

## Chapitre 13

# Génome des cellules et des virus, transmission de l'information génétique

## Cours

### PLAN DU CHAPITRE

- 1 La connaissance des génomes
- 2 Le génome des bactéries
- 3 Le génome des eucaryotes
- 4 Les virus, diversité structurale et génomique
- 5 Cycle cellulaire et réplication du matériel génétique
- 6 Les divisions cellulaires : mitose et méiose

### ZOOM

- 1 Le séquençage par la méthode de Sanger
- 2 Les courbes de Cot
- 3 Le bactériophage lambda
- 4 Le virus de la mosaïque du tabac
- 5 Les coronavirus
- 6 Durée et contrôle des cycles cellulaires
- 7 Le principe de la PCR

### INTRODUCTION

L'ADN est le support de l'information génétique dans les cellules procaryotes et eucaryotes où il constitue les chromosomes ; l'ARN peut aussi être l'unique support de l'information chez certains virus. Les séquences portées par ces molécules sont de différents types dont certaines correspondent aux gènes. Chez les eucaryotes, cette information génétique est transmise d'une génération à l'autre par deux types de divisions, la mitose et la méiose.

- ➔ Comment sont organisés les supports de l'information génétique ?
- ➔ Comment sont-ils transmis d'une cellule à l'autre au cours du cycle cellulaire, d'une génération à l'autre au cours de la reproduction sexuée ?

## 1 La connaissance des génomes

Le **génome** correspond à l'ensemble des molécules d'ADN contenues dans une cellule. Il contient l'information génétique nécessaire au développement et à la réalisation des fonctions biologiques. Le génome est représenté par les chromosomes présents dans la cellule. Dans le cas de certains virus le génome est représenté par de l'ARN.

## 1.1 Localisation des génomes

### a. Chez les Eucaryotes

L'information génétique des eucaryotes est compartimentée : on distingue le **génom**e nucléaire du **génom**e extra-nucléaire présent dans la matrice des mitochondries et, pour les cellules chlorophylliennes, dans le stroma des chloroplastes. La localisation de l'information génétique dans des organites à double membrane favorise sa protection vis-à-vis d'agents mutagènes et donc sa stabilité.

Dans le noyau, le support de l'information génétique est sous forme de plusieurs chromosomes linéaires, chaque chromatide de chromosome étant une molécule d'ADN double brin.

Une cellule **haploïde** ne contient qu'un lot de chromosomes tous différents ; une cellule **diploïde** renferme deux lots de chromosomes homologues, morphologiquement identiques par paire. Certaines cellules, **polyploïdes**, renferment le lot haploïde en un plus grand nombre pair de copies.

Dans les mitochondries et les chloroplastes, le génome est porté par des molécules d'ADN double brin généralement circulaires, le plus souvent présentes en plusieurs exemplaires. Ce génome propre peut être transcrit et traduit sur place ce qui donne à ces organites leur caractère semi-autonome.

### b. Chez les bactéries

Le génome des bactéries est localisé dans le cytosol de la cellule. Il est formé d'un **chromosome** et éventuellement de **plasmides**.

L'unique chromosome bactérien est constitué d'une molécule circulaire d'ADN bicaténaire et se localise dans une région cellulaire appelée nucléoïde. Quelques rares bactéries présentent un chromosome linéaire (ex. : genre *Borrelia*). Quelques bactéries ont plusieurs chromosomes (ex. : genre *Vibrio*).

Les plasmides sont des molécules d'ADN circulaires de petite taille (quelques centaines à plusieurs milliers de bases) et à répllication autonome.

### c. Chez les virus

Les virus sont des entités de petite taille (entre 10 et 400 nm) composées de protéines et d'une molécule d'acide nucléique, constituant leur génome.

Les génomes viraux sont de nature très variée. Le matériel génétique peut être de l'ADN, selon les cas circulaire ou linéaire, sous forme bicaténaire ou monocaténaire. Chez d'autres virus, le génome est porté par des molécules d'ARN, ce qui constitue une situation unique en biologie. Un virus ne contient jamais les deux types d'acides nucléiques.

## 1.2 Le séquençage de l'ADN

Toutes les parties du génome, quelle que soit leur nature, codante ou non, sont constituées par des séquences de nucléotides, monomères des acides nucléiques.

La connaissance des génomes et plus largement celle de l'information génétique qu'ils portent, font appel aux techniques de biologie moléculaire. Celles-ci désignent l'étude des acides nucléiques (ARN et ADN). Elles font intervenir une panoplie d'approches reposant notamment sur l'hybridation moléculaire entre séquences, la multiplication des séquences par la réaction de polymérisation en chaîne (PCR) et la reconstitution de l'ordre des nucléotides ou séquençage des acides nucléiques.

Les méthodes de **séquençage**, apparues dans les années 1970, ont connu des évolutions spectaculaires permettant aujourd'hui de recueillir des données sur l'ensemble du génome en quelques heures et à moindre coût.

#### ZOOM 7

Le principe de la PCR

## ZOOM 1

Le séquençage par la méthode de Sanger

### a. La méthode de Sanger

Une méthode historique est la méthode de Sanger (1977) et qui a permis l'obtention des premières séquences. Elle est fondée sur la polymérisation d'un brin d'ADN complémentaire au brin que l'on souhaite séquençer.

### b. Les techniques de séquençage à haut débit (*Next generation sequencing* ou NGS)

Apparues en 2007, ces techniques permettent des séquençages beaucoup plus rapides et en plus grand nombre que le séquençage par la méthode Sanger. Différentes méthodes ont été proposées, avec cependant des points communs :

- coupures de l'ADN en fragments relativement courts (30 à 250 pb) aux extrémités desquels sont fixées de courtes séquences, les adaptateurs ;
- amplification par PCR des fragments, guidée par les adaptateurs ;
- séquençage en parallèle de millions de fragments puis dénaturation de l'ADN. Les fragments simple brin servent de matrice pour la synthèse de brins complémentaires.

Chaque nucléotide est apporté un à un dans le milieu réactionnel. Son incorporation éventuelle est susceptible de générer un signal spécifique, dont la nature diffère selon la technique utilisée. L'enregistrement en continu des signaux livre directement la séquence. Différentes technologies sont utilisées mais elles ne requièrent plus la séparation des fragments par électrophorèse. Ces techniques atteignent une vitesse de séquençage de plusieurs dizaines de Gb par jour.

### c. Le séquençage de troisième génération (TGS : *Third generation sequencing*)

La **technologie nanopore** (*Oxford Nanopore Technologies* (ONT)) est apparue en 2014. Elle n'utilise pas de polymérase pour la synthèse d'une copie de la molécule à séquençer mais fait passer des molécules d'ADN dans un nanopore enchâssé dans une membrane qui sépare deux compartiments contenant des solutions ioniques. Ce passage entraîne des perturbations d'un signal électrique qui est fonction du nucléotide traversant le pore. La séquence est déduite de la succession des perturbations.

Ces technologies sont dites aussi de **lectures longues** (ou *long read*), du fait de leur capacité à séquençer des molécules uniques dont la longueur peut atteindre des centaines de kb. Le facteur limitant devient la longueur des fragments d'ADN que l'on réussit préalablement à obtenir lors de l'extraction de cette molécule. Le problème devient donc de trouver des méthodes d'extraction qui laissent l'ADN le plus intact possible.

## 1.3 Construction des génomes et annotation

Les **techniques NGS** (*Next Generation Sequencing*) fournissent les séquences de très nombreux fragments d'ADN mais ceux-ci restent de faible longueur. Elles ne permettent donc pas l'accès direct à un génome sous forme de séquences continues mais n'en livrent que des fragments dispersés. Ceci reste une difficulté majeure pour la reconstitution des génomes qui requiert alors l'assemblage des séquences identifiées. Celui-ci repose sur des analyses bioinformatiques visant par exemple à reconnaître des séquences chevauchantes entre fragments pour construire des séquences plus longues. Certaines parties du génome, notamment les séquences répétées, peuvent par ailleurs ne pas avoir été séquençées ou être difficilement mises en relation avec d'autres. Les techniques de séquençage de troisième génération, de lectures longues (TGS : *Third generation sequencing*) sont développées pour pallier ces difficultés.

L'**annotation** d'un génome consiste à analyser la séquence nucléotidique qui constitue l'information brute pour en extraire l'information biologique, avec deux objectifs : d'abord localiser les gènes et les régions codantes, puis, une fois ces gènes localisés, prédire leur fonction biologique. Ces deux étapes reposent sur l'utilisation d'algorithmes ou de méthodes relevant de la bioinformatique.

La localisation des gènes se fonde sur des méthodes statistiques qui identifient les régions codantes en recherchant des motifs particuliers (séquences promotrices, terminatrices, sites d'épissage...). Pour prédire la fonction potentielle de ces gènes, on utilise des programmes de recherche d'homologie de séquence : on compare le produit potentiel du gène avec une protéine connue, pour identifier dans la séquence prédite des motifs caractéristiques de certaines classes de protéines.

La branche de la génétique qui étudie les génomes est la **génomique**. Elle se divise en **génomique structurale** qui s'intéresse à la structure et à l'organisation physique du génome, dressant un inventaire des gènes de l'organisme, et en **génomique fonctionnelle** dont l'objectif est d'attribuer des fonctions aux gènes et d'étudier leur régulation et leurs interactions avec les produits d'expression (ARN et protéines).

### 1.4 Taille des génomes : nombre de paires de bases et de gènes

*Saccharomyces cerevisiae* a été le premier génome eucaryote publié en 1997 et le premier séquençage du génome humain a été réalisé en 2004. En janvier 2020, plus de 19 000 génomes étaient séquençés et le nombre de génomes en cours d'analyse dépassait 327 000.

Le séquençage permet d'accéder à la taille du génome, appelée **valeur C**, généralement exprimée en nombre de paires de bases (pb). Chez les eucaryotes, les valeurs sont données pour un lot haploïde de chromosome. L'annotation permet par ailleurs d'estimer le nombre de gènes.

Tableau 13.1 Taille de quelques génomes.

Organisme	Type	Nombre de chromosomes par jeu haploïde et nombre de lots haploïdes	Taille du génome (Mb)	Nombre de gènes
<i>Escherichia coli</i>	Bactérie	1	4,6	4 300
<i>Saccharomyces cerevisiae</i>	Ascomycète	16	13	6 200
<i>Arabidopsis thaliana</i>	Angiosperme (Brassicacées)	5 × 2	120	25 500
Riz	Angiosperme (Poacée)	12 × 2	430	Environ 40 000
Blé	Angiosperme (Poacée)	7 × 6	17 000	
<i>Drosophile</i>	Insecte	4 × 2	180	13 400
<i>Mus musculus</i>	Mammifère	20 × 2	3 000	Environ 20 000
<i>Homo sapiens</i>	Mammifère	23 × 2	3 400	21 000

Les génomes eucaryotes les plus simples comme les levures sont plus grands que les génomes procaryotes.

Les génomes eucaryotes montrent cependant des tailles très variées, sans relation avec la position phylogénétique ou avec le degré de complexité : le génome de l'homme est par exemple dix fois moins grand que celui de certains amphibiens et comprend moins de gènes que celui d'*Arabidopsis*. Des variations très importantes s'observent également au sein d'un seul taxon (variation d'un facteur de l'ordre de 120 chez les amphibiens).

Cette absence de corrélation est traduite sous le terme de « **paradoxe de la valeur C** ». Ces valeurs suggèrent l'existence de parties plus ou moins importantes des génomes qui ne sont pas exprimées, les gènes n'occupant que 1 à 5 % de leur taille.

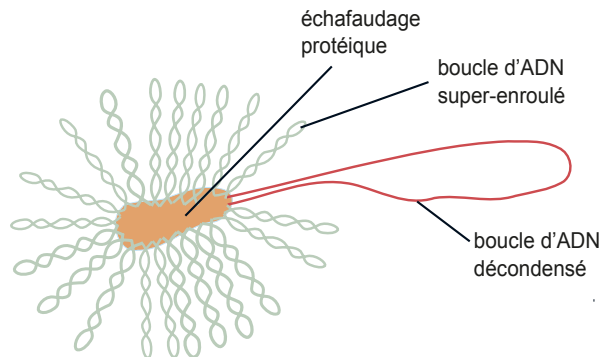
La taille des génomes viraux est très variable et s'étend de quelques centaines de nucléotides pour les plus petits jusqu'à 300 kb, les gènes restant toujours peu nombreux.

## 2 Le génome des bactéries

### 2.1 Chromosome bactérien et plasmides

#### a. Structure du chromosome bactérien

Le **nucléotide**, zone où se localise le chromosome bactérien, contient 40 à 50 boucles d'ADN enchâssées sur un noyau protéique. L'ADN est sous forme compactée grâce à des enzymes qui permettent son enroulement sur lui-même (surenroulement de l'ADN). L'ensemble est stabilisé par des protéines qui exercent des fonctions analogues à celles des histones chez les eucaryotes (figure 13.1).



**Figure 13.1** Organisation de l'ADN dans le chromosome bactérien.

(D'après T. Brown : Genomes - second edition, Wiley-Liss, 2003).

#### b. Contenu informatif du génome

L'ADN du chromosome bactérien porte les gènes indispensables au fonctionnement cellulaire. Leur nombre varie de quelques centaines à environ 6 000, pour un génome compris entre 1 000 et 5 000 kb.

Chaque gène n'est présent qu'à un seul exemplaire : le génome est haploïde.

Les **plasmides** sont des molécules d'ADN bicaténares de petite taille, non indispensables au fonctionnement normal de la cellule. Elles portent des gènes pouvant coder pour des protéines impliquées dans la résistance aux antibiotiques ou conférant la virulence et le pouvoir pathogène de certaines bactéries. Les plasmides peuvent se répliquer indépendamment du chromosome bactérien. Ils sont des vecteurs mobiles susceptibles parfois de réaliser des transferts d'information génétique d'une bactérie à une autre.

L'ADN bactérien, qu'il soit du chromosome ou d'un plasmide, est essentiellement codant (près de 90 % de l'ADN d'*Escherichia coli*). Le génome des bactéries est compact. Les séquences des gènes sont continues, sans être interrompues par des séquences non codantes (absence d'introns au sein des gènes).

Les séquences non codantes de l'ADN correspondent principalement à des séquences contrôlant l'activité des gènes.

## 2.2 Une organisation en opérons

De nombreux gènes bactériens (59 % chez *E. coli*) sont regroupés en **opérons**.

Ceux-ci correspondent à un ensemble de gènes accolés et placés sous le contrôle d'une même séquence régulatrice comprenant un promoteur et un opérateur. Les gènes accolés codent chacun pour une protéine spécifique, qui participe avec les autres à l'accomplissement d'une même fonction, toutes s'inscrivant dans une même voie métabolique.

