

- Profil de recrutement ASC -

INTITULE DU SUJET DE THESE :

Extraction de connaissances à partir de puces à ADN à haute densité : application à la compréhension des mécanismes d'expression et de régulation des gènes par les nutriments

OBJECTIFS SCIENTIFIQUES :

Le projet de recherche proposé vise à identifier, dans le muscle squelettique et le tissu adipeux, les réseaux de régulation de gènes impliqués dans l'adaptation métabolique à des changements de l'environnement nutritionnel et à vérifier si des anomalies dans ces processus régulateurs sont impliquées dans le développement de pathologies nutritionnelles telles que l'obésité et le diabète de type 2 chez l'homme.

Cette problématique biologique sera traitée par l'analyse globale du transcriptome grâce notamment à la technologie des puces à ADN de haute densité. Les données obtenues seront analysées avec les techniques de " data mining " (extraction de connaissances dans de grands volumes de données par des techniques non supervisées comme la découverte de règles d'association, ou le " clustering ") afin de faire émerger des hypothèses biologiques intéressantes. Ce travail repose sur un partenariat étroit entre une équipe de biologistes et une équipe d'informaticiens spécialisés dans ces techniques de " data mining " non supervisées

COMPETENCES RECHERCHEES (type d'école doctorale; discipline de formation initiale) :

Le/la candidat(e) doit être informaticien(ne) de formation, connaître les bases théoriques et technologiques de l'ECD) et avoir une réelle motivation pour rejoindre un projet multi-disciplinaire Biologie-Informatique. Une implication forte dans la coopération avec les biologistes est en effet indispensable (des connaissances préalables dans le domaine - génome, transcriptome, etc - ne sont pas indispensables). Enfin, il/elle doit être motivé(e) pour travailler dans un environnement scientifique exigeant

RESUME DU SUJET DE THESE :

Le projet repose sur le développement de méthodes d'apprentissage pour extraire automatiquement des informations pertinentes de données générées par les puces à ADN. L'approche de "data mining" proposée est la découverte de règles d'association puis le traitement des règles fréquentes et valides dans un processus interactif entre informaticien et biologiste. Les réseaux de gènes étudiés sont ceux de l'adaptation métabolique à des changements nutritionnels en vue de l'étude de leur implication dans le développement de l'obésité et du diabète de type 2.

REFERENCES	UNITE D'ACCUEIL 1	UNITE D'ACCUEIL 2
Intitulé : Directeur : Adresse :	INSERM U449 "mécanisme moléculaire du diabète" Jean-Paul RIOU U449-Faculté de Médecine Laënnec. 8 rue Guillaume PARADIN 69372 LYON Cedex 08	Ingénierie des Systèmes d'Information (LISI) INSA Bâtiment Blaise Pascal 69621 Villeurbanne cedex - France
EQUIPE D'ACCUEIL (objectifs, partenaires, moyens à disposition):	Nous générons les données biologiques par l'utilisation de puces à ADN de haute densité qui permettent l'analyse simultanée d'un très grand nombre de gènes (en théorie, tout le génome humain). L'équipe de recherche comprend 4 chercheurs permanents INRA et INSERM qui interagiront avec le candidat pour valider et orienter les hypothèses biologiques.	Les ressources du LISI pourront être utilisées (sur site) par le doctorant et, notamment, les prototypes de recherche qui y sont actuellement développés (extracteurs de règles d'associations, extracteurs de motifs séquentiels, etc) ou les systèmes commerciaux qu'il possède. Avec environ 40 permanents et 30 doctorants, c'est un laboratoire raisonnablement équipé pour conduire des recherches en informatique

RESPONSABLE A CONTACTER avant tout acte de candidature:

Nom : Sophie Rome INSERM-U449 / Jean-François Boulicault LISI-INSA

Téléphone : 04 78 77 86 29/04 72 43 89 05

Fax : 04 78 77 87 62/04 72 43 87 13

e-mail : srome@univ-lyon1.fr et jfboulic@lisisun1.insa-lyon.fr