

Thèse: Precision phenotyping by modelling for genetic improvement of disease resistance and resilience, with focus on mastitis in cattle - PrePheRR

Résumé de la thèse

Cette thèse propose de développer, dans une optique de sélection génétique, de nouveaux caractères de santé reflétant les capacités génériques de résistance et de résilience des individus face à une infection. Ce développement sera réalisé à l'aide d'un modèle dynamique sur la base de données zootechniques haut-débit. Tout l'enjeu de ce projet consiste à intégrer la nature dynamique et continue de l'état de santé des individus en prenant en compte la répétition des événements cliniques. Résolument génétique, cette thèse apportera de nouvelles perspectives à la sélection pour favoriser la transition vers un élevage plus durable et agroécologique.

Candidature sous forme de CV et de lettre de motivation à envoyer à :

Pauline.martin@inrae.fr ; Nicolas.friggens@inrae.fr ; Rafael.Munoz-Tamayo@inrae.fr

Présentation INRAE - Environnement de travail

L'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE) est un établissement public de recherche rassemblant une communauté de travail de 12 000 personnes, avec 272 unités de recherche, de service et expérimentales, implantées dans 18 centres sur toute la France. INRAE se positionne parmi les tout premiers leaders mondiaux en sciences agricoles et alimentaires, en sciences du végétal et de l'animal. Ses recherches visent à construire des solutions pour des agricultures multi-performantes, une alimentation de qualité et une gestion durable des ressources et des écosystèmes.

Cette offre de thèse sera réalisée au sein de l'unité de Génétique Animale et Biologie intégrative (GABI - <https://gabi.jouy.hub.inrae.fr/>), dans l'équipe GBOS : Génétique pour la durabilité de l'élevage bovin, sur le site INRAE de Jouy-en-Josas (78). Elle sera effectuée en très proche collaboration avec les unités de Modélisation systémique appliquée aux ruminants (MoSAR - <https://mosar.versailles-saclay.hub.inrae.fr/>) et de Physiologie, Environnement et Génétique pour l'Animal et les Systèmes d'Élevage (Pegase - <https://pegase.rennes.hub.inrae.fr/>).

Enjeux & missions

Lorsqu'il s'agit d'enregistrer des données de santé pour les animaux d'élevage, celles-ci sont généralement considérées de manière binaire : l'animal est sain ou malade au regard d'une affection donnée à un moment donné. Cette manière de phénotyper l'état de santé ne laisse place à aucune subtilité et on peut se demander si une description plus fine de la variabilité des réponses des animaux aurait un intérêt pour la sélection génétique de la sensibilité aux maladies. Ces dernières années, le déploiement de capteurs et de technologies haut-débit dans un nombre croissant d'élevages permet d'accéder à des séries chronologiques de nombreuses mesures automatisées au niveau de l'animal (Delaval OCC, Herd Navigator, O-CMT, etc.) (Sørensen et al., 2016; Deng et al., 2020). Ces données ouvrent la possibilité de suivre l'évolution dans le temps d'indicateurs de santé. Une meilleure intégration de la dimension dynamique de l'état de santé dans une mesure quantitative constituerait une réelle avancée en matière de phénotypage, mais représente également un défi méthodologique quant à l'extraction d'indicateurs biologiquement pertinents à partir de mesures longitudinales.

Quelques études ont déjà commencé à explorer cette question (par exemple, Detilleux et al., 2006) et, récemment, un modèle représentant le degré d'infection comme un caractère continu au cours d'une perturbation de santé et intégrant une composante génétique propre à l'animal a été proposé par les chercheurs encadrant cette thèse. Le modèle décrit l'évolution dans le temps du degré d'infection (DOI) à l'aide de trois coefficients qui rendent compte des variations individuelles entre les animaux en termes de capacité à éliminer l'infection, c'est-à-dire de résistance. De plus, en intégrant l'évolution dans le temps des performances (par exemple, de la production laitière), le modèle décrit également l'effet de l'infection sur les performances, permettant en ce sens d'estimer des variations individuelles en matière de résilience face à l'infection.