**L'opéron lactose** (figure 13.2) associe ainsi trois gènes codant pour trois protéines impliquées dans l'utilisation du lactose ou de molécules voisines :

- *lacZ* : gène de la  $\beta$ -galactosidase qui hydrolyse le lactose en glucose et galactose ;
- *lacY* : gène de la lactose-perméase qui permet l'entrée du lactose dans la cellule ;
- *lacA* : gène de la thiogalactoside-transacétylase, qui participe au métabolisme de molécules voisines du lactose, les thiogalactosides.

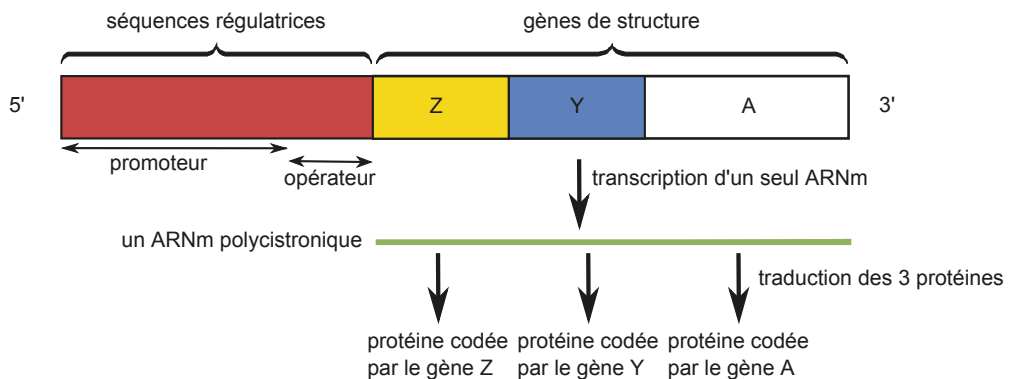


Figure 13.2 Organisation et fonctionnement d'un opéron (opéron lactose).

L'activation de l'opéron conduit à la transcription d'un seul ARNm couvrant les trois gènes : cet ARNm est dit **polycistronique**. Les protéines correspondant sont toutes synthétisées en même temps : l'opéron constitue ainsi une **unité coordonnée d'expression génétique**.

Le site promoteur est le site de fixation de l'ARN polymérase qui permet la transcription. Le site opérateur est un site régulateur sur lequel est susceptible de se fixer un répresseur.

Le fonctionnement de l'opéron garantit une plasticité phénotypique rapide des bactéries à un changement de l'environnement avec par exemple la synthèse de ces enzymes lors d'une introduction de lactose. Inversement, la répression supprime l'expression de l'ensemble de ces protéines : un opéron fonctionne donc sur un mode « on-off » qui permet une économie d'énergie en ne synthétisant les protéines que quand la cellule en a besoin. Par contre, cela ne permet pas une modulation indépendante de l'expression de chaque gène d'un opéron.

## 3 Le génome des eucaryotes

### 3.1 Le génome nucléaire

#### a. L'empaquetage et la condensation de l'ADN

Le génome nucléaire est constitué de chromosomes linéaires, chacun renfermant une molécule d'ADN (en absence de toute réplication). Ces molécules très longues font l'objet de replis de divers ordres en association avec des protéines.

Dans la cellule interphasique, le matériel chromosomique associant ADN et protéines constitue la chromatine présente sous deux états :

- l'**hétérochromatine** (ou chromatine dense, très opaque aux électrons) localisée essentiellement en périphérie du noyau contre la face interne de l'enveloppe nucléaire et autour du, ou des, nucléole(s) ;
- l'**euchromatine** dispersée dans le nucléoplasme.

Deux ensembles de protéines peuvent être extraits de la chromatine :

- les **histones**, les plus abondamment liées à l'ADN, sont de petites protéines basiques très conservées entre espèces. Elles assurent la condensation de l'ADN (rôle structural) ;
- les **protéines non histones** sont très variées et à rôle divers : ce sont des protéines régulatrices ou des enzymes participant aux grands processus génétiques : transcription, réplication, réparation de l'ADN.

Différents niveaux de condensation peuvent être observés (figure 13.3) :

- **La fibre nucléosomique de 11 nm** (ou **nucléofilament**)

Celle-ci n'est observée qu'*in vitro* dans certaines conditions salines. Elle correspond au premier niveau de compaction et montre des structures répétitives, les nucléosomes. Chaque **nucléosome** comprend un noyau formé d'un assemblage de 8 **histones** (2 H2A, 2 H2B, 2 H3, 2 H4) autour duquel s'enroule l'ADN et une région de liaison entre les noyaux dans laquelle l'ADN est lié à un autre type d'histones (H1). C'est pourquoi cette fibre est souvent dite « en collier de perles ». La structure des noyaux est très conservée d'une espèce à l'autre avec une longueur d'ADN enroulé d'environ 140 pb (146 pb). Les régions de liaisons sont de longueur plus variable, un nucléosome s'étendant sur environ 200 pb. Le degré de compaction par rapport à la molécule d'ADN seule est d'un facteur 6. La liaison ADN – histones est assurée par l'attraction des charges positives des histones et des charges négatives des phosphates de l'ADN. Les régions N-terminales des histones n'interagissent pas directement avec l'ADN et sont susceptibles de modifications (méthylation, acétylation, phosphorylation) affectant la compaction et par suite la transcription.

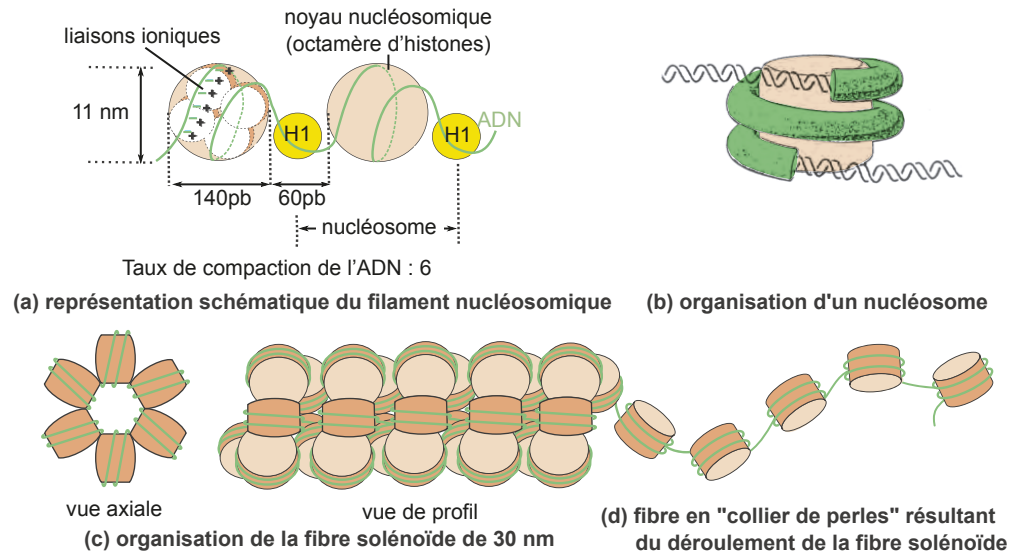


Figure 13.3 Structure de la chromatine.

- **La fibre chromatinienne de 30 nm**

Elle correspond au deuxième degré de compaction. L'histone H1 fixée à l'ADN entre deux nucléosomes associe les noyaux nucléosomiques en une hélice (fibre solénoïde) : chaque tour

comprend 6 nucléosomes pour un diamètre de 30 nm et une longueur de 11 nm, donnant un facteur de compaction de 40. Cette organisation de la fibre de 30 nm, classiquement décrite, fait cependant l'objet de nombreuses discussions : certaines techniques microscopiques (cryo-microscopie électronique) n'ont pas permis de l'observer dans de nombreuses cellules. Certaines hypothèses considèrent alors la fibre de 30 nm comme un artefact lié à des traitements de fixation particuliers. Des repliements de la fibre de 11 nm s'effectuent bien, mais selon des modes différents, générant des assemblages beaucoup plus irréguliers.

Ces structures se rencontrent dans l'**euchromatine** où se réalise le déroulement en nucléofilament, ce qui permet l'expression des gènes situés à ce niveau. L'euchromatine représente donc la chromatine susceptible d'être transcrite (sans l'être toutefois forcément).

Les assemblages précédents peuvent dessiner des replis d'ordre supérieur se rencontrant dans l'hétérochromatine qui n'est pas transcrite, même si elle peut contenir des séquences codantes. Dans un organisme pluricellulaire, on distingue l'**hétérochromatine constitutive** commune à toutes les cellules et l'**hétérochromatine facultative** qui diffère selon le type cellulaire et n'est pas vide de régions codantes. Du fait de la compaction, celles-ci ne sont cependant pas accessibles aux minisatellite à la transcription.

### b. Les différents types de séquences

#### ZOOM 2

Les courbes de Cot

Le suivi de la renaturation d'un mélange des fragments d'un ADN génomique, dénaturé, conduit au tracé des **courbes de Cot** qui traduisent la cinétique de réassociation (voir zoom 2). Ce suivi permet de distinguer trois types de séquences.

#### Les séquences hautement répétées (figure 13.4)

Ces séquences ne sont pas codantes. Elles représentent 10 à 15 % du génome. Elles sont constituées de séquences d'hétérochromatine que l'on retrouve généralement dans la partie centrale des chromosomes (centromères) ou à leurs extrémités (télomères). On distingue :

- L'**ADN satellite**, distribué principalement au niveau des centromères et des télomères avec :
  - les séquences CEN (centromères) : répétition de nombreuses fois chez l'homme d'un motif de 171 bp dans la zone qui servirait d'ancrage aux microtubules kinétochoriens du fuseau mitotique (voir § 6.2) ;
  - les séquences TEL (télomères) : très nombreuses répétitions d'un motif riche en A et C qui pourraient intervenir dans la protection de l'extrémité des chromosomes au cours des réplifications successives.
- L'**ADN mini-satellite** (VNTR : *variable number of tandem repeat*) : ce sont des fragments de 1 à 5 kb formés de motifs de 15 à 100 pb répétés en tandem (c'est-à-dire côte à côte). Ces fragments, dispersés sur l'ensemble du génome, présentent un très grand polymorphisme.
- L'**ADN microsatellite** (STR : *short tandem repetition*) : ce sont des fragments courts (inférieurs à 150 pb) formés de motifs de quelques pb (< 4), répétés en tandem de très nombreuses fois dans l'ensemble du génome.

L'ADN hautement répété est notamment utilisé comme empreintes génétiques (DNA fingerprints) dans l'identification des individus (police criminelle, recherche de paternité).

#### Les séquences moyennement répétées (figure 13.4)

Elles représentent une large part du génome et montrent des séquences non codantes et codantes. Ces dernières, transcrites, correspondent aux gènes.

- Les séquences non codantes : une forte proportion du génome eucaryote est constituée d'éléments répétitifs qui se sont propagés dans le génome en fabriquant des copies d'eux-mêmes, capables de gagner d'autres positions. Certains sont des **transposons**, le transposition n'ayant fait intervenir que de l'ADN. D'autres sont des **rétrotransposons**, plus largement répandus et dans lesquels la transposition a impliqué un intermédiaire ARN et une transcriptase réverse.

Ces éléments rétrotransposables (46 % du génome humain) comprennent notamment les :

- **LINE** : *long interspersed nuclear elements* : longs éléments dispersés de longueur de l'ordre de 5 à 7 kb, avec quelques milliers de copies sous la forme complète et plusieurs dizaines de milliers sous forme de copies tronquées.
- **SINE** : *short interspersed nuclear elements* : courts éléments dispersés. Le SINE le plus commun dans le génome humain est la séquence Alu, longue d'environ 200 pb, et présente à près d'un million de copies. Cette séquence répétitive est ainsi appelée car elle contient un site de coupure spécifique de l'enzyme de restriction Alu.
- Les séquences codantes (transcrites) : certaines sont seulement transcrites sans être traduites : ces séquences sont répétées en copies présentes côte à côte sur le chromosome, disposition dite **en tandem**. C'est le cas des gènes codant pour les ARN avec :
  - les gènes des ARNr : les différents ARNr du ribosome (petite et grande sous-unité) sont transcrits en une seule molécule initiale, ensuite scindée. Les gènes sont disposés en tandem de plus de 200 copies et ce, sur 5 chromosomes différents. Ces ensembles de gènes correspondent aux nucléoles dans le noyau interphasique et aux constriction secondaires des chromosomes mitotiques. Une telle redondance est un moyen de garantir une grande quantité d'ARNr par cellule ;
  - les gènes des ARNt regroupés également en tandem, avec de nombreuses copies (jusqu'à 20 000) ;
  - les gènes d'autres ARN : ARNsn (*small nuclear ARN*), ARNmi (microARN), ARN interférents.

D'autres sont traduites comme les gènes des histones : les gènes des différentes histones sont regroupés en unités, elles-mêmes répétées en tandem.

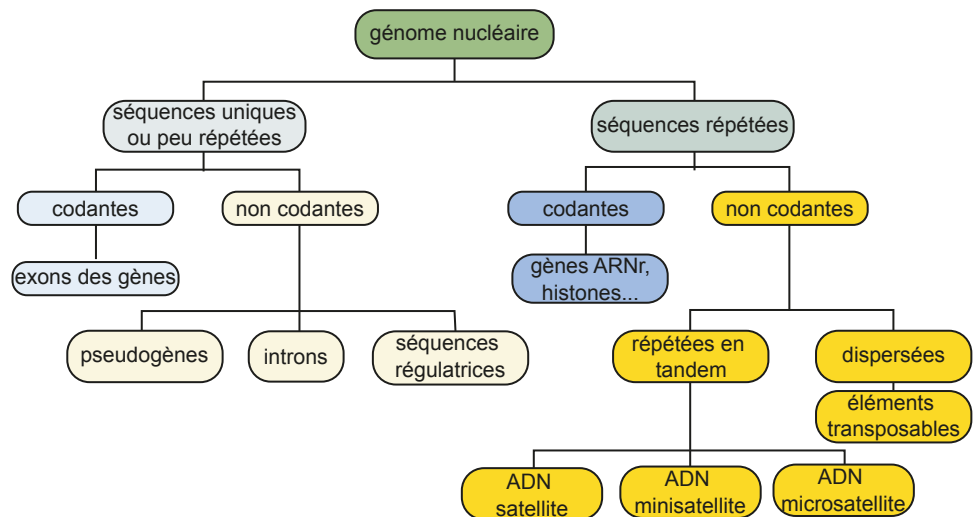


Figure 13. 4 Organisation de l'ADN eucaryote.

