

Prédiction de l'expression des gènes à partir de données épigénomiques chez le *Fusarium* à l'aide de méthodes de machine learning et de Deep Learning

lundi 18 mars 2024 11:00 (20 minutes)

Le *Fusarium* est un bioagresseur du blé, connu pour ses capacités d'adaptation rapide aux changements environnementaux, dues à des variations épigénétiques. À des intervalles de temps courts, les variations épigénétiques peuvent permettre des changements phénotypiques dus à la modification de l'expression génique, sans changement de la séquence génomique. Ainsi, les génomes et épigénomes interagissent avec l'environnement et contribuent à l'adaptation des organismes et à leurs capacités de résilience. Néanmoins, d'autres variations génétiques entrent en jeu dans les phénomènes d'adaptation, c'est pourquoi ici nous utilisons des données sur le *Fusarium* qui est un organisme se reproduisant de manière clonale et où l'adaptation par variation génétique est moindre. L'étude des variations d'expression est donc envisagée ici par l'étude des mécanismes épigénétiques, notamment par la variation de l'accessibilité de la chromatine qui joue un rôle majeur dans la transcription d'un gène et donc dans son expression. On obtient ces informations grâce à un séquençage MAINeSeq et l'expression des gènes est obtenue par séquençage RNAseq. Une première approche statistique vise à l'apprentissage de la structure épigénomique le long du gène afin de traduire son expression à l'aide d'une grande variété de méthodes de machine learning (régression linéaire, random forest, et deep learning), ce qui a conduit à des résultats peu concluants. Partant de l'observation que le signal d'expression dans l'amorce de la partie codante du gène est à la fois bien corrélé au signal épigénomique et à l'expression complète du gène, nous proposons une nouvelle méthode d'apprentissage où les signaux de motifs épigénétiques sont utilisés pour prédire le signal d'expression, cette fois le long de la séquence nucléotidique, dans une démarche de prédiction dite function-to-function.

Thématiques

Auteurs principaux: M. CAUSEUR, David (IRMAR, Institut Agro); M. LE TRIONNAIRE, Gaël (IGEPP); Mme HOUÉ-BIGOT, Magalie (Institut Agro Rennes Angers); Mme BRUGUET, Mathilde (Institut Agro, MycSA); Mme PONTS, Nadia (MycSA); M. TAVENARD, Romain (Université de Rennes 2); Mme LEROY, Romane (Institut Agro); Mme KIREZI, Santa (Institut Agro)

Orateur: Mme BRUGUET, Mathilde (Institut Agro, MycSA)

Classification de Session: Statistiques appliquées : Biologie