

haut débit, soulignent aussi une biodiversité élevée en eau douce. Dans le cadre d'une collaboration avec Michael DuBow, de l'Université Paris-Sud à Orsay, nous étudions la diversité des virus « de type T7 », un groupe majoritaire parmi les virus de bactéries à ADN double brin. Nous avons détecté, dans les eaux superficielles du lac d'Annecy, plus de 6000 séquences génétiques différentes dans un même échantillon (une date, une profondeur) – soit autant de génotypes ou « espèces » virales potentielles.

Des virus à ARN, à un brin d'ADN ou mixtes

L'énorme quantité d'informations issues de la biologie cellulaire et moléculaire a permis, ces dix dernières années, une remise à plat de nos connaissances sur les virus aquatiques. Tout d'abord, ils arborent des structures moléculaires bien plus variées qu'on ne l'imaginait. On connaissait surtout les virus bactériophages – ou « phages » –, majoritaires à la surface de la planète, qui infectent les bactéries et les archées (micro-organismes qui, comme les bactéries, sont unicellulaires et sans noyau, et représentent un des trois règnes du vivant). De fait, aujourd'hui encore, plus de 90 pour cent des virus connus sont regroupés dans l'ordre des *Caudovirus*, c'est-à-dire des phages à ADN double brin. Cette configuration à ADN double brin est aussi celle de nombreux virus du phytoplancton que l'on a rassemblés en une famille, les Phycodnaviridés. Néanmoins, des virus à ARN et, plus récemment, d'autres à ADN simple brin, ont aussi été observés.

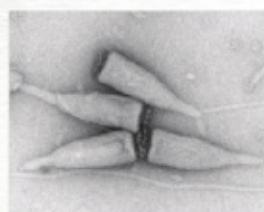
Les virus à ARN ont été découverts entre 2000 et 2006. Plusieurs études ont détecté de nombreux virus à ARN différents infectant le plancton, mais seuls quelques-uns ont été isolés et caractérisés, à cause de la lourdeur des techniques nécessaires. Plus récemment, la métagénomique, c'est-à-dire l'analyse du génome global de tous les micro-organismes présents dans un échantillon d'eau de mer, a mis en lumière l'existence de virus appartenant à la superfamille des Picornaviridés, des minuscules virus à ARN longtemps passés inaperçus.

En 2006, une étude publiée par Alexander Culley, de l'Université de Hawaï, a indiqué qu'aucun des virus à ARN recueillis dans diverses eaux côtières n'appartenait



PHAGE

FAMILLE : Myoviridés
HÔTE : bactéries et archées
TAILLE : 200 nm (queue comprise)
MILIEU : mers et lacs
GÉNOME : ADN double brin



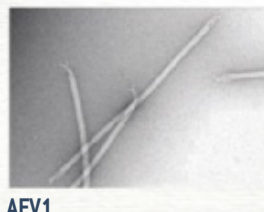
ABV

FAMILLE : Ampullaviridés
HÔTE : archées sulfobales
TAILLE : 200 nm
MILIEU : hypersalé, chaud, acide
GÉNOME : ADN double brin



ATV

FAMILLE : Bicaudaviridés
HÔTE : archées sulfobales
TAILLE : 100 nm (corps)
MILIEU : hypersalé, chaud, acide
GÉNOME : ADN double brin



AFV1

FAMILLE : Lipothrixiviridés
HÔTE : archées sulfobales et thermoprotéales
TAILLE : 200 nm
MILIEU : sources hydrothermales
GÉNOME : ADN double brin

à la longue liste des phages. Cette étude corroborait l'idée selon laquelle la plupart des virus à ARN ont pour hôtes non pas des procaryotes (bactéries ou archées), mais des eucaryotes, c'est-à-dire des organismes cellulaires à noyau bien différencié tels que les algues, les mollusques, les arthropodes, les poissons et les mammifères. De fait, la diversité virale observée ressemblait à celle de virus infectant des plantes ou des animaux terrestres et était déjà connue et répertoriée dans les bases de données génomiques.

