

SV-F-1 Génome des cellules et des virus, transmission de l'information génétique (BCPST 1)	
SV-F-1-1 Organisation des génomes	
Savoirs visés	Capacités exigibles
Les virus ou particules virales sont des entités nucléoprotéiques comprenant un acide nucléique (sous forme d'ADN ou d'ARN) constituant le génome viral, et des protéines. On distingue des protéines à rôle structural, formant la capsid, et parfois des protéines à rôle enzymatique. Les virus sont très divers et possèdent parfois une enveloppe lipoprotéique.	- Comparer l'organisation du génome des bactéries, des Eucaryotes et des virus. - Estimer la proportion de séquences codantes et non codantes dans les génomes des Eucaryotes, des bactéries et des virus. - Illustrer la diversité structurale et la diversité d'hôte des virus.
<b>Précisions et limites :</b> <i>Aucune monographie de chaque virus n'est attendue. Il s'agit de montrer la diversité structurale (organisation structurale, taille, présence ou non d'une enveloppe, nature de l'information génétique) et la diversité d'hôte à l'aide de trois exemples, sans rentrer dans les détails des cycles de multiplication : bactériophage lambda, VMT, un coronavirus zoonotique.</i>	
SV-F-2 L'expression du génome (BCPST 1)	
Savoirs visés	Capacités exigibles
Les virus se multiplient en détournant la machinerie d'expression de l'information génétique de la cellule hôte.	
<b>Précisions et limites :</b> <i>On présente une vue générale de la multiplication sur un exemple au choix en se limitant aux modalités de réplication de l'information génétique, de synthèse des protéines et de formation de nouveaux virus. On se limite à distinguer la synthèse des polymérase et des protéines de la capsid, sans détailler les différentes protéines virales et les mécanismes moléculaires impliqués.</i>	

En 1883, Mayer étudie la mosaïque du tabac et constate que l'aspersion de sève d'une plante malade contamine une plante saine. Pensant qu'il s'agissait de bactéries, un autre chercheur filtra cette sève avec un filtre ne laissant pas passer les bactéries. La sève était toujours contagieuse.

En 1897, Beijerinck asperge les plants avec la sève filtrée puis récupère leur sève pour contaminer d'autres plants. L'organisme pathogène ne se développait pas sur milieu gélosé et n'était pas inactivé par l'alcool (contrairement aux bactéries)

En 1935, Stanley cristallisa la particule infectieuse ou VMT puis en 1950 le cycle du batériophage est découvert.

## I. CARACTÈRES GÉNÉRAUX DES VIRUS

Les virus sont des objets biologiques particuliers, infectieux, constitués d'un acide nucléique et de protéines. Ils dépendent de cellules vivantes pour être répliqués et peuvent perturber leur information génétique. Ce sont des **parasites intragénétique obligatoires**.

On peut dégager quelques caractères communs :

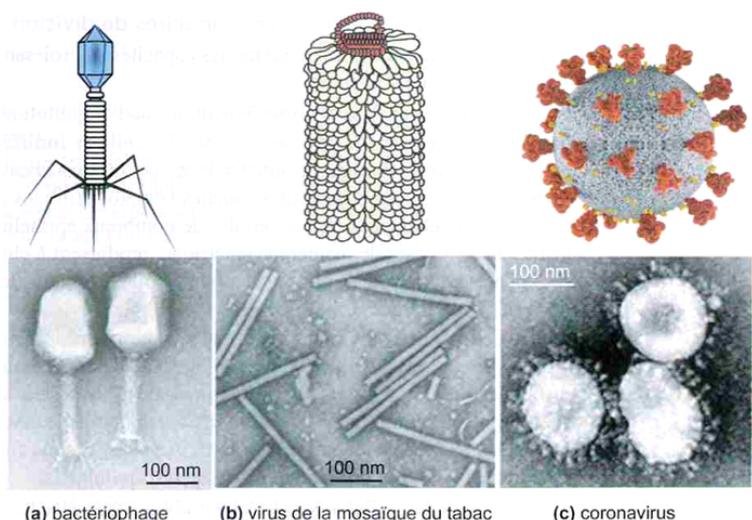
- Une taille, très variable (10 à 400 nm), inférieure à celles des cellules.
- Un génome sous forme d'ADN ou d'ARN, généralement de petite taille très inférieure au génome cellulaire
- Une réplication uniquement au sein de cellules hôtes, procaryote (bactériophage) ou eucaryote animal (coronavirus, VIH..) ou végétal (VMT)
- À l'extérieur des cellules, ils sont sous forme de particules virales libres où **virions**, biologiquement inerte. Ces virions sont formés d'une capsid protéique parfois entourée d'une enveloppe lipidique et renferment de l'acide nucléique avec éventuellement quelques protéines enzymatiques.