### Les séquences non répétées (figure 13.4)

Elles correspondent essentiellement aux gènes qui codent pour des protéines. L'appellation de séquences uniques peut apparaître abusive car les duplications géniques ont généré des séquences proches (exemple des familles multigéniques comme les globines). La répétition reste de toute façon très faible au regard des séquences considérées plus haut.

Les gènes issus d'une duplication se trouvent groupés au niveau d'une zone chromosomique. Un **pseudogène** est une séquence nucléotidique de l'ADN qui ressemble beaucoup à celle d'un gène fonctionnel mais qui n'est pas transcrite. Elle résulte d'un processus de duplication soit incomplète soit ayant accumulé des mutations qui ont endommagé la copie.

Le nombre de gènes codant des protéines est estimé aujourd'hui chez l'homme à près de 20 000 avec une distinction entre :

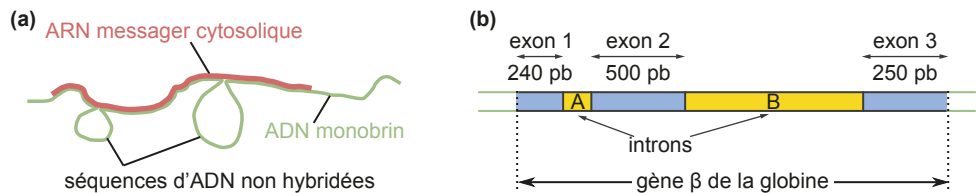
- les **gènes de ménage** impliqués dans les fonctions essentielles de la cellule et qui sont exprimés dans toutes les cellules de l'organisme : gènes codant pour les enzymes de la glycolyse, de la traduction, de la réparation de l'ADN... ;
- les **gènes tissu-spécifiques**, exprimés seulement dans certains tissus et à l'origine de la différenciation cellulaire.

### c. La structure des gènes eucaryotes

#### • La structure morcelée des gènes

L'identification du caractère morcelé découle de l'observation de l'hybridation moléculaire entre un segment d'ADN génomique et l'ARN cytosolique codé par ce même segment d'ADN (formation d'un **hétéroduplex**). La première observation a été réalisée sur le gène de l'ovalbumine et l'ARNm cytosolique correspondant. La [figure 13.5](#) représente le résultat obtenu sur le gène de la  $\beta$ -globine.

L'ADN génomique est rendu monocaténaire par dénaturation thermique puis mis en présence de l'ARN. Le résultat de l'appariement a été observé en microscopie électronique à transmission.



**Figure 13.5** Hybridation ADN-ARNm cytosolique et gènes morcelés.

Cas du gène  $\beta$  de la globine des mammifères avec 3 exons et 2 introns. Les nombres de bases sont indiqués. **(a)** figure d'hybridation ; **(b)** structure du gène.

L'hybridation révèle que :

- l'ADN du gène est beaucoup plus long que l'ARNm ;
- les deux extrémités de l'ADN n'ont pas d'équivalent sur l'ARNm ;
- l'ADN montre des secteurs appariés à l'ARNm séparés par des boucles d'ADN dépourvues de segments complémentaires sur l'ARNm.

Le gène présente ainsi des régions qui sont absentes sur l'ARNm cytosolique traduit et qui sont ainsi non codantes. Seules les séquences codantes du gène ont leur séquence complémentaire sur l'ARNm cytosolique.

Sur le plan moléculaire, la synthèse d'un **transcrit primaire**, ou **ARN pré-messager**, complémentaire de l'ADN sur la totalité du gène est réalisée dans le noyau. Elle est suivie de l'élimination, ou excision, des séquences non codantes grâce à des endonucléases et de l'association des séquences codantes grâce à des ligases, ou épissage.

Ces modifications de l'ARNm par **excision** et **épissage** correspondent à la **maturation** des ARN, spécifique des eucaryotes. Elle se produit dans le noyau avant exportation des ARN matures.

Les séquences non codantes du gène sont appelées des **introns** et les séquences codantes, finalement exprimées, des **exons**.

La présence d'introns et d'exons définit les **gènes morcelés** ou gènes en **mosaïque**. Ils sont très nombreux chez les eucaryotes, même s'il peut exister des gènes non morcelés (exemple des gènes des histones).

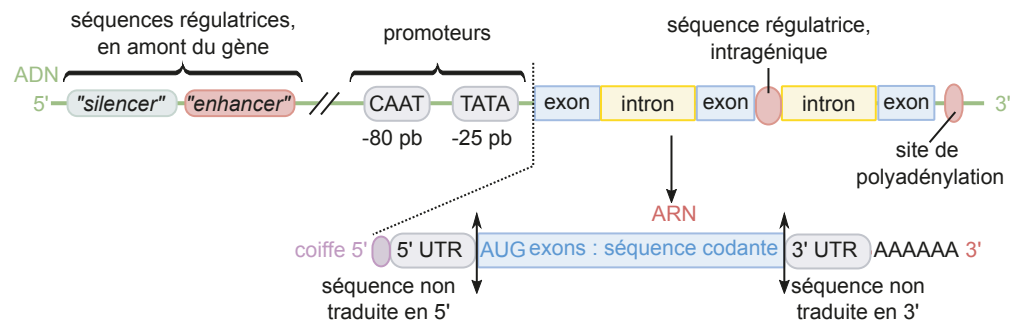
• **Les séquences impliquées dans le fonctionnement du gène**

Les **introns** et les **exons** se situent entre un codon initiateur et un codon stop qui constituent des signaux de début et de fin de traduction.

L'expression du gène et la maturation de l'ARN transcrit mettent en jeu d'autres signaux parmi lesquels :

- des signaux nécessaires à la transcription : séquence promotrice en amont du gène ;
- des signaux de modifications de l'ARN transcrit : signal d'ajout de la coiffe en 5' et signal de polyadénylation en 3' (ajout d'une queue poly A) ;
- des signaux d'excision et d'épissage aux bornes des introns.

Le **gène** peut alors être défini comme un fragment d'ADN transcrit, sa définition fonctionnelle associant le segment transcrit, la région promotrice et des séquences régulatrices. Ces dernières peuvent être situées en 5' ou en 3' et sont parfois loin du gène régulé (figure 13.6).



**Figure 13.6** Organisation fonctionnelle d'un gène eucaryote et de l'ARN mature associé. (UTR : *untranslated region*).

Au total, les séquences codantes ne représentent que 2 à 5 % de l'ADN total, ce pourcentage des séquences codantes (ou densité) restant variable d'un organisme à l'autre.

### 3.2 Le génome extra-nucléaire

Le génome extra-nucléaire est représenté par les génomes **mitochondrial** et **chloroplastique**. Dans les deux cas, il s'agit de molécules d'ADN circulaires, non associées à des histones et en plusieurs copies. Ce génome est en grande partie transcrit et traduit, avec cependant un code génétique légèrement différent du code génétique utilisé pour le génome nucléaire.

#### a. L'ADN mitochondrial

**Tableau 13.2** Caractères du génome mitochondrial.

<b>Taille</b>	<ul style="list-style-type: none"> <li>- <i>Saccharomyces cerevisiae</i> : 75 kb (35 gènes).</li> <li>- <i>Arabidopsis thaliana</i> : 367 kb (284 gènes).</li> <li>- Homme : 16,569 kb (37 gènes dont 13 codant des protéines de la phosphorylation oxydative ; 22 codant des ARNt et 2 des ARNr mitochondriaux).</li> </ul>
---------------	--

<b>Caractère du génome et information</b>	<ul style="list-style-type: none"> <li>- Le génome possède toute l'information nécessaire à la traduction des protéines (ARNm, ARNt, ARNr).</li> <li>- Le nombre d'ARNt (20) requis pour la synthèse de protéines mitochondriales est inférieur au nombre d'ARNt (32) requis pour traduire les ARN produits dans le noyau. Cette économie est permise par un flottement plus net de l'appariement des anticodons des ARNt, expliquant les différences avec le code génétique nucléaire.</li> <li>- Chez l'homme, tout l'ADN mitochondrial est codant mais chez certains végétaux et mycètes, il possède des séquences non codantes (introns).</li> </ul>
<b>Relation avec le génome nucléaire</b>	<ul style="list-style-type: none"> <li>- Le génome de petite taille ne permet le codage que de 5 à 10 % des protéines de la mitochondrie (protéines de la chaîne respiratoire).</li> <li>- Les autres protéines du système de phosphorylation oxydative sont codées par des gènes nucléaires. L'ARNm est traduit hors de la mitochondrie par les ribosomes cytosoliques.</li> <li>- Les protéines synthétisées sont transportées dans la mitochondrie et le système est assemblé dans la membrane mitochondriale interne.</li> </ul>

## b. L'ADN chloroplastique

**Tableau 13.3** Caractères du génome chloroplastique.

<b>Taille</b>	<ul style="list-style-type: none"> <li>- De 15 à 500 kb (155 kb pour <i>Arabidopsis thaliana</i>), avec de nombreuses copies.</li> <li>- 20 à 35 copies par chloroplaste pour <i>A. thaliana</i>, soit pour l'ensemble d'une cellule de 1 000 à 1700 copies</li> </ul>
<b>Caractère du génome et information</b>	<ul style="list-style-type: none"> <li>- Chez <i>Arabidopsis</i>, le nombre de protéines codées est de 87, dont la moitié est impliquée dans les processus photosynthétiques.</li> <li>- Les gènes portés codent pour des protéines et des ARN (ARNr et ARNt). Une partie des gènes chloroplastiques comporte des introns, produisant alors des ARNm épissés avant traduction.</li> </ul>
<b>Relation avec le génome nucléaire</b>	<ul style="list-style-type: none"> <li>- Une grande partie des protéines est codée par le génome nucléaire et est importée depuis le cytosol.</li> <li>- La Rubisco est constituée de sous-unités codées les unes par le génome nucléaire, les autres par le génome chloroplastique.</li> </ul>

Le séquençage des ARNr des ribosomes mitochondriaux et chloroplastiques fait apparaître des différences avec les ARNr constituant les ribosomes cytosoliques. Les premiers montrent des affinités avec les ribosomes procaryotes, ce qui constitue un argument majeur en faveur de l'origine endosymbiotique de ces organites semi-autonomes.

## 4

## Les virus, diversité structurale et génomique

### 4.1 Caractères généraux des virus

Les **virus** sont des entités biologiques qui se distinguent clairement du monde des cellules ; ce sont des parasites intracellulaires obligatoires. Ils constituent un ensemble infini avec des structures très diverses, au sein duquel il est toutefois possible de dégager des caractères communs :

- la taille des virus, très variable, est inférieure à la taille des cellules ;
- ils renferment un génome sous forme d'acides nucléiques, ADN ou ARN. Celui-ci reste généralement de petite taille, très inférieur aux génomes cellulaires ;

- ils ne peuvent se répliquer qu'au sein des cellules hôtes, procaryotes (bactériophages) ou eucaryotes (animales ou végétales).
- à l'extérieur des cellules, ils sont sous forme de particules virales libres, ou **virions**, biologiquement inertes (aucune activité métabolique). Ces virions sont formés d'une **capside** protéique, parfois entourée d'une enveloppe lipidique, et renfermant l'**acide nucléique**, avec éventuellement quelques protéines à rôle enzymatique.

La diversité structurale des virus réside à la fois dans les enveloppes et dans le matériel génétique qu'elles renferment. La nature de celui-ci conditionne le cycle de reproduction des virus au sein des cellules.

## 4.2 Diversité génomique des virus

La prise en compte du matériel génomique conduit à distinguer les **virus à ADN**, les **virus à ARN** et les **rétrovirus**. Les virus à ADN constituent l'immense majorité des virus affectant les bactéries (98 %) et la totalité de ceux parasitant les archées. Les cellules eucaryotes sont atteintes à parts à peu près égales par les virus à ADN et à ARN. Les familles de virus à ARN sont multiples et sont responsables de très nombreuses pathologies et d'épidémies chez l'homme. Une grande majorité de virus des végétaux sont à ARN.

La production de nouveaux virions impose toujours un stade de synthèse de protéines à partir du matériel génétique et une réplication du matériel génétique, ce qui conduit à distinguer différentes situations.

### a. Des génomes à ARN ou à ADN

#### • Les virus à ADN :

- Virus à ADN double brin. C'est par exemple le cas des **bactériophages** (virus n'infestant que les bactéries) : l'ADN linéaire dans le virion est injecté dans la bactérie dans laquelle il circularise. Les enzymes de la cellule assurent la transcription de l'ADN puis sa réplication. De nouvelles particules virales sont ainsi produites dans la cellule qui éclate (cycle lytique). L'infection par certains bactériophages, dits tempérés, comme le **bactériophage λ** conduit à l'intégration de l'ADN viral dans le chromosome bactérien, où il devient dormant (cycle lysogène). Les adénovirus, responsables de pathologies chez les vertébrés, sont également des virus à ADN.
- Virus à ADN simple brin : des réplications successives fournissent des doubles brins intermédiaires donnant, après séparation, les génomes des nouveaux virions. (exemple des parvovirus, parasites de cellules animales).

#### • Les virus à ARN

- Virus à ARN + : l'ARN introduit dans la cellule-hôte peut être immédiatement traduit par les ribosomes de celles-ci. Le génome est dans ce cas équivalent à des ARN messagers. C'est l'exemple du **virus de la mosaïque du tabac** et des **coronavirus**.
- Virus à ARN – : l'ARN introduit doit être copié en un ARN+ pour être traduit en protéines.
- **Rétrovirus** : l'ARN, bien que de sens (+), n'est pas traduit. Il sert de matrice à une transcriptase inverse, apportée avec la particule virale, et qui le copie en ADN double brin. L'ADN double brin est intégré dans le génome de l'hôte et sert de support d'information pour la synthèse de protéines virales. La réplication à l'aide des polymérases de l'hôte permet de générer de nouveaux virions. Un exemple de rétrovirus est constitué par le VIH.

### b. Des génomes compacts

De manière générale, les virus à ADN double brin sont ceux dont le génome est le plus long (en moyenne 44 kb) alors que les virus à ARN sont plus courts (aux alentours de 10 kb).

#### ZOOM 3

Le bactériophage lambda

#### ZOOM 4

Le virus de la mosaïque du tabac

#### ZOOM 5

Les coronavirus

Les coronavirus, virus à ARN positif, font exception par leur taille importante aux environs de 30 kb.

Les génomes des virus ne comprennent que peu de séquences non codantes (environ 10 % pour les virus à ADN et 6 % pour les virus à ARN). Ces capacités de codage importantes au regard leur taille en font des génomes **compacts**.

