

# Les coronavirus, biologie cellulaire et moléculaire

## *Cellular and molecular biology of coronaviruses*

Par Hubert LAUDE<sup>(1)</sup>

(communication présentée le 26 juin 2003)

### RÉSUMÉ

Un consensus semble maintenant établi selon lequel l'agent causal du Syndrome Respiratoire Aigu Sévère (SRAS) est un coronavirus nouvellement apparu. Pour tenter de répondre aux questions que pose l'émergence de ce pathogène inconnu jusqu'alors, l'auteur présente la classification et les caractéristiques des coronavirus : génome ARN simple brin de près de 30 000 bases (le plus grand génome ARN connu), originalité du mode d'expansion de l'expression de l'information génétique, organisation relativement simple des virions avec seulement 4 protéines structurales.

Les coronavirus ont un spectre d'hôte relativement étroit et un tropisme tissulaire restreint. Ils présentent une affinité particulière pour les cellules épithéliales de la sphère respiratoire ou digestive. L'auteur présente des données sur les mécanismes qui sous-tendent cette spécificité de tissu en prenant comme exemple les coronavirus du Groupe 1, auquel appartient l'agent de la gastro-entérite transmissible (TGEV). Ces virus utilisent comme récepteur l'aminopeptidase N pour pénétrer dans les cellules cibles. Cette molécule, exprimée à la surface des cellules épithéliales, est reconnue par les protéines de spicule du virus. La reconnaissance du récepteur est aussi chez ces virus un déterminant majeur de leur spécificité d'espèce. Ainsi, aucun des coronavirus animaux du groupe 1 n'est capable d'infecter des cellules humaines, dont ils ne reconnaissent pas l'aminopeptidase N.

L'auteur décrit ensuite la manière dont le coronavirus de la gastro-entérite transmissible (TGEV) a muté pour donner le coronavirus respiratoire porcin (PCRV), responsable d'une nouvelle virose respiratoire qui a diffusé de façon explosive dans l'espèce porcine. Une délétion d'environ 250 acides aminés dans la région N-terminale de la protéine de spicule a pu être associée à la perte de virulence observée et au changement de tropisme tissulaire.

Lors d'une infection par un coronavirus, s'observe une synthèse d'interféron alpha forte et précoce. Le mécanisme d'induction mis en évidence diffère de celui classiquement décrit chez les virus à ARN, puisqu'il est essentiellement indépendant de la réplication virale et met en jeu la protéine glycosylée M associée aux virions.

Le génome des coronavirus présente un taux de mutation relativement élevé, comme attendu pour un virus à ARN. Pour autant, dans le cas du virus du SRAS, la dérive génétique ne diffère guère de celle du TGEV. Les coronavirus sont également des spécialistes de la recombinaison intergénomique, faculté qui leur permettrait de reconstituer des génomes viables à partir de génomes altérés. Cependant, il ne semble pas que le virus du SRAS soit issu d'une recombinaison récente entre un virus humain et un virus animal parmi les coronavirus actuellement recensés.

Comment le virus du SRAS se positionne-t-il au sein des trois groupes présentement définis chez les coronavirus ? La comparaison de la séquence complète du génome du SRAS, maintenant disponible, avec celle du génome des autres coronavirus révèle que ce virus se situe dans une position intermédiaire entre le groupe des virus des oiseaux (par exemple le virus de la bronchite infectieuse ou IBV) et le groupe 2 des coronavirus de mammifères (ex.: coronavirus bovin). Toutefois, l'auteur fait remarquer que, sur la base d'un examen de l'organisation génomique globale du virus (incluant la distribution des gènes dits "accessoires"), le virus SRAS présente des similitudes, qui ne manquent pas d'intriguer, avec le groupe des virus aviaires. L'ensemble de ces caractéristiques justifierait l'inclusion de ce nouveau coronavirus dans un groupe distinct de ceux existants.

**Mots-clés :** coronavirus, virus du syndrome respiratoire aigu sévère, SRAS, virus de la gastro-entérite transmissible, coronavirus respiratoire porcin, tropisme, spécificité d'hôte, virulence, récepteur, génome.

(1) Unité de Virologie et Immunologie Moléculaires, Institut National de la Recherche Agronomique, 78350 Jouy-en-Josas

## SUMMARY

A consensus seems to have been reached on the pathogen responsible for SARS (Severe Acute Respiratory Syndrome) being a newly recognised coronavirus. To answer some of the questions raised by the emergence of this previously unknown pathogen, the author presents the classification and characteristics of coronaviruses: single positive RNA strand of approximately 30,000 bases (the largest known RNA genome), complex mode of expression of the genetic information, yet simple organisation of the virions with only 4 structural proteins.

Coronaviruses have a relatively small host spectrum and a narrow tissue tropism. They have a special affinity for epithelial cells in the respiratory and digestive tracts. The author presents data on the mechanisms of this tissue specificity, using as an example group 1 coronaviruses, which include the transmissible gastroenteritis virus (TGEV). These viruses use aminopeptidase N as their receptor to break into the target cells. This molecule, expressed at the surface of epithelial cells, is recognised by the virus spike proteins. Receptor recognition is a major determinant of the species specificity in these viruses. This is why none of the group 1 animal coronaviruses is able to infect human cells, whose aminopeptidase N they do not recognise.

The author describes the way the TGEV coronavirus has mutated into the Porcine Respiratory Coronavirus (PRCV), responsible for a new respiratory infection which spread explosively in swine populations. The deletion of approximately 250 amino acids in the N-terminal region of the spike protein may be associated with the reduced virulence and switch of tissue tropism.

Coronavirus infections induce early on a marked synthesis of interferon alpha. The identified induction mechanism partly differs from the mechanism generally described in RNA viruses, as it is essentially independent from viral replication, and involves the glycosylated M protein associated to the virions.

The genome mutation rate in coronaviruses is relatively high, as expected with RNA viruses. And yet, the genetic drift associated with the SARS virus is the same as in the TGEV. Coronaviruses are also specialists in intergenomic recombination, a characteristic which could allow them to reconstitute viable genomes from altered ones. However, the SARS virus does not appear to originate from a recent recombination between a human virus and one of the currently identified animal coronavirus.

To which of three known groups of coronaviruses does the SARS virus belong? The comparison of the complete genome sequence of the SARS virus, now known, with that of the other coronaviruses shows that its position is somewhere between the avian coronavirus group (e.g. Infectious Bronchitis Virus or IBV), and group 2 coronaviruses which affect mammals (e.g. bovine coronavirus). The author also points out that the organisation of the SARS virus genome (including the distribution of so-called "accessory" genes) shows intriguing similarities with that of the avian coronavirus group. All these characteristics would justify putting this new coronavirus in the group of its own, distinct from the existing ones.

**Key words:** coronavirus, severe acute respiratory syndrome virus, SARS, transmissible gastroenteritis virus, porcine respiratory coronavirus, tropism, host specificity, virulence, receptor, genome.

La plupart des chercheurs s'accordent à penser qu'un coronavirus est l'agent étiologique exclusif du SRAS. Même s'il ne s'agit pas d'une certitude absolue, il est certain que le virus associé à la pneumonie atypique est nouveau et que ce nouveau virus est un coronavirus. Cela soulève évidemment toute une série de questions : d'où vient-il ? quel est le potentiel d'évolution de ces coronavirus ? quelles sont les possibilités de transmission interspécifique ? quels sont les facteurs qui gouvernent leur pouvoir pathogène, leur tropisme de tissu, leur spécificité d'espèce ? Je vais tenter d'aborder ces questions en m'appuyant notamment sur un certain nombre de travaux réalisés à l'INRA il y a déjà quelques années et sur les données récentes du séquençage du virus du SRAS.

### • PRÉSENTATION DES CORONAVIRUS

La famille des *coronaviridae* est longtemps restée une famille monogénérique. Son nom provient de la couronne de projections tout à fait typique qui orne les virions examinés en microscopie électronique. Au fil des ans la famille s'est agrandie, en incorporant d'abord les *torovirus*, virus peu connus dont l'activité pathologique reste, pour l'instant, discrète. Puis, sur la base de considérations ayant trait à l'organisation du génome et au mode d'expression de son information génétique, il fut décidé d'y insérer les *arteriviridae*, groupe qui inclut l'un des virus émergents des années 90, le

virus du syndrome dysgénésique-respiratoire du porc (PRRSV, à ne pas confondre avec un autre virus respiratoire émergent du porc, le PRCV, dont il sera question plus loin), puis les *roniviridae*, l'ensemble formant l'ordre des *nidovirales*, le second ordre créé en virologie (figure 1).

