UNIVERSITE D'AIX-MARSEILLE

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

Unité de Recherche sur les Maladies Infectieuses Tropicales Emergentes (URMITE)

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline : Pathologie Humaine

Spécialité : Maladies Infectieuses

Sarah TEMMAM

Caractérisation des communautés virales de vecteurs & réservoirs de zoonoses :

Exemples des culicoïdes et de la viande de brousse.

Soutenue le 18/01/2016 devant le jury :

Mr le Professeur Philippe COLSON	Président du jury
Mr le Professeur Patrice MORAND	Rapporteur
Mr le Professeur Etienne THIRY	Rapporteur
Mr le Docteur Serafín GUTIERREZ	Examinateur
Mme le Docteur Christelle DESNUES	Directrice de thèse
Mr le Docteur Oleg MEDIANNIKOV	Co-directeur de thèse

Remerciements

Je tiens tout d'abord à remercier le Pr. Didier Raoult de m'avoir permis de réaliser ce travail dans son laboratoire et de m'avoir financée durant la dernière année de thèse.

Au Dr. Christelle Desnues (« ma chef », qui n'aime pas que je l'appelle comme ça !!!) je tiens à exprimer toute ma gratitude pour sa patience, son accueil, sa disponibilité, ses conseils et sa confiance. Grâce à toi, j'ai (enfin !) pu mener à bien la thèse que je voulais.

Je voudrais également remercier les Pr. Patrice Morand et Etienne Thiry pour avoir accepté d'être rapporteurs de cette thèse et pour leurs commentaires et corrections, ainsi que le Dr. Serafín Guttierez et les Pr. Philippe Colson et Oleg Mediannikov d'avoir accepté d'être examinateurs de ce travail.

Un grand merci à mon équipe : Maxime, Nicolás, Priscilla, Raja, Sébastien et Sonia mais aussi aux anciens : Dao, Emanuel, Ikram, Laura, Laure, Nikolay et les autres, pour leur aide, leur soutien et la bonne ambiance qu'ils mettaient au quotidien. Merci aussi à l'ensemble du personnel de l'URMITE : ingénieurs, techniciens, secrétaires, étudiants, que je ne peux pas tous citer mais sans qui tout aurait été plus compliqué !

Merci également au Dr. Bernard Davoust qui a réussi à me procurer des échantillons précieux et à Jean-Michel Bérenger (entomologiste) sans qui une partie de ce travail aurait été impossible. Merci enfin au Dr. Oleg Mediannikov, à Masse Sambou, Maxence Aubadie-Ladrix et à toute l'équipe de l'URMITE au Sénégal qui m'ont accueillie chaleureusement et m'ont fait découvrir ce beau pays.

A Nadia, Erwan, Laurence, Mag et les autres, j'aimerai dire merci : pour votre accueil, votre gentillesse et votre amitié. Ces courts séjours avec vous m'ont permis de décompresser et de prendre du recul. Vous me manquez ! Hervé : merci pour… tout ! Notre collaboration a fait de moi la virologue que je suis aujourd'hui.

Enfin et surtout, je tiens à exprimer toute ma gratitude et mon affection à ma famille et mes amis, d'ici et d'ailleurs… même si je ne le leur dis pas assez souvent !

Préambule

Le format de présentation de cette thèse correspond à une recommandation de la spécialité Maladies Infectieuses et Microbiologie, à l'intérieur du Master de Sciences de la Vie et de la Santé qui dépend de l'Ecole Doctorale des Sciences de la Vie de Marseille. Le candidat est amené à respecter des règles qui lui sont imposées et qui comportent un format de thèse utilisé dans le Nord de l'Europe permettant un meilleur rangement que les thèses traditionnelles.

Par ailleurs, la partie introduction et bibliographie est remplacée par une revue envoyée dans un journal afin de permettre une évaluation extérieure de la qualité de la revue et de permettre à l'étudiant de commencer le plus tôt possible une bibliographie exhaustive dans le domaine de cette thèse.

La thèse est présentée sur article publié, accepté ou soumis, associé à un bref commentaire donnant le sens général du travail. Cette forme de présentation a paru plus en adéquation avec les exigences de la compétition internationale et permet de se concentrer sur des travaux qui bénéficieront d'une diffusion internationale.

Professeur Didier RAOULT

Table des matières

Résumé

Les zoonoses constituent plus des deux tiers des pathologies virales qui concernent l'homme. Dans un monde en perpétuel changement, où les contacts entre l'homme et la nature sont de plus en plus fréquents, les cas d'émergence et de ré-émergence de virus zoonotiques sont rapportés de manière croissante en population humaine. Le développement et la démocratisation des outils de séquençage à haut débit et leur application en métagénomique en font de bons outils d'inventaire et de surveillance de virus potentiellement émergents. Dans cette optique, j'ai réalisé un inventaire des communautés virales de viande de singe illégalement importée en France et de culicoïdes, vecteurs arthropodes reconnus de nombreux virus d'intérêt médical et vétérinaire.

Bien que la majorité des virus zoonotiques possèdent un génome à ARN, il existe peu de méthodes standardisées d'isolement de tels viromes. J'ai donc tout d'abord développé et validé, au moyen d'un virome artificiel de poux, un protocole expérimental de purification des viromes à ARN qui permettait le maintien de l'infectivité des particules virales, ceci dans le but de pouvoir par la suite les manipuler après leur isolement en culture cellulaire. Ce protocole a ensuite été appliqué pour caractériser les communautés virales d'arthropodes hématophages et de prélèvements de faune sauvage.

L'une des sources les plus importantes de virus zoonotiques potentiellement transmissibles à l'homme étant les réservoirs de faune sauvage, j'ai, par la suite, réalisé l'inventaire des communautés virales à génome à ADN et à ARN de viande de brousse fumée, illégalement importée en France et confisquée par les douanes. Cette étude a révélé la présence de nombreux bactériophages, dont certains pourraient infecter des bactéries potentiellement pathogènes pour l'homme, mais n'a pas permis de montrer la présence de virus zoonotiques infectieux. Le traitement préalable de fumage de la viande ou son transport dans des conditions non conventionnelles pouvant expliquer ce résultat, des études similaires sur de la viande fraîche sont en cours afin d'évaluer le risque zoonotique d'infection par la

manipulation, le trafic et/ou la consommation de viande de gibier originaire de la faune sauvage.

Enfin j'ai caractérisé les communautés virales à génome ADN et à ARN de culicoïdes collectés au Sénégal, ce qui a permis de mettre en évidence la présence de nombreux virus géants à ADN infectant les amibes, et notamment d'un nouveau virus phylogénétiquement apparenté au *Faustovirus* et retrouvé à la fois chez les arthropodes et leurs hôtes vertébrés associés. Le séquençage des viromes à ARN a, quant à lui, révélé la présence, outre de virus d'insectes et de plantes, des arbovirus du genre *Orbivirus*, classiquement transmis par des culicoïdes et responsables de pathologies vétérinaires, mais aussi de nouveaux virus du genre *Thogotovirus* et de la famille des *Rhabdoviridae* qui pourraient constituer un risque d'émergence pour la santé humaine. Des investigations complémentaires, notamment au niveau de l'interface faune sauvage/ faune domestique qui sont en contact avec ces arthropodes et l'homme, permettront de mieux comprendre le cycle environnemental de ces virus et d'évaluer leur risque infectieux pour l'homme.

Du fait de nombreux facteurs intrinsèques et extérieurs à l'agent infectieux, la prédiction des futures émergences de virus zoonotiques est très compliquée voire utopique, mais elle reste un challenge crucial et d'actualité. La stratégie de réalisation d'inventaires des communautés virales présentes dans les différents acteurs des cycles de transmission zoonotique est un premier pas indispensable dans la connaissance des risques potentiels d'émergence en population humaine.