La découverte des virus aquatiques à ADN simple brin est encore plus récente. Connue pour infecter les plantes et les animaux – surtout les oiseaux, le porc et les bovins –, cette catégorie virale joue un rôle économique important en détruisant les récoltes ou en infectant le bétail. En analysant par métagénomique la communauté virale de divers écosystèmes marins, des biologistes de l'Université de San Diego ont identifié, il y a six ans, de tels virus en milieu aquatique, qui ressemblaient aux phages de la famille des Microviridés.

Plus récemment, en 2009, d'autres virus à ADN simple brin de la famille des Circoviridés ont été trouvés dans les eaux douces d'un lac en Antarctique : ils dominent la communauté virale au printemps, lors de la fonte des glaces, quand les eucaryotes – leurs hôtes potentiels – sont les plus abondants. À la fin de l'été, ils sont d'ailleurs l'une des deux populations majoritaires, avec les bactéries, ce qui suggère qu'ils jouent un rôle primordial dans l'équilibre des écosystèmes en contrôlant certaines populations eucaryotes.

Enfin, certains virus géants associent, à l'instar des cellules de mammifères, de l'ADN et de l'ARN. C'est le cas de *Marseillevirus*, isolé il y a trois ans dans une amibe d'eau douce. Le génome de ce virus ressemble à une mosaïque de gènes issus des trois règnes du vivant – bactéries, eucaryotes et archées.

Comme ses cousins giris (*Mimi*-, *Mama*-, *Mega*-, *LausanneVirus*, *CroV*), *Marseillevirus* se distingue aussi par sa taille, bien supérieure à la « normale » : son génome est le sixième plus grand génome viral séquencé (368 000 paires de bases) et son diamètre est d'environ 250 nanomètres. Aujourd'hui le record est détenu par *Megavirus chilensis*, isolé des eaux côtières du Chili en 2010 : son génome compte près de 1260 millions de paires de bases et le diamètre de sa capsid – son enveloppe –

est de 440 nanomètres sans sa couronne de fibrilles (une bactérie compte entre 100 000 et plusieurs millions de paires de bases et mesure en général quelques micromètres, soit quelques milliers de nanomètres).

La plupart des virus aquatiques mesurent entre 30 et 60 nanomètres, bien que ceux présents en eau douce aient une taille moyenne légèrement supérieure (entre 50 et 90 nanomètres). Des exceptions existent dans les marigots du Danube, par exemple, où près de 84 pour cent des virus ont des capsides de diamètre supérieur à 60 nanomètres. *A contrario*, dans le lac Supérieur, aux États-Unis, plus de 50 pour cent des pathogènes viraux ont des têtes (la partie qui contient le matériel génétique) dont le diamètre est inférieur à 30 nanomètres.

Des virus géants en abondance

Toutefois, depuis une vingtaine d'années, des virus géants, dont les dimensions ont fait exploser les critères de taille, ont été retrouvés un peu partout en eau douce : dans certains réservoirs eutrophes (surabondants en matières nutritives), dans des lacs de l'Antarctique ou dans ceux des Alpes. En milieu marin, le virus de *Cafeteria roenbergensis* (CroV), parasite de l'espèce microzooplanctonique éponyme, a été découvert il y a deux ans. Avec 730 000 paires de bases codant 500 protéines, dont celles nécessaires à la réplication du virus, son génome

est le plus grand jamais décrit chez les virus marins après *Megavirus*. Le fait qu'il parasite une espèce d'un groupe important du plancton unicellulaire, les straménopiles, suggère que ce virus et ses cousins sont très répandus dans les océans.

Les virus géants présentent d'autres particularités. D'une part, contrairement aux virus « classiques », certains, tel *Minivirus*, ne détournent pas la machinerie cellulaire de l'hôte pour que celui-ci synthétise les protéines nécessaires à leur multiplication. Leur génome contient des gènes qui leur permettent d'assurer cette fonction (voir l'article de J.-M. Claverie et C. Abergel page 30). D'autre part, ce type de virus peut lui-même être infecté par un autre virus. Depuis trois ans en effet, une nouvelle espèce virale a été identifiée : le virus de virus. Le premier de ce genre, *Sputnik*, a été trouvé dans *Mamavirus*, dont il empêche le développement normal, mais dont le matériel génétique lui est nécessaire pour proliférer. De même qu'un bactériophage infecte une bactérie, ce pathogène et *Mavirus*, un virus du virus *CroV* découvert en 2011, sont les premiers membres d'une catégorie virale inédite : les virophages.