Le génome des coronavirus est un ARN messager de polarité positive donc à simple brin, dont la taille est remarquable, près de 30 000 nucléotides (ou "bases"), ce qui en fait le plus grand génome ARN connu à ce jour (figure 2). Ce génome est néanmoins relativement compact, l'essentiel étant occupé par des séquences codantes. Les deux tiers du génome, soit près de 20 kilobases (la figure 2 n'est pas à l'échelle), sont occupés par un seul "gène", constitué en fait de 2 cadres de lecture chevauchants. Dans le tiers restant du génome se trouvent rassemblés tous les autres gènes, notamment ceux des protéines structurales des virions, dont il sera question dans un instant. Dans l'interstice laissé par ces gènes – dont l'ordre est immuable au sein des coronavirus – il existe un certain nombre de gènes additionnels qui ne sont pas indispensables *ex vivo* pour la propagation en culture, mais qui ont probablement un rôle chez l'hôte naturel. Ces gènes varient en distribution et en nombre selon les groupes et même au sein de certains groupes. Nous verrons que ce polymorphisme peut aider à situer le virus du SRAS au sein des coronavirus. On note également une autre région variable chez les coronavirus, relativement inatten-

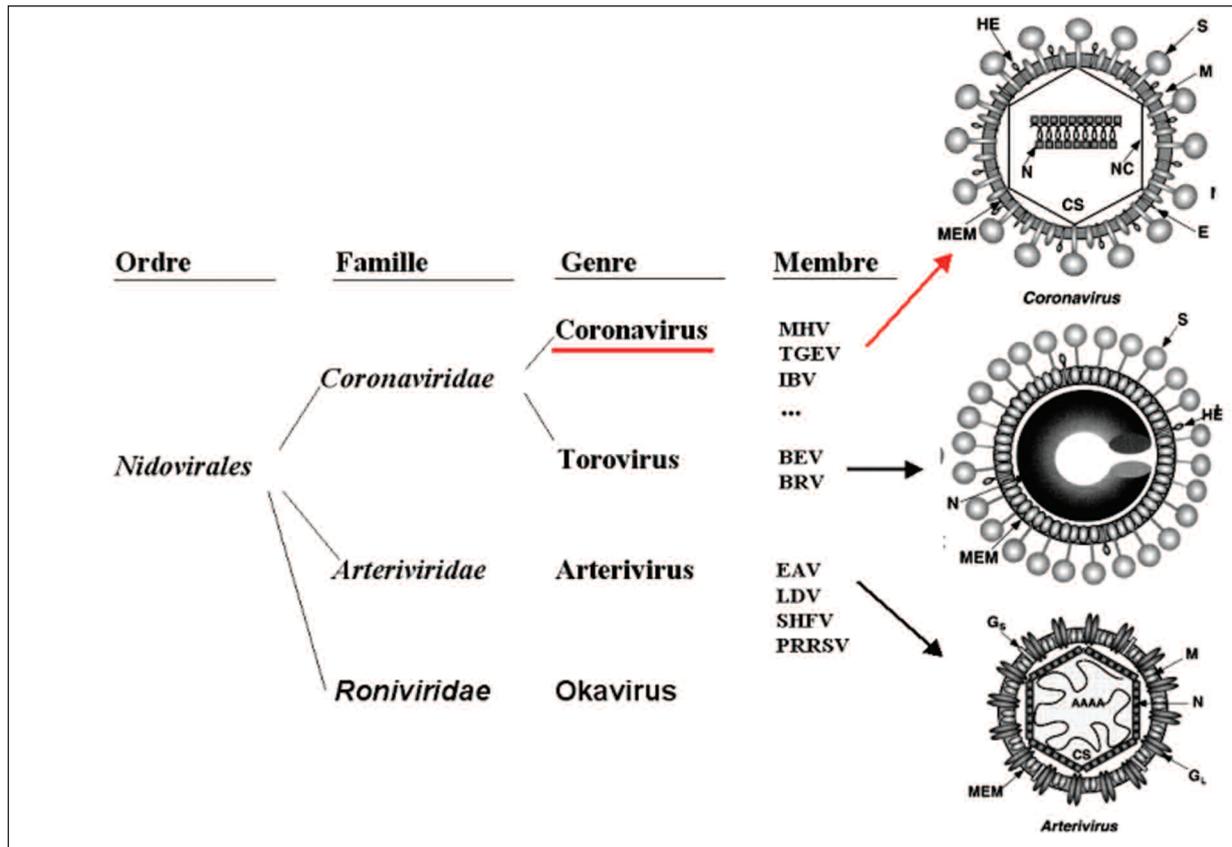


Figure 1 : Classification et aspect des nidovirales. Les virus membres sont désignés par leur acronyme international : MHV: mouse hepatitis virus ; TGEV : transmissible gastroenteritis virus ; IBV : infectious bronchitis virus ; BEV : Berne enteritis virus ; EAV: equine arteritis virus ; LDV: lactico-deshydrogenase virus ; SHFV: simian haemorrhagic fever virus ; PRRSV: porcine respiratory reproductive syndrome virus.

due puisqu'il s'agit d'un gène codant une protéine structurale majeure, la protéine du spicule S, qui peut tolérer d'importantes délétions sans pour autant que la viabilité du virus en soit notablement affectée.

Nonobstant la taille imposante de leur génome les virions des coronavirus ont une organisation relativement simple, avec seulement 4 protéines structurales (figure 3) : la **protéine S** qui par association de trois molécules, forme les spicules très proéminents et renflés à leur extrémité, deux autres protéines de membrane, la **protéine M** qui est essentiellement enfouie dans la membrane au sein du virion et la

**protéine E**. Cette protéine E, longtemps passée inaperçue en raison de sa faible taille mais aussi par sa présence en peu d'exemplaires sur les virions, joue pourtant un rôle crucial dans le bourgeonnement des particules virales. La seule protéine interne des virions est la **protéine N** (nucléoprotéine) qui est associée avec l'ARN génomique viral.

Un trait caractéristique des coronavirus est leur tropisme sinon exclusif, du moins électif pour les cellules épithéliales, donc des cellules polarisées (figure 4). Leur cycle viral se déroule entièrement à l'intérieur du cytoplasme.

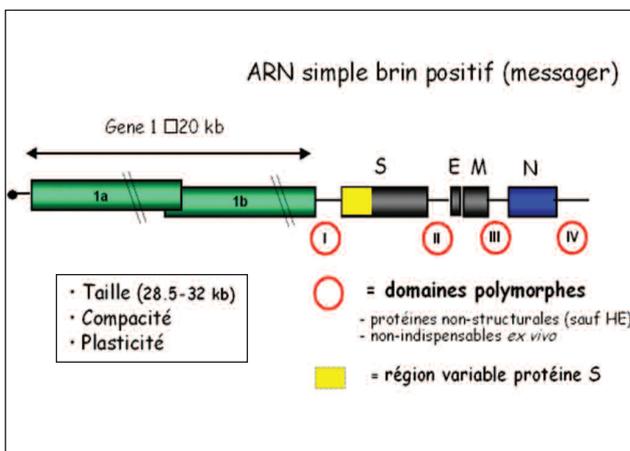


Figure 2 : Organisation du génome des coronavirus.

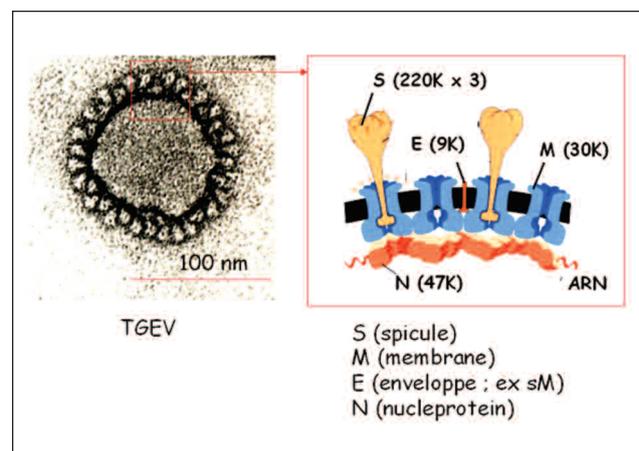


Figure 3 : Structure protéique d'un Coronavirus (Gastro-entérite transmissible).