Abstract

Zoonoses are responsible of more than two thirds of human viral infections. In addition, with increasing contacts between humans and the wildlife and the domestic fauna, the emergence and reemergence of zoonotic viruses is accelerating. The development and democratization of high-throughput sequencing tools and their application in metagenomics allow inventorying the viral communities of various reservoirs in order to detect the emergence of viruses before their infection to humans. In this context, I have characterized the viral communities of simian bushmeat illegally imported into France and of *Culicoides* biting midges, recognized vectors of several viruses of human and veterinary medicine importance.

Since the majority of zoonotic viruses are constituted of a RNA genome and since no standardized protocol is available to isolate and analyze RNA viromes, I have first developed and validated a protocol for the purification of RNA viromes which allowed maintaining the infectivity of viral particles and their future isolation. This protocol was subsequently applied to characterize viral communities of bloodsucking arthropods and wildlife samples.

In a second part I realized the inventory of the DNA and RNA viral communities of smoked simian bushmeat illegally imported into France and confiscated by the French customs. Although this study did not report the detection of zoonotic viruses potentially pathogenic for humans, it revealed the presence of a wide diversity of bacteriophages, in which some of them could infect bacteria potentially pathogenic for humans. Since the treatment of the meat could explain this result, similar studies conducted on fresh meat are under process to characterize the potential risk of zoonotic infection by the manipulation, the trade and/or the consumption of bushmeat originating from wildlife.

Finally I have characterized the viral communities of *Culicoides* biting midges collected in Senegal. The analysis of the DNA virome highlighted the presence of sequences related to several giant DNA viruses infecting amoeba, including a new virus related to Faustovirus and detected both in arthropods and their vertebrate hosts. Sequencing of the RNA virome revealed the

presence, in addition to insect-specific and plant-infecting viruses, arboviruses belonging to the genus *Orbivirus* usually transmitted by these arthropod vectors and responsible for several veterinary pathologies; and also the identification of new viruses belonging to the *Thogotovirus* genus and *Rhabdoviridae* family that could constitute a risk of emergence of zoonoses for humans. Further investigations, particularly in wildlife and domestic animals in close contact with these arthropods and humans, will help to better understand the environmental cycle of these viruses and assess the risk of infection for humans.

The prediction of future emerging zoonotic viruses is very difficult, if not impossible. However the characterization of viral communities present in the different actors of zoonotic transmission cycle is a first step to evaluate potential risks of transmission to humans.

Liste des figures / tables

Liste des abréviations

ADN : acide désoxyribonucléique
ARN : acide ribonucléique
DOP-PCR : PCR utilisant des oligonucléotides dégénérés (degenerate oligonucleotide primed PCR)
ECP : effet cytopathique
ELISA : enzyme-linked immunosorbent assay
MERS-CoV : Middle East respiratory syndrome - coronavirus
MIE : maladie infectieuse émergente
NCLDV : grands virus à ADN à réplication nucléo-cytoplasmique (NucleoCytoplasmic Large DNA viruses)
NGS : séquençage à haut débit (next-generation sequencing)
PCR : réaction en chaîne de la polymérase
RCA : amplification en cercle roulant (rolling circle amplification)
RVFV : virus de la fièvre de la vallée du Rift
SARS : syndrome aigu respiratoire sévère
SIA : amplification séquence-indépendante
SIDA : syndrome de l'immunodéficience acquise
SISPA : amplification séquence-indépendante utilisant une amorce unique (sequence-independent single primer amplification)
VIH : virus de l'immunodéficience humaine

Contexte de l’étude

Au cours de ce travail de thèse j’ai été amenée à étudier les communautés virales de certains arthropodes hématophages et de prélèvements de viande de singe afin de réaliser un inventaire, aussi exhaustif que possible, des virus potentiellement émergents et pouvant présenter un risque pour la santé humaine. Dans le cadre des recommandations de la spécialité "Pathologie Humaine", l'introduction bibliographique a été valorisée sous la forme d'une revue scientifique publiée qui est incluse à la fin de ce chapitre. Afin d'éviter les redondances avec cette revue qui porte sur la métagénomique comme outil de détection des zoonoses avant leur passage en population humaine, j'ai complété l'introduction de cette thèse en insistant tout particulièrement sur le contexte des maladies infectieuses émergentes, puis en présentant brièvement les virus zoonotiques d’origine arthropode ou animale, et les apports de la métagénomique virale dans ce type d’étude.

I. Les maladies infectieuses émergentes : définitions, exemples et facteurs d’émergence

1. Le concept d’émergence

En 1908, le Pr. Charles Nicolle, prix Nobel de médecine en 1928, déclarait « Il y aura donc des maladies nouvelles. C’est un fait. Un autre fait est que nous ne saurons jamais les dépister dès leur origine. Lorsque nous aurons connaissance de ces maladies, elles seront déjà toutes formées » et « Depuis toujours, l’homme a vécu avec des maladies infectieuses, de nouvelles pathologies apparaissent, d’autres disparaissent, dans une permanente dynamique évolutive ». Au cours du siècle dernier, de nombreux agents infectieux ont été découverts et incriminés dans des pathologies humaines. Cependant, vers la fin des années 1970, du fait des nombreux progrès de la médecine (vaccins, antibiotiques, antiviraux…) et de l’amélioration des conditions d’hygiène, le concept de « fin des maladies infectieuses » est apparu. Cependant, au cours de la décennie suivante, avec le « retour » de certaines maladies infectieuses qu’on pensait disparues, est apparu le concept

de maladies émergentes et/ou ré-émergentes. Ces maladies sont, pour la plupart, d'origine infectieuse, mais pas seulement, comme par exemple pour le cas du diabète en augmentation, suite à la sédentarisation des personnes et à l'augmentation des cas d'obésité.

Il existe de nombreuses définitions des maladies infectieuses émergentes (MIE). Morse en proposa une comme étant « des infections nouvellement apparues dans une population donnée ou déjà existantes mais dont l'incidence ou la répartition géographique a rapidement augmenté » [1,2]. Les MIE sont donc des maladies dues à des agents infectieux (viraux, bactériens, fongiques, parasitaires) jusqu'alors inconnus ou connus mais dont les caractéristiques (ex : hôte cible, distribution géographique, virulence, etc.) changent, résultant en une augmentation significative de l'incidence de ces maladies dans une population donnée à un temps donné par rapport à la situation épidémiologique basale. On distingue alors deux grands types de MIE : les maladies purement émergentes, correspondant souvent à des zoonoses (c'est-à-dire des maladies animales transmises à l'homme) et les maladies ré-émergentes, correspondant à des maladies s'étant déjà manifestées et ayant plus ou moins disparu, puis se manifestant de nouveau suite à des changements environnementaux ou intrinsèques à l'agent infectieux. En 2013, Chua et Gubler ont proposé 4 sous-types de MIE [3], à savoir :

- des agents infectieux connus colonisant de nouvelles niches écologiques (nouvelles répartitions géographiques, nouveaux hôtes)
- des agents infectieux connus mais avec des caractéristiques les distinguant des variants connus (ex : acquisition de résistance aux antibiotiques, nouvelles combinaisons de segments pour le virus *Influenza*)
- l'émergence d'un agent infectieux jusqu'alors inconnu pour l'homme (le plus souvent d'origine animale)
- une maladie préalablement décrite pour laquelle l'agent étiologique n'était jusqu'alors pas identifié.

La figure 1 ci-dessous présente quelques exemples d'émergence ou de ré-émergence de maladies infectieuses de 1977 à 2007 rapportés par Morens [4], les virus représentant ainsi la majorité d'entre elles [5,6]. Ainsi, Woolhouse rapporta que sur les 1399 espèces microbiennes capables d'infecter l'homme, les virus ne représentent que 13.5% mais constituent plus des deux tiers des

pathogènes émergents, et parmi eux, les virus à génome ARN représentent 84.5% des virus émergents [7].

