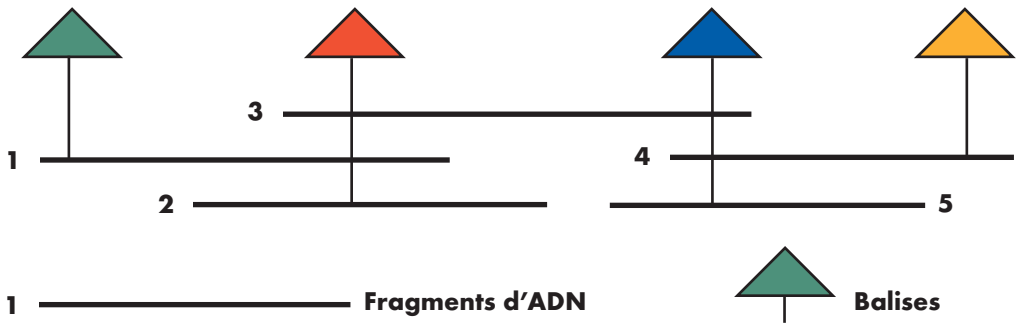
La cartographie

3



Les génomes des bactéries et des autres microorganismes comprennent en géréral moins d'une dizaine de millions de nucléotides. Il est possible de réassembler ces "petits" génomes directement à partir des données de séquence. Les génomes des animaux et des plantes, quant à eux, ont des tailles considérables, allant d'une centaine de millions à une centaine de milliards de nucléotides. Pour ces "grands" génomes, il est nécessaire de procéder en deux temps.

On commence par couper l'ADN en grands fragments – de 50 000 à 200 000 nucléotides, voire plus – dont on constitue une collection sous la forme d'une "banque" de clones bactériens (*voir Clonage*). Les grands fragments clonés de cette collection sont ensuite ordonnés (cartographiés) les uns par rapport aux autres au moyen de courtes séquences d'ADN qui servent de points de repère. Lorsque plusieurs fragments ont l'un de ces points de repère, ou balise, en commun, on en conclut qu'ils ont une partie d'ADN en commun. On dit que les fragments sont partiellement recouvrants ou chevauchants.



En analysant l'ensemble des fragments d'ADN pour leur contenu en balises, on peut reconstituer l'enchaînement des balises et des fragments d'ADN, tels qu'ils existent dans la molécule d'ADN de départ. Dans la pratique, on procède aussi à un ordonnancement indépendant des balises par d'autres méthodes de cartographie.

La reconstitution de la molécule d'ADN de départ sous la forme d'un ensemble de grands fragments chevauchants constitue la carte physique. C'est à partir de cette carte que sera choisi l'ensemble minimal de grands fragments assurant la couverture complète du génome à séquencer. Dans l'exemple ci-dessus, l'ensemble des fragments 1, 3 et 4 suffit à couvrir la portion de génome contenant les 4 balises verte, rouge, bleue et jaune.

Pour séquencer ces fragments, on redécoupera chacun d'eux en fragments plus petits *(voir la fiche* Séquençage*)*.