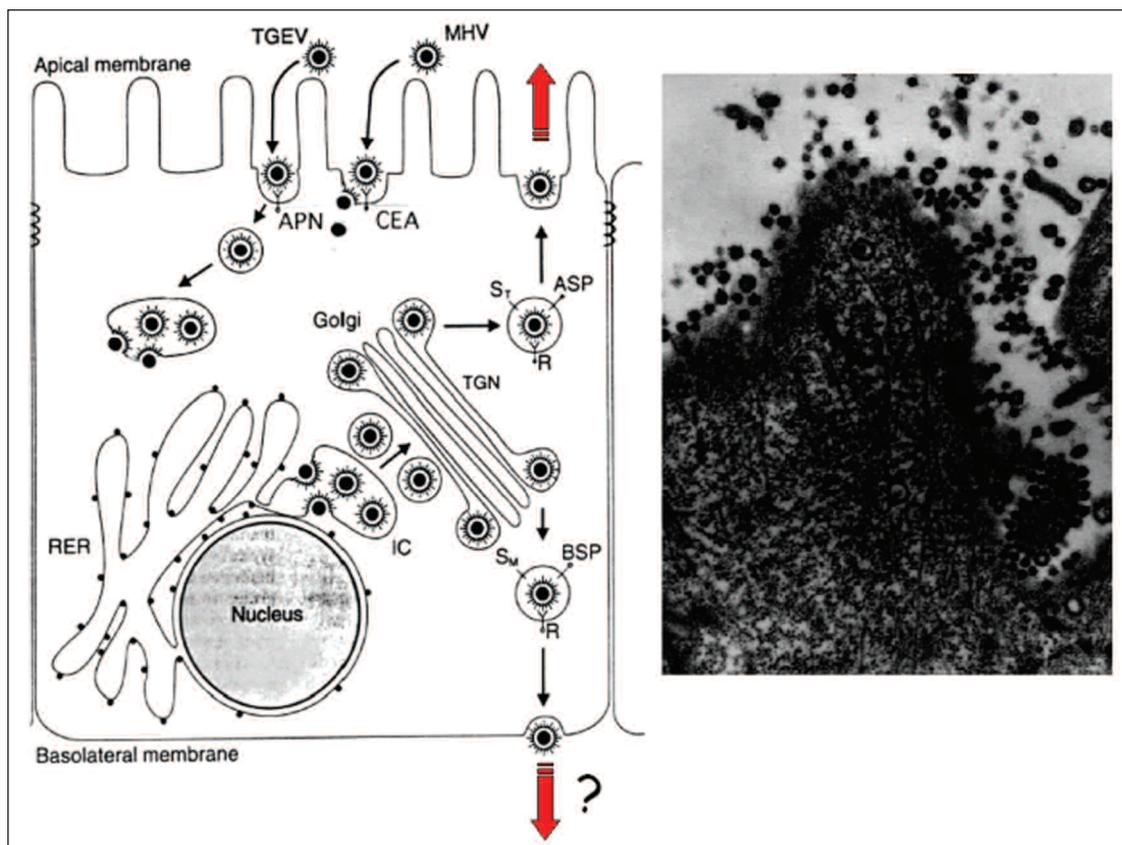


Figure 4 : Réplication des coronavirus dans les cellules cibles. À gauche sont schématiquement décrites les grandes étapes du cycle cellulaire des coronavirus (reconnaissance de la cellule cible, internalisation, assemblage et libération des particules). À droite, microphotographie qui montre une cellule épithéliale infectée par le virus TGEV et libérant de nombreuses particules dans le milieu extérieur.

La première étape de l'infection, la reconnaissance des cellules cibles et la pénétration dans ces cellules, met en jeu une molécule de surface de la cellule cible appelée récepteur. Deux récepteurs ont été identifiés à ce jour chez les coronavirus : d'une part, pour le coronavirus murin, l'antigène CEA (carcino-embryonnaire) qui appartient à la superfamille des immunoglobulines, souvent utilisée par les virus comme récepteur et d'autre part, une molécule reconnue par le virus de la gastro-entérite transmissible (TGEV) et par tous les virus apparentés au TGEV, l'aminopeptidase N (APN). Le génome étant devenu accessible et "lisible" par les protéines cellulaires, de nouvelles molécules de protéine et de génome sont synthétisées puis assemblées au niveau d'un compartiment intermédiaire entre le réticulum endoplasmique et l'appareil de Golgi. Les particules neo-assemblées ont alors deux voies de sortie possibles (avant la destruction éventuelle de la cellule) : soit par la face apicale par où elles sont entrées (figure 4, droite), soit par le pôle basolatéral. Cette dernière possibilité pourrait avoir un lien avec la faculté de certains coronavirus à engendrer des désordres autres que ceux liés au simple épithéliotropisme, tels que les désordres nerveux provoqués par les virus HEV et MHV (cf. intervention de Philippe Vannier).

Le mode d'expression de l'information génétique chez les coronavirus ne sera pas développé ici. Je soulignerai seulement qu'il est assez unique dans les annales de la virologie animale, non par la nature des mécanismes d'expression

mais par la diversité des mécanismes auxquels les coronavirus font appel pour exprimer cette information (figure 5).

Pour un virus dont le génome est constitué d'un ARN messager, le problème est toujours le même : la lecture des gènes situés en aval du premier gène. Il existe deux stratégies de base : l'une consiste à lire tous les gènes en une seule fois de A jusqu'à Z, puis ensuite à "tronçonner" cette polyprotéine en sous-unités fonctionnelles. Dans le cas des coronavirus, c'est ce qui se passe pour le gène 1. L'autre stratégie fait appel à un ensemble d'ARN messagers de taille subgénomique : s'opère ainsi l'expression génétique de tous les gènes situés en aval du gène 1.

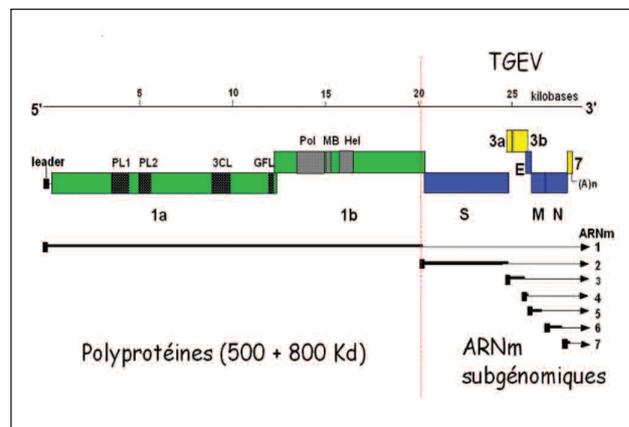


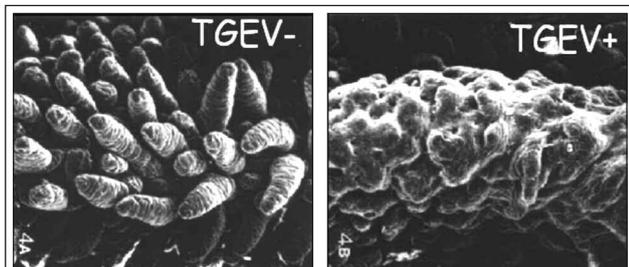
Figure 5 : Mécanismes de l'expression de l'information génétique chez les coronavirus (exemple du TGEV).

• DÉTERMINISME DE LA SPÉCIFICITÉ D'ESPÈCE ET DU TROPISME DE TISSU CHEZ LES CORONAVIRUS

Que sait-on des bases moléculaires de la spécificité d'espèce et du tropisme de tissu chez ces virus ? Je prendrai un certain nombre d'exemples qui concernent les coronavirus du Groupe 1, dont le prototype est le TGEV et sur lesquels mon équipe a plus particulièrement travaillé. Rappelons tout d'abord que le virus TGEV a un spectre d'hôte étroit et un tropisme tissulaire restreint : l'intestin, seul siège des lésions observées dans la gastro-entérite transmissible. La **figure 6** montre des villosités intestinales atrophiées chez le porc infecté, comparées à celles d'un individu indemne.

Nos travaux nous ont permis d'identifier comme récepteur du TGEV une molécule déjà mentionnée, l'aminopeptidase N (APN), qui est abondamment exprimée dans la bordure en brosse des entérocytes. Nous avons aussi démontré que le domaine qui interagissait avec le virus, et plus précisément avec les spicules, est situé dans un domaine de l'APN distinct de celui de la sous-unité catalytique (**figure 7**).

Nous avons proposé que l'interaction du virus avec cette molécule était une étape indispensable au bon déroulement du cycle infectieux notamment à l'internalisation, la rentrée des particules dans les cellules cibles. En effet, lorsqu'on fait exprimer un gène qui code l'APN porcine



- Spectre d'hôte étroit
- Tropisme tissulaire restreint (épitéliotropisme)
- > Déterminisme moléculaire ?

Figure 6 : Epithéliotropisme du virus TGEV : la multiplication du virus dans les entérocytes provoque une desquamation accélérée de ces cellules, qui aboutit à une atrophie des villosités intestinales (microscopie à balayage).

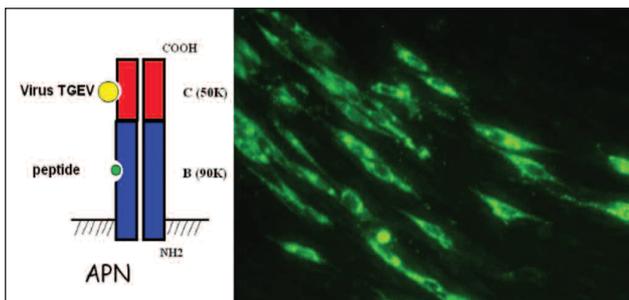


Figure 7 : L'aminopeptidase N, récepteur du virus TGEV. À gauche : organisation schématique de la molécule d'APN montrant les deux domaines fonctionnels (enzymatique et récepteur). À droite : cellules de hamster (lignée BHK) exprimant l'APN porcine et devenues permissives au virus TGEV, dont la multiplication est visualisée à l'aide d'anticorps fluorescents.

dans des cellules non infectées, naturellement réfractaires au virus TGEV, on peut les rendre sensibles à ce virus, résultat qui s'obtient avec des cellules de diverses espèces. Ce point est illustré dans la **figure 7** qui montre des cellules de hamster ainsi génétiquement modifiées et devenues sensibles au virus TGEV.

Par ailleurs, nous avons pu établir une corrélation assez étroite entre la faculté d'un coronavirus du Groupe 1 donné à interagir avec l'APN d'une espèce donnée et sa capacité à infecter des cellules en culture de cette espèce. Par exemple, le virus félin reconnaissant essentiellement l'APN de chat n'infecte que des cellules de chat (**figure 8**). En revanche, le virus porcine a un tropisme plus large, puisqu'il est capable d'infecter non seulement des cellules de porc, mais également de chat et de chien, car il reconnaît l'APN de ces espèces comme l'APN porcine. Le virus humain est également capable d'infecter des espèces animales, mais aucune flèche partant d'un coronavirus animal, sur la **figure 8**, ne se dirige vers l'homme, ce qui constituait à l'époque une donnée rassurante.