La diversité structurale des virus réside à la fois dans les enveloppes et dans le matériel génétique enfermé dont la nature conditionne le cycle de reproduction du virus au sein des cellules.

Chaque sorte de virus ne peut infecter et parasiter qu'une gamme limitée de cellules hôtes que l'on appelle **spectre d'hôtes**.

Cette spécificité d'hôtes est due à des mécanismes de reconnaissance entre particules virales et cellules hôtes.

Figure 1 : Les différents types de virus (in Biologie-Géologie 1 Dunod, 2021)



## II. DIVERSITÉ GÉNOMIQUE DES VIRUS

On distingue des virus à **ADN**, à **ARN** et les **rétrovirus**.

Les virus à ADN affectant les bactéries (98 %) et la totalité des archées. Les cellules eucaryotes hébergent à peu près égale les virus à ADN et à ARN, ces derniers étant responsables de nombreuses pathologies et d'épidémies. Une grande majorité des virus de végétaux sont à ARN.

La production de nouveaux virions impose toujours une phase de synthèse de protéines à partir de matériel génétique et sa réplication ce qui conduit à distinguer différentes situations.

### II.1. Des génomes à ADN ou à ARN

#### Les virus à ADN

Virus à ADN double brin. C'est le cas des **bactériophages** (virus n'infectant que des bactéries) : l'ADN linéaire du virion est injecté dans la bactérie dans laquelle il se 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 : c'est le **cycle lytique**. Il dure environ 20 minutes à 37 °C et libère 200 phages /bactérie.

- L'infection par certains bactériophages dits tempérés comme le **bactériophage λ** conduit à l'intégration de l'ADN viral dans le chromosome bactérien (**prophage**) où il devient dormant : c'est le **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 de nouveaux virions (cas des parvovirus, parasite les cellules animales).

#### Les virus à ARN

- Virus à ARN + : l'ARN introduit dans la cellule hôte peut être immédiatement traduit dans les ribosomes de celle-ci. Le génome viral est alors équivalent à des ARNm : cas du **VMT** et des **coronavirus**.

- Virus à ARN - : l'ARN doit être transcrit en son complémentaire pour être traduit en protéines : cas du virus de la Grippe (influenza A).

- **Rétrovirus** : l'ARN est de type + mais il 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. Celui-ci est intégré dans le génome de l'hôte où il sert de support d'information pour la synthèse de protéines virales. La réplication par les polymérase de l'hôte permet de générer de nouveaux virions : cas du VIH.

### II.2. Des génomes compacts

Les virus à ADN ont le génome le plus long (44 kb), les virus à ARN sont plus courts (environ 10 kb) sauf les coronavirus (30 kb).

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

Ces capacités sont renforcées 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 3 cadres de lecture différents à l'origine de 3 protéines différentes

Le nombre de gène est variable : 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**, les plus nombreux, impliqués dans la synthèse des protéines de la capsid et des enveloppes.
- **Gènes informationnels** impliqués dans la réplication et l'expression génétique (transcription, traduction).
- **Gènes** plus accessoires impliqués dans les réactions métaboliques ; ils résulteraient de transfert horizontaux (THG) à partir des cellules hôtes.
- de nombreux gènes restent par ailleurs non caractérisés à l'heure actuelle.

Lors de la réplication d'un virus, la proportion d'erreurs est inversement proportionnelle à la longueur de son génôme. Cette proportion est considérable pour un virus. L'infection d'une cellule hôte commence classiquement avec 1 virus et il y en a 1 milliard au bout de 30 générations. Il se produit une erreur par génération en moyenne. Ainsi 2 virus prélevés au hasard dans un organisme infecté différent en moyenne par 30 nucléotides. Ceci a été observé sur le VIH à l'Institut Pasteur de Paris.