La présence de séquences génétiques associées au virus *Sputnik* dans des échantillons d'eau de mer, mise en évidence par l'équipe marseillaise de Didier Raoult, suggère que cette interaction virale originale n'est pas un cas isolé. Les biologistes se penchent à présent sur l'interaction du

LES AUTEURS

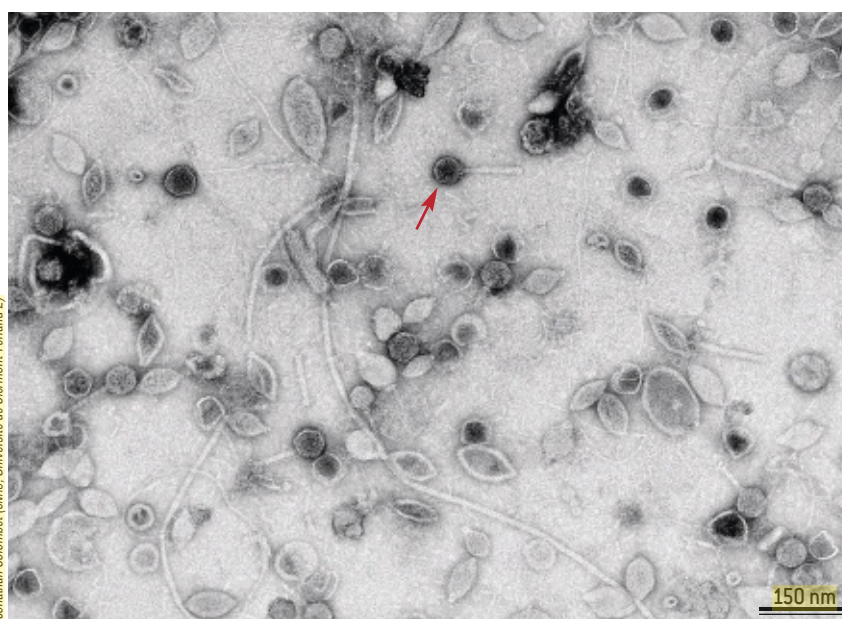


Stéphan JACQUET est directeur de recherche à l'Institut national de la recherche agronomique (INRA). Il travaille en écologie microbienne aquatique à Thonon-les-Bains.

Caroline DEPECKER est journaliste scientifique.