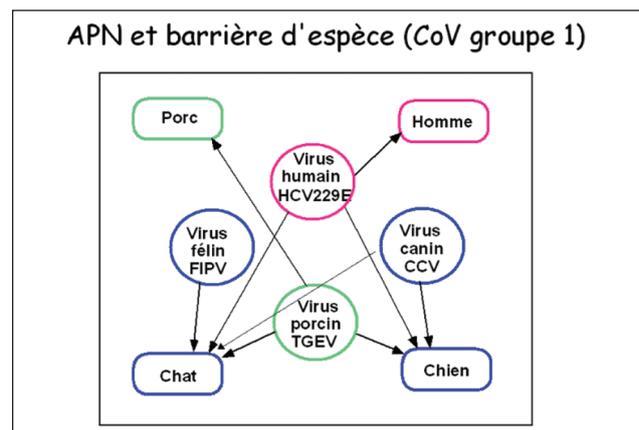


Figure 8 : Relation entre reconnaissance du récepteur APN et capacité à infecter des cellules d'espèces hétérologues.

En alignant la séquence en acides aminés de l'APN dans la région qui reconnaît la protéine de spicule virale on voit, matérialisées dans la **figure 9** par des flèches, un certain nombre de positions au niveau desquelles les acides aminés sont extrêmement conservés dans les différentes espèces animales, alors que chez l'homme est présent un acide aminé différent. On peut supposer que ces quelques différences expliquent pourquoi l'espèce humaine n'est pas infectée par ces coronavirus animaux. La localisation du virus TGEV dans les cellules intestinales est superposable à celle de l'aminopeptidase N. Sur des coupes d'intestin grêle, on met en évidence que le virus se réplique dans les entérocytes différenciés (**figure 10**, TGEV) et que la répartition des anticorps anti-APN fluorescents chez des animaux non infectés est semblable au marquage révélant la présence du virus chez des animaux infectés (**figure 10**, APN). On met très peu d'APN en évidence au niveau des cryptes de Lieberkühn et le virus TGEV ne se réplique pas ou peu à cet endroit. De même, il se réplique préférentiellement dans l'intestin grêle et pas dans le gros intestin, où l'APN est peu exprimée.

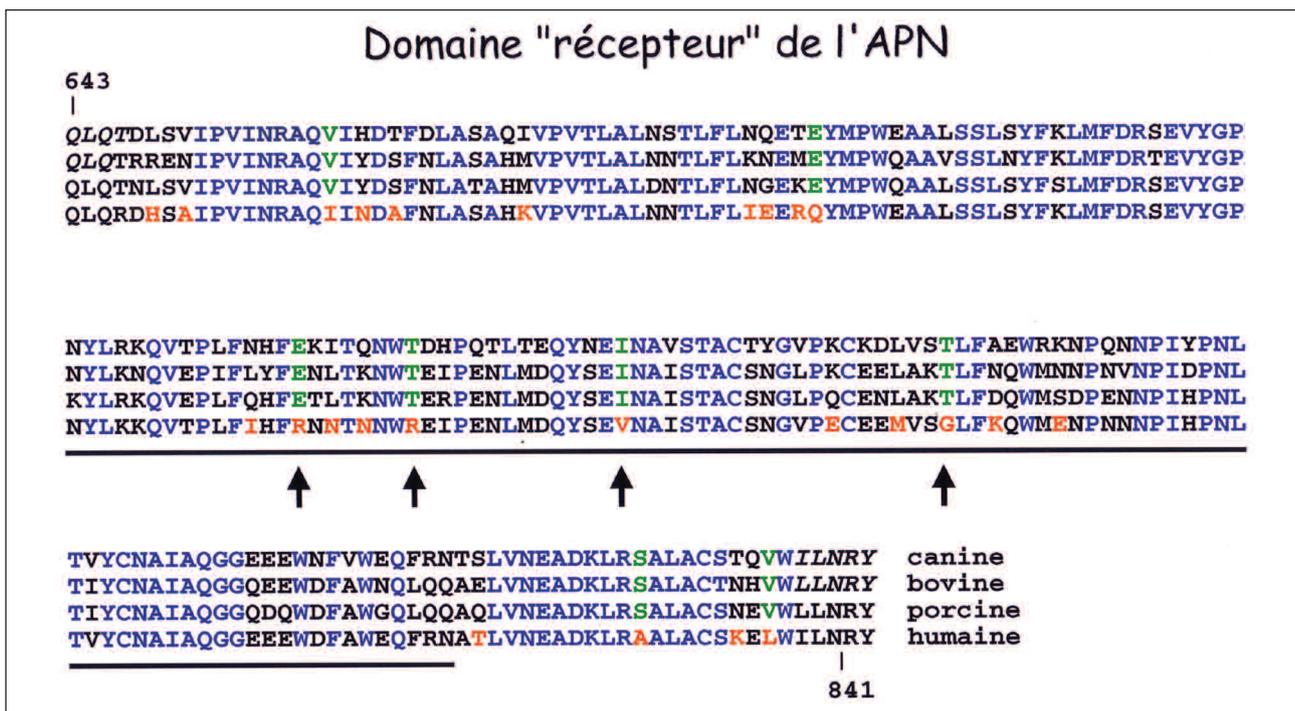


Figure 9 : Séquence en acides aminés de l'APN de quatre espèces dans la région (soulignée) du récepteur qui interagit avec les spicules virales.

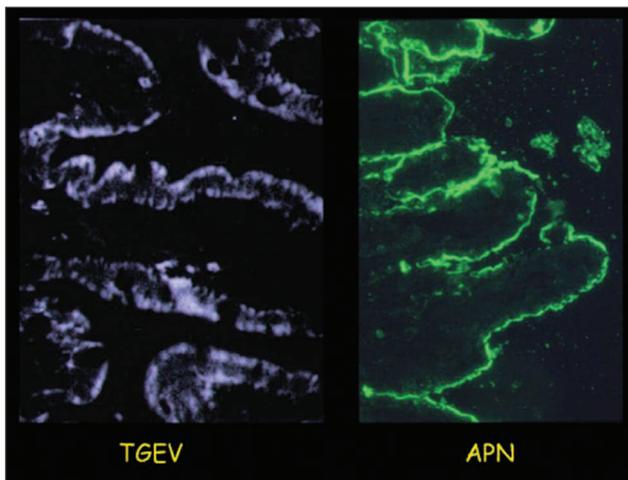


Figure 10 : Parallélisme entre la localisation du virus TGEV et de l'aminopeptidase N dans une coupe d'intestin grêle de porc (immunofluorescence)

- LE CORONAVIRUS RESPIRATOIRE PORCIN ; UNE CLÉ POUR LA COMPRÉHENSION DU POUVOIR ENTÉRO-PATHOGENE DU VIRUS TGEV

À partir de 1984, un virus respiratoire inconnu jusqu'alors, responsable d'une infection essentiellement inapparente, a rapidement et massivement diffusé au sein du cheptel porcin européen. Comme nous allons le voir, tout porte à croire que ce virus PRCV (Porcine Respiratory Corona Virus) dérive du virus entéritique TGEV. Dans la mesure où ces deux virus ne pouvaient être distingués par la sérologie classique, on pouvait penser qu'ils étaient extrêmement proches et donc que l'étude du virus PRCV pourrait nous aider à comprendre quels étaient les déterminants viraux impliqués dans

l'expression du pouvoir entéropathogène du TGEV. Nous avons donc réalisé le séquençage partiel du génome d'une souche de PRCV isolée pour la première fois à Gand par Maurice Pensaert, que nous avons comparé avec celui de souches classiques de virus TGEV.

Les deux virus sont effectivement très proches, avec moins de 1 % de divergence entre les séquences, du même ordre de ce que l'on trouve entre 2 isolats de gastro-entérite transmissible classique. On était par contre frappé de constater qu'il s'agissait d'un virus multidélétant, avec deux régions touchées: d'une part un gène codant une protéine non structurale parmi celles mentionnées précédemment et d'autre part une délétion assez importante dans le gène S, touchant la région N-terminale de la protéine de spicule (figure 11). Dans les années qui ont suivi, a été isolé aux États-Unis, en Indiana, un autre virus qui avait également perdu son pouvoir entéropathogène et qui présentait le même type d'événement, à savoir une importante délétion à l'extrémité N-terminale de la protéine S.

On pouvait supposer, compte tenu du très grand recouvrement observé entre les sites de réplication du virus entéropathogène (TGEV) et les sites d'expression de l'APN dans l'intestin (cf. figure 10), que la région de la protéine S perdue dans le virus respiratoire (PRCV) pouvait effectivement être impliquée dans la reconnaissance du récepteur. En fait on s'est rapidement aperçu que le virus respiratoire, comme le virus TGEV, dépendait de l'APN pour entrer dans les cellules cibles. Il est donc très probable que le virus respiratoire utilise également cette molécule *in vivo* pour se multiplier au niveau des cellules de l'appareil respiratoire.

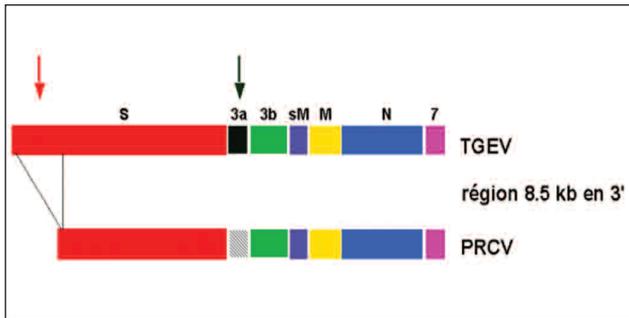


Figure 11 : Émergence du coronavirus respiratoire porcine. Les deux flèches indiquent les régions du génome du virus TGEV qui sont absentes dans celui du PRCV.

