

Projet M2 bio-info -- Laurent Modolo/Pascal Bernard

laurent.modolo@ens-lyon.fr/pascal.bernard@ens-lyon.fr

Implémentation de pipelines d'analyse de données pan-génomiques quantitatives permettant de mesurer l'impact des nucléosomes sur l'assemblage des chromosomes

Le stage sera réalisé au LBMC dans l'équipe de Pascal BERNARD sous la co-supervision de Laurent Modolo (Ingenieur Bio-Info) et Pascal BERNARD (Biologiste DR2 CNRS)

Introduction

Une limitation importante de la plupart des techniques d'analyse pan-génomiques (de type Chip-Seq) actuelles fondées sur le séquençage massif de nouvelle génération est qu'elles génèrent des données qualitatives, mais non quantitatives (Bonhoure et al., 2014; Chen et al., 2015; Hu et al., 2015). En effet, sans référence interne, il est impossible de quantifier des variations globales à l'échelle du génome. Cette limitation a récemment été contournée par l'introduction de spike-in controls dans les échantillons (Hu et al., 2015). La calibration des signaux NGS par rapport à un référentiel interne neutre nécessite l'implémentation des pipelines d'analyses spécifiques. L'objectif du stage sera d'implémenter un pipeline d'analyse ChIP-seq pour une analyse quantitative par ChIP-seq calibré, et de l'appliquer à l'étude des nucléosomes et du complexe condensine organisateur du génome.

Condensine est un complexe protéique ATPase en forme d'anneau qui conduit la condensation mitotique des chromosomes depuis les eucaryotes unicellulaires (levures) jusque chez l'Homme. Comment condensine réorganise les fibres de chromatine en chromosomes mitotiques demeure mal compris. Les données actuelles suggèrent que condensine opère au travers d'une activité motrice ATP-dépendante de type ADN-translocase. Condensine encercle une molécule d'ADN double brin, puis forme, agrandi et organise des boucles d'ADN d'environ 90-100 kbp, dont l'empilement hélicoïdal génère un chromosome compact en forme de bâtonnet. Ce modèle est porté par de nombreuses études réalisées *in vitro* ou *in vivo*. En revanche, l'on ignore dans quelle mesure l'activité ADN-translocase de condensine, clairement démontrée *in vitro*, sur de l'ADN nu, s'opère *in vivo*, quand l'ADN est empaqueté sous forme de chromatine. Notamment, l'impact des nucléosomes et des protéines associées à la chromatine sur la fixation de condensine à l'ADN et sur son activité d'extrusion de boucles demeurent très mal compris. La recherche menée dans l'équipe d'accueil (LBMC - P. Bernard) a pour objectif de tenter de répondre à ces questions.

Résultats préliminaires

En utilisant la levure *S. pombe* comme organisme modèle, nous avons montré (1) que condensine se lie à l'ADN au niveau de zones déplétées en nucléosomes, et (2) que l'éviction des nucléosomes facilite la fixation de condensine à l'ADN génomique (Toselli et al. EMBO J 2016). Nous avons ensuite observé que réduire la densité en nucléosomes à l'échelle du génome corrige les défauts de migration des chromosomes en anaphase causés par une déficience partielle de condensine, confortant l'hypothèse d'une entrave de condensine par les nucléosomes *in vivo*. En combinant des approches de protéomique et de génétique, nous avons découvert que le complexe FACT est présent à proximité de condensine sur les chromosomes mitotiques, et participe au fonctionnement de condensine en mitose. De façon importante, des expériences pilotes de ChIP-qPCR suggèrent que FACT n'a pas d'incidence sur la fixation de condensine à l'ADN.

FACT est une chaperonne d'histone assurant l'assemblage et désassemblage dynamique des nucléosomes indispensable à la fluidité de la fibre de chromatine. Collectivement, nos résultats indiquent (1) que les nucléosomes et FACT impactent condensine *in vivo*, (2) que FACT est proche de condensine sur la chromatine en mitose, et (3) que FACT opère en mitose une fonction qui pourrait freiner l'activité de condensine plutôt que son association à l'ADN.

Nous formulons l'hypothèse que les nucléosomes gênent l'extrusion de boucles d'ADN par condensine, et qu'un équilibre dynamique entre l'assemblage et le désassemblage des nucléosomes, établi en partie par FACT à proximité de condensine faciliterait son activité ADN-translocase en mitose.

La caractérisation des mécanismes qui lient condensine aux nucléosomes et à FACT permettra de montrer que des interactions fonctionnelles unissent l'activité de condensine à la dynamique de la chromatine *in vivo*, et de mieux comprendre comment ces interactions sous-tendent l'assemblage des chromosomes en mitose.