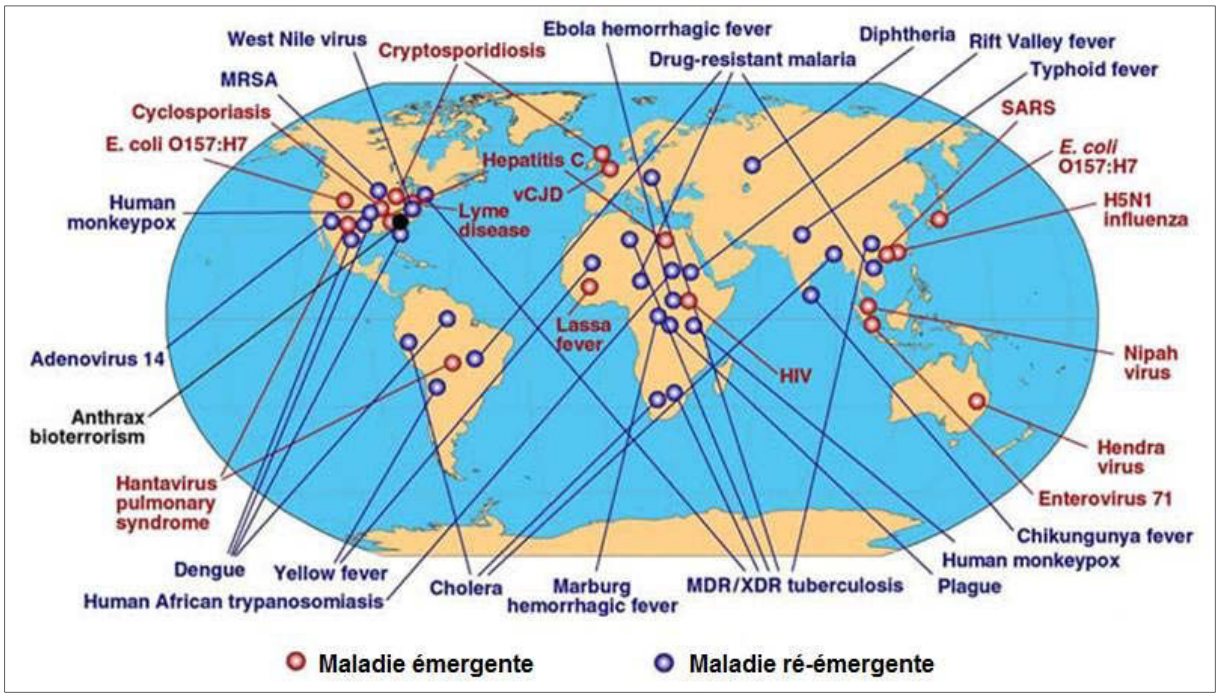


Figure 1. Exemples de maladies infectieuses émergentes ou ré-émergentes de 1977 à 2007, d'après [4,8].

2. Les facteurs d'émergence de nouvelles maladies

En 2008, Morens proposa 13 facteurs qui sous-tendaient à l'émergence d'une maladie infectieuse [4], ces facteurs pouvant être classés en trois grandes catégories : ceux liés à l'homme, à l'agent infectieux ou à l'environnement (Figure 2).

Parmi les facteurs liés aux activités humaines, on retrouve par exemple le trafic international de personnes et de denrées (alimentaires et non-alimentaires) qui peut favoriser l'introduction d'agents infectieux, de vecteurs, de réservoirs animaux, etc. dans des régions jusqu'alors indemnes ou des régions où l'agent infectieux avait été préalablement éradiqué [9-11]. La croissance de la population humaine a pour conséquence une augmentation de la promiscuité entre individus, accroissant ainsi les risques de transmissions inter-humaines, une augmentation de la pauvreté et la dégradation des conditions d'hygiène [12].

L'environnement dans lequel l'homme évolue et sur lequel il agit a aussi un impact sur l'émergence de nouvelles maladies, comme par exemple le

réchauffement climatique qui influence les aires de répartition géographiques de certains insectes vecteurs, ou encore la déforestation qui rend le contact homme/animal plus facile et augmente ainsi les possibilités d'échanges d'agents infectieux [13-14]. Deux hypothèses ont été proposées quant à la manière dont l'environnement influence l'émergence de maladies infectieuses : une hypothèse dite de « perturbation » (correspondant à une modification de l'écosystème de certaines espèces animales ayant pour conséquence une perturbation du cycle habituel de transmission d'un agent infectieux entre hôtes multiples et l'infection d'hôtes accidentels, tels que l'homme) et une hypothèse appelée « pathogen pool » (correspondant à l'exposition de nouveaux hôtes à un pool d'agents infectieux circulant dans des espèces auparavant jamais ou rarement en contact avec l'homme) [15].

Enfin, parmi les facteurs liés à l'agent infectieux, les principales causes d'émergence représentent des modifications génétiques résultant en une adaptation du micro-organisme en réponse ou non à une pression extérieure. L'usage intensif d'antibiotiques serait, par exemple, à l'origine de l'émergence ou de la sélection de souches bactériennes multi-résistantes, comme dans le cas de la tuberculose [16]. Les virus à génome ARN ont, quant à eux, une polymérase ayant un taux d'erreur très élevé par rapport aux virus à génome ADN [17], ce qui génère de nombreuses mutations dans les génomes et l'apparition de nouveaux variants (ou quasi-espèces) potentiellement capables de s'adapter à de nouvelles conditions (ex : nouveaux hôtes, résistance à une pression de sélection) [18-19]. Des évènements de recombinaison ou de réassortiment peuvent également avoir lieu, résultant en l'apparition de nouveaux virus, comme dans le cas du virus H1N1 pandémique en 2009, issu d'un triple réassortiment entre un virus porcin, humain et aviaire et qui est apparu chez le porc avant d'infecter l'homme [20].

Woolhouse [21] a proposé de classer ces facteurs par ordre d'importance (en fonction du nombre de pathogènes ayant émergé à cause de ces facteurs), tout en sachant qu'en réalité l'émergence est davantage due à une combinaison de ces facteurs :

1. changements de l'usage des sols et déforestation
2. changements démographiques, sociétaux et comportementaux
3. précarité des conditions sanitaires
4. liés à l'hôpital (maladies nosocomiales)
5. évolution intrinsèque des agents pathogènes
6. contamination des aliments et de l'eau, surtout dans les pays en développement
7. voyages et migrations de populations
8. défaut des systèmes de santé et de surveillance
9. échanges commerciaux de denrées et d'animaux
10. changements climatiques

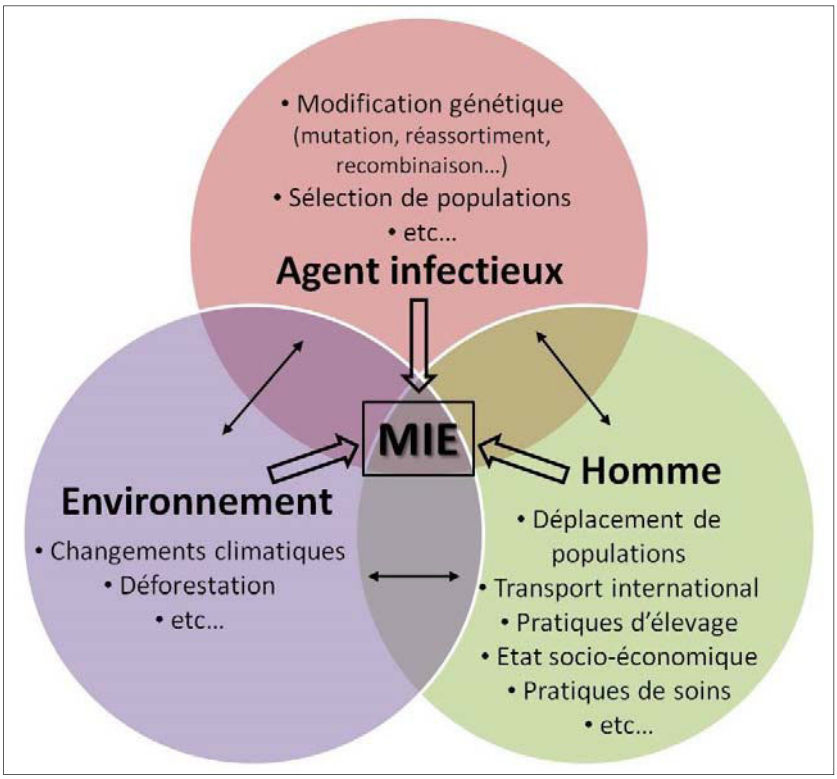


Figure 2. Exemples de facteurs d'émergence de maladies infectieuses.