L'approche de modélisation proposée repose sur des principes biologiques similaires à ceux utilisés par Detilleux et ses collaborateurs, mais au lieu de les appliquer dans un contexte épidémiologique (avec une modélisation de la transmission de la maladie au sein du troupeau), elle est conçue pour permettre l'extraction des paramètres génétiques de la réponse à la maladie. À notre connaissance, il s'agit du premier modèle de ce type destiné à être utilisé dans un contexte génétique. Dans le cadre d'une validation de principe, le modèle a été évalué à l'aide de 15 séries temporelles individuelles du nombre de cellules somatiques dans le lait associées à des épisodes de mammite, et les résultats sont prometteurs. Après cette première étape, il reste toutefois à adapter le modèle à une utilisation sur des jeux de données de grande taille et à l'évaluer ses performances face à des contextes plus variés de système d'élevage ou d'origine de données. Par ailleurs, la structure du modèle est conçue pour se prêter à l'estimation de paramètres animaux par des outils statistiques de modélisation dynamique. Cette extension fournirait le cadre statistique approprié pour traiter l'incertitude de mesure, un aspect qui n'a pas encore été abordé. Un autre aspect clé à évaluer concerne la question des infections répétées au cours d'une même lactation. Doit-on les considérer comme liées mais génétiquement différentes, ou comme des mesures répétées du même caractère sous-jacent ?

Au travers de cette thèse, vous réalisez un travail à l'interface entre la génétique et la modélisation afin de répondre aux questions suivantes :

- Est-ce qu'avoir des phénotypes issus de modèles mécanistes et ciblant directement la résistance ou la résilience apporte une finesse intéressante et utilisable en pratique pour la sélection de la santé par rapport aux caractères actuels ?
- Est-ce que des différences génétiques de résistance et résilience sont observées dans le cas d'infections répétées ou en fonction des pathogènes ou des conditions de phénotypage ?

Pour ce faire, vous pourrez vous appuyer sur les pseudo-phénotypes issus d'un modèle dont la base est décrite dans l'article *Phenotyping health status on a continuous scale using a degree of infection approach: a case study using mastitis* (Martin et al., en cours de révision). Vous serez amené à faire évoluer ce modèle pour le rendre mieux adapté aux données à grande échelle, et devrez aussi développer de nouveaux modules afin d'intégrer la répétition d'événements vécus par les animaux. Enfin, vous utiliserez des méthodes classiques de génétique quantitative (estimation de paramètres génétiques, analyses GWAS et post-GWAS) pour tester la pertinence de vos pseudo-phénotypes pour un usage en sélection. En fonction de l'avancée de la thèse, une évaluation pilote des caractères retenus pourra également être réalisée.

Différents types de données seront mis à votre disposition :

- des données issues de l'unité expérimentale du Pin (<https://uep.isc.inrae.fr/>) où il existe plusieurs milliers de lactations avec des analyses bi-hebdomadaires des cellules somatiques, dont environ 200 lactations où les pathogènes à l'origine des mammites sont identifiés. Ces données continuent de s'incrémenter dans le cadre de l'expérimentation GLOBAL actuellement en cours, l'analyse des cellules ayant maintenant lieu 3x/semaine et l'identification des pathogènes étant maintenant systématique. Cette nouvelle expérimentation apporte également une large connaissance sur les animaux et notamment sur de nombreux critères sanguins liés à l'immunité.
- des données à grande échelle issues de fermes commerciales équipées de robot de traite et apportées par la profession française. Des données internationales pourraient également être partagées dans le cadre de cette thèse.

Formation & compétences recherchées

Un Master 2 ou un diplôme équivalent (Ingénieur Bac+5) est indispensable. Une formation en zootechnie, en génétique ou en modélisation est vivement recommandée.

Très motivé par la science, vous faites preuve d'autonomie, de curiosité, de rigueur et d'un réel sens d'initiative. Une aisance en anglais est indispensable, la connaissance du français est optionnelle.

Références bibliographiques

Deng, Z., Hogeveen, H., Lam, T.J.G.M., van der Tol, R., Koop, G., 2020. Performance of Online Somatic Cell Count Estimation in Automatic Milking Systems. *Frontiers in Veterinary Science* 7. doi:10.3389/fvets.2020.00221

Detilleux, J., Vangroenweghe, F., Burvenich, C., 2006. Mathematical model of the acute inflammatory response to *Escherichia coli* in intramammary challenge. *Journal of Dairy Science* 89, 3455–3465. doi:10.3168/jds.S0022-0302(06)72383-9

Martin P., Foucras G., Muñoz-Tamayo R., Friggens N.C.. Phenotyping health status on a continuous scale using a degree of infection approach: a case study using mastitis, *Animal*, under review.

Martin P., Foucras G., Muñoz-Tamayo R., Friggens N.C. (2024). Exploring modelling approaches to address the dynamic nature of animal health. Presented at: 75. Annual Meeting of the European Federation of Animal Science (EAAP), Florence, Italy (2024-09-01 - 2024-09-04).

Sørensen, L.P., Bjerring, M., Løvendahl, P., 2016. Monitoring individual cow udder health in automated milking systems using online somatic cell counts. *Journal of Dairy Science* 99, 608–620. doi:10.3168/jds.2014-8823

West, M., Harrison, J., 1997. Bayesian Forecasting and Dynamic Models, Springer. ed, Springer Series in Statistics. New York, USA.