

Les virus à nouveau sur le devant de la scène

Patrick Forterre^{1,2}

¹ Université Paris-Sud, Institut de Génétique Microbiologie, CNRS UMR 8621, 91405 Orsay Cedex, France

² Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France

Auteur correspondant : Patrick Forterre, forterre@pasteur.fr

Reçu le 2 mai 2013

Résumé – Les virus ont longtemps été considérés comme des sous-produits de l'évolution biologique. Cette vision est en train de changer à la suite d'une série de découvertes récentes, la plus frappante étant celle des virus géants. Les virus apparaissent maintenant beaucoup plus vieux, complexes, et divers que l'on a pu l'imaginer. Les virus peuvent être définis comme des organismes qui se multiplient en produisant des virions. Cette définition exclut les ARN infectieux (parfois officiellement classés avec les virus) et implique l'existence d'au moins une protéine virale pour pouvoir parler de virion. Les virus sont donc probablement apparus après les cellules à ARN et protéines, mais avant LUCA, le dernier ancêtre commun à toutes les cellules modernes. Par ailleurs, l'assimilation classique du virus au virion a été remise en cause. L'infection entraîne la formation d'une « cellule virale » (*virocell*) qui possède tous les attributs d'un être vivant. Le concept de *virocell* met l'accent sur la phase intracellulaire du cycle viral, pendant laquelle s'exprime la « créativité » du virus, pouvant conduire à l'apparition de nouvelles protéines virales. Les virus sont la source principale de variation et de sélection, les deux piliers du schéma d'évolution darwinienne. Ces données placent les virus au centre du processus évolutif (moteur de l'évolution) et conduisent à les considérer comme des organismes à part entière.

Mots clés : Origine des virus / cellule virale / virions / arbre universel du vivant / bactériophages

Abstract – The great virus comeback.

Viruses have been considered for a long time as by-products of biological evolution. This view is changing now as a result of several recent discoveries. Viral ecologists have shown that viral particles are the most abundant biological entities on our planet, whereas metagenomic analyses have revealed an unexpected abundance and diversity of viral genes in the biosphere. Comparative genomics have highlighted the uniqueness of viral sequences, in contradiction with the traditional view of viruses as pickpockets of cellular genes. On the contrary, cellular genomes, especially eukaryotic ones, turned out to be full of genes derived from viruses or related elements (plasmids, transposons, retroelements and so on). The discovery of unusual viruses infecting archaea has shown that the viral world is much more diverse than previously thought, ruining the traditional dichotomy between bacteriophages and viruses. Finally, the discovery of giant viruses has blurred the traditional image of viruses as small entities. Furthermore, essential clues on virus history have been obtained in the last ten years. In particular, structural analyses of capsid proteins have uncovered deeply rooted homologies between viruses infecting different cellular domains, suggesting that viruses originated before the last universal common ancestor (LUCA). These studies have shown that several lineages of viruses originated independently, *i.e.*, viruses are polyphyletic. From the time of LUCA, viruses have coevolved with their hosts, and viral lineages can be viewed as lianas wrapping around the trunk, branches and leaves of the tree of life. Although viruses are very diverse, with genomes encoding from one to more than one thousand proteins, they can all be simply defined as organisms producing virions.

Virions themselves can be defined as infectious particles made of at least one protein associated with the viral nucleic acid, endowed with the capability to protect the viral genome and ensure its delivery to the infected cell. These definitions, which clearly distinguish viruses from plasmids, suggest that infectious RNA molecules that only encode an RNA replicase presently classified among viruses by the ICTV (*International Committee for the Taxonomy of Viruses*) into families of *Endornaviridae* and *Hypoviridae* are in fact RNA plasmids. Since a viral genome should encode for at least one structural protein, these definitions also imply that viruses originated after the emergence of the ribosome in an RNA-protein cellular world. Although virions are the hallmarks of viruses, viruses and virions should not be confused. The infection transforms the ribocell (cell encoding ribosomes and dividing by binary fission) into a virocell (cell producing virions) or ribovirocell (cell that produces virions but can still divide by binary fission). In the ribovirocell, two different organisms, defined by their distinct evolutionary histories, coexist in symbiosis in the same cell. The virocells or ribovirocells are the living forms of the virus, which can be *in fine* considered to be a living organism. In the virocell, the metabolism is reorganized for the production of virions, while the ability to capture and store free energy is retained, as in other cellular organisms. In the virocell, viral genomes replicate, recombine and evolve, leading to the emergence of new viral proteins and potentially novel functions. Some of these new functions can be later on transferred to the cell, explaining how viruses can play a major (often underestimated) role in the evolution of cellular organisms. The virocell concept thus helps to understand recent hypotheses suggesting that viruses played a critical role in major evolutionary transitions, such as the origin of DNA genomes or else the origin of the eukaryotic nucleus. Finally, it is more and more recognized that viruses are the major source of variation and selection in living organisms (both viruses and cells), the two pillars of darwinism. One can thus conclude that the continuous interaction between viruses and cells, all along the history of life, has been, and still is, a major engine of biological evolution.

Key words: Virus origin / virocell / virion / universal tree of life / bacteriophages

Introduction : un regard neuf sur le monde viral

Les virus se sont révélés particulièrement utiles pour les pionniers de la biologie moléculaire, comme modèles d'études simples pour mettre en évidence les principaux mécanismes universels du vivant (voir l'article de Laurent Debarbieux, Emilie Sausseureau et Damien Maura dans ce numéro). Cependant, une fois ces mécanismes fondamentaux bien établis, les virus ont perdu beaucoup de leur intérêt pour la plupart des biologistes. En particulier, la quasi totalité d'entre eux ne se sont plus posé de questions sur leur nature et leur origine, adoptant une fois pour toutes les concepts proposés autour des années 60–70 par André Lwoff et Howard Temin, basés sur leur découverte des prophages et des protovirus (Lwoff, 1957, 1966; Temin, 1971). Les biologistes se sont ainsi habitués à considérer les virus comme des sous-produits de l'évolution biologique, des entités à la limite du vivant. Les virologistes, qui auraient pu continuer à réfléchir sur la nature et l'origine des virus, avaient suffisamment de travail pour traquer et tenter de détruire les virus responsables de pathologies humaines, animales ou végétales.

Heureusement, les temps sont en train de changer, et les virus reviennent aujourd'hui sur le devant de la scène. En particulier, la question de leur nature et de leur origine se pose à nouveau. Plusieurs éléments sont à l'origine de ce revirement. Le premier est certainement la découverte en 2003 par l'équipe marseillaise du Professeur Raoult du plus gros virus connu à ce jour, le Mimivirus (voir l'article de Philippe Colson dans le prochain numéro). Les particules virales du Mimivirus sont en effet visibles en microscopie optique et son génome, long de 1,2 mégabases, est deux fois plus grand que celui de certaines bactéries (Raoult *et al.*, 2004). Ce génome code pour plus de mille protéines, une centaine d'entre elles étant localisées dans le virion lui-même. Des virus produisant des virions de grandes tailles et possédant de grands génomes étaient déjà connus (par exemple le virus de la variole), mais, dans l'imaginaire des biologistes, ce sont les petits virus qui en étaient venus depuis longtemps à symboliser l'ensemble du monde viral. En cause, l'utilisation comme virus modèles, pour des raisons pratiques, de petits virus, tel M13, Φ X174 ou λ chez les bactéries, de SV40 et du Polyoma virus chez les eucaryotes, et le retentissement médiatique des virus à ARN (SIDA, grippe, SRAS), tous de petites

tailles. L'arrivée du Mimivirus nous amenait à repenser notre vision du monde viral, en nous interrogeant sur les points communs entre les plus petits virus connus (un gène pour le virus satellite de la nécrose du tabac, SNTV) et les plus gros (plus de 1000 gènes pour le Mimivirus).

Un second facteur, qui a joué un rôle important pour replacer les virus au centre de la biosphère a été l'émergence, au tournant des années 90, d'une nouvelle discipline : l'écologie virale (Suttle, 2005 ; Rohwer & Thurber, 2009). En quelques années, plusieurs études de terrain montraient que les particules virales étaient les « entités biologiques » les plus abondantes de la biosphère. Ainsi, elles sont en moyenne dix fois plus nombreuses que les cellules bactériennes dans l'océan. L'extraordinaire abondance des virus a été depuis confirmée par la métagénomique, et plus précisément par l'analyse des « viromes » (échantillons d'ADN ou d'ARN viraux prélevés de l'environnement). Ces travaux ont montré que les génomes viraux représentent sans doute la plus grande source d'information biologique présente sur notre planète (Kristensen *et al.*, 2010).

Le séquençage des génomes et la génomique comparée ont également contribué à relancer l'intérêt pour les virus et leur origine en révélant l'extraordinaire diversité des séquences virales et leur contenu qui diffère largement de celui des séquences cellulaires. Chaque séquençage de nouveaux génomes viraux indiquait une abondance particulièrement élevée de gènes orphelins (de 50 à 100 %) et de gènes typiquement viraux, et la grande rareté de gènes viraux ayant des homologues cellulaires (de 0 à 20 %). Ces observations ne cadraient pas avec la vision traditionnelle des virus comme véhicules passifs de gènes cellulaires et comme organismes se construisant essentiellement à partir de la capture de tels gènes.

La division traditionnelle des organismes entre procaryotes et eucaryotes avait conduit à une division du monde viral entre bactériophages (virus infectant les bactéries) et « vrais » virus infectant les cellules eucaryotes. Là encore, cette vision traditionnelle allait être mise à mal à partir du milieu des années 90 avec la découverte et l'étude des virus infectant les archées, le troisième domaine du vivant (voir l'article d'Ariane Bize, Guennadi Sezonov et David Prangishvili sur les virus d'archée ou archéovirus) (Prangishvili *et al.*, 2006 ; Pina *et al.*, 2011). La plupart des archéovirus produisent en effet des virions atypiques qui leur sont spécifiques. Une minorité correspond à des virus tête-queue ressemblant à ceux des bactéries, mais de nombreux autres se distinguent par leur morphologie unique en forme de citron (avec ou sans queue), de bâtonnets terminés par des pinces, de bouteilles ou encore de ressort. Ces formes n'ont jamais été observées chez des virus infectant des bactéries (95 % des

bactériophages produisent des particules tête-queue) ou des eucaryotes (majoritairement sphériques, parfois filamenteux).