Ces capacités sont renforcées chez de nombreux virus par une organisation du génome en gènes chevauchants : un même fragment d'acide nucléique, appelé **cadre de lecture ouvert** (ou ORF : *open reading frame*) peut renfermer plusieurs codons initiateurs et correspondre ainsi à trois cadres de lecture différents, à l'origine de trois protéines différentes.

Le nombre de gènes est variable : il peut être très faible : le plus petit virus connu n'a que 4 gènes mais ce nombre peut atteindre plusieurs dizaines, voire centaines.

Les gènes portés par l'acide nucléique peuvent se rapporter à différents groupes fonctionnels :

- gènes **structuraux** impliqués dans la synthèse des protéines de la capsidite et des enveloppes, qui sont les plus nombreux ;
- gènes **informationnels** impliqués dans la réplication et l'expression génétique (transcription, traduction) ;
- gènes, plus accessoires, impliqués dans des réactions métaboliques ; ceux-ci pourraient provenir de transferts horizontaux à partir de cellules hôtes ;
- de nombreux gènes restent par ailleurs non caractérisés à l'heure actuelle.

### 4.3 Diversité des enveloppes

Le génome est toujours enveloppé et protégé dans une structure protéique : la **capsidite**. Celle-ci est constituée d'un grand nombre de copies d'un ou de plusieurs types de protéines organisées de manière très structurée. Ces protéines sont codées par le génome du virus. On distingue :

- des capsidites hélicoïdales à l'image de celle du virus de la mosaïque du tabac. Les particules virales sont allongées en filaments (de 300 à 500 nm de long pour un diamètre de 15 à 19 nm) dans lesquels l'acide nucléique est enroulé en hélice. Ces types sont fréquemment observés dans les cellules végétales ;
- des capsidites icosaédriques (polyèdres à 20 faces triangulaires équilatérales) à l'image des adénovirus. Certains virus présentent des capsidites plus complexes avec un nombre de faces supérieures.

Les **bactériophages** (dont le bactériophage  $\lambda$ ) montrent une organisation complexe associant une tête icosaédrique et une queue hélicoïdale. La tête renferme l'ADN ; la queue se termine par un plateau protéique porteur d'épines et de fibres protéiques. Celles-ci interviennent lors de la fixation du virion sur la paroi bactérienne.

Certains virus de formes sphériques, dont les coronavirus, sont dits **enveloppés**. La capsidite protéique est entourée d'une bicouche lipidique. Celle-ci provient de la cellule dans laquelle s'est formée la particule virale et dont elle est sortie par bourgeonnement membranaire. Ceci explique que la plupart des virus végétaux soient des virus nus du fait de la présence de la paroi squelettique et soient de structure plus simple que les virus infectant les animaux.

Cette enveloppe porte un certain nombre de glycoprotéines donnant des spicules impliquées dans la reconnaissance des cellules hôtes, la fixation et la pénétration des virions.

Les enveloppes ont donc deux rôles principaux : protéger le génome et permettre la fixation des virus aux cellules hôtes. La bicouche lipidique donne à ces virus une plus grande fragilité aux conditions du milieu (milieu extérieur ou tube digestif) et réduit leur pouvoir infectieux par rapport aux virus nus.

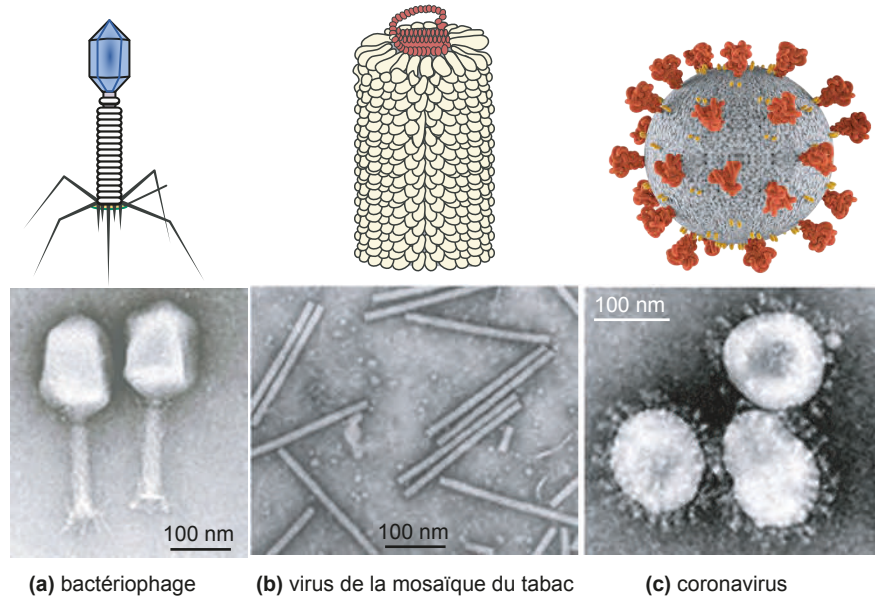


Figure 13.7 Les différents types de virus.

- (a) Phage HY01 (MET). Source : Lee H, Ku H-J, Lee D-H, Kim Y-T, Shin H, Ryu S, et al. (2016) « Characterization and Genomic Study of the Novel Bacteriophage HY01 Infecting Both Escherichia coli O157:H7 and Shigella Flexneri ».
- (b) Source :1994 Rothamsted Experimental Station <https://www.rothamsted.ac.uk/our-science>
- (c) Source : CDC/Dr. Fred Murphy Centers for Disease Control and Prevention's Public Health Image Library, #4814.

## 5 Cycle cellulaire et réplication du matériel génétique

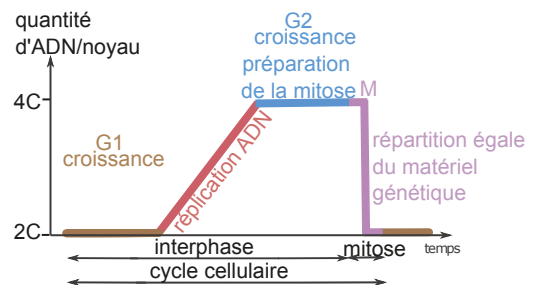
### 5.1 Le cycle cellulaire

#### a. Les étapes du cycle

Le cycle cellulaire des eucaryotes comprend deux étapes principales, l'**interphase** suivie d'une division cellulaire, la **mitose** (phase **M**) qui peut affecter les cellules haploïdes et diploïdes. (figure 13.8). L'interphase se divise en phases **G1**, **S** et **G2** (pour G, selon les auteurs : *Growth* pour croissance ou *Gap* pour intervalles, S : synthèse). La phase **S** est celle de la **réplication** du matériel génétique et du doublement du lot de chromosomes. La phase **G1** est le siège d'une synthèse protéique importante qui accompagne la croissance cellulaire. La phase **G2** est caractérisée par la synthèse de protéines nécessaire au bon déroulement de la division cellulaire. En phase **M**, le matériel génétique, sous forme de chromosomes condensés, est réparti dans les cellules filles en deux lots identiques.

Figure 13.8 Les étapes du cycle cellulaire.

Exemple d'une cellule diploïde contenant une quantité 2C d'ADN



### b. Durée des cycles cellulaires et capacités de division

Au sein d'un organisme pluricellulaire, les capacités de croissance et de division sont variables d'un type cellulaire à un autre.

Certaines cellules conservent normalement une activité mitotique élevée.

Chez les végétaux c'est le cas par exemple des cellules **indifférenciées** des **méristèmes** apicaux. Certains tissus animaux contiennent des cellules souches abondantes, assurant le renouvellement : c'est le cas des cellules souches hématopoïétiques à l'origine des hématies et des leucocytes, ou celles des couches basales de nombreux épithéliums. Les cellules souches sont capables de se diviser de manière asymétrique, produisant à chaque division une seule cellule qui s'engage dans la voie de différenciation assurant le renouvellement du tissu, alors que l'autre reste à l'état indifférencié.

D'autres cellules ne se divisent pas normalement, mais possèdent toujours une capacité de division si les conditions du milieu sont modifiées. C'est par exemple le cas des cellules du foie, qui peuvent proliférer après une ablation chirurgicale d'une partie de l'organe, ou des lymphocytes se développant en clones après interaction avec un antigène donné. Les cellules sont alors dites dans un état de **quiescence**, correspondant à une phase définie comme **G0**. La division peut ainsi se produire sous l'effet de signaux extra- ou intracellulaires.

La durée des cycles cellulaires est très variable et peut être très courte : des cycles très courts (30 minutes) peuvent par exemple s'observer en début de développement embryonnaire (étape de segmentation) chez les amphibiens où les phases S succèdent aux phases M, sans phases G1 et G2. De manière générale, c'est la durée de la phase G1 qui détermine celle du cycle cellulaire.

Enfin certaines cellules très spécialisées, comme les neurones ou les cellules musculaires, ont perdu la capacité de se diviser. Cette incapacité n'interdit cependant pas la croissance.

### c. Le contrôle du cycle cellulaire

Le déroulement du cycle cellulaire repose sur l'existence d'un certain nombre de **points de contrôle** (figure 13.9). Ceux-ci correspondent à des mécanismes de surveillance susceptibles de bloquer la progression du cycle.

Certains mécanismes peuvent détecter les éventuelles lésions de l'ADN (dues à des radiations ou à des agents chimiques) et bloquer le cycle, le temps nécessaire à sa réparation. L'impossibilité d'une réparation peut alors conduire selon les cas, à la mort de la cellule ou à un état d'arrêt permanent de la division (sénescence).

D'autres mécanismes contrôlent le bon accomplissement d'une phase avant le passage à la suivante. La fin de la phase G1 est marquée par un point d'engagement appelé **point de restriction** chez les mammifères dont le franchissement déclenche la machinerie de réplication de l'ADN. Les autres points de contrôle sont l'entrée en phase M après détection de la fin de réplication de l'ADN, et le déclenchement de l'anaphase au cours de la mitose qui requiert la bonne fixation préalable des chromosomes sur le fuseau mitotique.

Le franchissement des points de contrôle fait intervenir un certain nombre de molécules dont l'effet est essentiellement négatif, bloquant le cycle en cas d'anomalie et dont l'origine peut être intra- ou extracellulaire.

Le déroulement normal de toutes les étapes repose par ailleurs sur l'intervention de multiples facteurs protéiques agissant de manière cyclique et mettant en jeu de nombreuses interactions entre protéines, dont des kinases et des cyclines. Ces oscillations résultent à la fois de processus de synthèse et de protéolyse impliquant l'ubiquitine.

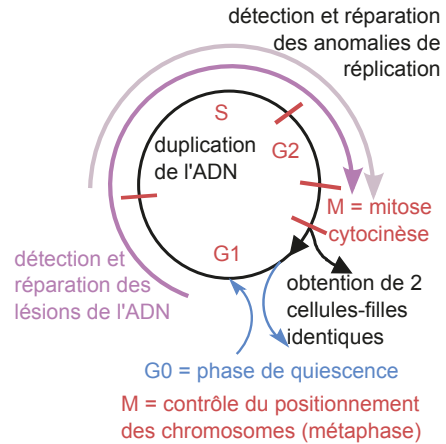
De nombreux résultats expérimentaux proviennent de travaux conduits sur les levures mais on montre que les systèmes de contrôle sont identiques chez tous les eucaryotes, avec des protéines remarquablement conservées au cours de l'évolution.

#### ZOOM 6

Durée et contrôle du cycle cellulaire

**Figure 13.9** Points de contrôle du cycle cellulaire.

Les points de contrôle sont indiqués par les barres rouge brique.



## 5.2 La réplication de l'ADN

Chez les eucaryotes, le matériel génétique est dupliqué au cours de la phase S interphasique qui précède toute division cellulaire.

### a. Réplication semi-conservative et complémentarité des bases

En même temps qu'ils proposaient un modèle de structure de l'ADN en double hélice, Watson et Crick envisageaient un modèle de réplication basé d'une part sur le désenroulement de la double hélice grâce à la séparation des deux brins complémentaires par rupture des liaisons hydrogène unissant les bases azotées et d'autre part sur la synthèse de deux nouveaux brins d'ADN chacun complémentaire d'un des brins de la molécule-mère servant alors de matrice.

Ce modèle est dit **semi-conservatif**. Les deux molécules formées (ou molécules – filles) sont identiques entre elles et identiques à la molécule-mère. Elles sont constituées d'un brin **hérité** de la molécule-mère et d'un brin **néoformé** complémentaire du brin hérité.

Ce modèle a été vérifié expérimentalement par des travaux conduits chez les procaryotes et les eucaryotes.

Au cours de la phase S, les chromosomes linéaires des eucaryotes présentent de nombreuses zones de réplication actives, définissant chacune un **œil de réplication** ou **réplicon**. La réplication progresse simultanément à partir du site d'initiation sur les deux brins de la molécule et dans les deux directions : la réplication est **bidirectionnelle**. Les deux extrémités d'un œil, en forme de Y, constituent les **fourches de réplication** : c'est à leur niveau que sont séparés les deux brins de la molécule mère et sont synthétisés les deux brins néoformés (figure 13.11). Elles sont le lieu d'activité de très nombreuses protéines se fixant à l'ADN : certaines (hélicases) sont impliquées dans le déroulement de la double hélice alors que d'autres (protéines SSB, *single-stranded DNA-binding*) assurent la stabilisation des brins séparés. La polymérisation des brins néoformés est effectuée par des **ADN polymérases**, formant un complexe de réplication ou **réplisome** (figure 13.11).

### b. L'action des ADN polymérases

#### • Une polymérisation orientée et la nécessité d'une amorce

Les ADN polymérases sont des enzymes qui catalysent la formation d'une chaîne polynucléotidique par additions successives de désoxyribonucléotides triphosphates (dNTPs) en présence d'une **matrice** qui sert de modèle. Chaque nucléotide entrant forme d'abord une association par établissement de liaisons hydrogène entre sa base azotée et la base azotée libre de la matrice. Le brin néoformé est ainsi complémentaire du brin matrice grâce à l'**appariement** des bases azotées (A et T d'une part, C et G d'autre part) (figure 13.10).

