Epigenetic marks and the diversity of human transcripts

Visualisation of Roadmap Epigenomics data

Guillaume Devailly1 and Anagha Joshi2

1 GenPhySE, UMRxxxx CNRS, *Université de Toulouse, INRA, INPT, ENVT, Castanet Tolosan, France*

guillaume.devailly@inra.fr

2 University of Bergen, Department of Clinical Science, *Haukeland universitetssykehus, Laboratoriebygget, Norway*

1. Résumé

Les marques épigénétiques présentes le long du génome sont cartographiées avec une précision et un débit en constante augmentation. Divers consortiums ont générés de grandes quantités de données épigénomiques, couvrant différentes marques et différents types cellulaires et tissus humains. C’est le cas du projet Roadmap epigenomics qui propose des données pour une trentaines de marques (WGBS, H2A.Z, acétylations et méthylations d’histones, Dnase1) dans une trentaines de types cellulaires, le tout processé de manière uniforme.

Nous utilisons ce riche jeu de données pour évaluer systématiquement les liens entre marques épigénétiques et transcription, au travers diverses analyses. Notamment :

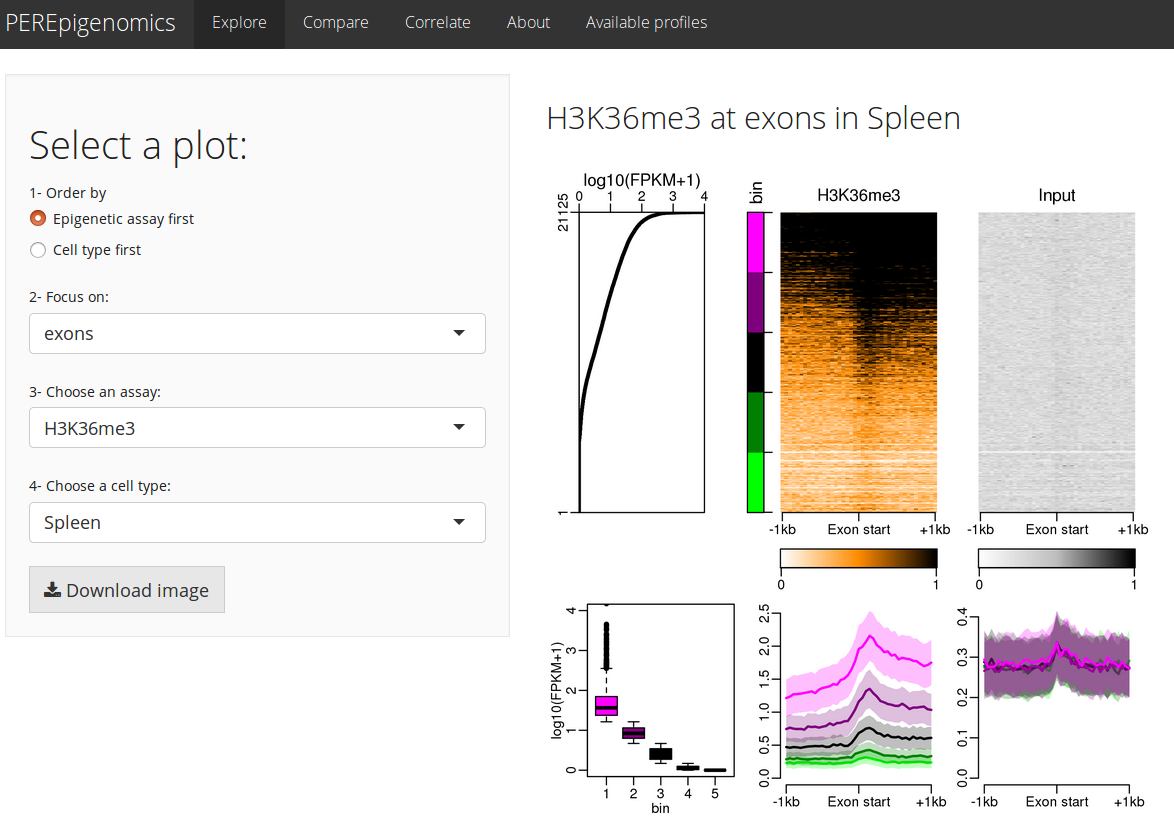
- corrélations entre marque épigénétique et niveau de la transcription autour des sites d’initiations de la transcriptions (TSS), des sites de terminaisons (TTS), et des exons intermédiaires.

- étude différentiée par types de gènes : codant pour des protéines, lincRNA, miRNA, pseudogènes, pour mettre en évidence des régulations spécifiques.

- pour la méthylation, analyses parallèle du ratio de méthylation (ou pourcentage de CpG méthylé), de la densité de méthylation (ou nombre de CpG méthylé) et de la densité en sites CpG, afin de démêler les liens entre îlots CpG et méthylation.

- corrélations entre ratio d’inclusion des exons et marque épigénomique sur ces exons pour confirmer ou infirmer les liens entre épissage alternatif et marque épigénétiques.

Ces analyses sont mises en valeur au travers d’un portail interactif d’exploration des résultats (en cours de construction) : [www.perepigenomics.roslin.ed.ac.uk](http://www.perepigenomics.roslin.ed.ac.uk/). Nous espérons confirmer ou infirmer à grande échelle des corrélations déjà connues, mais aussi en identifier de nouvelles. Dans un second temps nous espérons exploiter des expériences de perturbations de marques épigénétiques afin d’identifier des relations causales. Enfin nous proposons de mettre en place ce types de visualisations et d’analyses pour les données générées par le consortium FAANG.

1. 
2. L’application web [PEREpigenomics](http://www.perepigenomics.roslin.ed.ac.uk/) permet d’explorer les données du consortium Roadmap Epigenomics de manière interactive.

Acknowledgements

*The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme (FP7/2007-2013) under REA grant agreement No PCOFUND-GA-2012-600181.*