Quel pourrait être le rôle biologique de cette région de la protéine S qui est tronquée chez les virus respiratoires ?

La figure 12 représente de façon schématique la région N-terminale de la protéine S, soit environ 250 aminoacides, avec la région perdue dans le virus respiratoire par rapport au virus TGEV. Dans cette région existe un domaine antigénique reconnu par plusieurs anticorps monoclonaux neutralisants, qui nous ont permis de sélectionner une série de mutants d'échappement à la neutralisation. Nous avons ensuite examiné le pouvoir pathogène de ces mutants chez des porcelets nouveau-nés en allaitement artificiel, en comparant leur gain de poids et leur mortalité à ceux d'animaux infectés par d'autres mutants ou par le virus parental. Seuls les mutants ayant une substitution touchant un acide aminé situé au cœur de la région manquante dans la protéine S du PRCV ont montré un phénotype très fortement atténué. La mutation, dans cette région, d'un seul aminoacide parmi les quelques 1 430 que compte la protéine S suffit à abolir le pouvoir entéropathogène du TGEV. La région N-terminale de la protéine S apparaît donc jouer un rôle crucial dans l'expression de ce dernier.

La protéine de spicule comprend deux domaines : le domaine S2 qui correspond à la région d'ancrage et au

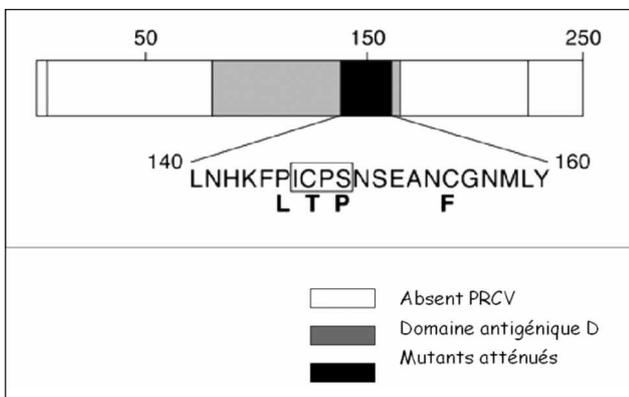


Figure 12 : Domaine N-terminal de la protéine S. Les positions des substitutions en acides aminés dans le domaine N-terminal de la protéine de spicule et associées à une perte de virulence du virus TGEV sont indiquées (code à une lettre). Ces dernières sont toutes regroupées dans une brève séquence (Proline --> Leucine, Cystéine --> Thréonine, Sérine --> Proline, Cystéine --> Phénylalanine).

socle des spicules, et la région globulaire qui correspond à la partie S1. On est frappé de constater que, par rapport au TGEV et à tous les virus entéropathogènes, les virus respiratoires, dont le virus respiratoire humain HCV229E (un coronavirus pourtant plus proche du virus de la diarrhée épidémique porcine -PEDV- que du virus PRCV), présentent une protéine de spicule tronquée à l'extrémité N-terminale de la région S1 (figure 13). Il existe manifestement un lien entre la présence de cette extra-séquence et la capacité d'exercer un pouvoir entéropathogène chez l'hôte. Pour l'instant, nous n'avons que des hypothèses concernant le rôle biologique de cette région, mais toutes tournent autour de l'idée que cette extra-séquence pourrait contribuer à maintenir l'intégrité structurale de la protéine de spicule dans l'environnement physico-chimique particulièrement agressif du tractus digestif.

Le séquençage d'isolats du PRCV provenant de différentes régions d'Europe a permis de constater qu'ils ont non seulement tous les mêmes modifications génétiques, mais aussi la même séquence. Cependant, l'examen d'une série de virus respiratoires isolés en Amérique du Nord durant une période de temps similaire, et qui sont des équivalents du PRCV européen, a révélé l'existence de génotypes distincts, avec des modifications voisines mais non identiques à celles du PRCV européen: une altération de la partie du génome codant la protéine non structurale 3A et une troncature de l'extrémité N-terminale du gène S codant la protéine de spicule. On aurait donc affaire à des événements convergents, mais indépendants. Tout se passe comme si des porcs infectés par le virus TGEV étaient potentiellement capables de "fabriquer" en permanence des mutants respiratoires tronqués. Pour une raison que l'on ignore l'un d'entre eux, le virus PRCV européen, a "réussi" une émergence spectaculaire. Cela tient sans doute à certaines caractéristiques moléculaires encore ignorées, peut-être aussi à la structure différente des élevages de porcs en Europe et aux USA.

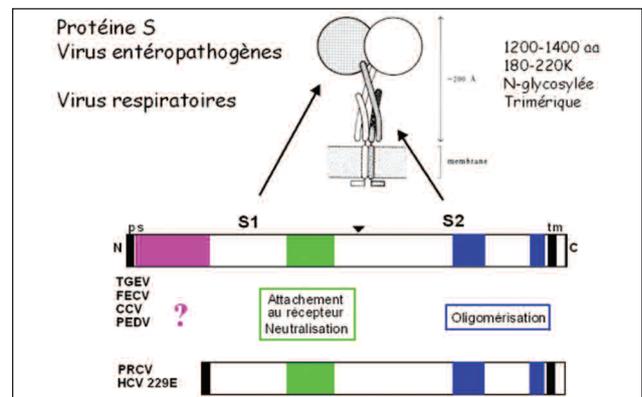


Figure 13 : Comparaison de la structure de la protéine de spicule entre les virus à tropisme intestinal et les virus à tropisme respiratoire. S1 : région globulaire, S2 : tige du spicule. L'extrémité S1 comporte un domaine présent chez tous les virus entéropathogènes (type TGEV) et absent chez les virus respiratoires (type PRCV).

On peut au moins tirer un enseignement de cet épisode, édifiant à bien des égards : il faut se méfier de la notion intuitive selon laquelle une perte d'information génétique rend un virus potentiellement moins dangereux que le virus parental. Dans le cas présent, le virus était certes moins pathogène, mais il avait acquis un pouvoir de diffusion tout à fait extraordinaire, ce que l'on peut tout de même considérer comme étant une propriété indésirable.

• POUVOIR INTERFÉROGÈNE DES CORONAVIRUS : MÉCANISMES MOLÉCULAIRES

Les coronavirus sont des virus puissamment interféro-gènes (cf. l'exposé de Bernard Charley). Dans les quelques heures qui suivent l'infection d'un porc par le virus TGEV, les titres d'interféron augmentent dans le sang et dans l'intestin à des niveaux spectaculaires. La synthèse de l'interféron alpha fabriqué lors d'un épisode viral est réalisée, selon le mode d'induction classique, par des ARN double

brin qui sont des intermédiaires dans le cycle viral des virus ARN simple brin, ce qui implique une réplication du virus à l'intérieur de la cellule. Un tel mécanisme existe aussi dans le cas des coronavirus, mais un autre mécanisme, particulièrement puissant, intervient. Des cellules infectées par ces virus, puis fixées (le virus est donc inactif) et mises en contact avec des leucocytes, induisent une synthèse très précoce et très forte d'interféron alpha ; le même résultat s'observe avec des virions purifiés, inactivés par des radiations ultra-violettes. (figure 14)

Il était intéressant d'identifier le déterminant inerte, indépendant de la réplication du virus, capable d'induire une réponse interféron alpha aussi intense. À cette fin, nous avons tiré avantage d'une propriété remarquable des coronavirus : des cellules génétiquement modifiées qui ne co-expriment que les protéines M et E libèrent dans le cytoplasme des pseudo-particules qui ressemblent beaucoup aux particules du virus normal (figure 15) : Ces dernières sont évidemment dépourvues de spicules (la protéine S étant absente) et d'ARN génomique. Or, le pouvoir interféro-gène spécifique de ces pseudo-particules est proche de celui du virus sauvage.

Cette observation impliquait au premier chef la protéine M, dont la topologie caractéristique est représentée figure 16, avec 3 segments qui forment une sorte d'épingle à cheveux dans la membrane du virion, l'une des "marques de fabrique" de l'ordre des *Nidovirales*. L'essentiel de la protéine est enfoui dans la membrane ou à l'intérieur de la particule virale ; il ne subsiste à l'extérieur qu'un très bref ectodomaine qui porte un site de glycosylation unique. Des virus mutés sur ce site perdent tout ou partie du pouvoir interféro-gène (figure 16). Ce résultat était en faveur de l'implication des résidus glucidiques greffés sur l'ecto-domaine. Plus surprenant, quand on exprime une chimère de la protéine M

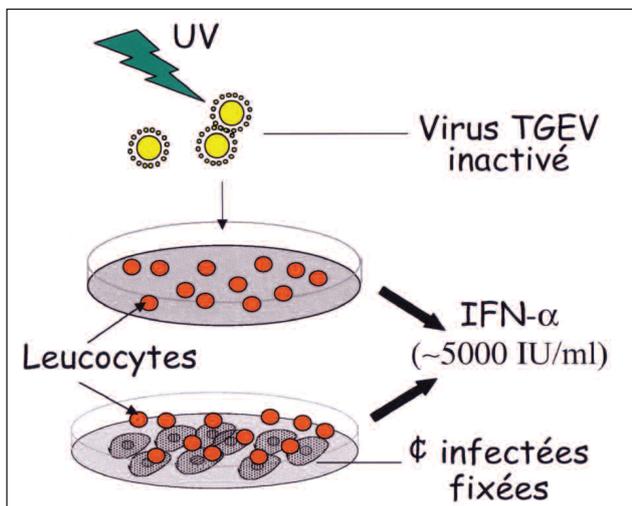


Figure 14 : Des coronavirus dont le pouvoir infectieux est inactivé sont capables d'induire une production d'interféron dans des leucocytes, en l'absence de réplication virale. Des leucocytes périphériques sont incubés en présence de virions purifiés et soumis à une irradiation aux rayons UV, ou d'un tapis de cellules infectées et fixées avec la paraformaldéhyde.