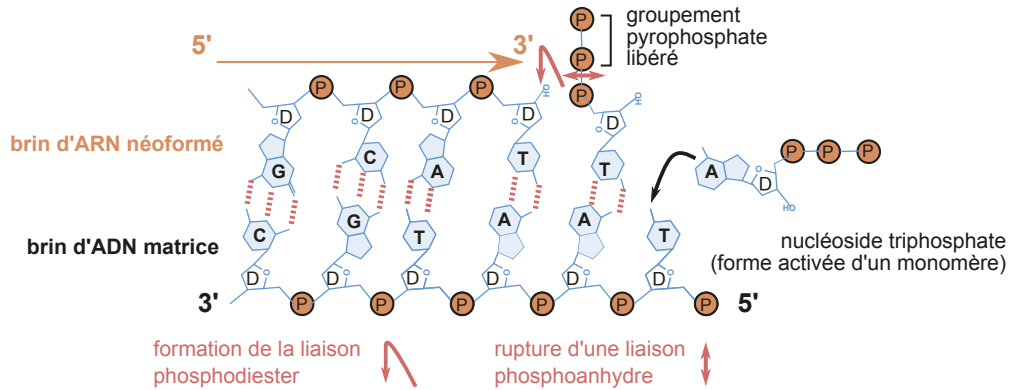
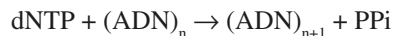


Figure 13.10 Polymérisation des nucléotides et synthèse du brin néoformé.

L'enzyme lie alors le nucléotide au nucléotide de l'extrémité de la chaîne par une **liaison phosphodiester** en les transformant en désoxyribonucléotides monophosphates par élimination du groupement pyrophosphate.

Elles catalysent la réaction globale, exergonique et irréversible suivante :



Cette polymérisation est orientée et ne s'opère que dans le **sens 5' → 3'**, la liaison phosphodiester s'établissant entre un groupement 5' du nucléotide entrant et le groupement 3' du dernier nucléotide de la chaîne en élongation.

Elle requiert toujours une amorce, constituée par un oligonucléotide simple brin déjà associé au brin matrice (figure 13.10).

L'action catalytique des ADN polymérases est rapide, avec l'ajout de 1 000 nucléotides par seconde à la matrice. Ces enzymes sont mises à profit dans les techniques de polymérisation in vitro (**PCR : Polymerase Chain Reaction**) et de séquençage (méthode de Sanger).

- **La correction des erreurs : l'activité exonucléasique**

Le faible taux d'erreur de la réplication repose sur une double activité des ADN polymérases, exercée par deux domaines distincts de la protéine : l'un exerce l'activité polymérase dans le sens 5' → 3' alors que l'autre présente une activité **exonucléasique** dans le sens 3' → 5'.

Pour corriger les erreurs de réplication, l'enzyme utilise un mécanisme de relecture de l'épreuve : grâce à l'activité exonucléasique 3' → 5', les nucléotides incorrectement appariés à l'extrémité 3' du brin d'ADN en cours d'élongation sont éliminés par hydrolyse. Le domaine à activité polymérase de l'enzyme effectue ensuite la correction du segment d'ADN.

Ce mécanisme de **correction** contraint le sens de polymérisation du brin néosynthétisé : seul ce sens permet que le nucléotide excisé soit remplacé par un nouveau nucléotide sous une forme dNTP nécessaire à la réalisation de la réaction de polymérisation.

- **Une réplication semi-discontinue**

La polymérisation ne s'opérant que dans le sens 5' → 3' alors que les deux brins matrices sont **antiparallèles**, la réplication des deux brins s'effectue selon des mécanismes différents.

La progression de la fourche de réplication et celle de la réplication du brin matrice orienté 3' → 5' se font dans le même sens : la synthèse de ce brin se déroule de façon continue et rapide, d'où son nom de **brin précoce** ou **brin direct**.

La progression de la fourche de réplication et celle de la réplication du brin matrice orienté 5' → 3' se font en sens inverse. La réplication est alors **discontinue** et se fait par de courts fragments d'environ 150 nucléotides, les **fragments d'Okasaki** (du nom de Reili Okasaki, leur

**ZOOM 7**

Le principe de la PCR

**ZOOM 1**

Le séquençage par la méthode de Sanger

découvreur). Leur polymérisation se fait bien dans le sens 5' → 3' mais en s'éloignant alors de la fourche, ce qui impose qu'un segment suffisant du brin matrice soit exposé pour que le fragment puisse être synthétisé.

Le brin matrice exposé dessine alors une boucle. Celle-ci permet aux ADN polymérases associées dans le réplisome de travailler de manière synchrone sur les deux brins et de progresser de manière conjointe en direction de la fourche (figure 13.11).

La production de chaque fragment d'Okasaki nécessite une courte amorce d'ARN synthétisée par une enzyme, la primase. Cette amorce est ensuite coupée par une endonucléase et les fragments d'Okasaki sont ensuite liés par une ligase.

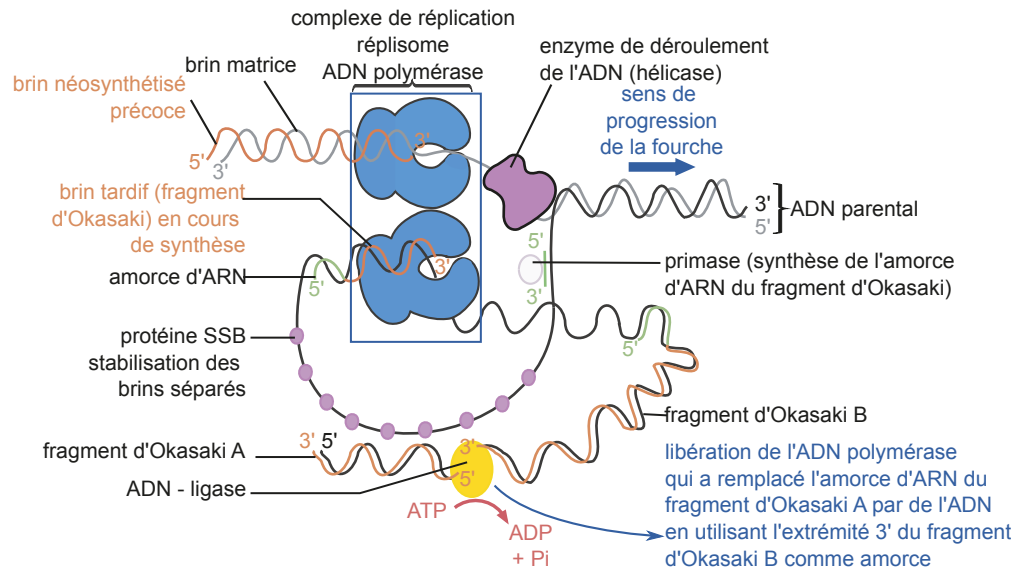


Figure 13.11 Fonctionnement d'une fourche de réplication.

Les observations montrent que les molécules-filles présentent des nucléosomes dès le passage de la fourche de réplication. Leur assemblage est donc très rapide, avec un mélange d'histones héritées du chromosome parental et d'histones néosynthétisées.

## 6 Les divisions cellulaires : mitose et méiose

### 6.1 Deux types de divisions

Une **division cellulaire** conduit à l'obtention de deux cellules-filles à partir d'une cellule-mère. Chez les eucaryotes, deux types de divisions peuvent s'observer, différant par la transmission du matériel génétique.

La **mitose** est une reproduction conforme qui assure une égale répartition des chromosomes entre les deux cellules-filles. Celles-ci sont identiques entre elles et à la cellule-mère en termes d'information génétique (aux mutations près). La mitose peut affecter des cellules haploïdes comme diploïdes.

La **méiose** est une série de deux divisions successives précédée par une interphase. Elle n'affecte que des cellules diploïdes particulières, les méiocytes : cellules-mères des **gamètes** (chez les animaux) et cellules-mères des **méiospores** (chez les végétaux). Elle constitue une étape majeure des cycles de reproduction sexuée où elle assure le passage d'une phase diploïde à une phase haploïde, dans la formation des gamètes et des méiospores.

## 6.2 La condensation des chromosomes

Les deux divisions sont précédées par une réplication préalable de l'ADN, conduisant à des chromosomes à deux chromatides (chromosomes **bichromatidiens**).

La première phase des divisions, ou prophase, est marquée par une condensation du matériel génétique, maximale dans le chromosome **métaphasique** (figure 13.12).

Les chromosomes présentent alors une morphologie caractéristique. C'est à ce stade de condensation que peut être établi le caryotype.

Les deux molécules d'ADN sont repliées dans deux chromatides-soeurs, maintenues au niveau du **centromère** (ou constriction primaire). Chacune de ces molécules s'organise en boucles disposées autour d'un squelette protéique central. Le centromère contient des séquences très répétées d'ADN (ADN satellites) servant de sites de fixation à des ensembles protéiques. Ceux-ci, organisés en plaques internes et externes, constituent les **kinétochores**.

Ces kinétochores sont le siège de la fixation de microtubules participant au **fuseau mitotique** et de protéines motrices nécessaires à l'activité de ces derniers. Ils participent également au point de contrôle guidant la transition métaphase - anaphase.

Les extrémités des bras des chromosomes sont les **téломères** au niveau desquels l'ADN présente de nombreuses séquences répétées, associées à des protéines particulières. Les téломères peuvent participer à la stabilisation et à la protection des chromosomes contre des nucléases et prévenir les fusions entre ces derniers.

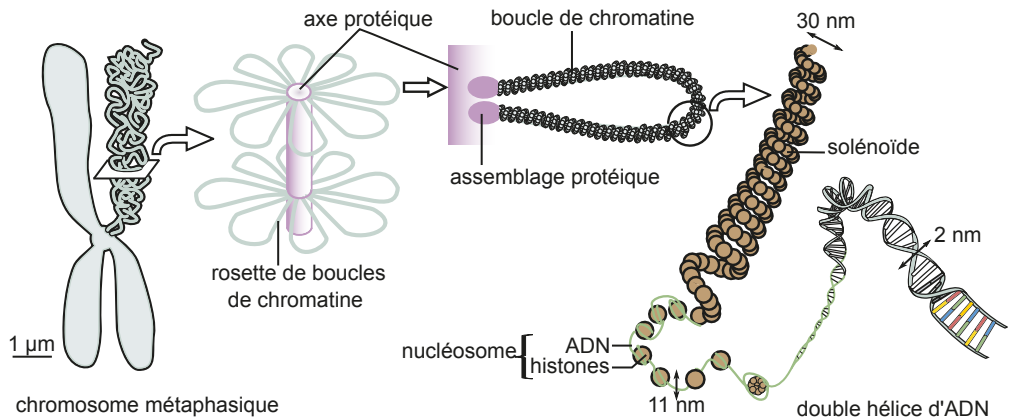


Figure 13.12 Structure du chromosome métaphasique.

## 6.3 Répartition conforme des chromosomes au cours de la mitose

### a. Les différentes étapes de la mitose

L'aspect des chromosomes permet de distinguer cinq phases : en **prophase**, des filaments de chromatine se raccourcissent et s'épaississent : cette condensation (d'un facteur 10 000) fait apparaître les chromosomes dupliqués avec deux chromatides. La **prométaphase** est initiée par la dissociation de l'enveloppe nucléaire et est marquée par la migration des chromosomes vers l'équateur de la cellule. En **métaphase**, les chromosomes se disposent sur ce plan équatorial, fixés par leur centromère à un réseau de microtubules, le fuseau mitotique. En **anaphase**, les chromatides de chaque chromosome se séparent et les chromosomes sont répartis en deux lots identiques aux deux pôles de la cellule. À la **télophase**, les deux

enveloppes des deux cellules apparues se mettent en place alors que les chromosomes se décondensent (tableau 13.4).

étape de la mitose (durée)	chromosomes	cytosquelette
prophase (20 à 30 min)	compaction d'un facteur 10 000 ; 2 chromatides par chromosome	séparation des centrosomes mise en place du fuseau mitotique grâce au COMT
prométaphase (5 min)	chromosomes rendus accessibles aux MT ; migration vers l'équateur	fixation des chromosomes aux MTK
métaphase (20 à 40 min)	disposition sur le plan équatorial de la cellule	mise sous tension des chromosomes par les MTK
anaphase (5 à 10 min)	clivage des centromères ; séparation des chromatides sœurs ; ascension polaire	dépolymérisation des MTK : écartement des pôles
télophase et cytodierèse (10 à 30 min)	2 lots identiques de chromosomes à une chromatide ; décondensation	sillon de division perpendiculaire au fuseau mitotique, dû à la formation d'un anneau contractile phragmoplaste

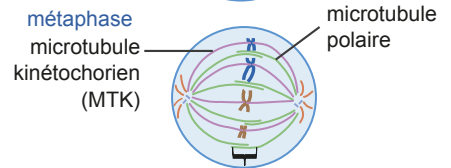
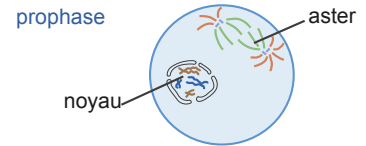
en noir : caractères communs aux cellules animales et végétales

en rouge : caractéristiques des cellules animales

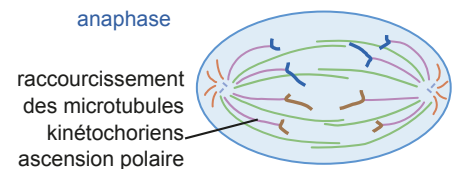
en vert : caractéristiques des cellules végétales

COMT : centre organisateur des microtubules

MT : microtubules



plaque métaphasique à  $2n = 4$  chromosomes



télophase et cytokinèse

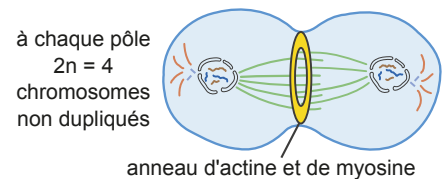


Tableau 13.4 les différentes phases de la mitose.

### b. La mise en place et le fonctionnement du fuseau mitotique

Dans la cellule animale, il existe en interphase deux **centrioles** orientés perpendiculairement l'un à l'autre et situés au cœur d'une région cellulaire très structurée, appelée **centrosome**. Cette zone correspond au centre organisateur des microtubules (COMT ou MTOC : *Micro-Tubules Organizing Center*), constituants majeurs du cytosquelette (tableau 13.4).

Au cours de la prophase, les centrioles et le centrosome se dédoublent, l'un des centrosomes migre et se positionne à l'opposé de son vis-à-vis. Ce mouvement est accompagné de la mise en place des microtubules du fuseau mitotique qui s'établit ainsi d'un pôle à l'autre.