À chacun des trois domaines : Archées, Bactéries, Eucaryotes, correspond donc un ensemble de virus très différents (voir le site ViralZone pour la description exhaustive des différentes familles de virus dans les trois domaines du vivant) (Hulo *et al.*, 2011). Avec la découverte des virus d'archées, le monde viral, qui semblait un peu monotone, reprenait des couleurs, une nouvelle fenêtre s'ouvrait sur la diversité virale, les virus connus jusque là ne constituaient finalement que le sommet d'un iceberg qui restait à explorer.

Les virus sont sans doute apparus avant LUCA (*the Last Universal Common Ancestor*)

L'association de trois espaces viraux distincts à chacun des trois domaines cellulaires aurait pu renforcer l'idée d'une origine indépendante des virus appartenant à ces trois espaces, une origine qui aurait pris sa source dans les domaines eux-mêmes (les bactériophages provenant par exemple de gènes bactériens, ancêtres communs aux prophages et aux phages). Cette vieille idée, déjà battue en brèche par la rareté des gènes cellulaires présents dans les génomes viraux, allait subir le coup de grâce de la biologie structurale. En contradiction apparente avec la diversité des trois ensembles viraux, la résolution de la structure des protéines de capsides et le séquençage des ATPases permettant l'empaquetage de l'ADN viral ont mis en évidence des caractères homologues (dérivés d'un ancêtre commun) présents chez des virus infectant des organismes aussi éloignés que l'Homme et la bactérie. Ces travaux ont permis de ramener l'extraordinaire diversité virale à un nombre relativement restreint de lignées évolutives qui regroupent des virus en apparence très différents (Abrescia *et al.*, 2012).

On compte déjà aujourd'hui quatre grandes lignées de virus définies par la structure de leurs virions, en particulier par celle des protéines de capsides, deux pour les virus à ARN, et deux pour les virus à ADN (Abrescia *et al.*, 2012). La résolution de nouvelles structures et la découverte de nouveaux virus devraient conduire à l'identification d'un bien plus grand nombre de lignées virales dans les années à venir. Les quatre lignées majeures déjà définies ne présentent pas d'homologies entre elles, ce qui montre que les virus ne sont pas monophylétiques. En clair, cela signifie que le mode de vie associé aux virus est apparu plusieurs fois indépendamment au cours de l'évolution du vivant sur notre planète.

Les deux grandes lignées de virus à ADN déjà identifiées sont universelles, en ce sens qu'elles comptent

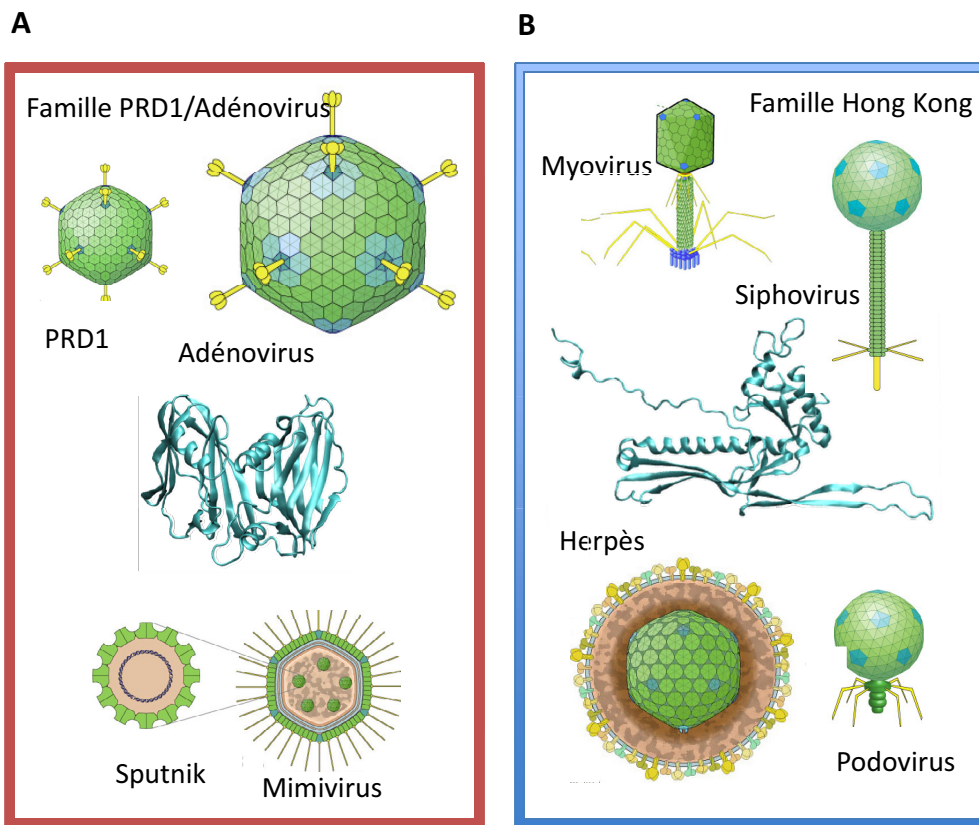


Fig. 1. Les deux grandes lignées évolutives de virus à ADN, à gauche, la lignée PRD1/Adénovirus caractérisée par le repliement protéique de type « double jelly roll » (au centre) présent dans leurs protéines majeures de capsid. À droite, la lignée HK97, caractérisée par le repliement protéique de type « Hong Kong » (au centre) présent dans leurs protéines majeures de capsid. Les schémas, qui ne sont pas tous à la même échelle, proviennent du site ViralZone (www.expasy.org/viralzone Swiss Institute of Bioinformatics) (Hulo *et al.*, 2011).

des représentants dans chacun des trois espaces viraux correspondant aux trois domaines cellulaires (on ne connaît pas actuellement d'archéovirus à ADN) (Abrescia *et al.*, 2012). La première (identifiée à la fin du siècle dernier) correspond à une lignée baptisée à partir de deux de ses représentants, PRD1/Adénovirus. Les virions de cette lignée possèdent tous une membrane lipidique interne (à l'exception des adénovirus!) et utilisent des ATPases homologues (de la famille FtsK/HerA) pour emballer leur ADN. Par ailleurs, leurs protéines majeures de capsid sont toutes caractérisées par un repliement complexe de type « double jelly roll fold », ce qui peut se traduire en français par « double gâteau roulé » (figure 1A). Cette lignée comprend les bactériovirus (bactériophages) de la famille des *Tectiviridae*, les archéovirus STIV (*Sulfolobus Islandicus Turreted Virus*) et plusieurs familles de virus infectant des organismes eucaryotes, les adénovirus, mais aussi tous les Mégavirales, un ordre viral comprenant plusieurs familles de grands virus eucaryotes à ADN, dont les *Mimiviridae* (Colson *et al.*, 2012) et

les virus de virus (ou virophages) qui infectent certains de ces virus géants (La Scola *et al.*, 2008; Desnue & Raoult, 2012).

La seconde lignée universelle correspond aux virus tête-queue des bactéries et des archées (Caudavirales) et aux eucaryovirus de la famille de l'Herpès (figure 1B). Tous ces virus possèdent une protéine majeure de capsid caractérisée par un repliement de type « Hong Kong », d'après le nom du premier virus de cette famille dont la structure de la protéine de capsid a été résolue, le bactériophage HK97. Tous ces virus « Hong Kong » possèdent également des ATPases d'emballage de l'ADN homologues entre elles (de la famille TerL) et différentes (non homologues) de celles utilisées par les virus de la lignée Adénovirus/PRD1 (ces deux types d'ATPases appartiennent en effet à deux superfamilles de protéines différentes). La relation évolutive entre bactériophages tête-queue et virus de l'Herpès n'avait pas été soupçonnée avant l'analyse structurale de leurs protéines de capsid. Cette parenté, *a priori* surprenante, a été confirmée récemment par la découverte d'une petite queue présente sur la

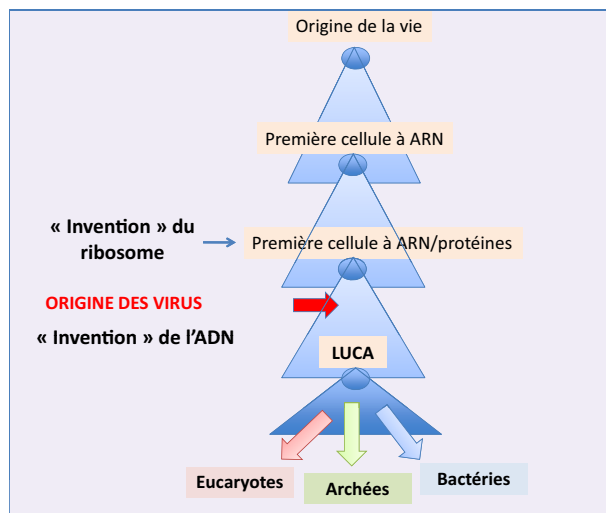


Fig. 3. Scénario évolutif schématisé des premières étapes de l'évolution : des origines de la vie jusqu'aux cellules modernes. Chaque grande transition correspond à un goulot d'étranglement, seuls survivent à terme les descendants de la cellule fondatrice. Les premiers virus à ARN apparaissent dans le monde à ARN protéines, également appelé deuxième âge du monde à ARN (Forterre, 2005). On ne sait pas si l'ADN est apparu avant ou après LUCA (Forterre, 2002).

