

Développement d'une procédure de test pour l'analyse quantitative de CHIP-seq calibrée, et analyse intégrative d'un jeu de données hétérogènes pour étudier l'impact des nucléosomes et de la transcription sur l'organisation 3D des chromosomes

Mots clés

Organisation 3D du génome, CHIP-seq calibrée

Nom, adresse de l'Unité d'accueil

Laboratoire de Biologie et Modélisation de la Cellule, CNRS/ENS-Lyon. 46 Allée d'Italie - 69364 Lyon cedex 07.

Nom, adresse de l'Equipe d'accueil / Nom du responsable d'équipe

Equipe Chromatin Dynamics in Mitotic Chromosome Assembly.

Responsable : Pascal BERNARD.

<http://www.ens-lyon.fr/LBMC/equipes/architecture-et-dynamique-fonctionnelle-des-chromosomes>. Nous cherchons à comprendre les mécanismes qui sous-tendent la réorganisation spatiale de la chromatine interphasique en chromosomes mitotiques. Nous étudions notamment l'impact de la dynamique de la chromatine sur la fixation de condensine à l'ADN et sur son activité de formation de boucles mitotiques. Nous abordons ces questions par des approches de génétique moléculaire et de génomique (CHIP-seq, MNase-seq et Hi-C) en utilisant la levure *Schizosaccharomyces pombe* comme modèle biologique(1-3).

Noms, tels et e-mails des encadrants de stage

Encadrant (partie biologie): Pascal BERNARD. 04 72 72 81 97.

pascal.bernard@ens-lyon.fr

Encadrant (partie Bio-info) : Laurent MODOLO 04 72 72 80 48. laurent.modolo@ens-lyon.fr

SUJET

Dans les trois domaines du vivant, l'organisation 3D du génome est intrinsèquement liée à ses fonctions de stockage, d'expression et de propagation de l'information génétique. La condensation mitotique des chromosomes, conduite par le complexe protéique condensine, illustre cette relation structure/fonction. Durant la mitose, condensine replie les fibres de chromatine en une série de boucles de ~ 100 kbp, formant une structure 3D (le chromosome mitotique) façonnée pour la transmission fidèle du génome en anaphase (4). Comment condensine s'associe à l'ADN en mitose et le replie sous forme de boucles, in vivo, demeure mal compris. Notre équipe a mis en évidence l'existence d'interactions fonctionnelles positives et négatives entre condensine, les nucléosomes et la transcription(2, 3, 5). Aujourd'hui, nous cherchons à comprendre le rôle joué, à l'échelle d'un génome entier, par la dynamique de la chromatine dans l'activité de condensine et la formation des chromosomes mitotiques. **Pour cela, nous avons implémenté la technique de CHIP-seq calibrée, qui permet une mesure semi quantitative de l'association pan-génomique d'un facteur d'intérêt, et développé un pipeline d'analyse dédié permettant de calculer base par base l'occupation normalisée dudit facteur (1).** Pour questionner les liens entre condensine et la dynamique

des nucléosomes et la transcription, nous avons généré des jeux de données de CHIP-seq calibrée portant sur condensine, les histones H2B et H3 et la RNA Pol II. Le ratio H2B/H3 calculé base par base nous renseigne sur l'intégrité structurale des nucléosomes. Les objectifs du stage sont de deux ordres :

(1) Refactoriser notre pipeline d'analyse

- Refactoriser le code du pipeline nextflow existant en suivant les recommandations du consortium nf-core (DSL2)
- Réimplémenter les étapes de normalisation du pipeline existant dans un programme dédié (Python ou Rust) base par base à partir des fichiers bam ou bigwig bruts, en favorisant une utilisation parcimonieuse de la RAM.
- Ajouter une étape de Peak-calling adaptée aux données de CHIP-Seq calibré. La détection de marques le long du génome suivit du test de leur enrichissement entraîne un problème d'inférence post-sélection qui interfère avec le contrôle du risque de la procédure de test. Korthauer **et al.** ont récemment développé une procédure de test prenant en compte ce problème pour des données de méthylation (6). Il s'agira d'adapter la procédure dmrseq (6) aux données de Chip-seq calibrée.
- Ajouter une fonctionnalité permettant l'obtention d'un rapport au format texte référençant les pics identifiés, leur classement et leurs coordonnées en fonction des paramètres (cutoffs) utilisés.
- Ajouter une fonctionnalité pour la production de graphiques métagènes à partir d'une liste de pics.

La principale difficulté de cet objectif sera de s'approprier les développements proposés par la procédure dmrseq (6) et de l'adapter aux particularités des mesures de CHIP-Seq calibrée.

(2) Analyses des données

L'objectif est de rechercher si des corrélations existent entre les niveaux d'occupation de l'ADN par condensine, par H2B/H3 et/ou par la RNA Pol II. Pour cela il s'agira :

- De réanalyser le premier jeu de données de CHIP-seq calibrée pour s'assurer que le nouveau pipeline fourni les mêmes résultats.
- D'analyser les données de CHIP-seq calibrée condensine, H2B/H3 et RNA Pol II.
- Classer les gènes par niveau d'occupation (quantiles).
- Étudier les covariations entre le niveau d'occupation de condensine et de celle des histones et de la RNA Pol II.

La principale difficulté de cet objectif sera d'intégrer des types de données hétérogènes et d'analyser leurs covariations spatiales le long du génome de *Schizosaccharomyces pombe*.

Publications d'intérêt

1. L. Colin, C. Reyes, J. Berthezene, L. Maestroni, L. Modolo, E. Toselli, N. Chanard, S. Schaak, O. Cuvier, Y. Gachet, S. Coulon, P. Bernard, S. Tournier,

- Condensin positioning at telomeres by Taz1 promotes telomere disjunction in anaphase (2022), *BioRxiv*, doi:10.1101/2022.03.18.484892.
2. E. Toselli-Mollereau, X. Robellet, L. Fauque, S. Lemaire, C. Schiklenk, C. Klein, C. Hocquet, P. Legros, L. N'Guyen, L. Mouillard, E. Chautard, D. Auboeuf, C. H. Haering, P. Bernard, Nucleosome eviction in mitosis assists condensin loading and chromosome condensation. *EMBO J.* **35**, 1565–1581 (2016).
 3. C. Hocquet, X. Robellet, L. Modolo, X.-M. Sun, C. Burny, S. Cuylen-Haering, E. Toselli, S. Clauder-Münster, L. Steinmetz, C. H. Haering, S. Marguerat, P. Bernard, Condensin controls cellular RNA levels through the accurate segregation of chromosomes instead of directly regulating transcription. *eLife.* **7**, e38517 (2018).
 4. I. F. Davidson, J.-M. Peters, Genome folding through loop extrusion by SMC complexes. *Nat. Rev. Mol. Cell Biol.* **22**, 445–464 (2021).
 5. J. Rivosecchi, D. Jost, L. Vachez, F. D. Gautier, P. Bernard, V. Vanoosthuyse, RNA polymerase backtracking results in the accumulation of fission yeast condensin at active genes. *Life Sci. Alliance.* **4**, e202101046 (2021).
 6. K. Korthauer, S. Chakraborty, Y. Benjamini, R. A. Irizarry, Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. *Biostatistics.* **20**, 367–383 (2019).