Ce **fuseau** est formé de différents types de **microtubules** tous ancrés au niveau du centrosome par leur extrémité (-). Les microtubules sont des structures dynamiques capables de s'allonger ou de se raccourcir rapidement par polymérisation ou dépolymérisation de la tubuline à leurs extrémités. Lorsqu'un microtubule en allongement rencontre un kinétochore, il s'y fixe et constitue un microtubule kinétochorien. On distingue ainsi :

- des microtubules rayonnants autour de chaque centrosome, l'ensemble définissant alors un **aster** ;
- des microtubules **polaires** dont les extrémités (+), libres, dépassent l'équateur et viennent s'intercaler avec celles de leurs vis-à-vis ;
- des microtubules **kinétochoriens** dont les extrémités (+) sont fixées aux kinétochores des chromosomes.

Lors de la métaphase, le fuseau mitotique maintient les chromosomes sous tension à l'équateur de la cellule et organise leur disposition en plaque équatoriale. À l'anaphase, après clivage des centromères, les chromatides se divisent en deux lots tractés à chaque pôle de la cellule. Cette **ascension polaire** se réalise selon deux mécanismes le plus souvent synchrones :

- éloignement des chromosomes par raccourcissement des microtubules kinétochoriens (anaphase A) ;
- allongement des microtubules polaires qui éloignent les pôles l'un de l'autre (anaphase B).

Cette migration assure la séparation des chromosomes en deux lots identiques.

Chez les angiospermes, il n'existe pas de centrioles ; par contre il existe des sites de formation des microtubules disposés tout autour de l'enveloppe nucléaire : ils constituent des structures équivalentes au centrosome, jouant le rôle de centre organisateur des microtubules.

### c. Le processus de cytokinèse

Celui-ci permet la séparation des deux cellules filles.

Dans les cellules animales, la **cytokinèse** est réalisée grâce à un sillon de division par mise en place d'un **faisceau annulaire** contractile d'actine associée à de la myosine et qui provoque l'étranglement de la cellule ([tableau 13.4](#)).

Dans les cellules végétales, la cytokinèse fait intervenir la migration de vésicules d'origine golgienne qui se disposent au centre de la cellule dans un plan situé entre les noyaux fils. Ces vésicules contiennent à l'état de précurseurs les constituants des macromolécules pariétales. Leur fusion forme une plaque cellulaire qui progresse de façon centrifuge jusqu'à la membrane plasmique latérale, ce qui sépare alors les deux cellules. Leur migration est guidée par un réseau de microtubules, le **phragmoplaste**, issu des restes du fuseau mitotique, et orienté perpendiculairement à la plaque cellulaire. Les membranes fusionnées des vésicules forment les membranes des deux cellules alors que leur contenu élabore la première lamelle moyenne. Cette fusion ménage des communications entre les deux cellules, qui constituent des **plasmodesmes** et qui assurent la continuité cytoplasmique du **symplasma**.

## 6.4 La méiose : passage d'une phase diploïde à une phase haploïde

### a. Évolution de la quantité d'ADN et des chromosomes

Au cours de la phase S précédant la méiose, la quantité d'ADN est doublée passant de  $2C$  à  $4C$ . ( $C$  : quantité d'ADN du lot haploïde). La cellule contient alors  $2n$  chromosomes, chacun étant dupliqué en 2 chromatides.

À l'issue de la première division méiotique, cette quantité est de  $2C$  dans chacune des deux cellules-filles. Les deux cellules-filles contiennent  $n$  chromosomes, chacun à deux chromatides. La seconde division s'engage sans nouvelle réplication de l'ADN.

À l'issue de la seconde division, chacune des 4 cellules formées contient  $C$  ADN. Celles-ci renferment  $n$  chromosomes à une seule chromatide ([figure 13.13](#)).

La première division correspond à la séparation des **chromosomes homologues dupliqués**, sans division des centromères. La seconde division correspond à la séparation des **chromatides** de chaque chromosome.

On observera qu'après la première division les cellules contiennent une quantité égale d'ADN à celle de la cellule initiale ( $2C$ ) mais dont la signification en termes d'information génétique est différente.

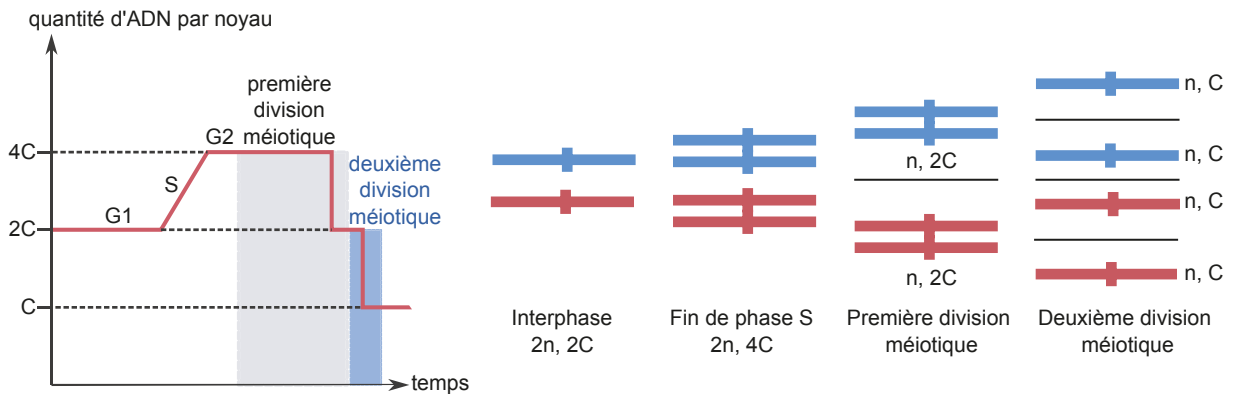


Figure 13.13 Évolution de la quantité d'ADN et du nombre des chromosomes au cours des deux divisions méiotiques.

### b. Les différentes étapes de la méiose

Un certain nombre de phénomènes cytologiques sont identiques à ceux décrits pour la mitose, dont la mise en place d'un fuseau de microtubules (figure 13.14).

Seule la première division méiotique présente des différences majeures au plan du comportement des chromosomes.

- **Prophase** de première division méiotique (prophase I). Elle est caractérisée par :
  - l'individualisation et la condensation progressive des chromosomes qui apparaissent formés de deux chromatides ;
  - l'appariement des chromosomes homologues maternel et paternel de chaque paire, chaque paire constituant un **bivalent** ou une **tétrade** ;
  - la réalisation éventuelle de **recombinaisons** génétiques entre des fragments de chromatides homologues et qui correspondent aux **crossing-over** ;
  - la séparation progressive des chromosomes homologues qui restent associés au niveau d'**enjambements** ou de **chiasmats**. Ceux-ci sont localisés aux endroits où se sont préalablement produits des crossing-over. Ces derniers sont à l'origine des **brassages intrachromosomiques**.

La prophase I s'achève avec la dissociation de l'enveloppe nucléaire.

- **Métaphase** de première division méiotique (métaphase I)
 

Les bivalents se disposent sur la plaque équatoriale. À la différence d'une métaphase de mitose, les kinétochores des deux chromatides d'un chromosome dupliqué sont reliés au même pôle du fuseau. Les centromères des deux homologues ne sont donc pas situés dans le plan équatorial mais de part et d'autre de celui-ci, qui contient en revanche les chiasmats.

La dissolution de chiasmats par clivage protéolytique de molécules d'adhésion permet la séparation des homologues de chaque paire à l'anaphase I, alors que les deux chromatides de chaque chromosome restent fermement liés. La disposition aléatoire des chromosomes maternels et paternels de chaque paire à la métaphase et la **disjonction indépendante** de chacune à l'anaphase déterminent un **brassage interchromosomique**.
- La **télophase** I est marquée par la séparation des deux cellules filles et la reconstitution de l'enveloppe nucléaire autour du lot de chromosomes.

La seconde division méiotique se déroule conformément à ce qui a été décrit pour la mitose.

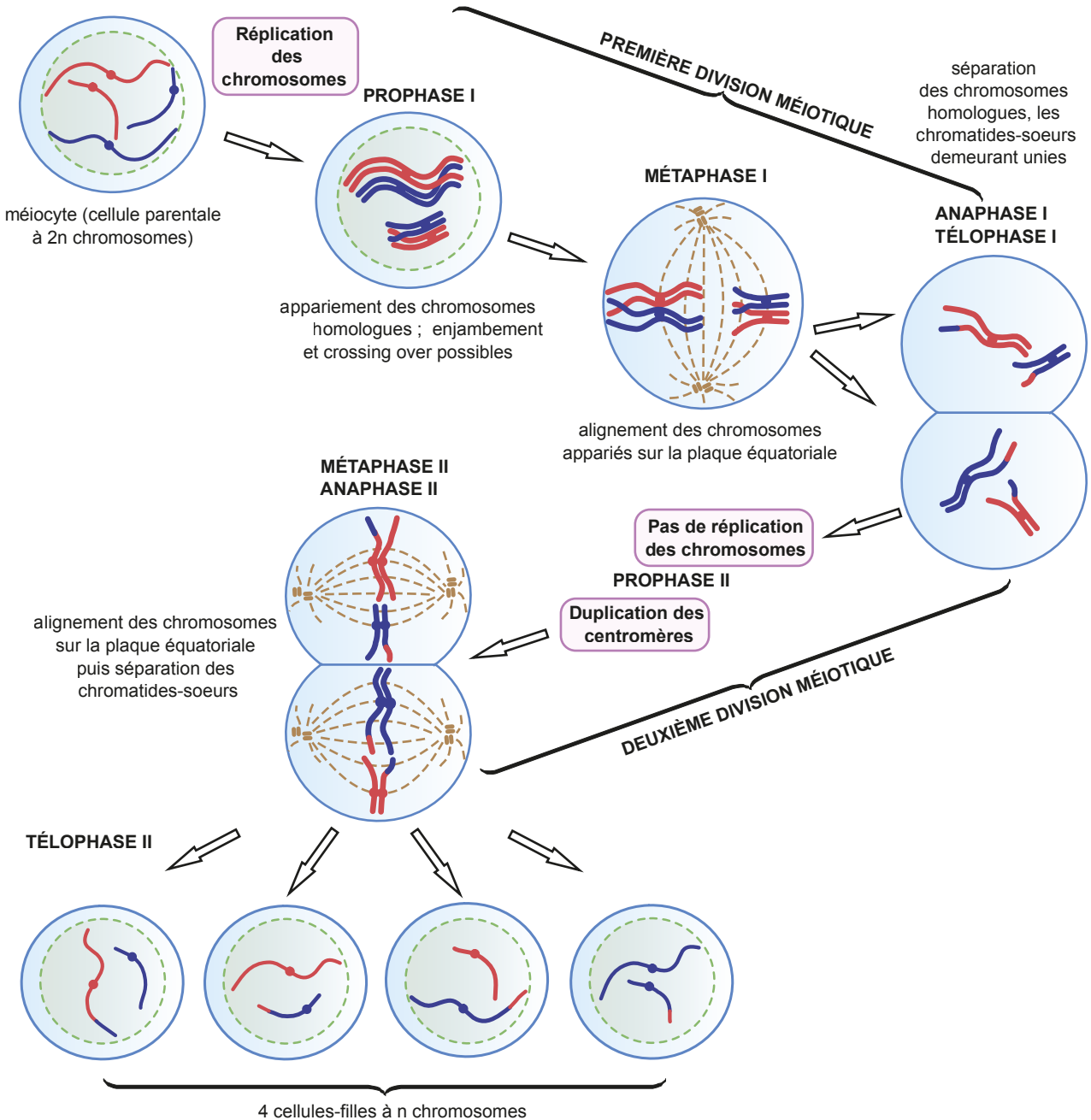


Figure 13.14 Les caractères de la méiose.

## ZOOM 1

**Le séquençage par la méthode de Sanger (1977)**

Cette méthode, historique, a permis de réaliser les premiers séquençages. Elle est dite « par terminaison de chaîne ». La longueur des fragments séquencés est de l'ordre de 500 à 700 pb.

Le fragment d'ADN à séquencer est d'abord amplifié par PCR (zoom 7) puis dénaturé, ce qui produit le fragment simple brin dont on cherche à déterminer la séquence (figure a). Ce fragment est utilisé comme matrice de brins alors néosynthétisés.

La synthèse de ces brins néosynthétisés complémentaires au brin matrice est effectuée en utilisant une amorce. L'élongation se fait dans le sens 5' → 3' grâce à une ADN polymérase. On introduit dans le milieu les 4 types de désoxyribonucléotides-triphosphates (dNTP : dATP, dCTP, dGTP, dTTP).

On introduit également dans le milieu réactionnel des nucléotides modifiés appelés didésoxyribonucléotides (ddNTP) qui ont un atome d'oxygène à la place du groupement OH sur le carbone 3' du ribose. Ils peuvent donc être incorporés dans un brin d'ADN en cours de synthèse, mais ils ne permettent pas qu'un autre nucléotide soit incorporé après eux. Lorsqu'un ddNTP est incorporé à la place d'un dNTP, l'ADN polymérase ne peut plus continuer sa polymérisation et l'élongation s'arrête (figure b).

Statistiquement au cours de l'élongation, au moins une fois un ddNTP complémentaire sera incorporé à la place d'un dNTP, cette incorporation dépendant notamment du rapport des concentrations entre les dNTP et les ddNTP introduits dans le milieu de réaction. La terminaison s'effectue donc de manière statistique sur toutes les portions possibles et on obtient en fin de réaction un ensemble de copies de longueurs différentes, susceptibles d'être interrompues au niveau de chacun des nucléotides.

Ces fragments sont alors séparés selon leur taille par électrophorèse, deux bandes consécutives exprimant une différence de taille d'un seul nucléotide.

Lors des premiers temps de la méthode, les ddNTP étaient introduits dans 4 milieux séparés : quatre réactions de séquençages se produisaient en parallèle dans 4 tubes distincts contenant chacun un seul ddNTP et l'amorce était rendue radioactive pour localiser la position des différents fragments. La lecture se faisait alors en 4 pistes séparées, chacune révélant les interruptions au niveau d'un type de nucléotide donné (figure c).

Aujourd'hui on utilise des ddNTP dont chacun est marqué par un fluorophore spécifique. La réaction de séquençage est alors conduite en un seul milieu avec le ddCTP fluorescent bleu, le ddATP fluorescent vert, le ddGTP fluorescent jaune et le ddTTP fluorescent rouge. Les fragments interrompus à un nucléotide donné porteront donc un fluorophore terminal de couleur spécifique à ce nucléotide. Ce sont ces fluorophores qui seront détectés lors de l'électrophorèse, rendant inutile le traceur radioactif (figure d).

L'électrophorèse se fait alors dans des séquenceurs capillaires : ceux-ci sont constitués de fins tubes de verre, de quelques microns de diamètres sur plusieurs dizaines de longueur (de 30 à 50 cm) et dans lesquels sont introduits l'ensemble des fragments.