Projet

L'objectif du stage est d'identifier, implémenter et utiliser les outils bio-informatiques appropriés à l'analyse des données de génomiques fonctionnelles générées dans le cadre de cette étude.

1. Développer un pipeline d'analyse ChIP-seq quantitative en utilisant le gestionnaire de pipeline nextflow. L'impact de FACT sur l'association de condensine à l'ADN sera mesuré par des expériences de ChIP-seq calibrées (quantitatives) contre condensine chez *S. pombe*. De nombreux travaux ont démontré que la ChIP-seq ne permet pas de mesurer des variations qui s'exercent sur l'entièreté du génome (Bonhoure et al., 2014; Chen et al., 2015; Hu et al., 2015). Sur la base d'une méthode récemment publiée (Hu et al., 2015), nous avons développé un protocole de ChIP-Seq calibré en utilisant de la chromatine de *S. cerevisiae* comme référentiel de calibration (spike-in control). Le stagiaire aura pour objectif d'implémenter les méthodes d'analyses de ChIP-seq pour les rendre quantitatives, comme décrit dans plusieurs publications récentes (Hu et al., 2015; Jeronimo et al., 2019).

2. Développer un pipeline nextflow pour l'analyse quantitative des nucléosomes par MNase-X-ChIP-seq calibrée. Un aspect central du projet consistera à déterminer le rôle joué par FACT dans l'occupation de l'ADN génomique par les nucléosomes en mitose. L'occupation de l'ADN par les nucléosomes sera déterminée par des expériences de MNase-X-ChIP-seq calibrées. Cette méthode offre une résolution optimale dès lors qu'il s'agit de cartographier les nucléosomes. La chromatine est digérée à la MNase en fragments mono-nucléosomaux, ces derniers sont immunoprécipités et l'ADN co-immunoprécipité identifié par séquençage massif (Skene and Henikoff, 2015). Nous avons développé un protocole de MNase-X-ChIP-seq calibré, adapté à *S. pombe* dans lequel la chromatine de *S. cerevisiae* sert de « spike-in control ». L'analyse de données pan génomiques de positionnement des nucléosomes de type MNase-seq est fréquemment réalisée avec le pipeline DANPOS (Chen et al., 2013). Le stagiaire aura pour objectif d'implémenter DANPOS afin de permettre une analyse quantitative.

Enfin, l'équipe est aussi entrain de générer des données de type Hi-C dans le cadre de ce sujet, lesquelles pourront être analysées par le candidat dans le cadre du M2 si le temps le permet, où d'une thèse si cette option apparaît pertinente.

Références Bio-info

- MNase-seq data processing and detection of dynamic nucleosomes and NDR analyses

1. Chen, K. et al. DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res* 23, 341-351, doi:10.1101/gr.142067.112 (2013).
2. Promoter nucleosome dynamics regulated by signalling through the CTD code. Materne et al. *Elife*. 2015 Jun 22;4:e09008. doi: 10.7554/eLife.09008.

- MNase-X-ChIPseq analyses

1. A simple method for generating high-resolution maps of genome-wide protein binding. Skene PJ, Henikoff S. *Elife*. 2015 Jun 16;4:e09225. doi: 10.7554/eLife.09225.

Biblio

Bonhoure, N., Bounova, G., Bernasconi, D., Praz, V., Lammers, F., Canella, D., Willis, I.M., Herr, W., Hernandez, N., Delorenzi, M., et al. (2014). Quantifying ChIP-seq data: a spiking method providing an internal reference for sample-to-sample normalization. *Genome Res*. 24, 1157–1168.

Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., and Li, W. (2013). DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res*. 23, 341–351.

Chen, K., Hu, Z., Xia, Z., Zhao, D., Li, W., and Tyler, J.K. (2015). The Overlooked Fact: Fundamental Need for Spike-In Control for Virtually All Genome-Wide Analyses. *Mol. Cell. Biol*. 36, 662–667.

Hu, B., Petela, N., Kurze, A., Chan, K.-L., Chapard, C., and Nasmyth, K. (2015). Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq. *Nucleic Acids Res*. 43, e132.

Jeronimo, C., Poitras, C., and Robert, F. (2019). Histone Recycling by FACT and Spt6 during Transcription Prevents the Scrambling of Histone Modifications. *Cell Rep*. 28, 1206-1218.e8.

Skene, P.J., and Henikoff, S. (2015). A simple method for generating high-resolution maps of genome-wide protein binding. *Elife* 4, e09225.