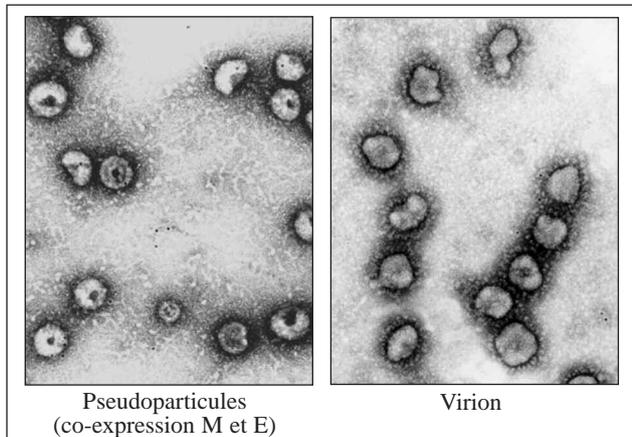


Figure 15 : Similitude morphologique entre des particules virales normales (à droite) et des pseudo-particules virales (à gauche) obtenues par transformation génétique cellulaire (expression de deux protéines du virus).

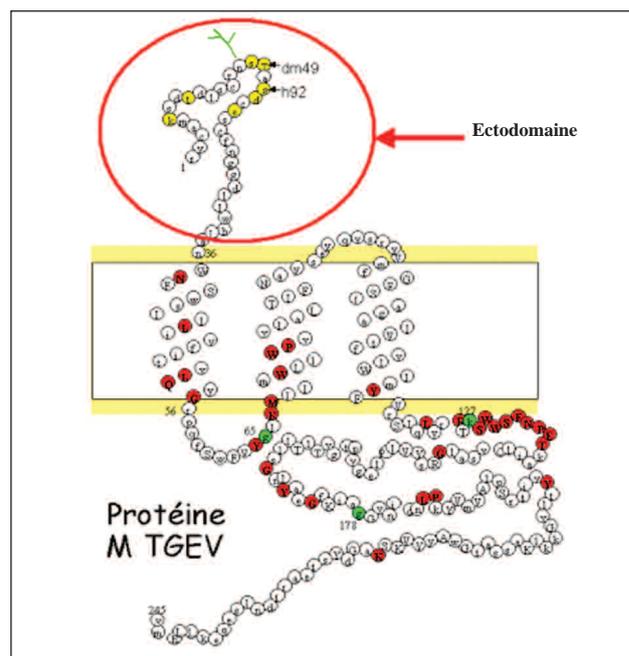


Figure 16 : Topologie de la protéine M du coronavirus dans la membrane virale.

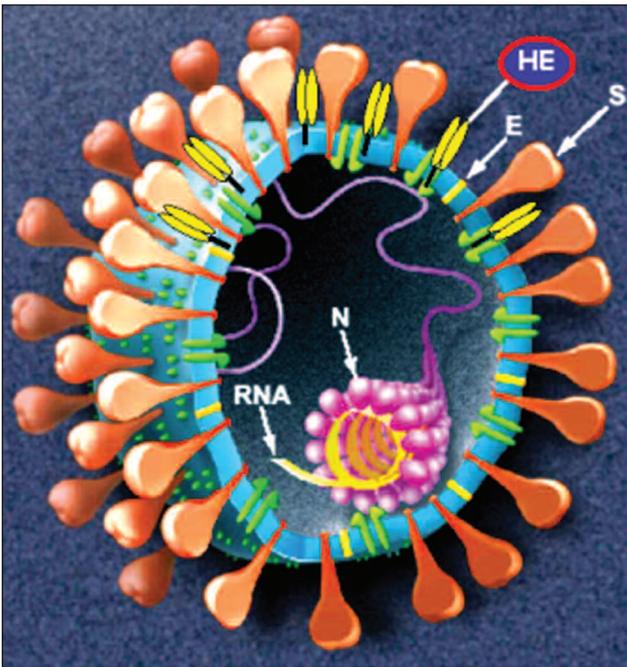


Figure 17 : L'hémagglutinine-estérase (HE) : un exemple de "capture" de gène.

dont l'ecto-domaine a été remplacé par celui du virus BCV, virus d'un autre groupe avec un type de glycosylation différent, on obtient à peu près le même résultat. Cette capacité d'induction de l'interféron ne dépendrait donc pas d'un motif aminoacide bien spécifique ou d'un type de glucide bien particulier, mais plus probablement d'un patron de motifs régulièrement répétés, capable de faire reconnaître ces objets comme étant du « non soi ».

• **ÉVÉNEMENTS GÉNÉTIQUES ASSOCIÉS A L'ÉVOLUTION DES CORONAVIRUS ET POSITIONNEMENT TAXONOMIQUE DU VIRUS DU SRAS**

Plusieurs types d'événement génétique sont impliqués dans l'évolution des Coronavirus, tels que mutation, délétion et recombinaison inter-génomique, légitime ou non. Dépourvus de système de correction d'erreurs de la polymérase, les virus à ARN ont un taux de mutation élevé, et les coronavirus ne font pas exception. Toutefois, la dérive génétique (les mutations fixées) qui en résulte n'est pas si importante : d'après le séquençage d'isolats sur une trentaine d'années de distance, le virus TGEV présente une dérive de  $7 \times 10^{-4}$  substitution par génome et par an. S'agissant du virus du SRAS, dont on a dit parfois qu'il mutait à grande vitesse, on s'aperçoit en prenant en compte d'une part les substitutions publiées sur les 15 isolats humains séquencés et d'autre part la durée de l'épidémie, que le taux de mutation est du même ordre que celui observé classiquement chez les coronavirus. Le virus SRAS n'est donc pas un "mutant fou".

Il se produit aussi chez les coronavirus des événements de délétion intragénomique, comme dans le cas du virus respiratoire PRCV (*vide supra*), ou du virus MHV murin. Les coronavirus sont aussi des "champions" de la recombinaison intergénomique : la probabilité que des erreurs s'accumulent est évidemment en proportion de la taille du génome, et on

- groupe 2 (MHV)
- optionnelle
- dimérique (2 x 65K)
- homologie protéine HE
  - Torovirus
  - Influenza C
- rôle biologique ?

suppose qu'ils font appel à cette stratégie de recombinaison pour reconstituer régulièrement des génomes fonctionnels à partir de plusieurs génomes défectifs. On peut observer ce phénomène de recombinaison dans un contexte expérimental mais aussi sur le terrain, par exemple dans le cas du virus de la bronchite infectieuse aviaire – plusieurs publications ont fait état d'évènements de recombinaison entre des souches vaccinales atténuées largement utilisées sur le terrain et des souches sauvages –, ou dans le cas du coronavirus félin, qui comprend en fait deux groupes, l'un exprimant une protéine de spicule plus proche de celle du virus canin que celle du virus félin originel.

Enfin il existe des phénomènes de capture de gène. Ainsi, chez plusieurs coronavirus du Groupe 2, dont le prototype est le virus murin MHV, on trouve un gène surnuméraire appelé HE (pour hémagglutinine-estérase), lequel code une protéine qui forme une deuxième frange de spicules, plus discrète, à la surface des particules virales (figure 17). Ce gène est également présent chez les torovirus, ce qui n'est pas étonnant puisque ce sont des virus de la même famille, mais on trouve un gène homologue chez le virus de l'influenza C, ce qui est moins attendu et suggère un événement de capture de gène.

Je terminerai en tentant de situer le virus humain du SRAS parmi les coronavirus connus actuellement chez les animaux et chez l'homme. La figure 18 présente les trois groupes génétiques répertoriés à ce jour : les groupes 1 et 2 qui infectent les mammifères et le groupe 3, avec notamment le virus de la bronchite infectieuse aviaire (IBV), qui infecte les oiseaux. Ces groupes se différencient par une protéine S, clivée ou non clivée, une protéine M, N- ou O-glycosylée, et la présence ou l'absence du gène hémagglutinine-estérase (HE).

SARSCoV ?			
<u>Sous-groupe</u>	<u>Protéine S</u>	<u>Protéine M</u>	<u>Autre C°</u>
1 TGEV FCoV CCV PEDV HCoV229E	Non-clivée	N-glycosylée	APN récepteur
2 MHV BCV HEV HCoV43	Clivée	O-glycosylée	HE gène (pseudogène)
3 IBV TCV	Clivée	N-glycosylée	

Figure 18 : Classification des coronavirus en trois groupes génétiques.

Sur cette base, le virus du SRAS pourrait très bien s'insérer dans le groupe 1. En fait il n'en est rien puisque l'examen des séquences publiées il y a déjà un mois révèle que le virus du SRAS est équidistant entre les groupes 2 et 3, c'est à dire entre les virus MHV et IBV (figure 19).

L'inspection du génome du SRAS ne laisse pas non plus penser que ce virus est le fruit d'un événement de recombinaison inter-génomique récent entre un coronavirus animal et un coronavirus humain connus. Cependant en s'intéressant de plus près à l'organisation génomique du virus du SRAS (figure 20), on peut constater que celle-ci ressemble à celle du virus aviaire (IBV) : en effet son génome comporte des gènes supplémentaires entre les protéines N et M comme dans le groupe 3, et il n'existe pas de gène en aval du gène N comme dans le groupe 1, ni de gène HE comme dans le groupe 2.