Par le passé, de nombreuses maladies infectieuses ont émergé, liées aux activités humaines telles que la sédentarisation et le développement de l'agriculture (-10 000 avant J.C.) qui ont mis en contact l'homme avec des pathogènes animaux tels que la tuberculose. A partir du XIVe siècle, les voyages et conquêtes des européens, comme par exemple la découverte de l'Amérique, ont conduit à l'importation des virus de la grippe, de la rougeole ou de la variole dans les populations indiennes naïves pour ces virus ou encore la peste, transportée par les rats sur les navires [22-24]. Mais depuis le début du XXe

siècle, les MIE sont en constante augmentation du fait d'une accentuation de certains facteurs d'émergence [7] (Figure 3). McMichael recensa ainsi 4 grandes « ruptures épidémiques » liées à différents facteurs ayant contribué à l'émergence d'agents infectieux [25]. Ainsi, de nos jours, l'accroissement important de la démographie humaine accompagnant des changements environnementaux sans précédent, les contacts entre l'homme et l'animal de plus en plus fréquents et le renforcement des échanges commerciaux de denrées, des voyages et des migrations de populations humaines peut expliquer la progression des émergences d'agents infectieux. Par exemple, la chasse et la manipulation de viande de brousse issue de la faune sauvage à des fins de consommation (primates, chauve-souris…) pourraient être à l'origine des premiers cas de VIH (virus qui se serait, par la suite, parfaitement adapté à la population humaine) [26] ou des précédentes épidémies d'Ebola [27-28]. L'épidémie en cours d'Ebola serait, quant à elle, due au contact d'un enfant de 2 ans avec des chauves-souris frugivores dans un arbre et non à la chasse [29-30]. La colonisation de l'Europe par le moustique *Aedes albopictus*, arrivé par avion ou par bateau dans des pneus, a conduit à l'émergence de cas de dengue ou de chikungunya en Europe [10].

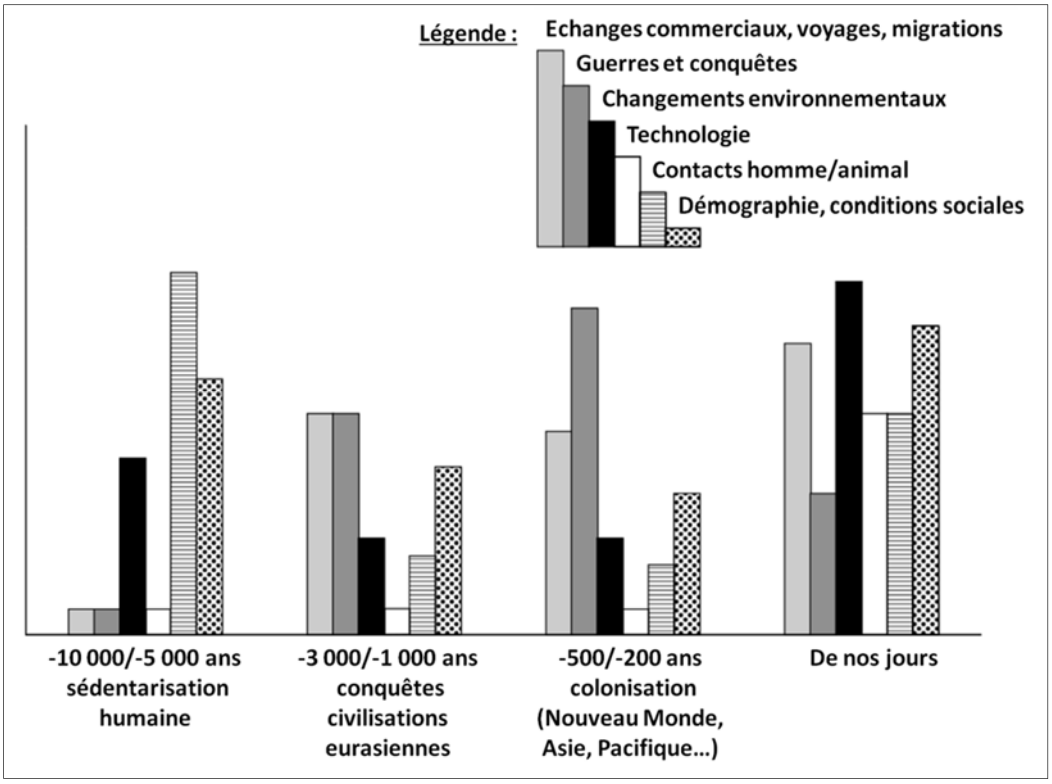


Figure 3. Importance relative de différents facteurs d'émergence au cours des 4 grandes ruptures épidémiques, d'après [25].

II. Virus zoonotiques et santé humaine

1. Définitions et exemples d'actualités

Soixante-dix pour cent des virus capables d'infecter l'homme ont une origine zoonotique [6]. On définit une zoonose comme étant une « maladie et infection dont les agents d'origine virale, bactérienne, fongique ou parasitaire se transmettent naturellement des animaux vertébrés à l'homme, et vice-versa » (OMS, 1959). Le terme, créé au XIXe siècle par Rudolf Virchow du grec « zôon » (animal) et « nosos » (maladie), englobe les zoo-anthroponoses (transmission de l'animal à l'homme) et les anthropo-zoonoses (transmission de l'homme à l'animal). Dans la suite du propos, nous appellerons « zoonoses » les maladies dues à des agents d'origine animale transmises à l'homme (anthropo-zoonoses) [31]. Il existe 4 types de zoonoses, selon leur mode de transmission :

- Orthozoonose : transmission directe d'un hôte vertébré à l'homme (ex : rage).
- Cyclozoonose : transmission via un hôte intermédiaire (ex : SARS).
- Métazoonose : transmission via un arthropode hématophage (ex : dengue).
- Saprozoonose : transmission via l'environnement (ex : leptospirose).

Dernièrement, de nombreux cas d'infections zoonotiques ont été rapportés dans la presse. Le *MERS-coronavirus* (MERS-CoV), un virus animal probablement issu de chauves-souris et infectant l'homme, possiblement via les chameaux, a déjà fait 567 morts sur 1 589 cas dans le monde depuis 2012 [32]. Le virus a été exporté en Corée du Sud puis en Chine cette année [33-35]. Depuis fin 2013, une épidémie sans précédent de fièvre hémorragique *Ebola* a touché principalement trois pays d'Afrique (Guinée, Sierra Leone et Libéria) et fait déjà 28 424 cas dont 11 311 morts selon le dernier bilan de l'OMS en date du 30 septembre 2015 [36]. Cette épidémie aurait pour origine un contact avec des chauves-souris [29]. Depuis le début de cette année, deux personnes ont été infectées aux Etats-Unis par le Bourbon virus (un nouveau thogotovirus) suite à une morsure de tique, et l'une d'elles est décédée [37].

2. Modes de transmission à l'homme et conditions de franchissement de la barrière d'espèce

Les virus zoonotiques sont des virus dont le cycle de réplication implique un réservoir animal, qu'il soit issu de la faune sauvage ou de la faune domestique, et qui se transmettent par contacts directs entre l'homme et l'animal (ex : chasse et consommation de viande de brousse [38], inhalation d'urines de rongeurs infectés, [39-40]) ou via la piqûre d'arthropodes hématophages entre animaux infectés et animaux sains (Figure 4, [41]). Dans la majorité des cas, l'homme n'est qu'un hôte accidentel.