La détection de la fluorescence se fait par un rayon laser qui traverse le capillaire qui excite les ddNTP fluorescents incorporés à l'ADN. Une caméra mesure l'émission de fluorescence, les différents ddNTP se distinguant par la longueur d'onde émise.

On obtient ainsi un enregistrement analogue à celui présenté page suivante.

Un capillaire permet la lecture d'environ 300 kb en 3 heures. Les séquenceurs utilisés actuellement comprennent de très nombreux capillaires.

On obtient ainsi la séquence du brin néosynthétisé :

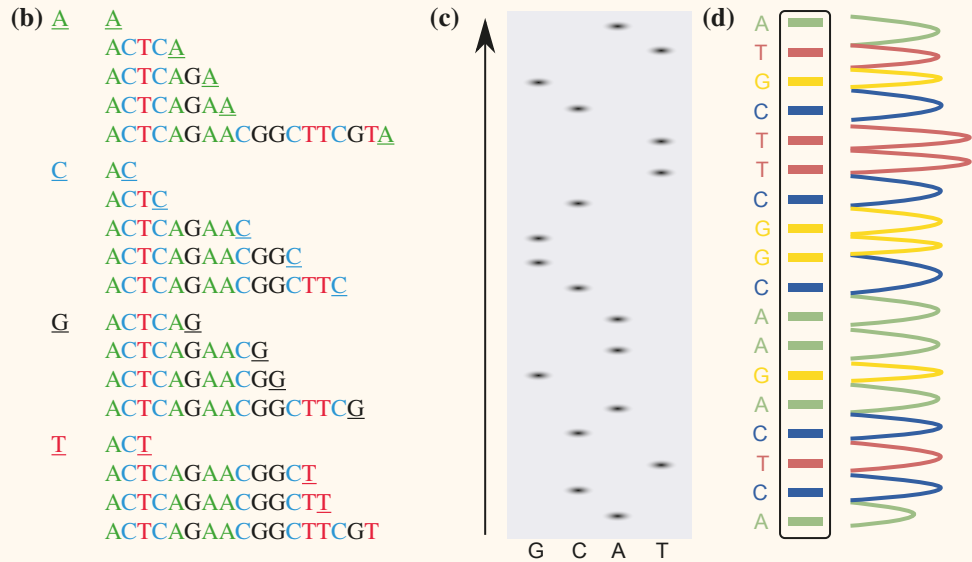
5' ACTCAGAACGGCTTCGTA 3'

Celle-ci est complémentaire de la séquence matrice recherchée (figure a) qui est alors déterminée.

(a) Fragment à séquencer :

3' TGAGTCTTGCCGAAGCAT 5'

(b) obtention de brins néosynthétisés de longueurs différentes après introduction de ddNTP des différents types (A, C, G, T)



(c) séparation des fragments par amorce radioactive

(d) séparation des fragments par détection des fluorophores

Les différentes étapes de séquençage par la méthode de Sanger.

## ZOOM 2

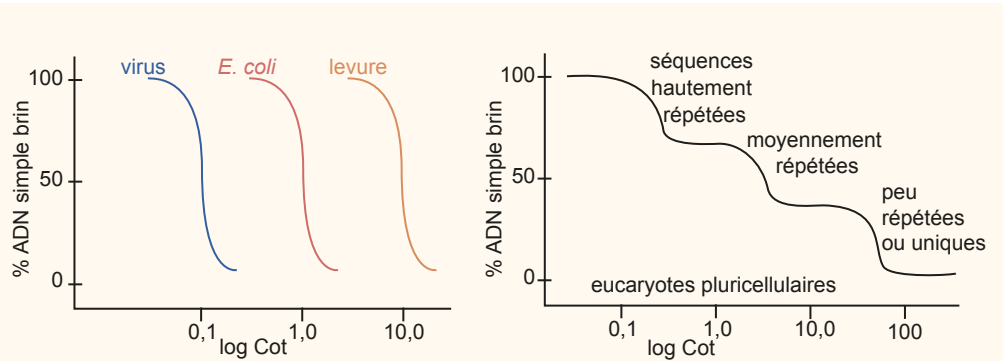
### Les courbes de Cot : les différents types de séquences

Ces courbes traduisent le suivi de la renaturation de l'ADN après coupure de celui-ci en fragments et dénaturation par la température (séparation en fragments monobrin)

L'axe des ordonnées traduit le pourcentage d'ADN sous forme monobrin (initialement 100 %). L'axe des abscisses correspond au produit, en échelle logarithmique, de la concentration initiale d'ADN ( $C_0$  en moles.L<sup>-1</sup>) par le temps écoulé (t), la réassociation dépendant de ces deux paramètres.

La courbe permet de suivre la renaturation du mélange initial de fragments, placé dans des conditions ménagées de refroidissement (65 °C). La renaturation est suivie par l'absorption aux UV, plus forte lorsque les bases azotées ne sont pas appariées et peut être ainsi quantifiée.

La renaturation découle de la probabilité qu'un fragment d'ADN retrouve dans le mélange un fragment complémentaire avec lequel s'apparier. Elle est d'autant plus rapide qu'elle est acquise pour un Cot faible.



Courbes de Cot pour les virus et les organismes unicellulaires (a) et pour les organismes pluricellulaires (b).

Dans le cas des procaryotes et des virus, la renaturation apparaît rapide et se produit de manière régulièrement continue. Elle apparaît plus rapide pour les virus que pour les bactéries, ce qui traduit des génomes de petite taille.

Dans le cas des eucaryotes pluricellulaires, la courbe de Cot présente des caractères très différents, avec trois paliers révélant trois ensembles de séquences, de cinétiques de réassociations différentes :

Le premier palier correspond à des séquences de renaturation rapide : un fragment donné retrouve aisément un autre brin de séquence complémentaire. Ceci atteste de l'existence de séquences très nombreuses constituant un ADN hautement répété.

Le second palier traduit une renaturation plus lente de fragments se rattachant à de l'ADN moyennement répété.

Le troisième palier signifie une renaturation très lente affectant alors des séquences peu répétées ou uniques.

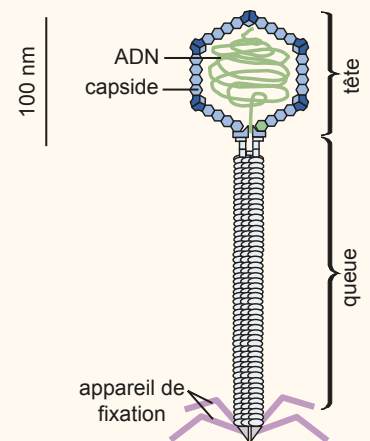
### ZOOM 3

#### Le bactériophage lambda

Le bactériophage lambda (bactériophage  $\lambda$ ) est un virus qui affecte la bactérie *Escherichia coli*. Son génome est constitué d'un ADN double brin linéaire de 48,5 kb contenant environ 70 gènes et dont les extrémités sont cohésives (forment des bouts collants). Elles permettront la circularisation de l'ADN une fois celui-ci injecté dans la bactérie. La capside comprend une tête icosaédrique prolongée d'une queue munie de fibrilles assurant la fixation sur la bactérie.

Après introduction dans la bactérie, le phage peut présenter deux types de cycles :

- Un cycle lytique : l'ADN circulaire se réplique de nombreuses fois et l'expression du génome produit toutes les protéines nécessaires à l'assemblage de nouveaux virions. La production de ceux-ci entraîne la lyse, et donc la mort, de la bactérie.



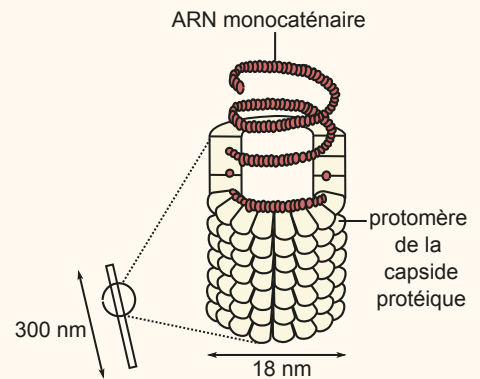
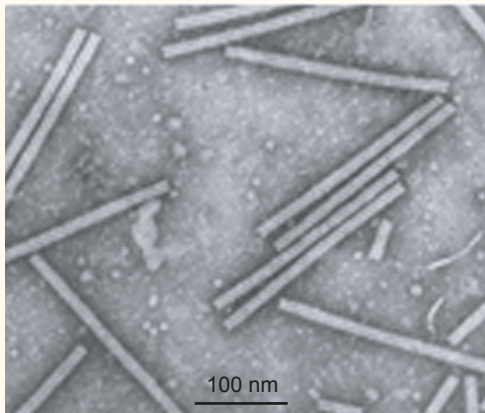
- Un cycle lysogène : le génome viral s'intègre alors dans le chromosome bactérien au niveau de séquences spécifiques par une intégrase virale et est transmis d'une bactérie à l'autre lors des divisions. Le phage reste à l'état latent au sein du génome bactérien (état de prophage). Sous l'effet d'une induction, perçue comme un stress par la bactérie (ultraviolets, carence nutritive...), le génome viral s'excise du chromosome bactérien et engage un cycle lytique qui s'achève avec la lyse bactérienne.

L'entrée en phases lytique ou lysogène dépend de l'équilibre entre des protéines virales, certaines constituant des répresseurs de la phase lytique.

## ZOOM 4

### Le virus de la mosaïque du tabac

Le virus de la mosaïque du tabac se marque au niveau des jeunes feuilles par l'apparition de marbrures ; des cloques déforment ensuite le limbe qui se recroqueville. Une infection précoce réduit la croissance de la plante. Le virus peut se maintenir dans le sol, sur les débris de feuilles et de racine pendant plusieurs années. Les contaminations peuvent se faire par l'intermédiaire d'outils ou par les racines des plants contaminés. Il peut affecter différents types de plantes.



#### (a) Electronographie du virus et (b) feuille de tabac infectée.

(a) Source : 1994 Rothamsted Experimental Station  
<https://www.rothamsted.ac.uk/our-science>. (b)  
 Cliché : D. Blancard. <http://ephytia.inra.fr/fr/C/10885/Tabac-Principaux-symptomes>

Le virion apparaît comme un bâtonnet de 300 nm de long, avec 2 300 protéines de capsid enroulées autour de l'ARN. Le génome est représenté par une molécule d'ARN simple brin positif de 6 395 nucléotides, 3 nucléotides étant alors liés à une protéine de la capsid. Il code pour un nombre limité de protéines : une réplicase, la protéine de capsid et une protéine de mouvement.

L'ARN simple brin positif correspond à un ARN messager : après entrée dans la cellule, des ribosomes de la cellule hôte se lient à l'extrémité 5' de l'ARN viral et facilitent le processus de désencapsidation. Le génome peut être traduit dès son entrée dans la cellule, avec la production d'une enzyme spécifiquement virale, l'ARN réplicase ou ARN polymérase ARN dépendante.

On obtient alors une population d'ARN simple brin négatifs qui permettent à leur tour la synthèse de nouveaux ARN+ qui servent pour la traduction des protéines et la constitution de nouveaux virions.

Deux types de protéines sont synthétisés : les protéines de la capsidie qui présentent ensuite la particularité de s'auto-assembler en structure hélicoïdale et une protéine de mouvement du virus. Celle-ci assure le mouvement des particules virales dans la cellule végétale, en interaction avec les structures cellulaires (cytosquelette, réticulum endoplasmique) et participe au franchissement des plasmodesmes. Le virus peut ainsi se déplacer de cellules en cellules par la voie symplasmique et gagner les cellules compagnes des tubes criblés avant de circuler ensuite à longue distance dans le phloème.

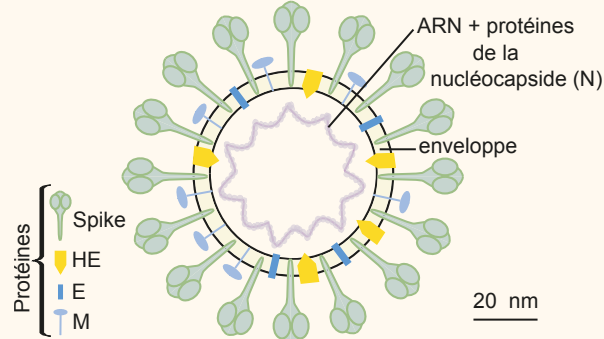
## ZOOM 5

### Les coronavirus

Les coronavirus sont des virus enveloppés et sphériques de 100 à 160 nm de diamètre. Leur nom est lié à l'aspect des virions en microscopie électronique qui présentent une couronne périphérique de protéines faisant saillie. Leur génome est à ARN simple brin positif de 27 à 32 kb, ce qui les situe parmi les plus gros génomes des virus à ARN.

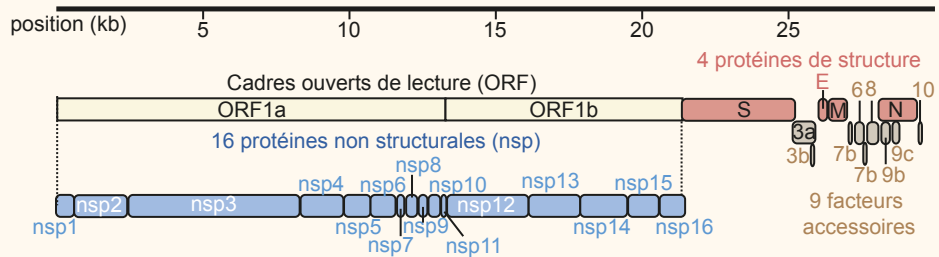
Les coronavirus sont susceptibles d'infecter l'homme ; quelques-uns (4) sont responsables de rhumes saisonniers, mais trois d'entre eux sont responsables de graves pneumopathies, à l'origine d'épidémies plus ou moins développées :

- le SARS-CoV, virus responsable du SARS (*Sever acute respiratory syndrome* – syndrome respiratoire aigu sévère) en Chine en 2002-2004 ;
- le MERS-CoV, virus responsable du MERS (*Middle East Respiratory Syndrome* – syndrome respiratoire) au Moyen-Orient en 2012 ;
- le SARS-CoV-2, responsable de la maladie CoviD-19 (*Corona Virus Disease 2019*) à l'origine d'une pandémie en 2020 et 2021.



Structure du SARS-coV-2.

Le génome du SARS-CoV-2 est un ARN de 30 kb qui code notamment pour 4 protéines de structure (S, E, M, N) dont la protéine S (pour spike) qui permet l'entrée du virus dans les cellules épithéliales en se fixant sur une molécule de celles-ci et la protéine N (nucléoprotéine) : elle se lie à l'ARN et forme autour de lui la nucléocapside.