### II.3. Diversité des enveloppes

Le génome est toujours enveloppé et protégé dans une structure protéique : la **capsid**. Celle-ci est constituée d'un grand nombre de copies d'un ou de plusieurs types de protéines organisées de manière très structurée. Codées par le génome viral, ces protéines présentent une certaine diversité :

- Des **capsides hélicoïdales** (VMT). Les particules virales sont allongées en filament (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 **capsides icosaédriques** (polyèdre à 20 faces triangulaires équilatérales) : cas des adénovirus.

Les sous-unités protéiques de la **capside** ou **capsomères** sont d'un seul type. Chez le VMT, la capsidie protéique est formée par plus de 2000 capsomères identiques. Cette monotonie peut être expliquée par la pauvreté de l'information génétique virale qui code peu de protéines différentes.

Les **bactériophages** montrent une organisation complexe associant une tête icosaédrique et une queue hélicoïdale. La tête renferme l'ADN ; la queue se termine par un plateau protéiques porteur d'épines et de fibres protéiques qui interviennent lors de la fixation du virion sur la paroi bactérienne.

Certains virus de forme sphérique dont les coronavirus sont dits enveloppés. La capsidie 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 sont des virus nus du fait de la présence de la paroi squelettique et soient de structure plus simple que les virus infectant des animaux.

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

Les enveloppes ont donc 2 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 l'extérieur ou du digestif et réduit le pouvoir infectieux par rapport au virus nus.

### Conclusion

Un virus est une particule infectieuse à ADN ou ARN. Son génome, pauvre en informations, code un petit nombre de molécules différentes. C'est un parasite génétique intracellulaire.

Il existe d'autres structures proches mais différentes :

- Transposons : éléments génétiques mobiles, infectieux mais non particuliers.
- Viroïdes : ARN nu sans protéine qui infectent certaines plantes.
- Prions : agents infectieux, particuliers et protéiques mais sans acide nucléique (maladie de la vache folle)

La connaissance des virus présente un certain nombre d'intérêts :

- Expérimental : introduction de gènes dans les manipulations génétiques (transfections).
- Médical : étude du processus d'oncogénèse (cancérisation) par amplification de la réplication de l'ADN de la cellule infectée en apportant un oncogène, gène qui contrôle le début de la mitose.

Quelle est l'origine des virus ?

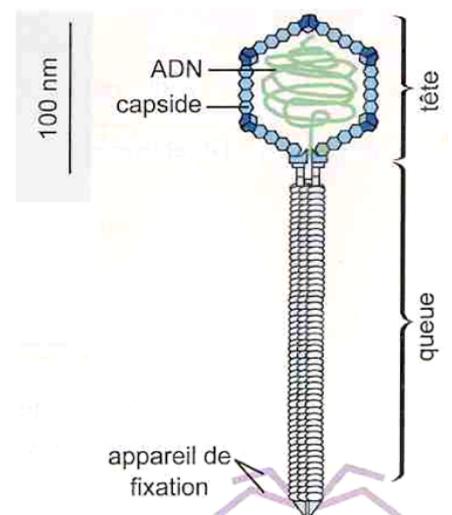
- Transposons dont la réplication ressemble à celle de certains virus ?
- Plasmides car leur réplication est indépendante de celle de l'ADN de la cellule ?

### FOCUS SUR CHAQUE VIRUS (Retenez un exemple au choix comme le demande le programme)

Le **bactériophage λ** c'est un virus qui affecte la bactérie *Escherichia coli*. Son génome est constitué d'un ADN double brin linéaire de 48,5 kb compteur environ 70 gènes et dont les extrémités sont cohésives (bouts collants). Elles permettront la circularisation de l'ADN une fois celui-ci injecté dans la bactérie. La capsidie 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 2 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. Leur production entraîne à 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 l'intégrase virale et être transmis d'une bactérie à l'autre lors de divisions. Le phage reste à l'état latent au sein du génome bactérien (prophage). Sous l'effet d'une induction perçue comme un stress par la bactérie (UV...), le génome viral s'excise du chromosome bactérien et engage ainsi un cycle lytique qui s'achève avec la lyse bactérienne. L'entrée en phase lytique ou lysogène dépend de l'équilibre entre les protéines virales, certains constituants des répresseurs de la phase lytique.



Le **virus de la mosaïque du tabac** (VMT) provoque des cloques qui déforment le limbe de la feuille ; une infection précoce réduit la croissance de la plante. Le virus peut se maintenir dans le sol sur les débris de feuilles et de racines pendant plusieurs années. Les contaminations peuvent se faire par des outils ou des racines de plants contaminés et affecter différents types de plantes.