pour expliquer leur origine et pour comprendre leur nature. Dans certains scénarios qui essaient d'expliquer les premières étapes de l'évolution, les virus jouent désormais un rôle de premier plan. Par exemple, deux auteurs ont proposé indépendamment il y a une dizaine d'années l'hypothèse selon laquelle le noyau des cellules eucaryotes était d'origine virale (Bell, 2001; Takemura, 2001) et j'ai proposé d'expliquer la transition des génomes à ARN vers les génomes à ADN par la compétition entre virus et cellules à ARN (Forterre, 2002, 2005, 2006). Cette compétition aurait conduit certains virus à ARN à modifier chimiquement leur génome (le transformant en ADN) pour contourner les mécanismes anti-ARN de leurs « hôtes ». De nouveaux concepts ont également été proposés concernant la nature des virus. Certaines propositions avancées peuvent parfois sembler contradictoires. Ainsi, Dennis Bamford, virologue à Helsinki, propose de considérer la capsid comme le « moi » (*self*) du virus (Bamford, 2003), alors que d'autres, tel Jean-Michel Claverie à Marseille, proposent de se focaliser sur l'étape intracellulaire du cycle viral (Claverie, 2006). J'ai suggéré de mon côté une nouvelle piste qui permet, à mon avis, de dépasser ces apparentes contradictions en mettant en avant le concept de cellule virale, définie comme cellule produisant des virions (Forterre, 2010a, 2010b). Avant de présenter ces nouvelles théories et ces nouveaux

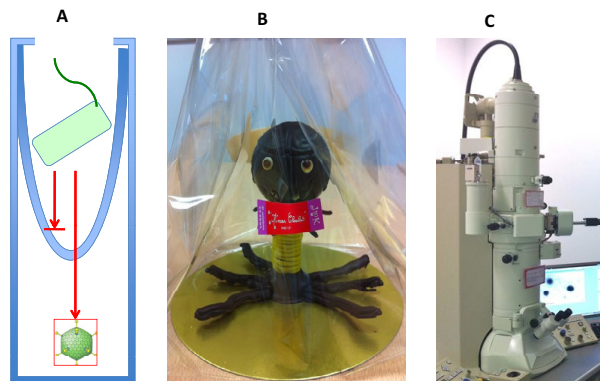


Fig. 4. Assimilation du virus au virion. Historiquement, le virus a été assimilé à l'agent infectieux qui traverse un filtre de porcelaine (A). Le filtre retient les cellules, y compris les cellules virales. Le virus a été définitivement assimilé au virion lorsque ce dernier est devenu visible en microscopie électronique. Le bactériophage a été visualisé comme un petit animal avec une tête, une queue, et parfois des pattes (*myoviridae*) (B).

concepts, je vais toutefois commencer par faire un bref rappel des événements qui ont conduit à la vision traditionnelle des virus telle qu'elle est encore défendue par certains auteurs (voir par exemple, Moreira & Lopez-Garcia, 2009; Lopez-Garcia, 2012; Lopez-Garcia & Moreira, 2012).

La confusion entre virus et virions : le virus comme entité biologique non vivante

Les virus ont été découverts à la fin du XIX^e siècle grâce à la propriété qu'ont leurs virions infectieux de traverser les filtres de porcelaine (Bos, 1999) (figure 4A). Cette observation a conduit dès le départ à associer le virus avec « ce qui passe à travers le filtre ». Par la suite, la ressemblance entre le virion des bactériophages et un petit animal (avec sa tête, sa queue et ses pattes) a renforcé l'identification du virus avec son virion (figure 4B). Pour le grand public, cette confusion est entretenue par la diffusion répétée sur les écrans télévisés d'images de virus représentés sous la forme d'une menaçante sphère épineuse.

Dans les années qui ont suivi leur découverte, les virus ont été considérés comme des êtres vivants, le filtrat étant assimilé parfois à un « fluide contagieux vivant ». En effet, les virus possédaient toutes les qualités de la vie : reproduction, multiplication, évolution. On pouvait encore, à l'époque, imaginer les virus sous forme de « bactéries » miniatures, invisibles aux yeux du microscope optique. Cette situation a changé au tournant du siècle dernier avec la visualisation des particules virales, rendue possible

par la mise au point du microscope électronique (figure 4C). La caractérisation biochimique de ces particules a mis en évidence leur simplicité, certaines ont même pu être cristallisées! Le caractère inerte du virion (absence de tout métabolisme) et la relative simplicité de son organisation ont donc amené logiquement les biologistes à considérer les virus comme des assemblages macromoléculaires plutôt que comme des êtres vivants. Les virus ont ainsi été officiellement considérés comme « non vivants » par l'ICTV (*International Committee on Taxonomy of Viruses*) (Van Regenmortel, 2004).

Bien sûr, les biologistes savent tous que le virus ne peut se résumer à sa particule virale et qu'il peut « exister » sous différentes formes : virion, génome activement répliqué et multiplié dans la cellule infectée ou passivement répliqué, une fois intégré dans le génome cellulaire (lysogénie). Mais seule l'étape « virion » permet de donner une forme au virus, de définir physiquement son identité. La phase du cycle viral correspondant à la cellule infectée a été baptisée « phase végétative » ou encore « phase d'éclipse », ce dernier terme montrant bien que la disparition du virion est assimilée à celle du virus « éclipsé ». *In fine*, tout en sachant que le virus ne peut se résumer à sa particule virale, les biologistes ont fini par assimiler dans la pratique le virus et la particule virale. Ceci explique par exemple pourquoi les virus ont été définis comme ne contenant qu'un seul type d'acide nucléique, ARN ou ADN (Lwoff, 1966). Il est frappant de constater que, dans cette définition, l'ARN messager des virus à ADN passe par pertes et profits. En fait, les virus à ADN ont bien deux types d'acides nucléiques, ADN et ARN, comme les organismes cellulaires, sauf si nous les identifions arbitrairement à leurs virions.

L'assimilation des virus à leurs virions a également conduit à l'idée implicite selon laquelle toute activité virale était issue *in fine* d'une activité cellulaire, ce qui revient à nier les capacités créatrices des virus. Dans cette vision réductrice, les virus ne pouvaient pas porter d'innovation évolutive en soi, indépendamment du monde cellulaire, ce qui explique pourquoi ils ont en grande partie disparu des scénarios évolutifs retraçant l'histoire du vivant (Koonin & Wolf, 2012). Dans la pratique, de nombreux évolutionnistes en sont venus à considérer que tout gène viral avait une origine cellulaire, ce qui a abouti au concept du virus « pickpocket » (Moreira & Lopez-Garcia, 2009). Pour ces biologistes, les génomes viraux ne sont en fin de compte qu'une mosaïque de gènes empruntés au cours du temps à leurs hôtes cellulaires (voir l'analyse du génome du Mimivirus par Moreira & Brochier-Armanet, 2009, et sa critique par Forterre, 2010a). Je pense que cette vision a été nourrie par l'utilisation intensive des « phages » transducteurs dans les années 1970–1980 par les généticiens, et par la découverte

de quelques gènes d'origine cellulaire dans certains génomes viraux. Plus récemment, elle a été entretenue par l'utilisation des virus comme porteurs de gènes cellulaires en thérapie génique. Dans tous ces cas, l'accent est mis sur le gène cellulaire, le virus étant conçu comme un véhicule passif au service de l'évolution des cellules (transfert de gènes) ou de la technologie humaine (thérapie génique).

Un nouveau concept : la cellule virale (*virocell*)

L'assimilation (confusion) du virus au virion a été critiquée au siècle dernier par Claudiu Bandea aux Etats-Unis, mais sans grand impact (Bandea, 1983). C'est Jean-Michel Claverie, responsable de l'équipe qui a séquencé le génome du Mimivirus à Marseille, qui a le premier émis une critique forte et pertinente de cette assimilation en faisant remarquer que confondre virus et virion revenait à confondre l'homme avec ses spermatozoïdes (Claverie, 2006). Le virion correspond en fait au messenger moléculaire qui permet au virus de transmettre son information génétique. Mais alors, où est le virus? Jean-Michel Claverie, en se basant sur le cas de figure du Mimivirus, a proposé d'assimiler le virus, non plus au virion, mais à l'usine virale, considérant cette dernière comme le véritable organisme viral. Les virus eucaryotes à ARN ou ADN qui se répliquent dans le cytoplasme produisent en effet, au cours de l'infection, une structure complexe, l'usine virale, parfois aussi appelée virosome ou mini-noyau (dans le cas du virus de la vaccine) et qui ressemble en effet à un noyau miniature (Miller & Krijnse-Locker, 2008). Dans le cas du Mimivirus, cette usine cellulaire est en fait aussi grosse, en taille, que le noyau de l'amibe infectée!