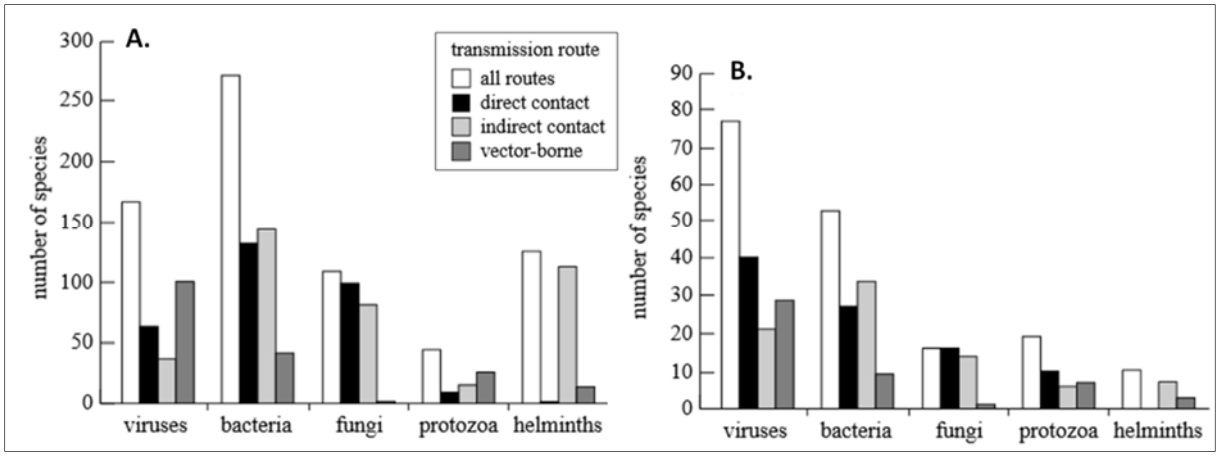


Figure 4. Nombre d'espèces d'agents infectieux zoonotiques (A, total = 868) ou émergents (B, total = 175) en fonction de leur mode de transmission, d'après [41].

Dans l'exemple du virus de la fièvre de la vallée du Rift (RVFV), le cycle naturel du virus implique un réservoir sauvage et des moustiques du genre *Aedes* ou *Culex* (cycle sylvatique du virus). L'homme est infecté de façon accidentelle après que le virus soit préalablement passé, toujours via la piqûre infectante d'un moustique, par un hôte intermédiaire, souvent un animal domestique (cycle domestique du virus) (Figure 5).

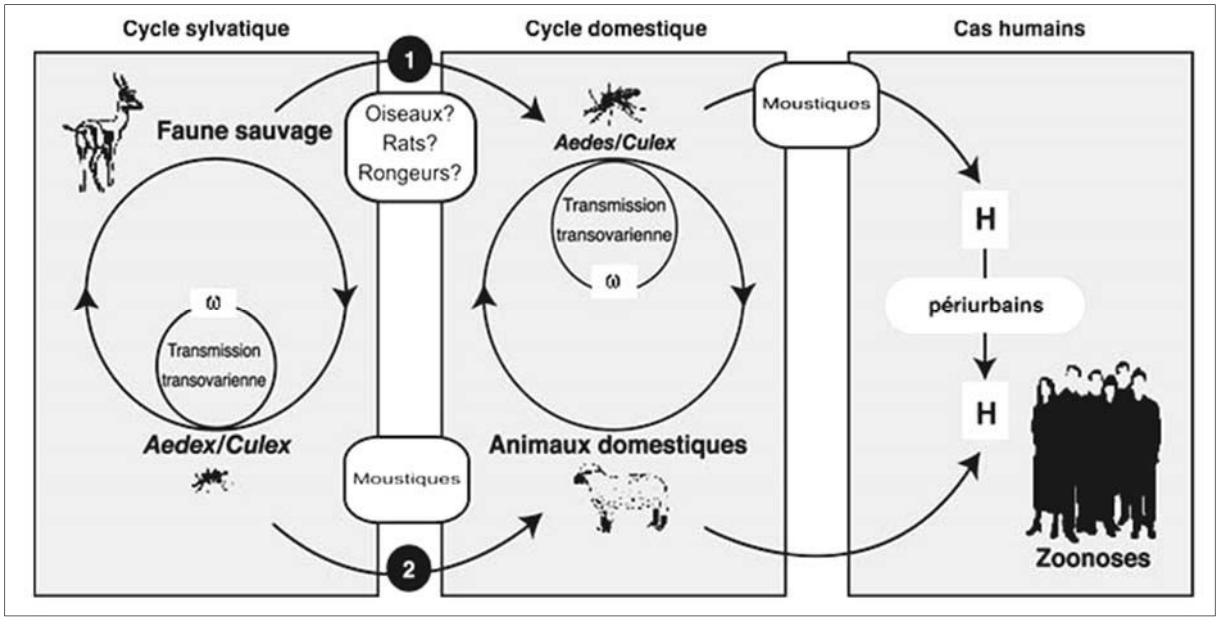


Figure 5. Cycle théorique de transmission du virus de la fièvre de la vallée du Rift entre l'animal et l'homme, d'après [42].

Ainsi, on distingue différents acteurs dans le cycle naturel de transmission d'un virus zoonotique [43] :

- le réservoir : le plus souvent d'origine animale, le réservoir est un hôte dans lequel se réplique l'agent infectieux sans que celui-ci ne présente de signes d'infection, due notamment à une réponse immunitaire de l'hôte : on parle alors de « tolérance » à l'infection [44].
- l'hôte intermédiaire / l'homme, des hôtes occasionnels : là encore souvent d'origine animale, l'hôte intermédiaire est lui sensible à l'infection, tout comme l'homme.
- le vecteur : d'origine arthropode hématophage, celui-ci n'est pas obligatoire dans le cycle de transmission. La transmission se fait alors en trois temps (1) l'arthropode prend un repas de sang infestant sur un hôte infecté (2) le virus se multiplie dans l'arthropode et migre dans les glandes salivaires (3) le virus est transmis à un hôte sain lors d'un nouveau repas de sang de l'arthropode.

On distingue donc différents facteurs pouvant impacter la transmission d'un virus zoonotique à l'homme, allant de la distribution géographique du réservoir animal au mode d'exposition et à la susceptibilité des hôtes occasionnels. Dans l'exemple de la transmission du virus *Hendra* décrit par

Plowright [45], virus de chauve-souris infectant directement les chevaux et l'homme sans intervention vectorielle, seule la conjonction de plusieurs facteurs aboutit à des cas humains : (1) les espèces de chauves-souris réservoirs du virus *Hendra* doivent être présentes ; (2) les chauves-souris doivent être infectées par le virus et l'excréter dans l'environnement ; (3) le virus doit être capable de survivre dans l'environnement ; (4) les hôtes accidentels (chevaux, hommes) doivent être en contact avec une quantité suffisante de virus pour que celui-ci puisse infecter l'hôte et (5) les hôtes accidentels doivent être sensibles au virus pour développer une infection.

Ainsi, pour qu'une transmission zoonotique soit efficace, il faut, entre autres, des contacts (directs ou indirects) entre l'homme et l'animal. Or ceux-ci sont de plus en plus fréquents. Comme vu précédemment, l'accroissement des activités humaines, l'augmentation démographique et la déforestation sont des exemples de facteurs d'émergence de zoonoses, en cela qu'ils favorisent les contacts directs entre l'homme et l'animal ou les contacts entre faune sauvage et faune domestique, elle-même en contact avec l'homme. Par exemple dans le cas du *SARS-coronavirus*, c'est le contact indirect de l'homme avec la chauve-souris, via la civette, qui serait à l'origine de l'émergence du virus en population humaine [46].