Le virion est un bâtonnet de 300 nm de long avec 2300 protéines de capsidie enroulées autour de l'ARN simple brin positif de 6395 nucléotides. Il code pour un employé limité de protéines : une réplicase, la protéine de capsidie et une protéine de mouvement.

L'ARN + 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 répliqueuse ou ARN polymérase ARN dépendante. Il en résulte une population d'ARN - 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 synthétisées : les protéines de la capsidie qui s'auto-assemblent en structure hélicoïdale et une protéine de mouvement du virus, qui assure le mouvement des particules virales dans la cellule végétale en interaction avec le cytosquelette et participe aux franchissements des plasmodesmes. Le virus peut ainsi se déplacer de cellule en cellule par la voie symplasmique et gagner les cellules compagnes des tubes criblés avant de circuler dans le phloème.

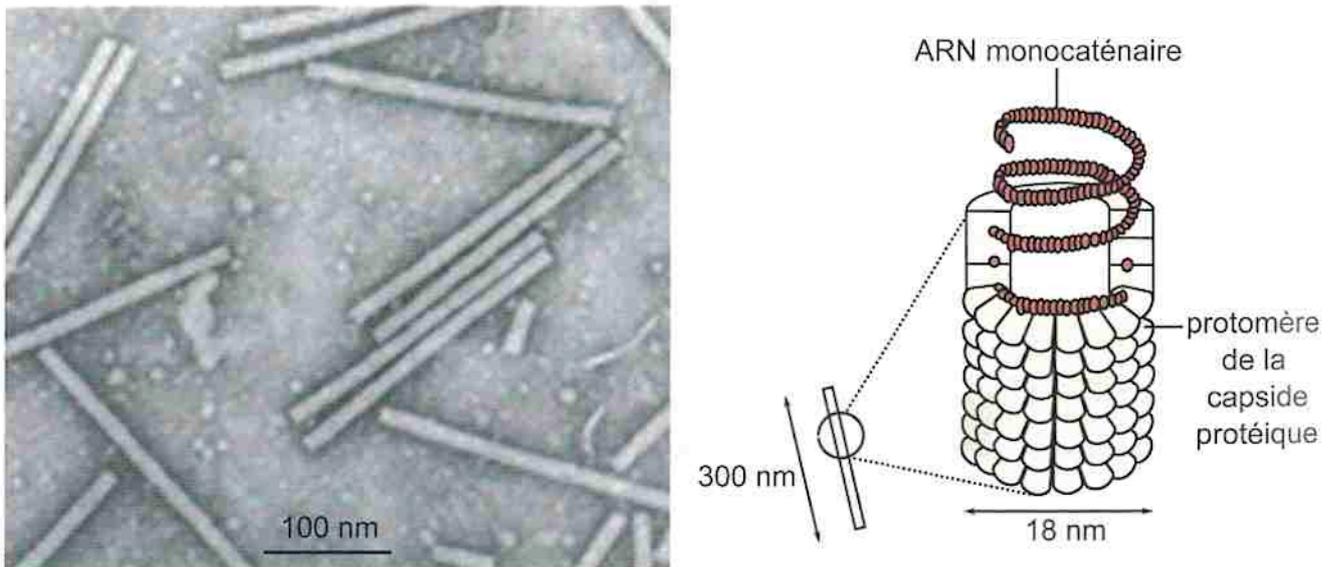


Figure 2 : Électronographie du VMT et schéma d'organisation (in Biologie-Géologie 1 Dunod, 2021).

**Les coronavirus** Ce sont des virus enveloppes et sphériques de 100 à 160 nm de diamètre. Leur nom est lié à l'aspect des virions au microscopie électronique qui présente une couronne périphérique de protéines en relief. Leur génome est à ARN +, de 27 à 32 kb, ce qui le situe parmi les plus gros génomes des virus à ARN. Les coronavirus sont susceptibles d'infecter l'homme ; quelques-uns sont responsables de Rémi saisonnier mais 3 d'entre eux sont responsables de graves pneumopathie à l'origine d'épidémie :

- Le SARS-CoV, virus responsable du syndrome respiratoire aigu sévère (Chine en 2002-2004).
- Le MERS-CoV, responsable du syndrome respiratoire (Moyen-Orient en 2012).
- Le SARS-CoV 2 responsable de la COVID-19 à l'origine de la pandémie en 2020 et 2021 .