Définir le virus comme l'usine virale pose toutefois un problème : en effet, de nombreux virus ne produisent pas de telles usines. C'est le cas des virus qui infectent les archées ou les bactéries. Pour avancer dans le débat, j'ai proposé récemment le concept de cellule virale (*virocell*) pour mettre en avant l'étape intracellulaire du cycle viral, la cellule virale correspondant à la cellule infectée (Forterre, 2010b, 2012, 2013). Cette proposition renvoie à une expression d'André Lwoff qui, en 1965, avait écrit dans son discours pour la remise de son prix Nobel : « la bactérie infectée par un bactériophage virulent devient une usine à virus » (Lwoff, 1966). Si l'on rejette l'assimilation virus/virion, on en déduit que le virus transforme la cellule en usine à virion, c'est-à-dire... en virus. Dans cette optique, le virus devient donc lui-même un organisme « cellulaire », certes atypique par rapport aux organismes cellulaires classiques (Archées, Bactéries et Eucaryotes) mais cellulaire pour la partie active de

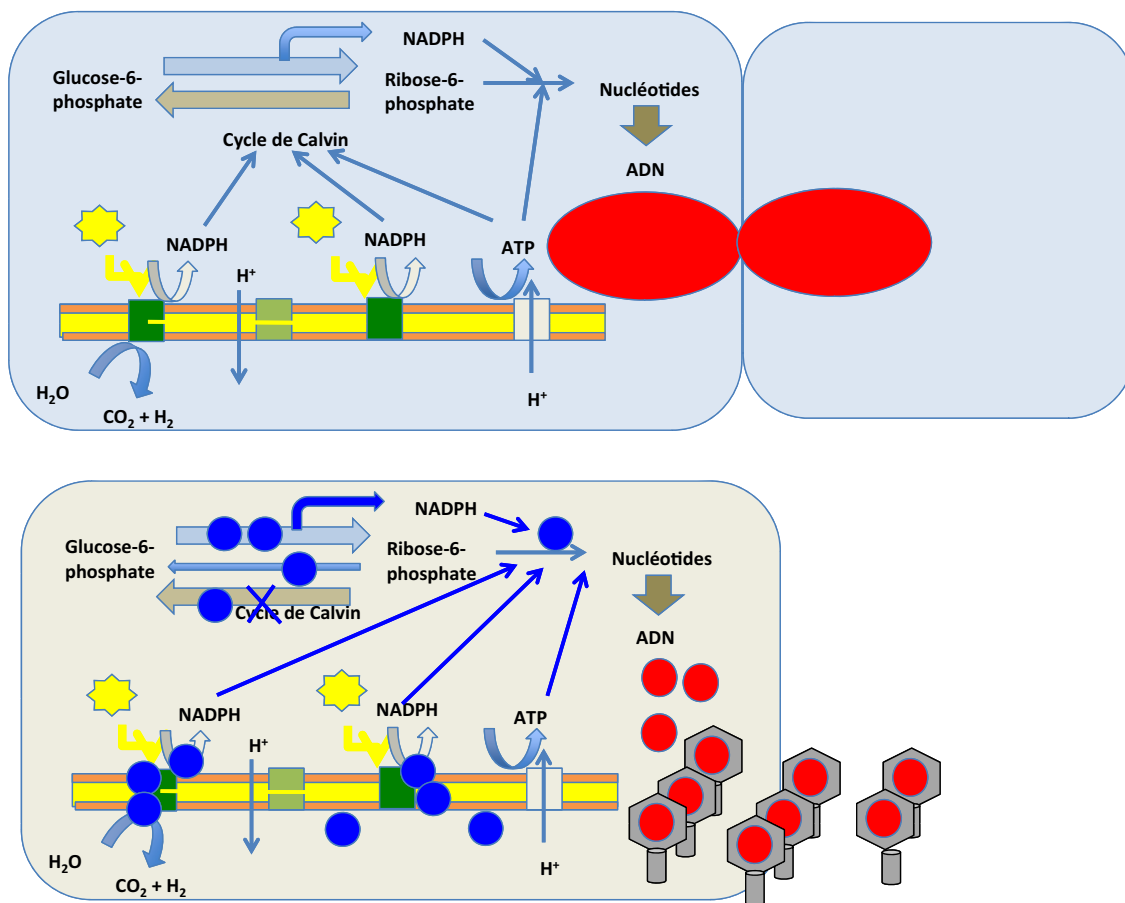


Fig. 5. Transformation d'une cyanobactérie en cellule virale photosynthétique par un cyanovirus. Le génome viral code pour plusieurs protéines (ronds bleus) qui vont permettre de poursuivre la photosynthèse après la destruction du génome cellulaire et de rediriger le métabolisme cellulaire vers la synthèse d'ADN viral (adapté de Thompson *et al.*, 2011).

son cycle de vie. Du coup, le concept de cellule virale permet de concilier l'idée selon laquelle les virus sont des organismes vivants et l'idée souvent exprimée selon laquelle tous les organismes vivants doivent être, par définition, cellulaires (Van Regenmortel, 2004).

Un argument souvent avancé pour refuser aux virus le qualificatif de vivant est l'absence de métabolisme. Les virus seraient incapables de transformer l'énergie en matière et de créer de l'ordre à partir du désordre (Van Regenmortel, 2004; Lopez-Garcia & Moreira, 2012). Cette remarque est valable si l'on continue d'assimiler le virus au virion. Par contre, elle ne tient pas si l'on réfléchit en termes de cellule virale. Dans cette dernière, l'information apportée par le génome viral réoriente le métabolisme cellulaire vers la production de virions, la cellule virale continue de transformer l'énergie en matière et de créer de l'ordre à partir du désordre. Il faut noter que de nombreux virus codent pour des protéines du métabolisme qui participent directement à ce processus, par exemple des protéines

de la photosynthèse dans le cas de certains virus qui infectent des cyanobactéries (Sharon *et al.*, 2009; Thompson *et al.*, 2011) (figure 5). Récemment, une étude exhaustive des protéines codées par les gènes viraux présents dans l'environnement a mis en évidence la grande diversité des enzymes métaboliques codées dans les génomes viraux, parmi lesquelles des ATP synthases ou encore des enzymes de la glycolyse (Mart Krupovic, communication personnelle). Les enzymes métaboliques codées par les virus sont souvent très divergentes de leurs homologues cellulaires, en termes de séquence. Elles ont pu être acquises par les virus depuis très longtemps, mais, du fait de cette divergence, on ne peut pas non plus exclure que, dans un certain nombre de cas, elles auraient pu apparaître en premier dans le monde viral avant d'être transmises aux cellules (Hamacher *et al.*, 2012)! La figure 6 présente un exemple frappant de transformation d'une *ribocell*, dans ce cas une cellule épithéliale humaine, en cellule virale infectée par le virus de la vaccine. Le comportement de la cellule infectée provoque la destruction

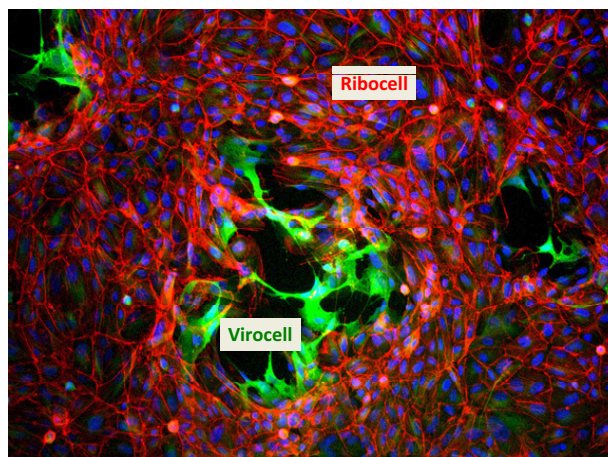


Fig. 6. Cellules épithéliales humaines infectées par le virus de la vaccine. Les cellules non infectées sont colorées en rouge, les cellules infectées en vert (*Cell Picture Show*, photo prise par Yoshiaki Arakawa et Michael Way, du *London Research Institute*, http://www.cell.com/cell_picture_show-Viruses). Les *virocells* ont perdu leurs contacts intercellulaires et deviennent capables de migrer en dehors de leur site d'origine.

de la structure régulière du tissu épithélial, ce qui va faciliter la diffusion des virions sur de longues distances. La cellule infectée ne travaille plus pour l'organisme eucaryote qui l'héberge, mais bien pour le virus qui en a pris le contrôle.

En replaçant les virus au sein du monde vivant, le concept de cellule virale résout une situation paradoxale. Nous avons vu que la plus grande partie de l'information génétique présente dans la biosphère est sans doute d'origine virale. Considérer les virus comme non vivants revient donc à dire que la plus grande partie de l'information génétique est produite par des entités « non vivantes » ! De même, il est de plus en plus admis que les virus ont joué et jouent toujours un rôle majeur dans l'évolution. Il paraît donc paradoxal de considérer que des entités « non vivantes » ont pu jouer un tel rôle.

Bien sûr, il ne faut pas plus confondre virus et cellule virale que virus et virion, le virion et la cellule virale sont deux états transitoires, chacun associé à une étape clef du cycle viral. Ces deux étapes doivent être incluses dans toute description exhaustive des virus. Le virus apparaît comme un processus qui se matérialise tout au long du cycle viral sous différentes formes, le virion et la cellule virale étant les deux plus caractéristiques.

Le concept de cellule virale a été violemment caricaturé par des chercheurs attachés au caractère non vivant des virus ; selon eux, si l'on suit cette idée, le poisson rémora, parasite du requin, transformerait le

requin en « requin rémora », ce qui est effectivement un non-sens (Lopez-Garcia & Moreira, 2012). Cette tentative de discréditer par l'absurde le concept de cellule virale renvoie à une question qui m'est régulièrement posée lorsque je parle du concept de cellule virale devant un public de scientifiques : « si l'on accepte le concept de cellule virale, doit-on en conclure qu'une bactérie parasite intracellulaire transforme la cellule eucaryote infectée en bactérie ? ». La réponse est bien entendu négative. Il est toutefois intéressant de réfléchir à cette question en prenant comme exemple un cas extrême, celui des *Chlamydiae*, bactéries parasites intracellulaires qui existent sous deux formes : une forme infectieuse (*the elementary body*) pendant laquelle la bactérie se condense pour former une structure pouvant être assimilée à un virion et une forme « végétative » (*the reticulate body*) pendant laquelle la bactérie ressemble... à une bactérie (Moulder, 1991) (figure 7). Après son entrée dans la cellule, le « *corps élémentaire* » se transforme en « *corps réticulé* » à l'intérieur d'une vacuole et la bactérie commence à se diviser sous sa forme « classique ». Les nouvelles bactéries ainsi formées vont s'accumuler dans la vacuole dont la taille augmente en conséquence. Elles vont ensuite se transformer en *corps élémentaires* infectieux qui vont être libérés de la cellule dans le milieu extérieur à la suite de la lyse cellulaire. Ce schéma, qui en apparence ressemble à un cycle viral, s'en distingue par plusieurs points. Tout d'abord, les protéines de la bactérie sont synthétisées tout au long du cycle par les ribosomes bactériens ; ensuite, la continuité physique de la bactérie est maintenue, elle aussi, tout au long du cycle par sa membrane qui peut se transformer, mais qui ne disparaît jamais. Ainsi la bactérie n'a pas besoin de « s'emparer » du corps de la cellule pour former son propre corps, elle l'occupe, tout en conservant son intégrité. Le comportement du virus est très différent, il se substitue en quelque sorte à la cellule infectée pour mettre à son service sa machinerie moléculaire en la transformant de l'intérieur. Un peu comme dans les films de science-fiction où le cerveau d'un personnage est remplacé par celui d'un tiers, le génome viral remplace celui de la cellule infectée. Les virus et les organismes « classiques », archées, bactéries et eucaryotes, sont bien deux types d'organismes radicalement différents, même s'ils font tout deux partie de la grande famille du vivant.

