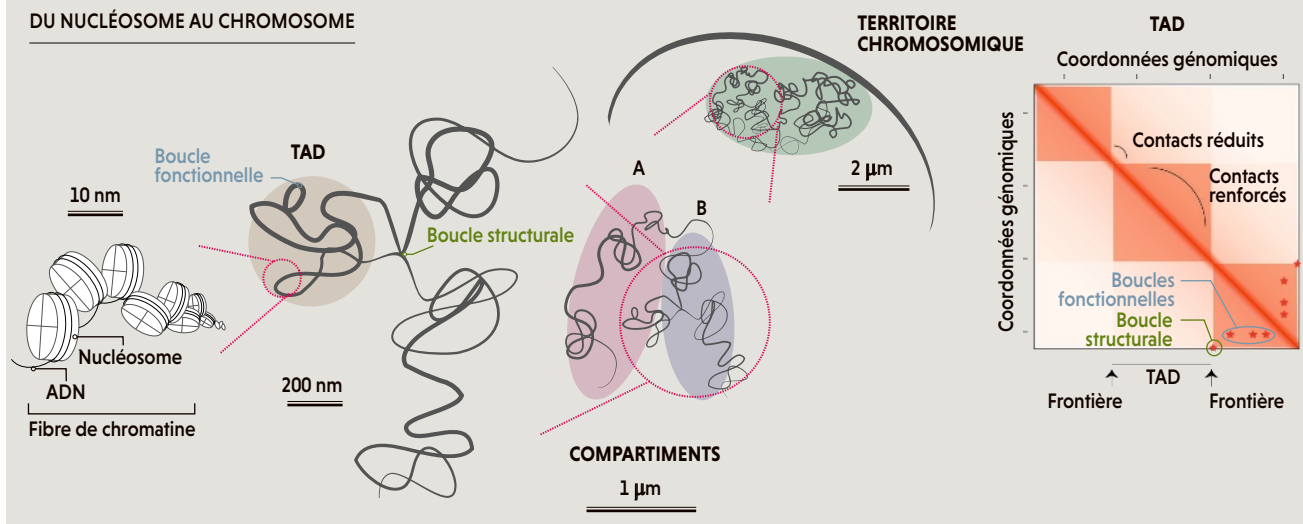


DU NUCLÉOSOME AU CHROMOSOME



UNE ORGANISATION À TOUTES LES ÉCHELLES

L'organisation spatiale du génome est loin d'être aléatoire au sein des noyaux de nos cellules. On connaissait le premier échelon de cette organisation, le nucléosome, dont la structure cristallographique avait été révélée en 1997. Le nucléosome permet d'enrouler la molécule d'ADN et représente donc un premier niveau de compaction de l'ADN pour former la fibre de chromatine (voir la figure ci-dessus à gauche). À l'autre bout de l'échelle, on savait, grâce à des approches cytologiques et de microscopie photonique, que les chromosomes s'organisent en territoires chromosomiques individuels, occupant des positions préférentielles dans les noyaux suivant leur activité génique, leur densité en gènes ou encore leur taille. En revanche, au milieu de l'échelle, les choses paraissaient beaucoup plus floues. On avait notamment observé des repliements à large échelle entre des régions génomiques éloignées, ainsi que des regroupements de gènes même lorsqu'ils étaient situés sur

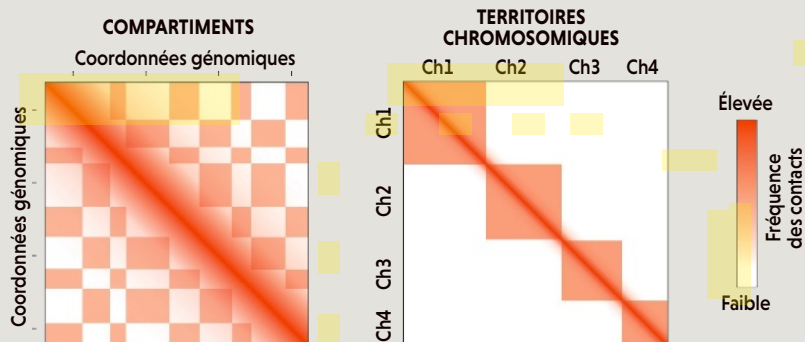
des chromosomes différents. Bien sûr, on soupçonnait aussi l'existence de repliements de la fibre de chromatine et, par là même, de rapprochements physiques entre les promoteurs des gènes et leurs séquences stimulatrices ou « enhancers » – les boucles qui font l'objet de cet article. Les enhancers sont de véritables boutons d'activation, parfois situés à plusieurs dizaines de milliers de paires de bases des gènes qu'ils régulent. Des avancées majeures dans la compréhension de l'organisation spatiale du génome, de la centaine de kilobases (100 000 paires de bases) jusqu'à la mégabase (1 million de paires de bases), ont vu récemment le jour grâce à l'application de nouvelles techniques de biologie moléculaire, telles que la méthode Hi-C.

En 2009, on a vu apparaître les premières cartes de Hi-C, qui saisissent l'ensemble des contacts entre régions génomiques dans l'espace tridimensionnel du noyau. Plus précisément, ces représentations cartographient la fréquence de chacun des contacts possibles sous la forme de points plus ou moins intenses (voir ci-dessus à droite). Les premières cartes avaient une résolution relativement faible, mais on y distinguait déjà des organisations préférentielles de larges régions génomiques et, notamment, l'organisation en compartiments actif et réprimé, appelés respectivement compartiments A et B. En 2012, l'application de la méthode Hi-C ou de dérivées avec une résolution accrue a permis à trois équipes – la nôtre, celle d'Edith Heard, à l'institut Curie, et celle de Bing Ren, à l'institut Ludwig pour la recherche sur le cancer à La Jolla,

aux États-Unis – de découvrir de nouveaux domaines d'organisation de la chromatine, qui sont désormais appelés TAD (pour *topologically associating domains*).

