

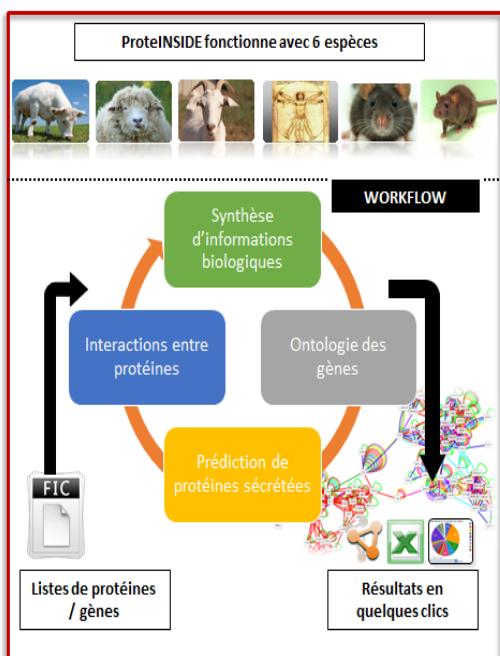
UMR1213 Herbivores

Equipe Animal Muscle Viande (Amuvi)

ProteINSIDE : un service web pour la fouille de données de génomique chez les ruminants

La génomique permet de mieux comprendre le vivant. Toutefois, à partir d'expérimentations de génomique, il n'est pas toujours aisé d'extraire ou de générer une information biologique pertinente. Nous avons créé ProteINSIDE, un service web pour la fouille des données de génomique chez le bovin, l'ovin et le caprin. L'outil utilise les connaissances disponibles chez les ruminants et complète les informations manquantes en utilisant, par orthologie, les annotations disponibles chez des espèces mieux annotées (souris, rat, homme).

Les expérimentations de transcriptomique et de protéomique produisent de grandes quantités de données. Une manière robuste et efficace de fouiller de telles données est d'utiliser des outils intégratifs, tels que DAVID (Huang et al. 2009), BioMyn (Ramirez et al. 2012), ToppGene (Chen et al., 2009), Expander (Ulitsky et al. 2010), PathwayStudio (Nikitin et al. 2003) ou Ingenuity Pathway Analysis® (www.ingenuity.com; Redwood City, CA, USA). De tels outils ne sont malheureusement pas conçus pour l'analyse des données issues des ruminants et actuellement, l'absence d'outils automatisés dédiés aux ruminants entrave l'exploration des données de génomique.



Nous avons conçu l'outil en ligne ProteINSIDE qui synthétise l'information biologique stockée dans les bases de données publiques (NCBI et UniProt) ou fournie par les annotations fonctionnelles issues de l'ontologie des gènes. Il prédit aussi les protéines qui sont sécrétées (secretome des tissus) et qui interviennent dans la signalisation entre les cellules ou tissus. Il lie les protéines selon leurs interactions moléculaires afin d'identifier et de visualiser celles qui contribuent à un même processus biologique et celles qui sont centrales à un processus biologique. ProteINSIDE est le seul outil, qui en une seule requête, permet une telle fouille et prédition de données non seulement chez le bovin, l'ovin, le caprin, mais aussi chez l'homme, le rat et la souris. L'interface web permet la création d'analyses et la visualisation des résultats en temps réel ou leurs téléchargement dans de nombreux formats (images, pdf, tableaux, données brutes, fasta...). La base de données de ProteINSIDE recueille et stocke l'information biologique nécessaire pour un fonctionnement efficace. Elle est mise à jour mensuellement afin de recueillir les dernières connaissances disponibles dans les bases de données publiques. ProteINSIDE a été testé avec des jeux de données de 1000 protéines par espèce et a été comparé avec succès à DAVID, BioMyn et AgBase, conçus pour la recherche d'information et l'annotation, ainsi qu'à PrediSi et Phobius qui prédisent les protéines sécrétées. ProteINSIDE est disponible gratuitement en utilisant un simple navigateur internet à l'adresse www.proteinside.org. Un exemple de résultats obtenus avec ProteINSIDE est fourni sur la page d'accueil du site <http://www.proteinside.org/>.

ProteINSIDE sera utilisé pour la fouille des données issues des nombreuses études de génomique conduites à travers le monde et rassemblées dans la base de données Fat&MuscleDB (Tournayre et al., 2015) pour identifier des gènes et des protéines à l'origine d'une variabilité de la composition corporelle des ruminants et donc de l'efficacité de production des carcasses et des viandes. L'intégration des données publiées favorisera la découverte de gènes ou des protéines clés dans le développement des tissus musculaires et adipeux, car la variabilité de leur expression ou de leur abondance sera systématiquement liée à une variabilité de la masse des tissus. Ainsi, l'exploitation de données existantes favorisera la compréhension des mécanismes contrôlant la composition corporelle tout en minimisant les redondances dans les efforts de recherche.



Valorisation :

Kaspric N., Picard B., Reichstadt M., Tournayre J., Bonnet M. 2015. ProteINSIDE to easily investigate proteomics data from ruminants: Application to mine proteome of adipose and muscle tissues in bovine foetuses. Plos One , 10 (5), e0128086

Kaspric N., Reichstadt M., Picard B., Tournayre J., Bonnet M. 2015. Protein Function Easily Investigated by Genomics Data Mining Using the ProteINSIDE Online Tool. Genomics and Computational Biology, (1)1, e16

Contact : Muriel Bonnet, muriel.bonnet@clermont.inra.fr, UMR1213 Herbivores, F-63122 Saint-Genès-Champanelle, France.