En parallèle au terme de cellule virale, j'ai proposé pour nommer les cellules « classiques » le terme de *ribocell*, cellules à ribosomes. Le fait de coder pour ses propres ribosomes est en effet la caractéristique majeure de ces organismes, par opposition aux virus. Par ailleurs, les trois grandes lignées cellulaires (Archées, Bactéries et Eucaryotes) ont été précisément définies par leurs ribosomes. Toutefois, le monde vivant ne se limite pas aux *ribocells* et aux *virocells*. On connaît

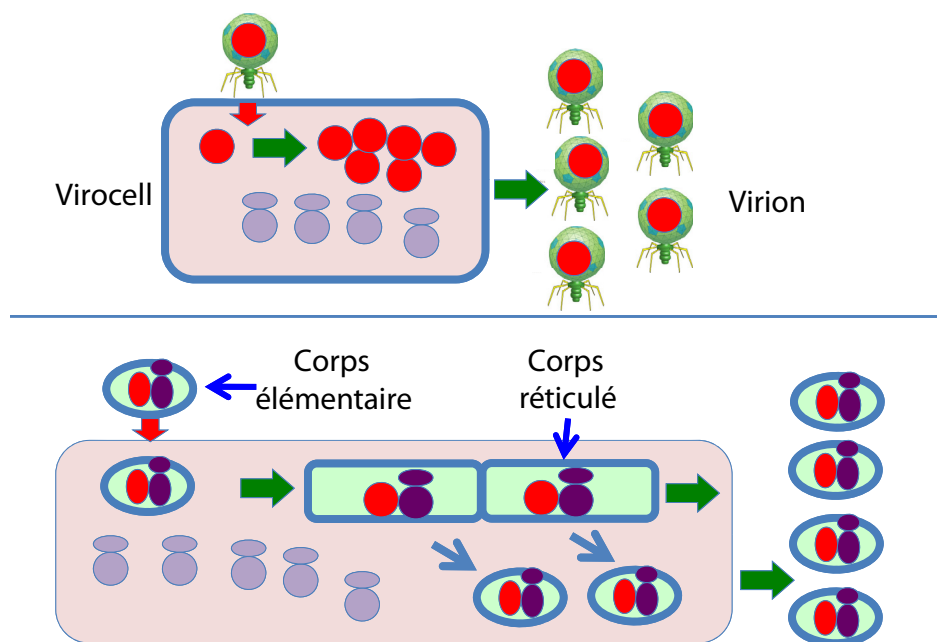


Fig. 7. Comparaison schématique du cycle d'un virus avec celui d'une *Chlamydiae*. En haut, le génome viral s'exprime et se réplique dans le cytoplasme et/ou le noyau de la cellule infectée et utilisent ses ribosomes, se substituant ainsi à la cellule infectée et la transformant en cellule virale. En bas, la bactérie garde son intégrité (assurée par sa membrane) à l'intérieur de la cellule infectée en dépit de profondes transformations structurales. Elle continue d'utiliser ses propres ribosomes. La cellule infectée est modifiée, mais elle garde un territoire distinct de celui occupé par les bactéries. La cellule eucaryote infectée ne devient donc pas une « cellule bactérienne ».

de nombreux exemples de « symbioses » entre virus et cellules ; par suite de l'infection, la cellule produit continuellement des virions (généralement en petite quantité) tout en restant capable de se diviser (un cas relativement fréquent, appelé état porteur, *carrier state*). Pour nommer ce type de cellules, j'ai proposé le terme un peu barbare de *ribovirocell* ! Ce cas de figure ne doit pas être confondu avec la lysogénie. À mon avis, les cellules qui possèdent des génomes viraux intégrés mais inactifs restent de simples *ribocells*, même si nous pouvons les considérer comme des cellules virales en puissance, attendant le bon agent inducteur pour se « réveiller ».

Le concept de cellule virale et ses conséquences

Le concept de *virocell* a de nombreuses implications. Au début de l'infection, la cellule est déjà une cellule virale tout en restant, pour quelque temps, une cellule de bactérie, d'archée ou d'eucaryote qui tente de se défendre contre l'attaque du virus. Deux organismes différents sont donc en lutte et coexistent au sein de la même cellule. Comme dans le cas des organismes multicellulaires, nous sommes donc amenés à déconnecter la notion de cellule et celle d'organisme.

Ces organismes qui cohabitent peuvent échanger des gènes, mais ils ne se confondent pas pour donner une chimère. Une amibe peut ainsi être infectée par un Mimivirus, lui-même infecté par un virus satellite (ou virophage) appelé « *sputnik* » (un virus de virus selon la dernière terminologie proposée par Didier Raoult) (Desnué & Raoult, 2012). Par ailleurs, cette amibe (en voie de disparition) peut héberger des bactéries endosymbiontes ! Ces différents organismes échangent des gènes, mais ils gardent chacun leur individualité. Nous avons bien à faire à une cellule qui héberge quatre organismes différents, un eucaryote (l'amibe), une bactérie et deux virus (Slimani *et al.*, 2013). Je pense que la notion d'organisme doit être étroitement associée à la notion de lignée évolutive. Dans le cas ci-dessus, nous avons bien en effet quatre lignées évolutives différentes (cinq si l'on compte les mitochondries) pour une seule lignée « cellulaire ». Matti Jäalasvuori propose de considérer la cellule comme un véhicule (ou encore un vaisseau) capable de transporter différents types de réplicons (Jäalasvuori, 2012). Au delà de simples réplicons (plasmides, génomes viraux), ce sont des organismes qui sont bel et bien transportés par le vaisseau cellulaire.

Une autre conséquence du concept de cellule virale concerne la quantification des virus dans l'environnement. Les écologistes qui comptent les virus

dans la nature, généralement par épifluorescence, énumèrent en fait les particules virales (les virions). Ils en concluent, en général, que les virus sont dix fois plus abondants que les cellules dans presque tous les biotopes étudiés jusqu'à présent. J'ai l'habitude de dire que cela revient à compter les poissons en comptant leurs œufs ! La conséquence dans ce cas serait l'abandon des quotas de pêche, ... vu l'abondance des poissons !! En fait, les virus ne peuvent pas être plus abondants que les cellules si l'on considère la cellule virale comme un élément constitutif essentiel du virus puisque cette dernière ne peut correspondre qu'à une fraction (plus ou moins grande) des cellules présentes dans l'environnement. Par la même logique, les cellules observées dans un environnement donné (bactéries, archées ou protistes) ne sont pas toutes de véritables cellules rêvant de se diviser, mais, dans une grande proportion des cas (jusqu'à 40 % dans l'océan) des cellules ... virales qui vont bientôt disparaître en relâchant leurs virions dans l'environnement !

Le concept de *virocell* correspond à une réalité biologique, la cellule infectée. Nous avons vu que dans la vision traditionnelle, cette phase du cycle viral est appelée « phase d'éclipse » car le virus, assimilé au virion, semble avoir disparu (!) ou encore phase « végétative » car le génome viral serait répliqué « passivement » par la cellule. Ces connotations négatives ont contribué à une sous-estimation de l'activité créatrice des virus. Pour moi, le véritable intérêt du terme *virocell* est précisément de mettre l'accent sur la phase du cycle viral pendant laquelle peuvent se former de nouveaux gènes viraux. Le concept de cellule virale permet ainsi de mieux comprendre l'extraordinaire potentiel de création associé aux virus. En effet, de nouveaux gènes apparaissent continuellement dans les génomes viraux par les mêmes mécanismes que ceux qui assurent l'apparition de nouveaux gènes dans les génomes cellulaires : erreurs de réplication (glissement de polymérase, mutations ponctuelles) ou de recombinaison (homologue et non homologue) aboutissant à des insertions, délétions, duplications, etc. De plus, ces processus sont amplifiés dans le monde viral, car les génomes viraux sont beaucoup plus nombreux et se répliquent beaucoup plus rapidement que les génomes cellulaires.

Le rôle des virus dans l'évolution

De nombreux gènes formés dans le monde viral sont ensuite transférés aux *ribocells* par suite de l'intégration fréquente de génomes viraux dans les génomes cellulaires. La génomique comparée a en effet montré que, si les gènes d'origine cellulaire sont relativement rares dans les génomes viraux (la taille des particules virales imposant une limite à cette

intégration), tous les génomes cellulaires sont remplis de génomes viraux intégrés ou de leurs reliques. Il faut rappeler par exemple que le génome humain comprend quatre fois plus de gènes correspondant à des rétrovirus endogènes que de gènes humains (Ryan, 2011). Chez les bactéries et les archées, c'est en moyenne 10 % à 15 % du génome qui est constitué de gènes viraux ou plasmidiques récemment intégrés, ce qui représente un très grand nombre de gènes (Cortez *et al.*, 2009). Si de nombreux gènes viraux sont éliminés au cours de l'évolution des génomes cellulaires, d'autres se fixent en permanence par sélection, s'ils répondent à certains besoins de la cellule. Les virus sont donc *in fine* à l'origine de nombreuses protéines aujourd'hui cellulaires (certains auteurs parlent de domestication à leur sujet). Ce flux continu de gènes viraux dans les génomes cellulaires explique le rôle très important que les virus ont joué et qu'ils jouent encore dans l'évolution (Brüssow, 2009 ; Forterre & Prangishvili, 2009).