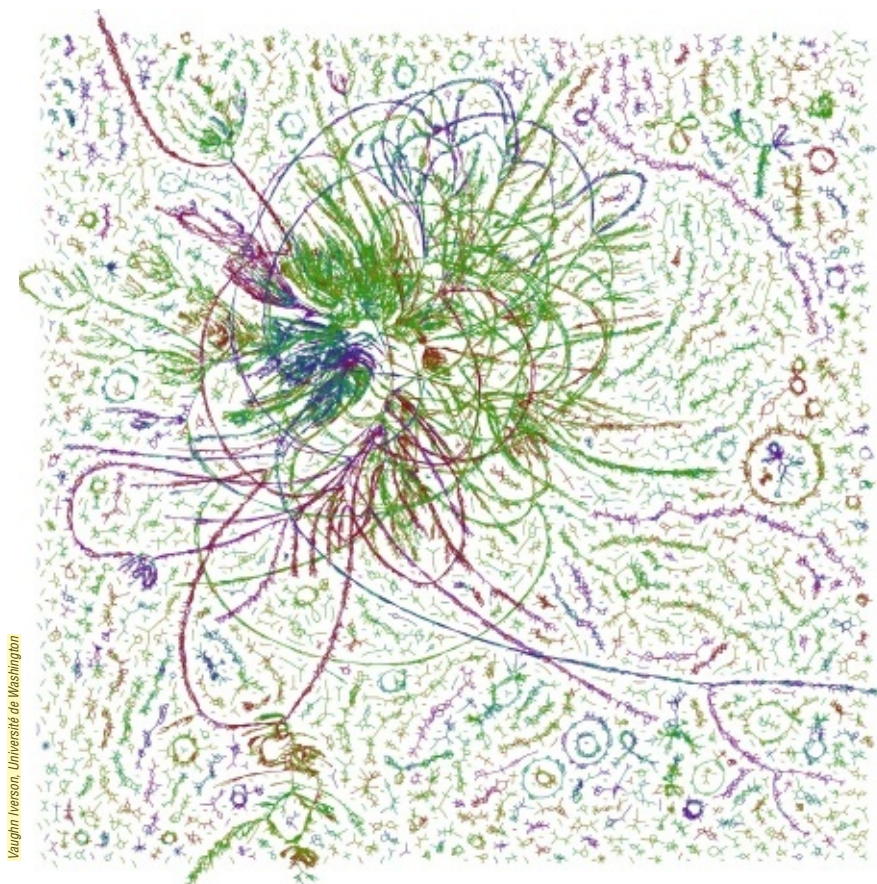


Stéphan Jacquet (INRA, Thonon)



Jonathan Colombeau (CNRS, Université de Clermont-Ferrand 2)

2. LE LAC ROSE, AU SÉNÉGAL (à gauche) et un échantillon viral du lac observé au microscope électronique à transmission (à droite). Les phages sont présents (flèche), mais peu nombreux, mêlés à de nombreuses autres formes virales.



Vaughn Iverson, Université de Washington

3. REPRÉSENTATION IMAGÉE DU MÉTAGÉNOME D'UN ÉCHANTILLON MARIN. Chaque chaîne représente une séquence ADN reconstituée par analyse génomique. Les longues chaînes représentent des séquences de procaryotes, et les petites circulaires, des séquences de virus ou de plasmides [des molécules d'ADN autres que l'ADN chromosomique, chez les bactéries].

plancton marin, des virus géants et de leur propres virus parasites.

Enfin, les virus se distinguent aussi par la diversité de leurs formes. Leur étude par microscopie électronique a montré, dans les systèmes aquatiques, la prédominance des phages à ADN double brin, constitués d'une tête icosaédrique et d'une queue de longueur variée (voir la figure 1). Néanmoins, dans les milieux hypersalés ou les sources géothermales, les virus d'archées prennent des formes inhabituelles. Bouteilles, ampoules, gouttes, épines, billes, filaments, à une ou deux queues... La quarantaine de virus isolés et décrits à ce jour – tous des virus à ADN double brin – ont été regroupés en quelques familles selon leur forme. Dans les eaux du lac Rose, au Sénégal, par exemple, un lac saturé en sel qui doit sa couleur à des bactéries photosynthétiques, nous avons dénombré plus de 100 millions de bactéries et d'archées par millilitre, et noté une teneur en virus dix fois supérieure à la valeur habituelle. Parmi ces virus, des chercheurs français (Institut Pasteur, Uni-

versité de Clermont-Ferrand, IRD) et américains ont observé que moins d'un pour cent disposent d'une tête et d'une queue, et de nouvelles formes ont été observées : filaments branchus, chaînettes, maïs, cannes, etc. (voir la figure 2).

On a constaté que les séquences génétiques virales analysées dans ce lac sont très proches de celles isolées dans des milieux hypersalés similaires, mais localisés sur d'autres continents. Pourquoi les virus de régions si éloignées présentent-ils de telles ressemblances génétiques ? Une diffusion des gènes sur de si longues distances est peu probable. Selon l'hypothèse la plus plausible, dans ce type de milieux, les virus se seraient adaptés de la même façon pour survivre.

De façon générale, la diversité des virus aquatiques suscite de nombreuses questions sur l'histoire évolutive des virus et celle de leurs hôtes. Quand et comment les virus sont-ils apparus ? La spécificité des virus pour leurs hôtes s'est-elle forgée avant ou après la séparation, sur l'arbre de l'évolution, des règnes du vivant ? On

a en effet constaté que généralement, chaque type de virus a un hôte privilégié et, inversement, que chaque procaryote est la cible d'un virus particulier. Notamment, ils n'infectent pas des organismes des autres règnes du vivant. L'étude phylogénique des virus aquatiques, de leur diversité et de ses liens avec la diversité du vivant apporte des pistes de réponses.