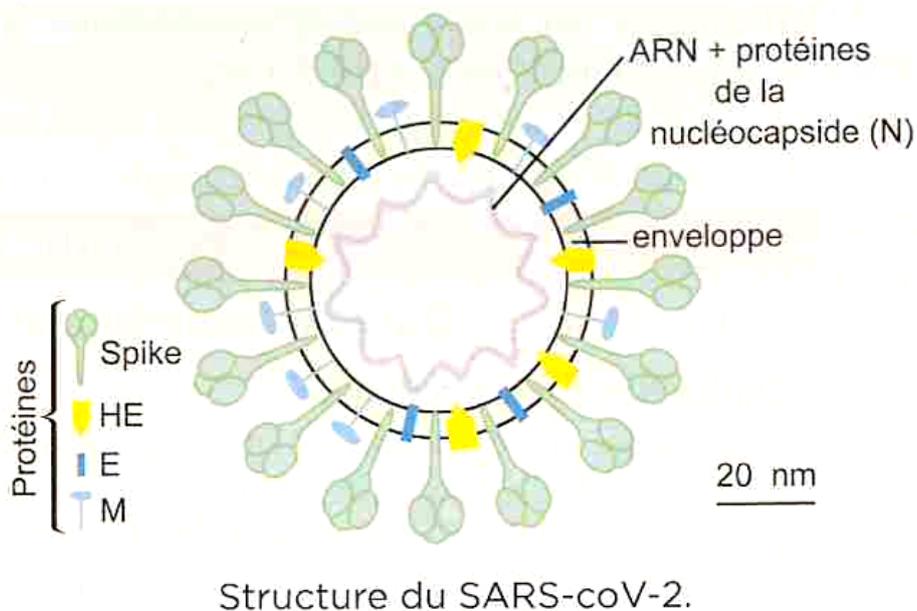


Figure 3 : Structure du SARS-CoV-2 (in Biologie-Géologie 1 Dunod, 2021).

Le génome du SARS-CoV 2 est un ARN de 30 kb qui code notamment pour 4 protéines de structure dont la protéine S (Spike) qui permet l'entrée du virus dans les cellules épithéliales et la protéine N (nucléoprotéine) : elle se lit à l'ARN et forme autour de lui la nucléocapside.

Le génome comprend aussi deux ORF, chacun correspond à une séquence codante non interrompue par un codon stop. La traduction produit 2 longs peptides ensuite scindés en protéines non structurales, impliqué dans la réplication du virus dont l'ARN polymérase ARN dépendante (nsp 12) qui synthétise les brins ARN – et ARN +. La vitesse de catalyse de cet enzyme est très rapide et la transcription est peu fidèle. Un autre polypeptide (nsp 14) représenté par une ARN exonucléase exerce une fonction de correction sur les ARN transcrits, caractère remarquable des coronavirus.

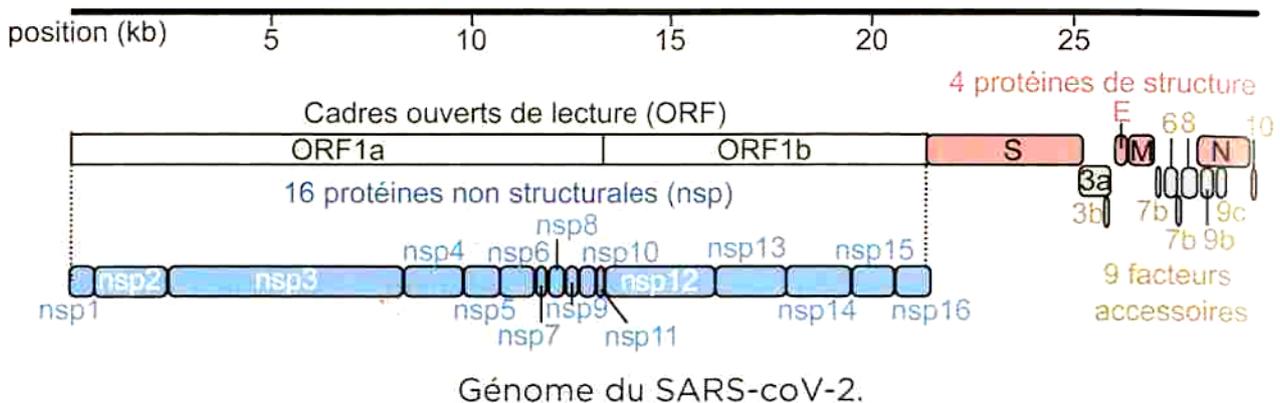


Figure 4 : Génome du SARS-CoV-2 (in Biologie-Géologie 1 Dunod, 2021).