La génomique comparée aboutit donc à une vision à l'opposé de la vision traditionnelle qui assimile le virus à un pickpocket de gènes cellulaires. En fait, ce sont les cellules qui sont des « pickpockets » très efficaces de gènes viraux, et non l'inverse. Dans certains cas, les gènes viraux remplacent des gènes codant des protéines qui possédaient la même fonction. Un événement de ce type s'est produit au cours de l'évolution, qui a conduit à la formation des mitochondries (Filée & Forterre, 2005). Les gènes codant pour l'ARN polymérase, l'ADN polymérase et l'ADN hélicase répliquative de la bactérie à l'origine des mitochondries ont été remplacés par les gènes d'un provirus qui était intégré dans le génome de cette bactérie. Aujourd'hui, ce sont donc des protéines d'origine virale (du type T3/T7) qui répliquent et transcrivent le génome mitochondrial. Dans d'autres cas, le gène viral apporte une véritable nouveauté qui peut changer le destin d'une lignée, ou même d'un organisme cellulaire. Ainsi, on sait depuis les années 2000 que la formation du placenta chez les mammifères implique des protéines codées par des rétrovirus endogènes (de Parseval & Heidmann, 2005). Ces protéines virales ont des propriétés immunosuppressives, ce qui expliquerait pourquoi le système immunitaire de la mère ne reconnaît pas l'embryon comme un corps étranger.

Finalement, le concept de cellule virale permet de mieux comprendre (et d'admettre, au moins à titre de spéculations) certaines hypothèses qui donnent aux virus un rôle décisif à certaines étapes de l'évolution biologique. Ainsi, nous avons vu que deux auteurs ont proposé au début de ce siècle l'existence d'un lien de parenté entre les Mégavirales et le noyau des cellules eucaryotes (Bell, 2001 ; Takemura, 2001). Cette hypothèse paraît ridicule si l'on s'imagine un virion se transformant en noyau ! Elle devient toutefois plus

crédible si l'on se tourne vers la cellule infectée. Ainsi, dans le cas du Mimivirus, on s'aperçoit que la cellule virale comprend une usine virale géante (son noyau ?) de même taille que le noyau de l'amibe infectée ! Cela ne veut pas dire que l'hypothèse d'une origine virale du noyau est correcte, mais cela la rend plausible. Les ressemblances entre les usines virales cytoplasmiques des Mégavirales et les noyaux des cellules eucaryotes sont frappantes. En particulier, dans les deux cas, les chromosomes et les usines de réplication associées s'entourent parfois d'une membrane formée à partir de la membrane du réticulum endoplasmique. Ces ressemblances peuvent au moins suggérer l'existence de relations évolutives entre ces deux structures. La structure des chromosomes eucaryotes elle-même a été déterminée par des éléments d'origine virale. Ainsi, les télomères sont apparentés aux rétrovirus (la télomérase étant homologue des transcriptases inverses) et les centromères sont riches en rétroéléments. Les virus (en particulier les Mégavirales et les rétrovirus) ont donc très certainement joué un rôle important dans l'origine et l'évolution des cellules eucaryotes.

L'hypothèse d'une origine virale de l'ADN est un autre exemple d'hypothèse qui peut paraître incongrue si l'on s'en tient à la vision classique du virus. Dans cette hypothèse, la transformation de l'ARN viral en ADN, modification chimique qui lui permet d'échapper aux nucléases de l'hôte, se produit bien dans la cellule virale, catalysée par des enzymes (ribonucleotide réductases, thymidylate synthases) recrutées à partir de la cellule infectée ou directement produites à partir du génome viral. Le concept de cellule virale permet ainsi de « penser » le virus autrement, en réhabilitant sa capacité d'innovation.

Comment définir le virus ?

Il nous reste à aborder la question de fond : Qu'est ce qu'un virus ? « C'est un virus ! », répondait André Lwoff, pour souligner leur originalité ; les virus n'étant ni procaryotes, ni eucaryotes. En 2008, Didier Raoult et moi-même avons proposé de définir les virus comme des « organismes codant des capsides » par opposition aux cellules, « organismes codant des ribosomes » (Raoult & Forterre, 2008). Cette définition met l'accent sur les deux caractères majeurs des virus – production de virions et absence de ribosomes – et donc sur la nécessité d'un parasitisme obligatoire. Par opposition au mode de division binaire des cellules (une cellule mère donne deux cellules filles), le fait de se reproduire en produisant des virions infectieux correspond bien à la marque de fabrique du virus. Si, comme le disait François Jacob, « *le rêve d'une cellule est de donner deux cellules* », le rêve d'une cellule virale sera

de produire autant de virions que possible pour assurer sa descendance (plus de 10 000 dans le cas d'une cellule humaine infectée par le virus du SIDA !).

Par rapport à notre proposition initiale, organismes « codant des capsides », je définirai plutôt aujourd'hui les virus comme des organismes « producteurs de virions ». En effet, le terme « capsid » peut paraître trop restrictif à certains auteurs qui le réservent aux particules virales icosaoédriques ; or de nombreux virions présentent des formes très différentes. Par exemple, nous avons récemment découvert à l'Institut Pasteur un virus dont la particule virale est formée par une hélice d'ADN simple brin recouvert de protéines. Cette hélice s'enroule pour former une spirale régulière semblable à un ressort très serré qui correspond au virion (Mochizuki *et al.*, 2012). *A minima*, on peut donc considérer le virion comme une structure dans laquelle l'acide nucléique viral est associé à une ou plusieurs protéines pour protéger son génome et faciliter sa dissémination. *A priori*, une seule protéine peut suffire. Il peut s'agir d'une protéine permettant la formation d'une capsid de géométrie régulière, d'une protéine se fixant dans la membrane d'une vésicule lipidique pour la stabiliser ou encore d'une protéine permettant de former un filament nucléoprotéique. Dans de nombreux cas bien sûr, les virions sont beaucoup plus complexes et formés de nombreuses protéines, comprenant parfois des membranes lipidiques.

La notion selon laquelle la définition du virus va impliquer l'existence d'au moins une protéine a été récemment brillamment illustrée par Dennis Bamford et Mart Krupovic, qui ont comparé le génome d'un virus ne possédant que deux gènes, l'un codant pour une protéine de capsid et l'autre pour une protéine de réplication, avec celui d'un plasmide qui ne code que pour une seule protéine, celle de sa réplication (Krupovic & Bamford, 2010). La seule différence entre le virus et le plasmide est bien la présence chez le virus d'au moins un gène codant pour une protéine impliquée dans la formation d'un virion. Cet exemple montre aussi combien il est facile de transformer un virus en plasmide, puisqu'il suffit pour cela que le virus perde par délétion un seul gène essentiel pour la formation du virion. Réciproquement, un plasmide peut participer à la formation d'un nouveau virus en capturant une cassette de gènes pour la production de virions. Ceci explique pourquoi les espaces de séquences virales et plasmidiques se recouvrent au moins partiellement (Cortez *et al.*, 2009).

Définir le virus par la production de virions a d'importantes conséquences pratiques et théoriques qui ne sont pas encore toujours appréciées. En particulier, cette définition permet d'éviter la confusion entre virus et éléments apparentés. Ainsi, si l'on prend cette définition au sérieux, on devra en conclure que certains

« virus » à ARN reconnus comme tel par l'ICTV, ne sont pas des virus ! C'est le cas des *Endornaviridae* et des *Hypoviridae*, qui ne codent pas pour une protéine de capsid et ne produisent pas de virions. Ces molécules d'ARN infectieuses ont été classées au sein des virus car leur ARN répliquent sont apparentées aux répliquases de « vrais » virus à ARN. Ceci crée un vrai problème. En toute logique, si l'on inclut ces ARN infectieux au sein de la famille virale pour cette raison, il faudrait également y inclure tous les plasmides des archées et des bactéries, puisque nombre d'entre eux sont apparentés à des virus par leurs protéines de réplication. Les *Endornaviridae* et les *Hypoviridae* ne devraient donc pas être considérés comme des virus mais comme des plasmides eucaryotes à ARN. Le fait de posséder une capsid, ou au minimum de coder une protéine qui protège son génome et facilite sa dissémination, semble donc bien correspondre à la marque de fabrique du virus. Par ailleurs, ce sont précisément ces protéines qui permettant aujourd'hui de classer les virus en grands groupes évolutifs.

Origine des virus : l'hypothèse de l'échappement ancien

Le fait de définir le virus comme organisme producteur de virions a une autre conséquence importante, plus théorique, sur le problème de leur origine. Si nous acceptons cette définition, les premiers virus n'ont pas pu apparaître avant l'émergence du ribosome, puisque tout virus doit coder au minimum une protéine. D'autre part, les virus n'ont pas pu apparaître avant les cellules, puisqu'eux mêmes sont des organismes cellulaires. Les premiers virus sont donc très probablement apparus dans un monde de cellules à ARN et protéines, comme un nouveau procédé d'amplification et de dissémination de l'information génétique. Nous avons vu par ailleurs que les virus ont précédé LUCA. Les premiers virus seraient donc apparus dans la période comprise entre l'invention du ribosome et l'époque de LUCA. Les premiers virus à ARN devaient infecter les cellules à ARN qui ont précédé LUCA. Si, comme je l'ai proposé, l'ADN lui-même a tout d'abord émergé dans le monde viral, les virus seraient donc apparus dans la période comprise entre l'émergence du ribosome et celle de l'ADN (figure 3).