Génome du SARS-coV-2.

Le génome comprend aussi deux ORF, cadres ouverts de lecture (*open reading frame*) : Chacun correspond à une séquence codante non interrompue par un codon stop. La traduction produit deux longs polypeptides ensuite scindés en protéines non structurales (nsp) impliquées dans la réplication du virus. Parmi celles-ci se trouve l'ARN polymérase ARN dépendante (nsp12) qui synthétise les brins ARN<sup>-</sup> et ARN<sup>+</sup>. La vitesse de catalyse de cette enzyme est très rapide, ce qui pourrait être mis en relation avec la grande taille du génome, inhabituelle pour un virus à ARN. La transcription est alors peu fidèle et un autre polypeptide (nsp14) est représenté par une ARN exonucléase, qui exerce une fonction de correction sur les ARN transcrits, ce qui est un caractère remarquable des coronavirus.

Lors de l'émergence de la maladie, les malades observés présentaient un taux d'identité du génome de 99,98 %, ce qui conduit à formuler l'hypothèse que cette souche virale ne résidait pas depuis longtemps chez l'homme. Ce génome apparaît par ailleurs identique à 96 % identique à celui d'un virus retrouvé chez les chauve-souris. Les chauves-souris pourraient constituer le réservoir initial dont aurait émergé récemment le virus. La Covid-19 serait due à un franchissement récent de la barrière d'espèce par le coronavirus ce qui correspond à une zoonose (coronavirus zoonotique). L'une des difficultés est alors de comprendre comment s'est effectué ce franchissement en l'absence de contact direct entre la chauve-souris et l'homme, ce qui suggère l'intervention d'autres organismes animaux intermédiaires (exemples de la civette pour le SARS-CoV, le dromadaire pour le MERS-CoV).

## ZOOM 6

### Durée et contrôle des cycles cellulaires

#### Durée des cycles cellulaires

La connaissance expérimentale des étapes dans lesquelles se trouve une population de cellules se fonde sur la mesure de la quantité d'ADN présente dans les cellules et qui double pendant la phase S. Cette estimation est permise par l'utilisation de colorants fluorescents se fixant à l'ADN, la fluorescence des cellules étant ensuite mesurée par cytométrie de flux.

Les cellules sont ainsi séparées en trois lots de fluorescence différente fonction de leur quantité d'ADN : celles qui ont un ADN complet non répliqué et qui sont en phase G1, celles qui ont un ADN entièrement répliqué et qui se trouvent en phases G2 ou M, et celles, en phase S, qui ont une quantité intermédiaire d'ADN.

Les quantités relatives des cellules de trois lots montrent qu'il y a davantage de cellules en phase G1 qu'en phases G2 et M, ce qui traduit une phase G1 plus longue que G2 + M.

### Contrôle des cycles cellulaires

Les mécanismes de contrôle peuvent être appréhendés par des expériences de fusion entre des cellules dont les noyaux sont à différentes étapes du cycle. On suit alors l'évolution des noyaux ce qui met en évidence des facteurs cytoplasmiques influant sur le passage d'une phase à l'autre.

États des noyaux des cellules fusionnées	Phase S		Phase G1	Phase M
	+ phase G1	+ phase G2	+ phase G2	+ phases G1, ou S ou G2
<b>Résultat</b>	Le noyau en G1 entre en S	Le noyau en G2 reste en G2	Le noyau en G2 reste en G2, le noyau en G1 suit son rythme en S	Les noyaux en G1, S ou G2 entrent prématurément en début de phase M
<b>Conclusion</b>	Le cytoplasme en S contient des facteurs déterminant l'entrée en S	Le noyau en G2 reste réfractaire aux facteurs produits en S	Les facteurs produits en S ont disparu en G2	Le cytoplasme en M contient un facteur promoteur de la phase M (appelé MPF)

## ZOOM 7

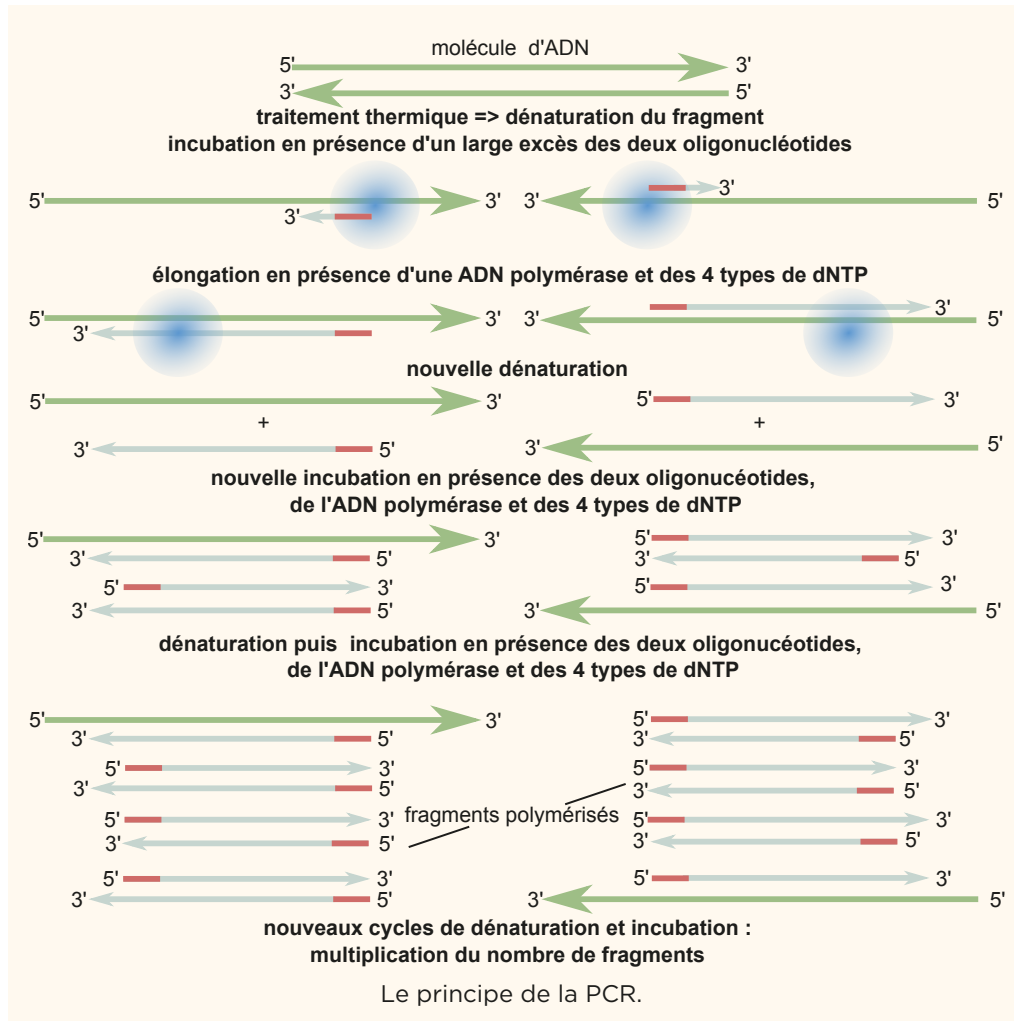
### Le principe de la PCR

La réaction de polymérisation en chaîne (PCR : *polymerase chain reaction*) est utilisée pour obtenir *in vitro* un très grand nombre de copies d'un fragment d'ADN.

La PCR comporte les étapes suivantes qui forment un cycle répété x fois.

- dénaturation thermique des fragments d'ADN vers 95 °C qui sépare les brins ;
- fixation vers 60 °C d'oligonucléotides spécifiques, complémentaires des extrémités 3' et 5' du fragment à amplifier et à valeur d'amorces pour l'ADN polymérase ;
- élongation (vers 72 °C) des brins complémentaires à l'aide de la Taq ADN-poly-mérase, enzyme isolée à partir de la bactérie thermophile *Thermus aquaticus*. Cette enzyme a une très bonne stabilité thermique, autorisant une succession de cycles sans refroidissement.

Les premiers fragments sont obtenus au troisième cycle de dénaturation-incubation. Ceux-ci permettent une amplification rapide et importante de leur nombre lors des cycles suivants.



## Réviser

### Résumé

L'ensemble des molécules d'ADN contenues dans une cellule forme le génome. Le génome des eubactéries est porté par le chromosome circulaire et les plasmides. Il est compact car de petite taille et riche en séquences codantes, avec souvent des gènes groupés dans un opéron. Chez les eucaryotes, existent le génome nucléaire et le génome extranucléaire des organites semi-autonomes. Le génome nucléaire est organisé en fibres plus ou moins compactes constituant l'eucaryote et l'hétérochromatine. Il est de grande taille et l'information génétique est donc diluée. Les gènes sont souvent morcelés. Les particules virales sont des entités associant un génome à ARN ou à ADN à une capsidie protéique, entourée ou non d'une enveloppe lipidique.

La réplication s'inscrit dans l'anabolisme de l'ADN. Elle est semi-conservative et se réalise au cours de la phase S du cycle cellulaire. La réaction, catalysée par les ADN polymérase est une polymérisation de désoxyribonucléotides, dont la séquence est établie par celle du brin matrice. Elle s'effectue dans le sens 5' → 3' avec des modalités différentes, continue ou discontinue pour les deux brins anti-parallèles de la molécule. La réplication est un processus fidèle, avec un faible taux d'erreur du fait de la complémentarité des bases et du système d'autocorrection.

La mitose répartit de manière égale les chromosomes dans les cellules-filles qui sont génétiquement identiques.

La méiose correspond à la succession de deux divisions permettant de passer d'une phase diploïde à une phase haploïde et produisant selon les cas, gamètes ou méiospores.

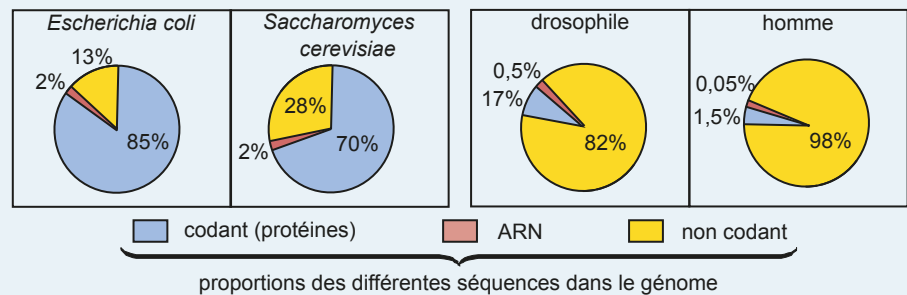


Figure de synthèse Estimations des séquences codantes et non codantes dans différents génomes.

## S'entraîner

### QCM de connaissances

- 1 Indiquez si les propositions suivantes sont exactes :
  - a. Le génome correspond à la fraction codante de l'ADN.
  - b. Le génome des bactéries est entièrement nucléaire.
  - c. L'euchromatine est un état du matériel nucléaire caractérisant l'interphase alors que l'hétérochromatine caractérise la mitose.
  - d. Le génome peut être représenté par de l'ARN.

Pour les propositions 2 à 4 sélectionnez la ou les réponses correctes (a, b et/ou c) :
- 2 À l'issue de la réplication, dans les deux molécules d'ADN produites on trouve :
  - a. A = T.
  - b. A + T = C + G.
  - c. C + T = A + G.
- 3 Il est possible de dire pour la mitose :
  - a. Il faut au moins deux chromosomes pour décrire une mitose.
  - b. La métaphase montre les paires de chromosomes homologues en vis-à-vis à l'équateur de la cellule.
  - c. Les processus de la mitose et de cytokinèse sont identiques chez les métazoaires et les embryophytes.

- 4 À la fin de la première division de méiose, les cellules obtenues contiennent :
- a.  $2n$  chromosomes à 1 chromatide.
  - b.  $n$  chromosomes à 1 chromatide.
  - c.  $n$  chromosomes à 2 chromatide.

### QCM à partir de documents

De jeunes racines de *Bellevalia* (Liliacée) sont cultivées sur un milieu de culture (milieu 1) contenant de la thymidine radioactive et la colchicine (substance qui interdit l'ascension polaire des chromosomes). La durée d'exposition dans ce milieu 1 est égale à la moitié de la durée d'une mitose. Les racines sont ensuite lavées et transférées dans un milieu 2 contenant de la thymidine non radioactive et de la colchicine.

On effectue deux prélèvements à deux moments différents : les cellules sont à chaque fois écrasées et observées par autoradiographie.

(a) plaque métaphasique d'une cellule de racine prélevée dans le milieu 2 après un temps correspondant à la durée de la fin de la mitose commencée dans le milieu 1.

(b) plaque métaphasique d'une cellule de racine prélevée dans le milieu 2 après un temps correspondant à la durée d'une mitose supplémentaire

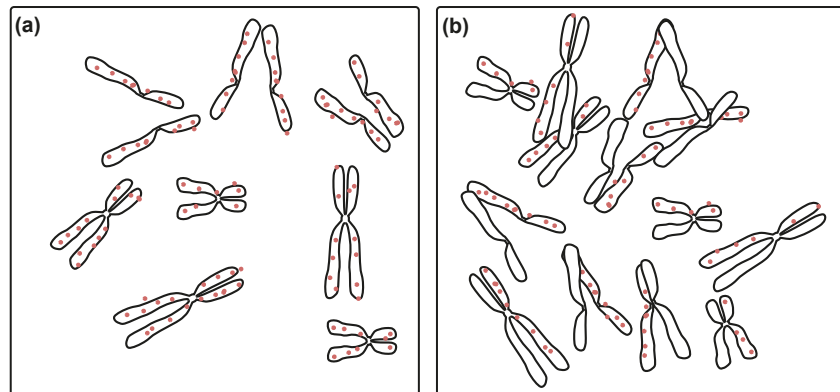


Figure 13.15 Autoradiographie.

Les points indiquent des traces de radioactivité.

- 1 Indiquez si les propositions suivantes sont exactes :
- a. À l'issue de la première mitose chaque chromatide renferme une molécule d'ADN formée d'un brin radioactif et d'un brin non radioactif.
  - b. À l'issue de la deuxième mitose, la chromatide radioactive renferme moins de radioactivité que les chromatides issues de la première mitose (on suppose que les temps d'observation sont beaucoup plus courts que la période de l'élément radioactif utilisé).
  - c. La chromatide non radioactive observée à l'issue de la deuxième mitose renferme deux brins non radioactifs issus de la mitose précédente.
  - d. Cette expérience permet de conclure à une réplication semi-conservative de l'ADN.