Notes RDV - David

Question – Niveau moyen d’un signal = niveau moyen de l’expression génomique ou transcriptomique ?

Biocondutor -

Séquençage – MAINE seq – Pas trop de méthode en matière d’analyse de données. A quel point on peut prédire le transcriptome par rapport à l’épigénomique ? Prédiction mais pas d’obnubilation sur performance parfaite. A terme l’idée c’est d’expliquer la transcription à partir de l’épigénétique. Déplétion avant l’ATG, favorise l’expression du gène.

**Hypothèse :** Travailler à l’échelle de 14000 gènes c’est un peu illusoire – il existerait certaines règles de classification pour des groupes de gènes. Idées :

* Groupe de fonction ?
* Territoire chromosomique
* …

On pensait que si classif sur signal épigénomique 🡺 permettrais d’associer à un « type » de signal transcriptomqiue. Sauf que ça marche pas.

Méthode de clustering – qui prend en entrée 2 types de tableaux.

Expression RNA seq = comptage de reads – sur l’ensemble du gène. Choix ici de garder le signal en entier car on se dit 🡪 concordance spatial parfaite entres les 2 signaux – épigénomique et transcriptomique.

Contestable aussi – Interpolation – distance à ATG = distance relative

Soit on considère que ce qui se passe est lié à la position de la nucléotique par rapport à ATG ou alors c’est relatif (perte d’information ?)

Interpolation – nouvelle méthode avec données cumulés.

Validation croisé R2 = 0.4, 0.5