Le long des chromosomes, les TAD représentent des zones du génome caractérisées par des interactions internes bien plus fortes que les interactions de TAD adjacents. Entre les TAD, on trouve de véritables frontières, qui conduisent à la partition du génome en une série de TAD. L'obtention de cartes de Hi-C à très haute résolution a permis de mettre en évidence de nombreuses boucles d'ADN, décrites dans cet article. Certaines, dites « structurales », séparent les TAD les uns des autres. D'autres, situées à l'intérieur des TAD, joueraient divers rôles fonctionnels. Certaines rapprocheraient des promoteurs et des enhancers, d'autres les deux extrémités d'un gène (ce qui renforcerait la direction selon laquelle le gène est transcrit), d'autres encore se formeraient entre des sites reconnus par des protéines impliquées dans la répression de nombreux gènes – les protéines du groupe Polycomb –, renforçant leur action. En d'autres termes, les TAD permettent de mieux confiner certaines régions du génome, ce qui facilite des repliements de la fibre chromatinienne et la régulation de l'expression des gènes.

La découverte de ces différents échelons de l'organisation des chromosomes a ouvert de nouvelles voies d'investigation sur la compréhension du fonctionnement du génome, ainsi que sur celle de maladies dont l'origine restait mystérieuse. Notamment, des réarrangements chromosomiques, du type



L'organisation de la chromatine aux différentes échelles (à gauche) apparaît lorsque l'on cartographie la fréquence des contacts entre les différentes régions du génome (à droite).

inversion ou délétion, surviennent au niveau des régions frontières des TAD. Cela a pour effet d'entraîner des défauts d'isolement entre les gènes. Ainsi, des séquences stimulatrices localisées dans le TAD voisin peuvent communiquer et activer de manière erronée des gènes dans des tissus où ils sont normalement éteints. C'est par exemple le cas de gènes du développement directement impliqués dans des malformations des membres (syndrome F, polydactylie...), des défauts neurologiques ou des dysplasies. De même, des réarrangements chromosomiques peuvent entraîner l'activation de gènes nommés « proto-oncogènes », normalement éteints dans les cellules et susceptibles de conduire à la transformation des cellules et à l'apparition de tumeurs. Il est donc primordial de prendre en compte l'organisation 3D du génome pour expliquer les nombreux défauts cellulaires liés à la dérégulation des gènes.

De façon plus générale, grâce à tous ces travaux, on comprend bien mieux l'organisation 3D des génomes. Néanmoins, des questions primordiales demeurent : quels mécanismes précis définissent la taille et les bordures des TAD ? Quel est leur rôle dans la régulation des fonctions du génome ? Enfin, sera-t-il possible de prédire, dans un type cellulaire donné, la régulation de son expression génique en connaissant la séquence de l'ADN et son organisation tridimensionnelle ? Les recherches dans ce domaine révéleront sans doute de nouvelles surprises...

**FRÉDÉRIC BANTIGNIES
ET GIACOMO CAVALLI**

directeurs de recherche au CNRS,
Institut de génétique humaine
(CNRS, université de Montpellier)

➤ toujours l'un vers l'autre et vers la boucle. Cette «règle de convergence» expliquait comment les histones évitaient les boucles: il leur suffisait de déterminer dans quelle direction le motif de liaison à CTCF pointait. En revanche, le fait que l'orientation des motifs ne soit pas aléatoire défiait notre entendement.

Pour mettre les choses en perspective, grossissons à nouveau le génome un million de fois. Les motifs mesurent à présent chacun 5 millimètres de long et sont séparés par 1 kilomètre de nouille génomique. Et pourtant, comme guidés par une boussole magique, ceux situés aux extrémités d'une boucle pointent toujours l'un vers l'autre. Comme tout bon tour de magie, la règle de convergence semblait impossible. De plus, elle contredisait l'idée que l'on avait alors de la façon dont les boucles se forment.

À l'époque, presque tout le monde – y compris nous-mêmes – pensait qu'elles apparaissaient par diffusion. Dans cette hypothèse, une protéine nécessaire à la formation d'une boucle se lie à une extrémité d'un segment d'ADN. Une autre fait de même à l'autre extrémité. Comme à son habitude, l'ADN se tortille. Si finalement sa danse réunit les deux protéines, elles forment un lien physique, créant ainsi une

Presque tout le monde pensait que les boucles apparaissaient par diffusion

boucle. Le problème est que la chaîne d'ADN a tellement de place pour s'agiter que, si le modèle de diffusion est correct, l'orientation relative des motifs de liaison à CTCF n'a plus d'importance. Or, pourtant, on assistait à une convergence. Et de fait, au cours de l'année, deux équipes, l'une dirigée par Suzana Hadjur, de l'University College de Londres, et l'autre par Yijun Ruan, du laboratoire Jackson, aux États-Unis, avaient confirmé la règle de convergence avec leurs propres données. La règle était là pour rester, et les boucles ne se formaient donc pas par diffusion. Comment apparaissaient-elles alors ? Et quels étaient les rôles de CTCF et de la cohésine ? Décontenancés, nous avons commencé à jouer avec les câbles de nos écouteurs. ➤