3. Arthropodes d'intérêt médical et pathogènes viraux associés

Les arthropodes sont des animaux invertébrés ayant un corps segmenté à symétrie bilatérale, des pattes articulées et une carapace rigide qui constitue leur exosquelette. Parmi eux, on distingue différentes classes d'arthropodes, en fonction du nombre de paires de pattes : les insectes en possèdent notamment 3 paires et les arachnides 4. En entomologie médicale, on distingue les arthropodes vecteurs des arthropodes venimeux, vénéneux, urticants ou allergisants pour lesquels aucun agent infectieux n'est incriminé. Dans la suite du propos, nous allons nous focaliser sur les arthropodes vecteurs.

On appelle « vecteur » un arthropode hématophage qui assure la transmission biologique active d'un agent pathogène d'un vertébré à un autre vertébré. De nombreux arthropodes non hématophages peuvent possiblement jouer un rôle dans le cycle infectieux d'une zoonose par simple transport mécanique, mais ils ne sont pas considérés comme « vecteurs » [47]. Parmi les arthropodes hématophages anthropophiles, on retrouve les ordres des *Diptera*

(ex : moustiques), *Hemiptera* (ex : punaises), *Anoploura* (ex : poux), *Siphonaptera* (ex : puces), *Mesostigmata* et *Trombidiformes* (ex : acariens) ou encore *Ixodida* (ex : tiques).

Non reconnue par l'International Commitee for Taxonomy of Viurses, la dénomination « arbovirus » correspond à un virus transmis par un arthropode. Il existe un très grand nombre d'espèces d'arthropodes capables de transmettre des arbovirus, notamment plus de 300 espèces de moustiques et 116 espèces de tiques, ce qui explique la répartition mondiale des arboviroses (Figure 6, [48-49]). De nombreux virus ont déjà été détectés dans d'autres arthropodes, et le rôle vecteur a déjà été clairement établi pour un certain nombre d'entre eux (Table 1). La détection d'un virus dans un arthropode n'est cependant pas suffisante pour pouvoir l'incriminer comme vecteur de cet agent infectieux. En effet, afin qu'un arthropode soit considéré comme vecteur, il doit être capable de transmettre activement un agent pathogène. On distingue donc deux notions : la compétence vectorielle et la capacité vectorielle. La compétence vectorielle correspond à l'aptitude d'un arthropode à ingérer un agent pathogène, en assurer la multiplication et le transmettre à un nouvel hôte. Cette compétence peut être évaluée en laboratoire, à la différence de la capacité vectorielle qui correspond à l'aptitude de l'arthropode à transmettre un agent infectieux en fonction de sa bio-écologie et des conditions environnementales. La capacité vectorielle est donc la résultante de la compétence vectorielle et de la bio-écologie du vecteur (ex : abondance, longévité, préférences trophiques).

Table 1. Exemples de familles virales préalablement retrouvées dans des arthropodes hématophages.

Règne	Classe	Ordre	Famille	Arthropode	Exemples de familles virales
Arthropoda	*Arachnida*	*Ixodida*	*Ixodidae*	Tique dure	*Flaviviridae*, *Bunyaviridae*, *Reoviridae*
			Argasidae	Tique molle	*Flaviviridae*, *Bunyaviridae*
	Insecta	*Anoploura*	*Pediculidae*	Poux	**non rapporté**
		Siphonaptera	*Pulicidae*	Puce	*Flaviviridae*
		Hemiptera	*Cimicidae*	Punaise de lit	*Bunyaviridae*, *Hepadnaviridae*
			Reduviidae	Triatome	**non rapporté**
		Diptera	*Simuliidae*	Simulie	**non rapporté**
			Tabanidae	Taon	*Bunyaviridae*
			Psychodidae	Phlébotome	*Bunyaviridae*, *Flaviviridae*, *Rhabdoviridae*
			Muscidae	Mouche tsé-tsé	**non rapporté**
			Culicidae	Moustique	*Flaviviridae*, *Togaviridae*, *Bunyaviridae*
			Ceratopogonidae	Culicoïde	*Bunyaviridae*, *Reoviridae*, *Flaviviridae*

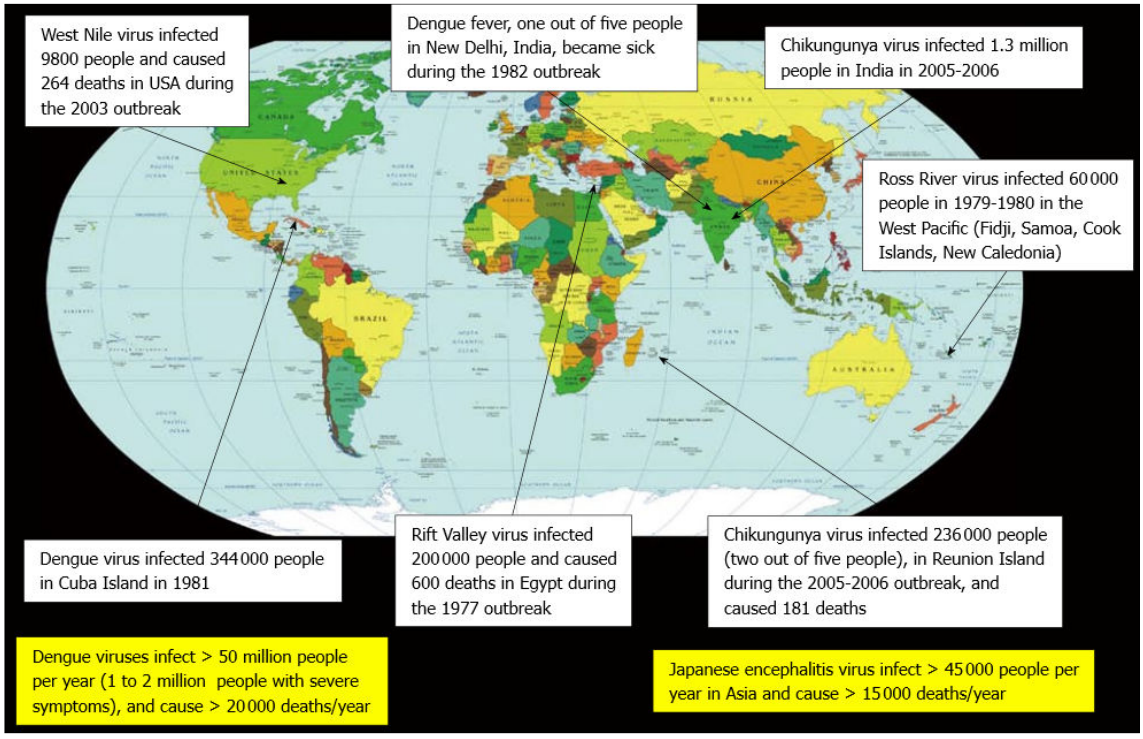


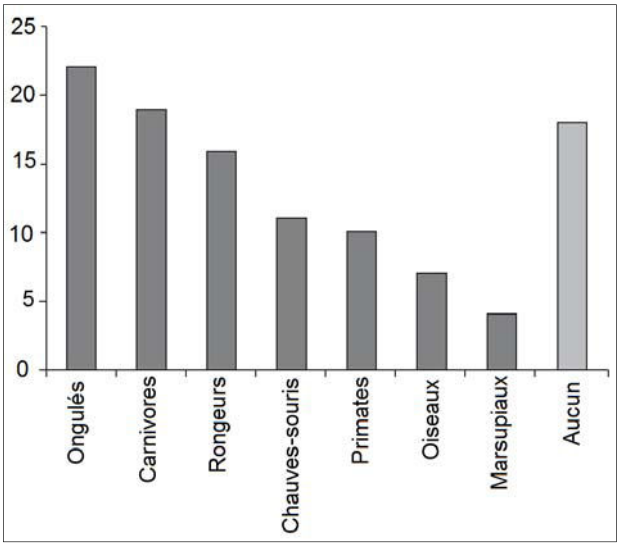
Figure 6. Exemples d'épidémies dues à des arbovirus, d'après [49].

