

ANR-MICRO-Obes

Profil de recrutement d'un IR pour un contrat à durée déterminée de 18 mois (à partir de Septembre 2008) à l'INRA / Unité MIA de Jouy-en-Josas (Yvelines, France)

Profil général :

Statisticien (Bac+5 au moins) avec une ouverture d'esprit et une curiosité scientifique marquées pour s'impliquer dans la problématique scientifique bioinformatique inhérente au sujet d'étude.

1. Le contexte scientifique

L'importance de la microflore intestinale (plusieurs centaines d'espèces bactériennes) dans la dégradation des aliments et la protection de la paroi intestinale en font une sorte d'organe à part entière. Les relations de cette flore avec la santé humaine sont mal comprises et ce n'est qu'avec les progrès récents du séquençage à haut débit qu'on peut espérer mieux connaître les métagénomes de systèmes aussi complexes (d'où l'émergence de plusieurs projets internationaux). Dans le cadre du projet franco-chinois Micro-Obes (Human Intestinal Microbiome in Obesity and Nutritional Transition, voir site <https://www.jouy.inra.fr/uepsd>), on cherche à mieux cerner les relations entre ce métagénome et des caractéristiques physiologiques, ici l'obésité. Pour cela, différents prélèvements sont effectués sur des populations dites saines, des populations obèses avant régime, des populations obèses après régime. Ces prélèvements sont préparés et analysés via la technique SOLEXA générant des millions de fragments de courte taille du métagénome. Ces fragments sont alors mis en relation avec des séquences génomiques et métagénomiques correspondant à des gènes bactériens séquencés et annotés. A chaque prélèvement, on associe *in fine* un profil métagénomique indexé en abscisse par un ensemble de gènes et en ordonnée par le nombre de fois où un court fragment correspondant à un gène a été repéré dans le prélèvement. L'objectif du travail proposé ici consiste alors à comparer ces populations de profils métagénomiques afin de mettre en évidence d'éventuelles relations entre un caractère phénotypique (obésité) ou la réponse à un régime (transition nutritionnelle) et des ensembles de gènes présents dans la microflore intestinale. D'un point de vue statistique, la situation est assez atypique avec un très grand nombre de variables explicatives potentielles, peu d'observations mais des corrélations connues ou inconnues entre ces variables.

2. Les compétences statistiques nécessaires

Après avoir mené une réflexion sur la nature des données qui seront générées, notamment par nos collègues bioinformaticiens de l'Unité MIG (Mathématiques Informatique et Genome) de Jouy, il nous est apparu qu'une certaine spécialisation était indispensable pour contribuer efficacement à la résolution du problème. Trois types de méthodes statistiques pourraient composer cette spécialisation:

- des méthodes classiques d'analyses unidimensionnelles et multidimensionnelles de données, typiquement des **méthodes de classification** (particulièrement sur les

variables que sont les gènes, les individus statistiques étant des profils de comptage ou des profils descriptifs biologiques),

- des méthodes d'analyses de données avancées, notamment celles basées sur **la régression PLS, et plus particulièrement la régression logistique PLS avec ou sans sélection de variables, l'analyse discriminante PLS avec ou sans sélection de variables,**
- des techniques de **validation croisée.**

3. Intérêt du travail pour le recruté

L'étude à mener présente pour le recruté un intérêt majeur que l'on peut décliner selon les trois volets suivants :

- En premier lieu, l'environnement scientifique sera pluridisciplinaire: des cliniciens, des biologistes, des bioinformaticiens, des informaticiens, et des statisticiens seront impliqués. Cet environnement sera donc une source très riche de connaissances de haut niveau que le recruté pourra facilement valoriser dans la suite de sa carrière.
- En second lieu, il faut insister sur la nature particulièrement nouvelle des données métagénomiques en cours de construction, et ceci au niveau international. Cette étude participera à la visibilité internationale de l'INRA dans l'étude des métagénomes microbiens associés à l'homme.
- Enfin, pour un statisticien, l'enjeu est particulièrement motivant si l'on souligne le côté innovant de l'adaptation, de la programmation et la mise en œuvre de méthodes statistiques avancées dans un cadre de plus en plus présent d'analyses statistiques de masses de données avec beaucoup de covariables.

4. Résultats attendus

Du point de vue du fond on attend l'élaboration d'une méthodologie pour l'analyse comparative de métagénomes ainsi que de procédures informatiques (en Matlab ou en R par exemple) la mettant en œuvre. Des résultats mathématiques justifieront la pertinence de cette méthodologie par rapport aux différentes sources d'incertitude. On souhaite également que les procédures informatiques puissent permettre un usage ultérieur dans d'autres études métagénomiques similaires. Du point de vue de la forme, cette étude devra conduire à la rédaction en bon français d'un rapport conséquent explicitant dans le détail tous les choix et l'adaptation des méthodes statistiques utilisées.

5. Contexte et aptitudes

Le recruté travaillera au sein de l'unité MIA (Mathématiques et Informatique Appliquées) du Centre de Recherche de l'INRA de Jouy-en-Josas. Il sera encadré par MM. J.-P. Gauchi et A. Trubuil, chercheurs dans cette Unité. Le travail à réaliser se fera donc au sein d'une équipe et dans le cadre d'un projet pluridisciplinaire dirigé par un biologiste. Cela suppose de la part du recruté une implication dans la problématique biologique et une capacité à instaurer un dialogue constructif avec les autres partenaires du projet. Le recruté devra montrer une capacité à s'appropriier les connaissances et outils statistiques utiles à l'avancement du projet et les adapter si nécessaire.

6. Salaire et conditions matérielles

Le salaire proposé est de 1700 euros nets mensuels, et la durée du contrat est de 18 mois; le début souhaité est septembre 2008. Le poste est localisé au Centre de Recherches de l'INRA, à Jouy-en-Josas, France.

Les courriels des personnes à contacter simultanément pour une information complémentaire et/ou postuler (CV et lettre de motivation) sont les suivants :

Jean-Pierre.Gauchi@jouy.inra.fr

Alain.Trubuil@jouy.inra.fr