Cette ressemblance va plus loin: l'examen attentif de la séquence d'un des gènes supplémentaires du virus SRAS et du virus IBV révèle en effet une homologie significative. De surcroît, l'extrémité 3' du génome du virus IBV, mais non des coronavirus de mammifères, conserve une séquence d'une trentaine de bases (présente chez nombre d'astrovirus : encore un phénomène de capture génétique), qui est également présente chez le virus du SRAS humain. Enfin, à l'autre extrémité du génome, on note des motifs protéiques "papain-like" (chargés notamment du découpage de la polyprotéine spécifiée par le gène 1) : on observe 2 motifs répétés chez la plupart des coronavirus, mais un motif unique chez le virus du SRAS, comme dans le cas de l'IBV. En résumé le virus du SRAS est certes équidistant en terme de séquences entre les groupes 2 et 3 mais il possède un "châssis" de virus aviaire. Le virus du SRAS représenterait donc un nouveau chaînon dans l'évolution des coronavirus.

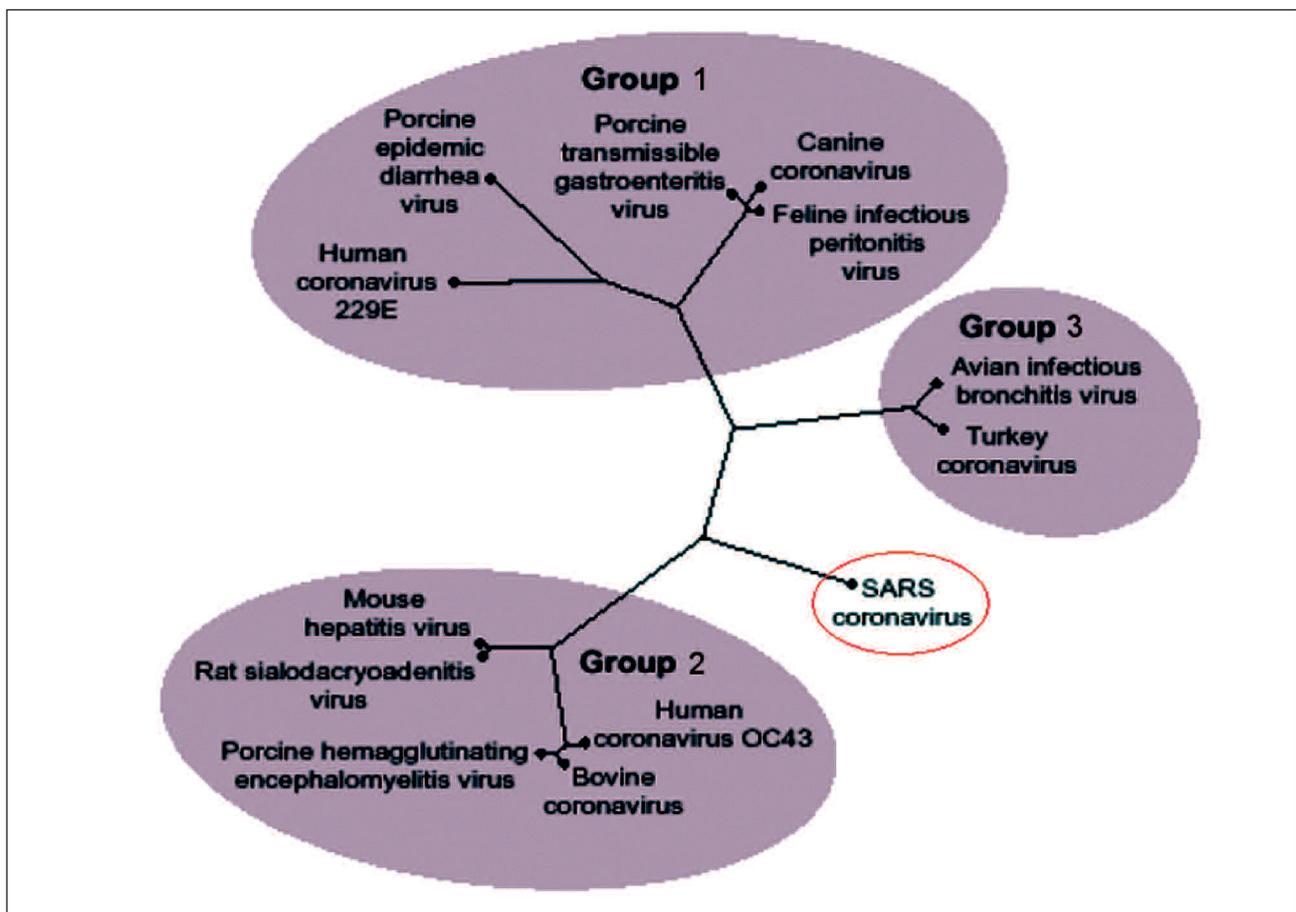


Figure 19 : Position du coronavirus du SRAS selon les séquences génétiques (405 nucléotides du gène de la polymérase ; d'après. ENSE-RING M, *Science*, avril 2003).

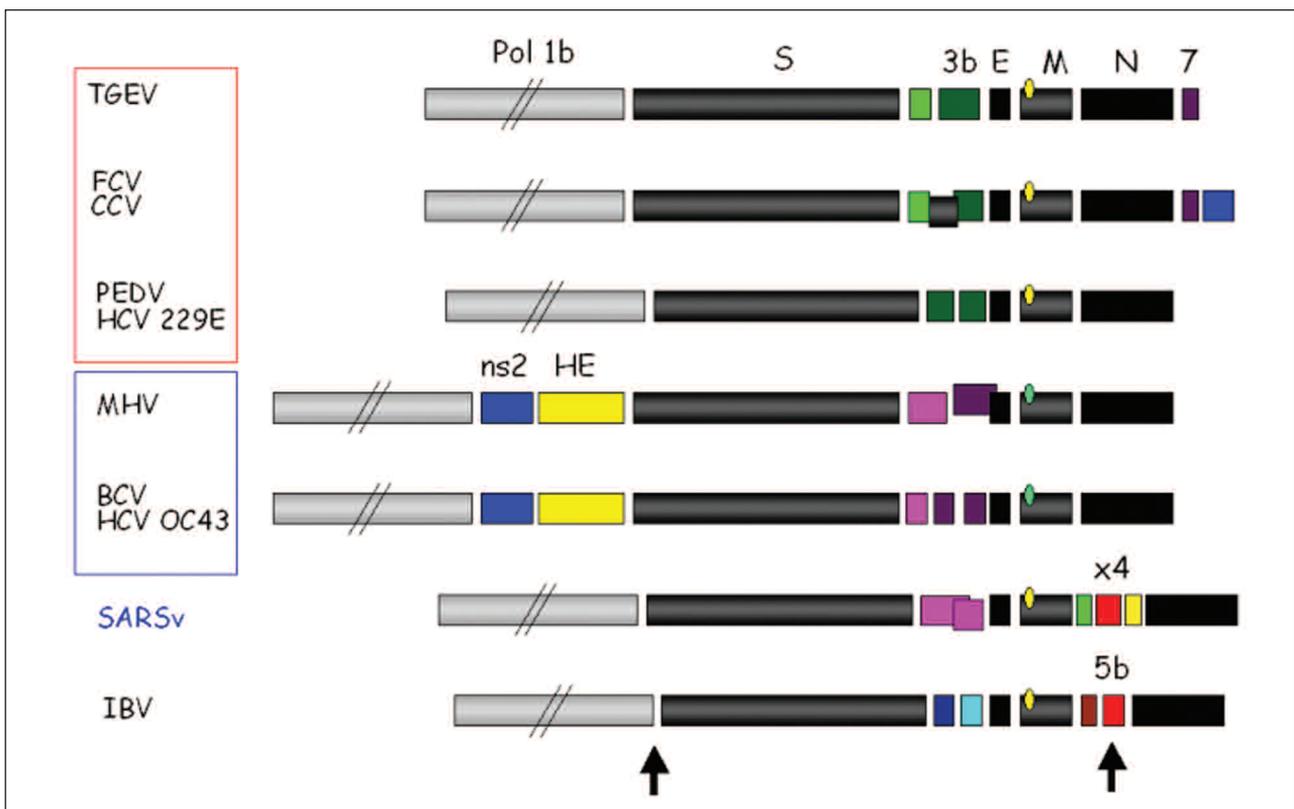


Figure 20 : Comparaison de l'organisation génomique de différents coronavirus animaux et du SRAS.

### REMERCIEMENTS

Je ne puis conclure cet exposé sans rappeler les noms des personnes qui ont participé au travail de l'équipe, en citant particulièrement les noms de Jacqueline Gelfi et de Bernard Delmas, qui ont fait le plus long bout de chemin à mes côtés et en indiquant également que ces travaux ont fait l'objet de collaborations non seulement à l'intérieur de l'INRA avec notamment Bernard Charley à Jouy et Serge Bernard à Tours mais aussi avec des groupes européens (Zürich, Marburg, Madrid et Copenhague).

Crédits iconographiques : les figures 2, 3, 5 et 11 à 13 sont extraites d'un article de l'auteur cité en référence, avec l'aimable autorisation de la revue *Virologie*.