4. Les faunes sauvage et domestique réservoirs de virus

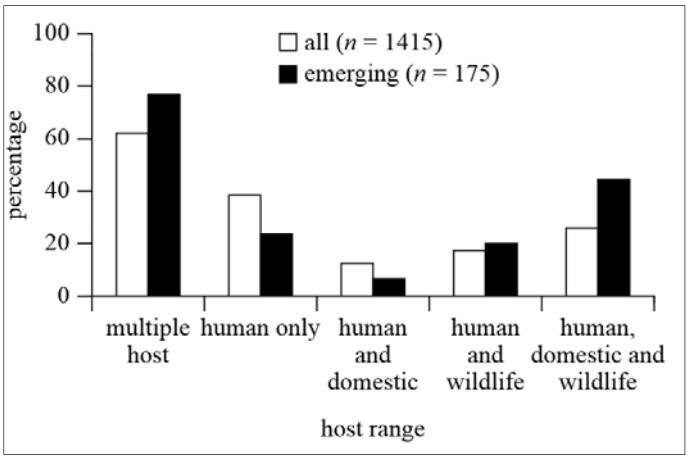
En 2009, une étude rapporta une classification des pathogènes émergents en fonction de leur origine zoonotique et a ainsi pu montrer que la faune sauvage constitue le principal réservoir d'infections zoonotiques (Figure 7A, [7,21]). Parmi eux, les chauves-souris et les rongeurs constituent une source importante de nouveaux virus. Les virus issus de la faune sauvage nécessitent souvent des hôtes intermédiaires de faune domestique avant de passer chez l'homme (Figure 7B, Table 2, [21,50]).

Table 2. Exemples de familles virales préalablement retrouvées dans la faune sauvage ou les animaux domestiques.

Réservoir	Zoonoses	Famille virale	Vecteur ?	2d hôte ?
Chauve-souris	Nipah/Hendra	*Paramyxoviridae*	non	Porc, cheval
	Fièvre hémorragique Ebola	*Filoviridae*	non	-
	SARS	*Coronaviridae*	non	Civette, chat
	Rage	*Rhabdoviridae*	non	Chien
Rongeur	Chorioméningite lymphocitaire	*Arenaviridae*	non	-
	Fièvre hémorragique Lassa	*Arenaviridae*	non	-
	Syndromes à hantavirus	*Bunyaviridae*	non	-
Oiseau	Encéphalite japonaise	*Flaviviridae*	moustique	Porc
	West Nile	*Flaviviridae*	moustique	Cheval
	Grippe aviaire	*Orthomyxoviridae*	non	Volaille, porc
Primate	Fièvre hémorragique Marburg	*Filoviridae*	non	-
	SIDA	*Retroviridae*	non	-
Bovin, ovin, caprin	Fièvre de la vallée du Rift	*Bunyaviridae*	moustique	-
	Vaccine	*Poxviridae*	non	-
Porc	Hépatite E	*Hepeviridae*	non	-



A. Réservoirs incriminés [7,21]



B. Types de contacts homme/animal [50]

Figure 7. Origine animale des principaux pathogènes émergents découverts depuis 1980.

III. Outils cellulaires et moléculaires d'identification de virus

Depuis le siècle dernier, le nombre de découvertes d'agents infectieux, et notamment de virus, a considérablement augmenté, passant de quelques dizaines à plusieurs milliers dont plus de 200 espèces virales décrites à ce jour et près de 2000 génomes viraux complets déposés dans GenBank [51]. Avec l'avènement des techniques de biologie moléculaire puis du séquençage à haut débit, ce phénomène tend encore à s'accélérer (Figure 8).

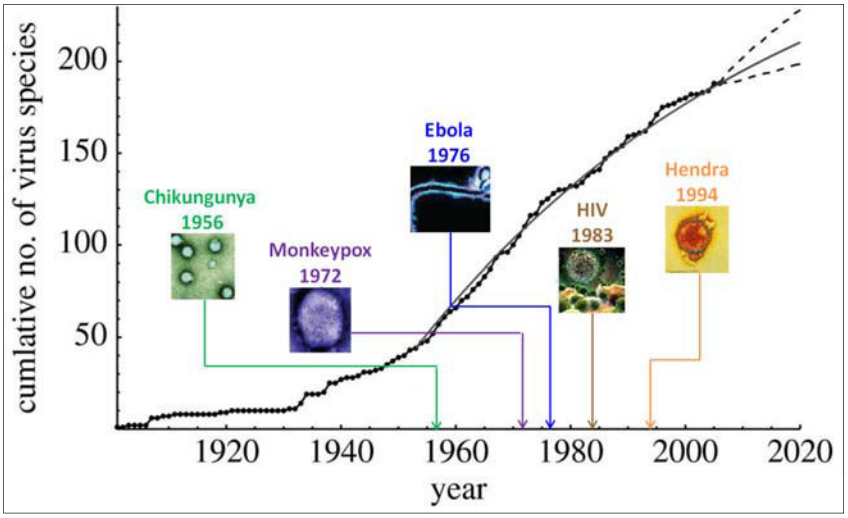


Figure 8. Rythme des découvertes de nouveaux virus de 1900 à 2008, d'après [51].

1. Les outils « historiques »

Les techniques classiquement utilisées pour identifier et caractériser des virus dans un échantillon complexe sont basées sur la détection de particules virales ou de leur constituant génomique (Table 3, [52-53]). Ainsi on retrouve par exemple l'isolement viral, technique de référence couramment utilisée jusqu'à l'avènement des techniques moléculaires. Les techniques basées sur la détection d'antigènes viraux au moyen d'anticorps sont encore utilisées de nos jours mais nécessitent une connaissance préalable du virus recherché et sont peu compatibles avec la recherche de nouveaux virus. Les techniques moléculaires, notamment celles basées sur l'utilisation de PCR pan-génériques ou de sondes couvrant tous les génomes viraux (microarray) sont, quant à elles, plus aptes à reconnaitre de nouveaux variants variants viraux mais, là encore elles ne sont pas compatibles avec l'identification de virus complètement nouveaux (Table 3).

Table 3. Principales techniques de détection et d'identification de virus.

		Technique	Principe	Principaux avantages	Principaux inconvénients
Détection de particules virales	**Biologie cellulaire**	**Culture cellulaire, isolement sur œuf embryonné ou cerveau de souriceau nouveau-né**	Isolement d'un virus par infection d'une cellule hôte *in vitro* ou *in vivo*	Virus disponible pour des caractérisations ultérieures, possible détermination de l'appartenance à une famille virale en fonction du type d'ECP, grande sensibilité	Technique longue, taux d'isolement très bas, non compatible avec virus non cultivables ou pour lesquels l'hôte est inconnu, modification des quasi-espèces
		Séroneutralisation, ELISA, immunofluorescence	Détection d'antigènes viraux	Rapidité, facilité de mise en œuvre, spécificité	Utilisation d'anticorps spécifiques d'un virus connu, ne convient pas pour des virus divergents, coût élevé
		Microscopie électronique	Visualisation de particules virales par coloration négative ou inclusion	Détermination +/- fine de l'appartenance à une famille virale, caractérisation morphologique d'un nouveau virus	Côut d'équipement, nécessité d'experts, faible sensibilité

	Technique	Principe	Principaux avantages	Principaux inconvénients
Détection de génomes viraux – Biologie moléculaire	**PCR spécifique**	Amplification d'une région spécifique d'un génome viral connu utilisant des amorces spécifiques	Bas coûts, grande spécificité, technique simple	Non compatible avec nouveaux variants ou nouveaux virus
	PCR pan-générique	Amplification d'une région conservée commune à un groupe de virus utilisant des amorces dégénérées	Bas coûts, technique simple	Faible sensibilité et spécificité, non compatible avec des virus trop divergents
	Microarray	Détection d'acides nucléiques viraux par hybridation sur des sondes basées sur des séquences conservées	Permet de rechercher des virus dans un échantillon complexe sans a priori des virus présents	Non compatible avec des virus trop divergents, hybridations non-spécifiques

2. La métagénomique : avantages et challenges

La métagénomique vise à étudier le contenu génétique d'un échantillon issu d'un environnement complexe (intestin, océan, sols, air, etc.) ainsi que les interactions entre les génomes présents. Cette approche permet donc d'avoir une description génomique du contenu de l'échantillon (« qui est présent ?») mais aussi un aperçu du potentiel fonctionnel d'un environnement (« qui fait quoi ? »). Le principal avantage de cette technique est qu'elle ne nécessite pas l'isolement préalable des virus, ni une connaissance préliminaire de virus apparentés. Ainsi, il est possible de caractériser la diversité microbienne d'un échantillon sans aucun a priori. On appelle « virome » l'ensemble des communautés virales d'un environnement complexe.

