

## Science de l'information géographique et génétique des populations

# Une nouvelle méthode pour identifier les régions du génome soumises à la sélection naturelle

*Comment gagner un prix prestigieux dans une école comme la nôtre dont l'ambition est d'entrer dans le club des prix Nobel en faisant une thèse sur les moutons, les ours et les grenouilles ? Stéphane Joost a réussi cet exploit en recevant le mois dernier à San Francisco le prix Garrison de l'Association of American Geographers (AAG) pour sa thèse de doctorat intitulée «The Geographic Dimension of Genetic Diversity: a GIScience Contribution for the Conservation of Animal Genetic Resources».*

William L. Garrison, désormais professeur émérite à l'Université de Californie à Berkeley, est un géographe âgé aujourd'hui de 83 ans. Il est considéré comme le «père» de ce que les géographes ont appelé la Révolution Quantitative initiée au milieu des années 50. Garrison rompa avec la géographie traditionnelle en introduisant une nouvelle approche basée sur la mesure, sur le recours à davantage de variables physiques et surtout sur un plus ample usage de la statistique et de l'informatique (alors à ses premiers balbutiements). Beaucoup de ses étudiants sont devenus des chercheurs et professeurs connus, notamment Duane Marble, Richard Morill, Ronald Boyce, John Nystuen, Arthur Getis, Waldo Tobler, tous présents lors de la remise du prix le 19 avril dernier à San Francisco. Les apports de Garrison ont profondément modelé, dans le monde anglo-saxon, les cursus académiques en géographie en donnant une plus large place à la physique et aux méthodes statistiques.

Le Dr Joost a pu réaliser sa thèse dans le cadre d'un projet européen (*Sustainable conservation of animal genetic resources in marginal rural areas: integrating molecular genetics, socio-economic and geostatistical approaches*) que le soussigné avait obtenu en tant requérant pour la participation du LaSIG. Ce projet, aux ambitions très larges, reposait sur la caractérisation du génome de chèvres et moutons sur une aire géographique allant du Moyen Orient jusqu'en Angleterre et au Portugal. L'un de nos objectifs était de tenter de détecter des régions du génome soumises à la sélection naturelle et d'identifier les facteurs environnementaux potentiellement responsables, avec l'aide de méthodes et d'outils développés par la science de l'information géographique.



De gauche à droite:  
Douglas Richardson,  
Duane F. Marble, Stéphane  
Joost, William Garrison

La méthode mise au point consiste à rechercher dans le génome une éventuelle signature des variables environnementales. La coïncidence spatiale permet de mettre en relation les caractéristiques génétiques des organismes étudiés avec des valeurs de paramètres environnementaux grâce à des coordonnées géographiques communes. Un simple calcul de corrélation n'est pas possible puisque nous nous trouvons dans une situation de présence-absence. Pour contourner l'obstacle, la statistique a développé le concept de régression logistique. La relation de «tout ou rien» est remplacée par une fonction de forme sigmoïdale de sorte qu'il devient possible d'associer une probabilité de présence-absence de marqueurs génétiques<sup>1</sup> selon le contexte environnemental. La régression logistique fournit la mesure du niveau d'association entre la fréquence des marqueurs génétiques et les valeurs des paramètres environnementaux. On calcule un niveau de confiance des modèles constitués par toutes les paires

possibles (marqueur génétique x paramètre environnemental) afin d'identifier les marqueurs impliqués dans les modèles les plus significatifs et qui jouent probablement un rôle dans le processus d'adaptation.

Les analyses ont été effectuées sur différentes races de chèvres et de moutons, sur la grenouille rousse et l'ours brun de Scandinavie, avec deux types de marqueurs génétiques, et confrontées à des données produites par la Climatic Research Unit (CRU) de l'Université East Anglia de Norwich, et le modèle numérique d'altitude SRTM30 de la NASA. Les résultats les plus intéressants ont été obtenus chez le mouton pour lequel deux marqueurs significatifs ont été identifiés, associés respectivement au taux d'humidité et à de grandes quantités de précipitations: le premier est lié à un gène impliqué dans un processus de résistance à un parasite, et le second au processus de sécrétion de cire de laine.

La méthode proposée produit ainsi des informations utiles pour le repérage des gènes et la compré-

hension de leur(s) fonction(s). Elle fournit notamment des pistes de recherche grâce à l'identification de possibles sources de pression environnementale. Parallèlement, l'étude des associations entre des régions du génome et des variables environnementales permet de procurer une capacité prédictive afin d'établir des cartes d'habitat potentiel pour des espèces végétales ou animales. Ce raisonnement a été appliqué avec succès à des populations d'Ours Bruns en Scandinavie (ill. 2).

Dans son rapport de jury de thèse, le Dr Pierre Taberlet, biologiste et directeur de recherche au CNRS, a relevé que «les résultats obtenus sont impressionnants, et ouvrent des perspectives complètement nouvelles dans le domaine de la génétique des populations, non seulement au niveau de l'analyse des données, mais aussi au niveau de la conception de nouvelles études.»

L'interdisciplinarité, souvent promue dans les discours mais si peu souvent soutenue dans les financements, fournit, par les résultats de cette recherche, la preuve qu'elle est efficace et source de découvertes scientifiques.

Régis Caloz,  
ENAC, INTER, Laboratoire de SIG

<sup>1</sup> Séquences d'ADN particulières qui balisent le génome.