## BIBLIOGRAPHIE SUCCINCTE

Travaux réalisés à l'INRA de Jouy-en-Josas (1990-1998) :

- DELMAS B, LAUDE H (1990) Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J. Virol.*, **64**, 5367-5375.

- RASSCHAERT D, DUARTE M, LAUDE H (1990) Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. *J. Gen. Virol.*, **71**, 2599-2607.

- DELMAS B, GELFI J, L'HARIDON R, VOGEL LK, SJÖSTRÖM H, NOREN O, LAUDE H (1992) Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. *Nature*, **357**, 417-420.

- GODET, M., L'HARIDON, R., VAUTHEROT, J.F., LAUDE, H. 1992. TGEV coronavirus ORF4 encodes a membrane protein that is incorporated into virions. *Virology*, **188**, 666-675.

- LAUDE H, GELFI J, LAVENANT L, CHARLEY B (1992) Single amino acid changes in the viral glycoprotein M affect induction of alpha interferon by coronavirus TGEV. *J. Virol.*, **66**, 743-749.

- DUARTE M, LAUDE H (1994) Sequence of the spike protein of porcine epidemic diarrhoea virus. *J. gen. Virol.*, **75**, 119-1200.

- DELMAS B, GELFI J, KUT E, SJÖSTRÖM H, NOREN O, LAUDE H (1994) Determinants essential for the transmissible gastroenteritis virus-receptor interaction reside within a domain of aminopeptidase-N that is distinct from the enzymatic site. *J. Virol.*, **68**, 5216-5224.

- GODET M, GROSCLAUDE J, DELMAS B, LAUDE H (1994) Major receptor-binding and neutralization determinants are located within the same domain of the transmissible gastroenteritis virus (coronavirus) spike protein. *J. Virol.*, **68**, 8008-8016.

- ELEOUET JF, RASSCHAERT D, LAMBERT P, LEVY L, VENDE P, LAUDE H (1995) Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology*, **206**, 817-822.

- BERNARD S, LAUDE H (1995) Site-specific alteration of transmissible gastroenteritis virus spike protein results in markedly reduced enterovirulence. *J. Gen. Virol.*, **76**, 2235-2241.

- BAUDOUX P, CARRAT C, BERNARDEAU L, CHARLEY B, LAUDE H (1998) Coronavirus pseudoparticles formed with recombinant M and E proteins induce interferon alpha synthesis by leukocytes. *J. Virol.*, **72**, 8636-8643.

- LAUDE H, RASSCHAERT D, DELMAS B, ELEOUET JF (1998) Le coronavirus respiratoire porcin : un virus émergent pas comme les autres. *Virologie*, **2**, 305-316.

Données sur le virus du SRAS :

- DROSTEN C *et al.* (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.*, **15**, 1967-1976

- KSIAZEK TG *et al.* (2003) A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.*, **15**, 1953-66

- MARRA MA *et al.* (2003) The Genome sequence of the SARS-associated coronavirus. *Science*, **30**, 1399-1404.

- ROTA PA *et al.* (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science*, **30**, 1394-1399.

## DISCUSSION

**Professeur Jeanne Brugère-Picoux :**

Merci beaucoup pour cette démonstration que la biologie moléculaire peut être passionnante. Y a-t-il des questions dans la salle sur ce sujet ? Monsieur Manuguerra.

**Jean-Claude Manuguerra :**

Lorsqu'un coronavirus passe d'une espèce à une autre, y a-t-il un accroissement de la fréquence de mutation ?

**Hubert Laude :**

C'est une notion classique : lorsqu'un virus s'installe dans une nouvelle niche écologique, une nouvelle espèce, on voit un taux de mutation plus important que quelques années plus tard lorsqu'il est définitivement installé dans cette espèce. Ce phénomène n'a pas été décrit chez les coronavirus, ce qui ne veut pas dire qu'il n'existe pas, il faut rester extrêmement prudent à cet égard. Dans le cas du virus du SRAS, on n'a pas l'impression qu'il soit en position de mutation accélérée parce qu'il vient de s'installer dans une nouvelle espèce.

**J.-C. Manuguerra :**

Cela irait plutôt dans le sens de ce que tu as dit au départ sur le coronavirus qui n'est peut-être pas d'origine animale, mais lorsque tu rapproches l'organisation du coronavirus du SRAS de celle du virus aviaire IBV, cela me fait quand même penser aux méta-pneumo-virus qui ont été décrits dans l'espèce humaine en 2000 et qui sont quand même très proches des méta-pneumo-virus C des oiseaux, n'aurait-on pas là l'équivalence, je ne sais pas ce que tu en penses ?

**H. Laude :**

J'ignorais cette information très intéressante, mais quand je me suis aperçu de cette similitude de l'organisation génomique (et non de la séquence encore une fois) entre le virus du SARS et les virus aviaires, je me suis dit pourquoi pas un passage de barrières d'espèces. Ce serait plutôt une ou des espèce(s) sauvage(s) bien sûr, plutôt que des espèces domestiques, et on pouvait penser aux oiseaux aquatiques ; en effet, si je ne me trompe pas, le premier cas est apparu chez un marchand de crevettes, le second chez un marchand de poisson. La seule réserve à ce raisonnement est qu'on n'a pas énormément d'exemple, mais là tu es beaucoup mieux placé que moi pour parler de virus qui passent directement et efficacement des oiseaux à l'homme.

**X. :**

Vous avez présenté le cycle du virus et montré qu'il était non lytique

**H. Laude :**

Il est lytique, le plus souvent mais c'est vrai qu'on observe une libération de particules virales importante bien avant que la cellule ne meure. Les coronavirus savent aussi être non lytiques, ils peuvent établir des infections persistantes. Ils sont souvent lytiques en culture de cellules en tout cas.

**X. :**

Et *in vivo* ?

**H. Laude :**

Je crois qu'on est incapable de dire si le phénomène d'atrophie des villosités intestinales chez les animaux infectés par le TGEV, qui reflète une incapacité de compenser par des cellules des zones de synthèse des cryptes, résulte d'une simple desquamation accélérée des entérocytes ou est simplement dû au fait que ceux-ci sont tués par le virus.

**Gérard Orth :**

Vous avez dit au début de votre exposé que ce virus dont on parle beaucoup est vraisemblablement associé au SRAS mais ce n'est pas une certitude pour l'instant. Pouvez-vous préciser ?

**H. Laude :**

Ce n'est pas à moi de me prononcer réellement sur ce sujet. Je constate que tout le monde n'est pas entièrement convaincu que ce soit l'agent étiologique exclusif.

**G. Orth :**

Que vous faudrait-il pour en être convaincu ?

**H. Laude :**

Il faudrait que des données permettant de montrer que le postulat de Koch est vérifié soient publiées. Je n'ai pas vu d'article à ce sujet.

**G. Orth :**

En fait, une courte note dans la revue *Nature* rapporte que chez le singe macaque, on peut reproduire les désordres observés chez l'homme.

**H. Laude :**

Alors qu'on ne le peut pas avec le méta pneumo-virus. C'est vrai qu'en tant que spécialiste des coronavirus, connaissant la sévérité des désordres que peuvent provoquer ces virus, je pensais effectivement qu'une association de malfaiteurs n'était pas nécessaire pour engendrer les désordres observés dans le SRAS. Néanmoins ne travaillant pas moi-même sur ce virus, je suis obligé de respecter l'opinion d'un certain nombre de collègues. Je pense qu'il vaudrait mieux discuter de ce point avec Jean-Claude Manuguerra.

**J.-C. Manuguerra :**

Toute petite question technique, nos collègues ont remarqué au laboratoire que lorsqu'on essaie de cultiver le virus du SRAS, on observe un effet cytopathogène avec lyse du tapis cellulaire relativement rapide et en quelques jours, une reconstitution du tapis, lequel a l'air normal au bout d'une semaine. As-tu déjà rencontré ce phénomène avec d'autres coronavirus sur lesquels tu as travaillé ?

**H. Laude :**

Non, et pourtant j'en ai eu pas mal entre les mains, mais je n'ai pas observé cela.

**Professeur J. Brugère-Picoux :**

Pas d'autre question ? Je poserai donc la dernière : vous avez montré justement qu'il y avait un spectre d'hôte étroit avec des flèches allant de l'homme vers le chat ou le chien mais pas dans l'autre sens. Ne peut-on pas penser à l'existence de réservoirs asymptomatiques du SRAS qui feraient vraiment craindre une épidémie à venir quand arriveront les rigueurs de l'hiver ?

**H. Laude :**

On peut malheureusement craindre que l'épidémie ne reprenne de la virulence lorsque novembre arrivera. Le problème que vous soulevez est tout à fait réel : la bibliographie, relativement ancienne, signale des phénomènes d'excrétion

inapparente, qui sont extrêmement bien documentés aussi bien pour les coronavirus du chat que ceux du porc. Quand un porc est infecté par le virus TGEV, dans la plupart des cas, il n'excrète plus de particules virales dans les fèces après quelques jours. Mais on trouve toujours quelques individus qui peuvent excréter le virus pendant plus d'une centaine de jours, alors qu'il n'y a plus aucun signe entérique, et je crois qu'une situation tout à fait analogue a été décrite dans des chatteries atteintes de virus de la péritonite infectieuse féline. Il existe donc effectivement, pour des raisons que l'on s'explique mal, des porteurs inapparents chez lesquels l'excrétion persiste longtemps et qui peuvent compliquer fortement la mise en place de la prophylaxie de cette maladie.