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% à celui d'un virus qui se retrouve chez les chauve-souris. Elles pourraient ainsi constituer le réservoir initial dont aurait émergé le virus.

La COVID-19 serait due à un franchissement récent de la barrière d'espèces par le coronavirus qui représente ce qui correspond à une zoonose (coronavirus zoonotique). L'une des difficultés est de comprendre comment s'est effectué ce franchissement en absence de contact direct entre la chauve-souris et l'homme, ce qui suggère l'intervention d'autres organismes animaux intermédiaires (civette pour le SARS-CoV et le dromadaire pour le MERS-CoV).

**CORONAVIRUS ET COURSE AUX ARMEMENTS** (Source : *Souvent bénins, parfois mortels : comment fonctionnent les coronavirus ? SIMON MAKIN, Thème n° 21, Pour La Science 25/03/2020 : La science face au défi du coronavirus, modifié*)

Une autre caractéristique importante des coronavirus est que des protéines « accessoires » semblent être impliquées dans l'évitement de la réponse immunitaire innée de l'hôte. La réponse immunitaire est déclenchée lorsqu'une cellule détecte un envahisseur et libère des protéines appelées interférons, qui interfèrent avec la réplication de l'agent pathogène. Les interférons déclenchent une cascade d'activité antivirale, allant de l'arrêt de la synthèse des protéines de l'hôte à l'induction de la mort de la cellule.

Bien sûr, la réponse immunitaire a pour but d'éliminer les envahisseurs, et les virus déploient des parades. Ce dernier trait semble être ce qui diffère le plus d'un coronavirus à l'autre. « Ces virus sont étroitement apparentés, mais ils ont des protéines accessoires différentes » explique Susan Weiss, ajoutant qu'ils « ont évolué pour stopper des aspects différents de la réponse immunitaire innée ». Selon certains chercheurs, les chauves-souris abritent des coronavirus parce qu'elles ne développeraient pas une réponse immunitaire aussi forte que celle des humains.

« Beaucoup de molécules de signalisation qui alertent notre système immunitaire n'existent pas chez les chauves-souris, si bien qu'elles ne tombent pas malades », avance Christine Tait-Burkard. Plutôt que de réagir énergiquement, l'organisme des chauves-souris maintient une réponse immunitaire faible mais constante, ce qui peut contribuer à l'évolution des virus. « Les chauves-souris présentent une expression constante d'interférons.

Ce processus sélectionne les virus qui échappent à cette réponse » précise la biologiste. « Les chauves-souris sont donc de très bons vecteurs de sélection pour les virus qui sont doués pour échapper au système immunitaire. »

Les protéines accessoires des virus sont cependant loin d'être totalement comprises. « On peut les supprimer de certains virus sans aucun effet sur leur capacité à se développer », dit Stanley Perlman. « On pourrait penser que si on enlève une protéine essentielle pour contrer la réponse immunitaire, celle-ci l'emportera, mais ce n'est pas nécessairement le cas. »

Certains chercheurs pensent que les protéines accessoires jouent sur la mortalité des coronavirus. Des études sur le SRAS ont montré que l'élimination d'une protéine accessoire ne modifie pas sa capacité de réplication, mais que cela le rend moins pathogène. « De nombreux virus sont encore fabriqués, mais ils semblent être moins nocifs », résume Burtram Fielding.

Enfin, les coronavirus peuvent corriger les erreurs de réplication dans une certaine mesure, mais ils « négligent » certaines régions de leur génome, explique Christine Tait-Burkard. Deux régions en particulier sont particulièrement sujettes aux mutations : celles qui codent la protéine S et les protéines accessoires. « Concernant ces deux éléments, les coronavirus évoluent donc rapidement et acquièrent la capacité de se fixer sur de nouveaux récepteurs et à échapper à de nouvelles réponses immunitaires », explique la biologiste. « C'est pourquoi les coronavirus passent si facilement d'une espèce à l'autre ».