Plusieurs mécanismes peuvent être imaginés pour l'apparition des premiers virions (figure 8). Par exemple, toutes les cellules (Archées, Bactéries ou Eucaryotes) produisent des vésicules membranaires capables de transporter de l'ARN ou de l'ADN (Deatherage & Cookson, 2010). Cette propriété, qui pourrait être liée à la nature même des membranes biologiques, existait donc très probablement déjà chez LUCA et chez les cellules à ARN qui l'ont

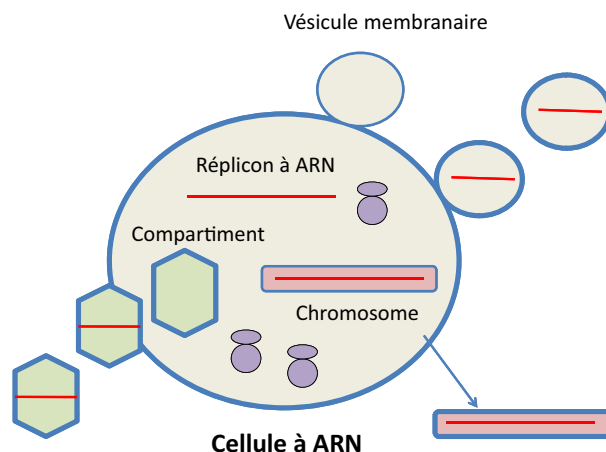


Fig. 8. Origine des virus. Dans ce schéma très simplifié, différents types de virions infectieux permettant le transfert de répliqueurs à ARN apparaissent indépendamment à partir de composants d'une cellule à ARN ancestrale : vésicule membranaire, structure chromosomique (pour les virions filamenteux) ou compartiment icosaédrique. Dans la réalité, ces processus ont dû apparaître indépendamment dans des cellules différentes donnant naissance à des lignées de virus non apparentées.

précédé. Certains virions ont pu se former à partir de telles vésicules, par incorporation de nucléoprotéines capables de protéger et de faciliter la dissémination de répliqueurs à ARN. D'autres particules virales ont pu se former à partir de structures chromatiniennes donnant naissance aux virions filamenteux. D'autres encore ont pu se former à partir de compartiments intracellulaires icosaédriques, similaire aux carboxysomes des bactéries modernes (utilisés pour le confinement de produits toxiques du métabolisme cellulaire). Mart Krupovic et moi-même avons récemment discuté cette nouvelle hypothèse sur l'origine des virus, que nous avons baptisée « théorie de l'échappement ancien » (Forterre & Krupovic, 2013) en référence aux théories classiques d'échappement dans lesquelles les virus sont issus de cellules modernes. Dans notre scénario, les premiers génomes viraux se sont formés à partir de répliqueurs à ARN présents dans les cellules du deuxième âge du monde à ARN, après l'invention du ribosome (pour une définition des deux âges du monde à ARN, voir Forterre, 2005).

Par opposition aux théories d'échappement, certains auteurs ont parfois avancé l'idée selon laquelle les virus avaient une origine cellulaire et auraient évolué par régression (Bandeia, 1983). J'ai moi-même caressé un temps cette idée en supposant que les virus pouvaient provenir d'anciennes lignées cellulaires aujourd'hui disparues (Forterre, 2005, 2006). Je me rends compte aujourd'hui que j'étais encore à l'époque influencé par la confusion entre virus et

virions (la cellule parasite ancêtre du virus se transformant en virion) lorsque j'envisageais une origine virale cellulaire. Si l'on pense aux petits virions formés d'une seule protéine, on voit bien que cette idée ne tient pas la route ! Plus récemment, Jean-Michel Claverie et Chantal Abergel ont repris l'idée d'une origine cellulaire pour les virus géants. Ces derniers codent pour quelques protéines impliquées dans la traduction, qui, dans cette hypothèse, correspondraient à des reliques d'une machinerie complète présente dans l'ancêtre cellulaire de ces virus (Claverie & Abergel, 2013). Dans le cas des Mégavirales, la théorie de l'origine cellulaire ne prend toutefois pas en compte la complexité (et la taille) de leurs virions. De tels virions n'ont pas pu apparaître d'un seul coup pour accommoder le génome de la cellule supposée être à l'origine de ces virus. Il est également difficile d'imaginer une cellule se transformant en virion, vu les différences de structure considérables entre ces deux types d'entités. D'autre part, dans ce scénario, l'origine cellulaire du virion qui devient le virus nous ramène à la confusion virus/virion.

En fait, comme nous l'avons vu, les Mégavirales font partie de la lignée PRD1/Adénovirus, et il est plus logique de penser qu'ils partagent avec les autres membres de cette famille un ancêtre viral qui produisait des virions de petite taille. Cela ne veut pas toutefois dire que les Mégavirales sont « récents ». Ces virus infectent des eucaryotes très divers, ce qui suggère qu'ils existaient déjà à l'époque du dernier ancêtre commun aux eucaryotes modernes, LECA (*The Last Eucaryotic Common Ancestor*). On peut penser que les Mégavirales ont augmenté de taille en même temps que leurs hôtes proto-eucaryotes, entre LUCA et LECA. En fait, la diversification des virus de la lignée PRD1/Adénovirus avait sans doute déjà commencé avant l'époque de LUCA.

Les virus dans l'arbre universel du vivant

La vision traditionnelle des virus a conduit à les considérer comme des sous-produits de l'évolution cellulaire. En conséquence, si les virus ont parfois été utilisés comme « boîte à outils » pour comprendre certains mécanismes évolutifs (voir l'article de Debarbieux, Saussereau et Maura dans ce numéro), ils ont été généralement ignorés par les évolutionnistes qui tentent de reconstruire l'histoire du vivant (Koonin & Wolf, 2012). Cette tendance a été renforcée à la fin du siècle dernier par l'utilisation de l'ARN ribosomique comme marqueur moléculaire « universel ». Les virus ne possédant pas ce marqueur, ils ont été exclus de fait de l'arbre « universel » construit par les évolutionnistes.

À la suite de la découverte du Mimivirus, Raoult et ses collaborateurs ont proposé de placer les

Mégavirales dans l'arbre universel, en tant que quatrième domaine, en se basant sur les phylogénies de protéines présentes à la fois chez ces virus géants et dans les trois domaines cellulaires (Raoult *et al.*, 2004, voir l'article de Philippe Colson dans ce numéro). Cette proposition s'est heurtée à une très vive opposition de la part des auteurs qui refusent de placer les virus sur un pied d'égalité avec les cellules (Moreira & Lopez-Garcia, 2009). Pour ma part, les domaines ayant été définis par la nature de leurs ribosomes, je pense effectivement préférable de limiter la notion de domaine aux organismes composés de *ribocells* et de classer les virus en grandes lignées basées sur les relations évolutives entre leurs virions. Limiter l'utilisation du terme « domaine » aux trois lignées cellulaires identifiées par Carl Woese et ses collaborateurs a l'avantage d'être cohérent avec la division du monde vivant en deux grands groupes d'organismes, les cellules à ribosomes d'un côté, les virus produisant des virions de l'autre (Raoult & Forterre, 2008 ; Koonin & Lwoff, 2012).

Un autre argument contre l'inclusion des virus dans l'arbre universel en tant que nouveau « domaine » se fonde sur la différence de nature entre les lignées évolutives cellulaires et virales : toutes les lignées évolutives cellulaires modernes descendent d'un même ancêtre commun, LUCA, formant un groupe monophylétique. Par contre, les grandes lignées évolutives virales proviennent d'ancêtres différents, correspondant chacun à des inventions indépendantes de différents types de virions (Moreira & Lopez-Garcia, 2009). Les virus sont donc polyphylétiques. De mon point de vue, ceci n'empêche pas de les introduire dans l'arbre universel du vivant, mais sous forme de trois ensembles de lignées qui coévoluent avec leurs hôtes à ribosomes : des lianes qui s'enroulent autour du tronc et des branches de l'arbre.

Conclusion

Comme nous l'avons vu, tout au long de leur histoire évolutive commune, l'interaction entre cellules et virus va être un moteur essentiel (peut-être le plus important) qui va conduire ces deux types d'organismes à se diversifier et donc à évoluer. On le comprend bien si l'on se réfère aux deux piliers du darwinisme : diversification et sélection (Forterre, 2012). Les virus imposent une énorme pression sur leurs hôtes. Ils modifient la structure des populations infectées, par exemple en conduisant à l'éradication des espèces majoritaires, ce qui favorise la diversité. En même temps, les virus et les éléments apparentés sont les principaux artisans de la diversification, par leur capacité à modifier les génomes (insertion, délétion, recombinaison, duplication) et par leur extraordinaire potentiel

d'innovation (formation de nouveaux gènes qui finissent, pour certains, par s'intégrer dans les génomes cellulaires). Dans ce contexte, le conflit entre virus et cellules, qui dure depuis au moins trois milliards d'années, est particulièrement propice à l'innovation, car il induit une course perpétuelle aux armements. La sélection des deux côtés de mécanismes d'attaque et de défense toujours plus efficaces, et de contre-feu à ces mécanismes, va être à l'origine de systèmes moléculaires toujours plus complexes. Les virus sont donc les « agents numéro un » à la fois de la variation et de la sélection, et l'une des causes majeures de la formation de structures de plus en plus complexes observée dans de nombreuses lignées évolutives. Ces considérations ne sont pas seulement importantes sur le plan théorique, elles invitent les biologistes à reconsidérer la place des virus dans leur propre recherche ; comme le dit Forest Rohwer, l'un des chercheurs les plus productifs dans le domaine de l'écologie virale : « *consider something viral in your search* » (Rohwer & Youle, 2011).