Virus et hôtes coévoluent

Tout d'abord, des similitudes (des protéines de la capside identiques, par exemple) ont été observées chez les virus infectant les trois règnes du vivant, ce qui suggère que les virus existaient avant la divergence de ces règnes. Au cours de l'émergence de ces trois domaines du vivant, des virus auraient été sélectionnés et auraient ensuite coévolué avec leurs hôtes. Trois populations virales auraient ainsi connu en parallèle des histoires différentes, associées à celles des trois domaines du vivant.

Plusieurs mécanismes permettent cette coévolution des virus et de leurs hôtes et, de façon plus générale, lient étroitement diversité virale et diversité du vivant. Nous citerons les deux principaux : la pression de sélection qu'exercent les virus sur les populations de procaryotes (bactéries et archées), et le transfert de gènes.

Un milieu donné comporte divers procaryotes dont le nombre fluctue en fonction de la pression de sélection exercée par les virus et ce, de diverses façons. En « tuant le meilleur », c'est-à-dire en parasitant l'espèce la plus compétitive (qui est souvent la plus abondante) dans un écosystème donné, les virus laissent vacante une niche écologique où les espèces plus rares – moins compétitives pour les ressources nutritives, notamment – peuvent alors se développer. Celles-ci profitent encore des éléments nutritifs relargués lors de la destruction cellulaire des espèces dominantes – la lyse virale. Enfin, la lyse libère des enzymes qui perturbent les cellules, détruisant les espèces les plus sensibles. La conjonction de ces effets accroît la diversité cellulaire et, par là même, celle des fonctions codées dans les génomes du milieu.

À Thonon-les-Bains, nous avons illustré ces différents processus lors d'expériences consistant à augmenter ou réduire le nombre de virus bactériophages ou cyanophages au sein d'échantillons d'eau douce mis en bouteille et issus de lacs péri-alpins.

Ces variations ont eu un impact mesurable sur le nombre de populations de micro-organismes, planctoniques notamment. Elles ont aussi modifié la quantité de matière organique libérée sous forme de carbone ou de phosphore, ou encore la structure de la communauté bactérienne et la proportion de bactéries dites lysogéniques. Ces bactéries portent un virus « dormant », c'est-à-dire qui ne se réplique pas, mais qui peut se « réveiller » à tout moment en fonction de son environnement.

Un autre mécanisme important par lequel la biodiversité virale influence le vivant est le transfert de gènes entre espèces. Un virus dormant dans un procaryote ou un eucaryote peut intégrer son matériel génétique dans celui de son hôte. Cette intégration confère au receveur de nouvelles propriétés, en général une plus grande résistance et une protection contre l'infection par d'autres virus, mais aussi des propriétés métaboliques, morphologiques, immunogénétiques, voire pathogènes inédites. C'est ainsi que la bactérie responsable du choléra, *Vibrio cholera*, a

acquis sa toxicité grâce à un gène d'origine virale, tout comme *Corynebacterium diphtheria* (diphtérie) et *Clostridium botulinum* (botulisme). De même, la syncytine, une protéine d'origine virale, semble indispensable à la formation du placenta des mammifères. Le transfert de gène a ici eu un effet positif.

Des gènes transférés du virus à l'hôte

En milieu aquatique, le transfert de gènes entre les cyanobactéries et leurs virus, les cyanophages, est particulièrement intéressant. Au fil de leur évolution, les cyanophages ont assimilé dans leur génome de nombreux gènes – probablement récupérés chez leurs hôtes – qui codent des composants clés de l'appareil photosynthétique bactérien. Exprimés pendant l'infection, ces gènes assurent que la photosynthèse et d'autres fonctions de la cyanobactérie perdurent pendant la réplication des virus, alors que les protéines bactériennes dont c'était le rôle en temps normal sont devenues

✓ BIBLIOGRAPHIE

S. Jacquet *et al.*, Viruses in aquatic ecosystems : Important advancements of the last 20 years and prospects for the future in microbial oceanography and limnology, *Advances in Oceanography and Limnology*, vol. 1, pp. 71-101, 2010.

A. Bize *et al.*, Les archéovirus, *Virologie*, vol. 14, n° 2, pp. 101-117, 2010.

M. G. Fischer *et al.*, Giant virus with a remarkable complement of genes infects marine zooplankton, *PNAS*, vol. 107, pp. 19508-19513, 2010.

Les virus, ennemis utiles, *Dossier Pour la Science*, n° 55, avril 2007.

