDR) à 5 %). Par une approche de régression linéaire, nous avons sélectionné pour chaque eQTL un SNP représentatif, évitant ainsi de rapporter plusieurs SNP d'une région de 5Mb ayant le même effet sur la variation d'expression d'une même sonde (Fairfax *et al.*, 2014).

2. RESULTATS

Le principe d'une analyse eGWAS repose sur celui d'une analyse GWAS ou « Genome Wide Association Study ». Il s'agit de tester une association statistique, par une approche de modèle linéaire, entre les allèles de SNP couvrant l'ensemble du génome et la variation de traits phénotypiques. Pour une analyse eGWAS, les caractères phénotypiques considérés sont les variations d'expression des gènes (ou sondes). Un eQTL est un locus génomique associé à la variation d'expression d'un gène. Dans notre analyse, un eQTL est représenté par un SNP grâce à une méthode de régression linéaire.

2.1. Mise en évidence d'associations en *cis* et en *trans*

Nous avons mis en évidence 4 591 associations significatives entre un SNP et une sonde, dont 62 % sont en *cis* (eQTL à moins d'1Mb de la sonde). La variation d'expression d'une sonde peut être associée à un ou plusieurs eQTL, jusqu'à six eQTL différents dans notre analyse. La majorité des sondes est associée à un (71 %) ou deux (25 %) eQTL (Tableau 1).

Tableau 1 – Nombre d'eQTL par sonde

eQTL	1	2	3	4	5	6
Sondes	2423	839	141	14	1	1

Parmi les 3 419 sondes dont la variation d'expression est associée à au moins un eQTL, 70 % ont une association en *cis*. Parmi ces sondes, 89 % ont pu être assignés à un gène porcine, soit l'équivalent de 2 554 gènes.

2.2. Répartition des associations sur le génome

Nous avons mis en évidence 3 195 eQTL. Un même eQTL est associé à une ou plusieurs sondes (jusqu'à 51). Les chromosomes 7 et 12 sont enrichis en eQTL. Les chromosomes 10 et 7 sont enrichis en gènes associés à un eQTL.

2.3. Enrichissement de la région du complexe majeur d'histocompatibilité (CMH) en eQTL ?

Les associations identifiées sur le chromosome 7 se concentrent dans la région du complexe SLA (« Swine Leucocyte Antigen »). 86 % des eQTL associés à la région SLA sont sur le chromosome 7, et 93 % des gènes associés à un eQTL de la région SLA sont sur le chromosome 7. Le complexe SLA est organisé en trois sous-régions. Les sous-régions de classe I et III sont sur le bras court et la sous-région de classe II est sur le bras long du chromosome 7. Parmi les 121 gènes annotés dans le complexe SLA, 65 sont associés à un eQTL, dont six gènes d'histocompatibilité de classe I et sept de classe II. Il est à noter que quelques eQTL sont associés à des gènes présents dans différentes régions du CMH. Par exemple, un SNP situé en 5' du gène NOTCH4, localisé dans la sous-région des gènes de classe III est associé à la variation d'expression de trois gènes : APOM localisé dans la sous-région de classe III et PSMB8 et SLA-DRB1 situés dans la sous-région de classe II.

CONCLUSION

Nos résultats, basés sur l'analyse de 243 individus, ont permis de mettre en évidence de nombreuses associations entre des polymorphismes et le profil d'expression de gènes dans le sang d'animaux sains. Le sang est un tissu intéressant pour le phénotypage immun puisqu'il permet des prélèvements peu invasifs et répétables, et que les cellules du sang sont des actrices majeures dans la réponse immunitaire. Les régions génomiques ainsi mises en évidence sont autant de régions candidates pour l'analyse fine de l'architecture génétique des variations de réponse immunitaire. Nous poursuivons actuellement ces analyses en y ajoutant les informations de phénotypage immunitaire des animaux afin de progresser dans l'identification de biomarqueurs sanguins et de marqueurs génétiques exploitables pour une sélection sur des critères immunitaires.

REFERENCES BIBLIOGRAPHIQUES

- Mach N., Gao Y., Lemonnier G., Lecardonnell J., Oswald I.P., Estellé J. and Rogel-Gaillard C., 2013. The peripheral blood transcriptome reflects variations in immunity traits in swine: towards the identification of biomarkers. *BMC Genomics*, 14, 894.
- Flori L., Gao Y., Laloë D., Lemonnier G., Leplat J-J., Teillaud A., Cossalter A-M., Laffitte J., Pinton P., De Vaureix C., Bouffaud M., Mercat M-J., Lefèvre F., Oswald I.P., Bidanel J-P., Rogel-Gaillard C. 2011. Immunity Traits in Pigs: Substantial Genetic Variation and Limited Covariation. *PLoS One*, 6(7), e22717.
- Y. Aulchenko, S. Ripke, A. Isaacs, C. van Duijn. GenABEL: an R library for genome-wide association analysis. 2007. *Bioinformatics*, 23, 1294-1296.
- Chen W.M., Abecasis G.R.. Family-Based Association Tests for Genomewide Association Scans. 2007. *Am. J. Hum. Genet.*, 81, 913–926.
- Fairfax B.P., Humburg P., Makino S., Naranbhai V., Wong D., Lau E., Jostins L., Plant K., Andrews R., McGee C., Knight J.C., 2014. Innate Immune Activity Conditions the Effect of Regulatory Variants upon Monocyte Gene Expression. *Science*, 343, 1246949.