Le retour des virus sur le devant de la scène est donc un fait majeur de la biologie en ce début du siècle, et nous pouvons prévoir que ce nouvel intérêt pour les virus va aller crescendo. La découverte de nouveaux mécanismes de défense antivirale sophistiqués chez les archées et les bactéries (le système CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*)) a encore récemment mis en lumière l'importance du conflit cellules-virus. Nous découvrons le rôle essentiel des virus dans les grands cycles géochimiques, en particulier pour le recyclage du carbone. Le regain d'intérêt pour la thérapie phagique nous renvoie à notre peu de connaissance sur les interactions entre virus et bactéries chez l'animal, et chez l'Homme en particulier (voir l'article de Debarbieux, Saussereau et Maura dans ce numéro). Nous n'en sommes qu'au début d'une histoire qui s'annonce passionnante sur un aspect du vivant qui est resté jusqu'à présent trop en marge du courant dominant de la biologie fondamentale.

Références

- Abrescia N.G., Bamford D.H., Grimes J.M., Stuart D.I., Structure unifies the viral universe. *Annu Rev Biochem*, 2012, 81, 1–23.
- Bamford D.H., Do viruses form lineages across different domains of life? *Res Microbiol*, 2003, 54, 231–236.
- Bandea C.I., A new theory on the origin and the nature of viruses. *J Theor Biol*, 1984, 105, 591–602.
- Bell P.J., Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus? *J Mol Evol*, 2001, 53, 251–256.
- Bos L., Beijerinck's work on tobacco mosaic virus: historical context and legacy. *Philos Trans R Soc Lond B Biol Sci*, 1999, 354, 675–685.
- Brüssow H., The not so universal tree of life or the place of viruses in the living world. *Philos Trans R Soc Lond B Biol Sci*, 2009, 364, 2263–2274.
- Claverie J.M., Viruses take center stage in cellular evolution. *Genome Biol*, 2006, 7, 110.
- Claverie J.M., Abergel C., Open questions about giant viruses. *Adv Virus Res*, 2013, 85, 25–56.
- Colson P., de Lamballerie X., Fournous G., Raoult D., Reclassification of Giant Viruses Composing a Fourth Domain of Life in the New Order Megavirales. *Intervirology*, 2012, 55, 321–332.
- Cortez D., Forterre P., Gribaldo S., A hidden reservoir of integrative elements is the major source of recently acquired foreign genes and ORFans in archaeal and bacterial genomes. *Genome Biol*, 2009, 10, R65.
- Deatherage B.L., Cookson B.T., Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. *Infect Immun*, 2012, 80, 1948–1957.
- Desnues C., Raoult D., Virophages question the existence of satellites. *Nat Rev Microbiol*, 2012, 10, 234.
- de Parseval N., Heidmann T., Human endogenous retroviruses: from infectious elements to human genes. *Cytogenet Genome Res*, 2005, 110, 318–332.
- Filée J., Forterre P., Viral proteins functioning in organelles: a cryptic origin? *Trends Microbiol*, 2005, 13, 510–513.
- Forterre P., The origin of DNA genomes and DNA replication proteins. *Curr Opin Microbiol*, 2002, 5, 525–532.
- Forterre P., The two ages of the RNA world, and the transition to the DNA world, a story of viruses and cells. *Biochimie*, 2005, 87, 93–803.
- Forterre P., The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res*, 2006, 117, 5–16.
- Forterre P., Giant viruses: conflicts in revisiting the virus concept. *Intervirology*, 2010a, 53, 362–378.
- Forterre P., Manipulation of cellular syntheses and the nature of viruses: the virocell concept. *CR Chimie*, 2010b, doi:10.1016/j.crci.2010.06.007.
- Forterre P., Virocell Concept. In: *eLS*, 2012, John Wiley & Sons Ltd, Chichester.
- Forterre P., Darwin's goldmine is still open: variation and selection run the world. *Front Cell Infect Microbiol*, 2012, 2, 106.
- Forterre P., The virocell concept and environmental microbiology. *The ISME J*, 2013, 7, 233–236.
- Forterre P., Prangishvili D., The great billion-year war between ribosome- and capsid-encoding organisms (cells and viruses) as the major source of evolutionary novelties. *Ann N Y Acad Sci*, 2009, 1178, 65–77.
- Forterre P., Krupovic M., The origin of virions and virocells: the escape hypothesis revisited. In Witzany G. (Ed.) *Viruses: Essential Agents of Life*, 2013, Springer Dordrecht Heidelberg, New York, London, pp. 43–61.
- Guliyev M., Yilmaz S., Sahin K., Marakli S., Gozukirmizi N., Human endogenous retrovirus-H insertion screening. *Mol Med Rep*, 2013, 7, 1305–1309.

- Hamacher K., Greiner T., Ogata H., Van Etten J.L., Gebhardt M., Villarreal L.P., Cosentino C., Moroni A., Thiel G., Phycodnavirus potassium ion channel proteins question the virus molecular piracy hypothesis. *PLoS One*, 2012, 7, e38826.
- Hulo C., de Castro E., Masson P., Bougueleret L., Bairoch A., Xenarios I., Le Mercier P., ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Res*, 2011, 39, 576–582.
- Jalasvuori M., Vehicles, replicators, and intercellular movement of genetic information: evolutionary dissection of a bacterial cell. *Int J Evol Biol*, 2012, 2012, 874153.
- Koonin E.V., Wolf Y.I., Evolution of microbes and viruses: a paradigm shift in evolutionary biology? *Front Cell Infect Microbiol*, 2012, 2, 119.
- Kristensen D.M., Mushegian A.R., Dolja V.V., Koonin E.V., New dimensions of the virus world discovered through metagenomics. *Trends Microbiol*, 2010, 18, 11–19.
- Krupovic M., Bamford D.H., Order to the viral universe. *J Virol*, 2010, 84, 12476–1247.
- La Scola B., Desnue C., Pagnier P., Robert C., Barrassi B., Fournous G., Suzan-Monti M., Forterre P., Koonin E., Raoult D., The virophage, a unique parasite of the giant Mimivirus. *Nature*, 2008, 455, 100–104.
- López-García P., The Place of Viruses in Biology in Light of the Metabolism- versus-replication-first Debate. *His Phil Life Sci*, 2012, 34, 391–406.
- López-García P., Moreira D., Viruses in biology. *Evo Edu Outreach*, 2012, 5, 389–398.
- Lwoff A., The concept of virus. *J Gen Microbiol*, 1957, 17, 239–253.
- Lwoff A., Interaction among virus, cell, and organism. *Science*, 1966, 152, 1216–1220.
- Miller S., Krijnse-Locker J., Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol*, 2008, 6, 363–374.
- Mochizuki T., Krupovic M., Pehau-Arnaudet G., Sako Y., Forterre P., Prangishvili D., Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. *Proc Natl Acad Sci USA*, 2012, 109, 13386–13391.
- Moreira D., Brochier-Armanet C., Giant viruses, giant chimeras: The multiple evolutionary histories of Mimivirus genes. *BMC Evol Biol*, 2008, 8, 12.
- Moreira D., López-García P., Ten reasons to exclude viruses from the tree of life. *Nat Rev Microbiol*, 2009, 7, 306–311.
- Moulder J.W., Interaction of chlamydiae and host cells *in vitro*. *Microbiol Rev*, 1991, 55, 143–190.
- Pina M., Bize A., Forterre P., Prangishvili D., The Archeoviruses FEMS. *Microbiol Rev*, 2011, 35, 1035–1054.
- Prangishvili D., Forterre P., Garrett R.A., Viruses of the Archaea: a unifying view. *Nat Rev Microbiol*, 2006, 4, 837–848.
- Raoult D., Forterre P., Redefining viruses: lessons from Mimivirus. *Nat Rev Microbiol*, 2008, 6, 315–319.
- Raoult D., Audic S., Robert C., Abergel C., Renesto P., Ogata H., La Scola B., Suzan M., Claverie J.M., The 1.2 megabase genome sequence of Mimivirus. *Science*, 2004, 306, 1344–1350.
- Rohwer F., Thurber R.V., Viruses manipulate the marine environment. *Nature*, 2009, 459, 207–212.
- Rohwer F., Youle M., Consider something viral in your search. *Nat Rev Microbiol*, 2011, 9, 308–309.
- Ryan F., Virus et Homme, un destin commun, 2011, Éditions du Pommier.
- Sharon I., Alperovitch A., Rohwer F., Haynes M., Glaser F., Atamna-Ismael N., Pinter R.Y., Partensky F., Koonin E.V., Wolf Y.I., Nelson N., Béjà O., Photosystem I gene cassettes are present in marine virus genomes. *Nature*, 2009, 461, 258–262.
- Schmid M.F., Hecksel C.W., Rochat R.H., Bhella D., Chiu W., Rixon F.J., A tail-like assembly at the portal vertex in intact herpes simplex type-1 virions. *PLoS Pathog*, 2012, 8, e1002961.
- Slimani M., Pagnier I., Raoult D., La Scola B., Amoebae as battlefields for bacteria, giant viruses, and virophages. *J Virol*, 2013, 87, 4783–4785.
- Suttle C., Crystal ball, The virosphere: the greatest biological diversity on Earth and driver of global process. *Environ Microbiol*, 2005, 7, 481–482.
- Takemura M., Poxviruses and the origin of the eukaryotic nucleus. *J Mol Evol*, 2001, 52, 419–425.
- Temin H.M., The provirus hypothesis: speculations on the significance of RNA-directed DNA synthesis for normal development and for carcinogenesis. *J Natl Cancer Inst*, 1971, 46, 3–7.
- Thompson L.R., Zeng Q., Kelly L., Huang K.H., Singer A.U., Stubbe J., Chisholm S.W., Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci USA*, 2011, 108, 757–764.
- Van Regenmortel M.H., Mahy B.W., Emerging issues in virus taxonomy. *Emerg Infect Dis*, 2004, 10, 8–13.