F. E. Angly *et al.*, The marine viromes of four oceanic regions, *PLoS Biology*, vol. 4, n° 11, e368, doi:10.1371/journal.pbio.004036, 2006.

LES OUTILS D'ANALYSE DE LA DIVERSITÉ VIRALE

La diversité génétique est difficile à évaluer pour deux raisons. D'une part, les virus se cultivent mal en laboratoire et, d'autre part, ils n'ont pas de gène universel conservé, ce qui ne facilite pas leur classification. Présent chez tous les procaryotes, le gène codant l'ARN ribosomique 16S a par exemple été très utile pour établir les liens de parenté entre diverses souches bactériennes, par l'examen des variations, d'une souche à l'autre, de la séquence du gène.

Les méthodes classiques dites d'empreintes génétiques comme la PFGE (électrophorèse sur gel en champ pulsé) ou la PCR-DGGE (réaction de polymérisation en chaîne - électrophorèse sur gel en gradient dénaturant) permettent d'analyser la structure de populations ou de communautés à partir de morceaux de génome ou du génome entier. Elles révèlent la diversité virale, mais présentent des limites : les populations doivent être concentrées et donc dominantes dans le milieu pour que leur empreinte apparaisse sur le gel (sous

la forme de bandes sombres). Par exemple, à partir d'un échantillon de 20 litres d'eau, on n'observe pas plus de 10 à 30 empreintes de populations par la technique de la PFGE (une population correspond à une bande du gel), chacune équivalant à 0,1 à 5 pour cent de la communauté détectée. En outre, la PFGE ne distingue pas les génotypes de même taille et ne détecte que les virus à ADN double brin (les virus à ARN traversent ses grilles d'analyse).

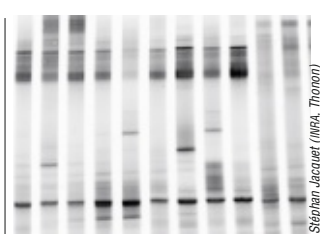
La métagénomique est une approche récente, puissante, qui permet d'analyser les génomes de tous les micro-organismes (virus, archées, bactéries, eucaryotes) présents dans un même échantillon sans qu'il soit nécessaire de les séparer – ni de les cultiver en laboratoire. On prélève plusieurs dizaines de litres d'eau dans l'environnement, que l'on filtre pour ne garder que les entités virales (de tailles inférieures à 450 nanomètres ou 200 nanomètres). L'échantillon est ensuite concentré, puis l'ADN ou l'ARN total en est extrait,

purifié et séquencé par des méthodes à haut débit.

Après reconstitution des séquences obtenues, l'analyse informatique fournit des informations sur la diversité virale et aide à identifier des fonctions importantes. Comment repère-t-on les génomes viraux parmi toute l'information génétique de l'échantillon ? À l'aide d'indices de régularité dans les séquences et des similarités de séquences. Ces éléments suffisent souvent à identifier la famille du virus si celle-ci est présente dans les bases de données existantes.

Au moins 50 métagénomes viraux d'eaux de mer ont été séquencés et ont ainsi révélé, par séquençage massif, l'étendue de la diversité virale. Pour autant, la métagénomique est une méthode lourde à utiliser, notamment pour déterminer comment varie cette diversité dans l'espace et dans le temps, et encore trop onéreuse pour analyser un grand nombre d'échantillons.

Un écueil majeur des analyses métagénomiques virales est



Stéphane Jacquet (INRA, Thonon)

Un gel DGGE montrant l'évolution des populations virales cyanophages [bandes noires] du lac du Bourget sur 12 mois (colonnes).

l'absence de résultats quantitatifs : les techniques utilisées pour amplifier l'ADN viral extrait du prélèvement (notamment l'amplification aléatoire de génomes entiers) engendrent des biais quantitatifs importants, ce qui limite l'analyse des abondances relatives des différents virus. En outre, il est impossible d'étudier dans un même métagénome les virus à ADN et ceux à ARN, ce qui ne facilite pas leur étude comparative.

Aujourd'hui, aucun consensus n'existe quant aux outils à utiliser pour étudier la diversité virale. Leur choix dépend des moyens disponibles et de ce que l'on recherche.