De nombreux virus issus de divers environnements ont déjà été découverts par métagénomique (Figure 9). Les premiers métagénomes viraux ont été réalisés sur des environnements marins et, par la suite, cette technique a été largement employée en écologie virale dans des études de diversité. Le premier métagénome viral associé à l'homme est néanmoins très vite apparu (en 2003), suivi du premier métagénome viral d'arthropodes en 2007 et du premier métagénome viral animal en 2008 [54]. Du fait de à sa capacité à

générer rapidement des millions de séquences en une seule réaction (« profondeur de séquençage ») pour un coût de plus en plus bas, cette technique s'est très rapidement imposée comme technique de choix dans des études d'inventaire des virus présents dans un échantillon donné, comparé aux techniques traditionnelles de découvertes de virus. Elle a notamment permis de découvrir des virus responsables de pathologies jusqu'alors inexpliquées (Table 4, [55]), que ce soit en santé humaine [56] ou en médecine vétérinaire [57].

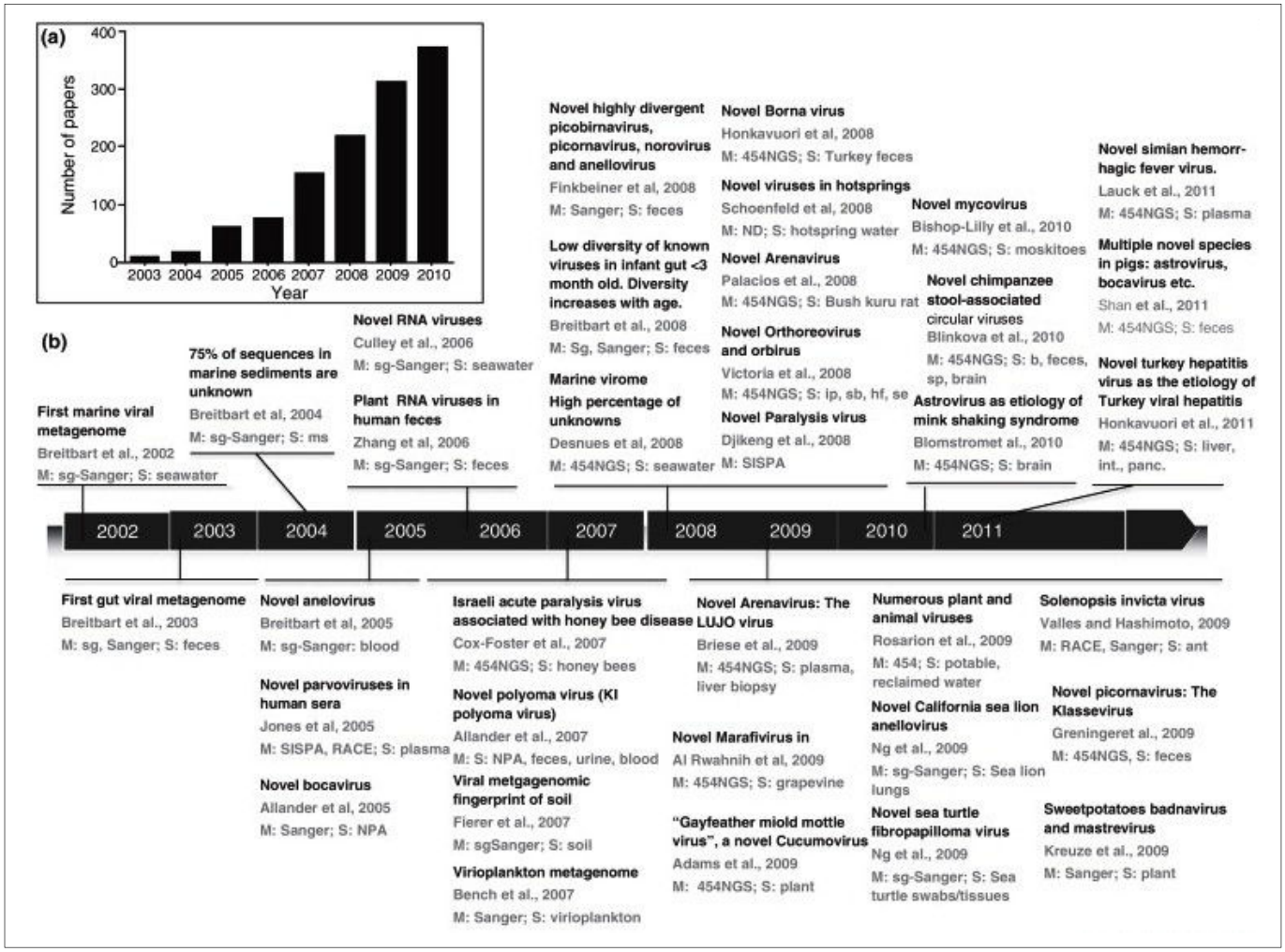


Figure 9. Nombre d'articles publiés (A) et principales découvertes de virus par métagénomique de 2002 à 2011 (B), d'après [54]. M : méthode de séquençage / S : échantillon.

Table 4. Exemples de virus découverts par NGS et leur implication dans des pathologies, d'après [55].

Virus	Technologie	Maladie associée		Corrélation virus / pathologie
Coronavirus-EMC	Roche 454	Pneumonie sévère	Homme	++
Severe fever with thrombocytopenia virus	Illumina	Fièvre et thrombocytopénie	Homme	++
Heartland bunyavirus	Roche 454	Syndrome febrile	Homme	++
Raccoon polyomavirus	RCA	Tumeur cérébrale	Raton laveur	++
Titi monkey adenovirus	Illumina	Pneumonie	Singe	++
Bas-Congo virus	Illumina	Fièvre hémorragique	Homme	++
Human astrovirus Puget Sound	Roche 454	Encéphalite	Homme	++

Il existe deux grands principes de métagénomique: la métagénomique basée sur des amplicons et le « whole-genome (shotgun) » (Figure 10). La métagénomique basée sur des amplicons utilise des amorces universelles, ciblant par exemple un gène conservé comme celui codant pour l'ARN ribosomal 16S procaryote, couplées à des adaptateurs qui serviront d'amorces pour le séquençage. Ainsi, une PCR universelle est réalisée puis le produit d'amplification est séquencé par NGS. Cette technique est largement employée pour la description des communautés bactériennes ; mais comme les virus ne possèdent pas de séquence conservée de type 16S, il n'est pas possible de réaliser une amplification universelle. En 2003 une étude de métagénomique portant sur les communautés virales marines a rapporté l'utilisation d'amorces pseudo-universelles ciblant la polymérase virale en vue de réaliser un séquençage par amplicons [58], cependant cette technique ne ciblait que les virus apparentés aux picornavirus chez qui la polymérase est hautement conservée. N'étant pas possible, à ce jour, de trouver des amorces réellement universelles, la technique dite de « shotgun » est la plus couramment employée. Cette technique consiste à extraire les acides nucléiques totaux d'un échantillon complexe, à les fragmenter et à liguer des adaptateurs de séquençage. Ainsi la totalité du contenu génomique de l'échantillon sera séquencée par NGS (Figure 10). Dans le cas des viromes à ARN, une étape de reverse-transcription est nécessaire.

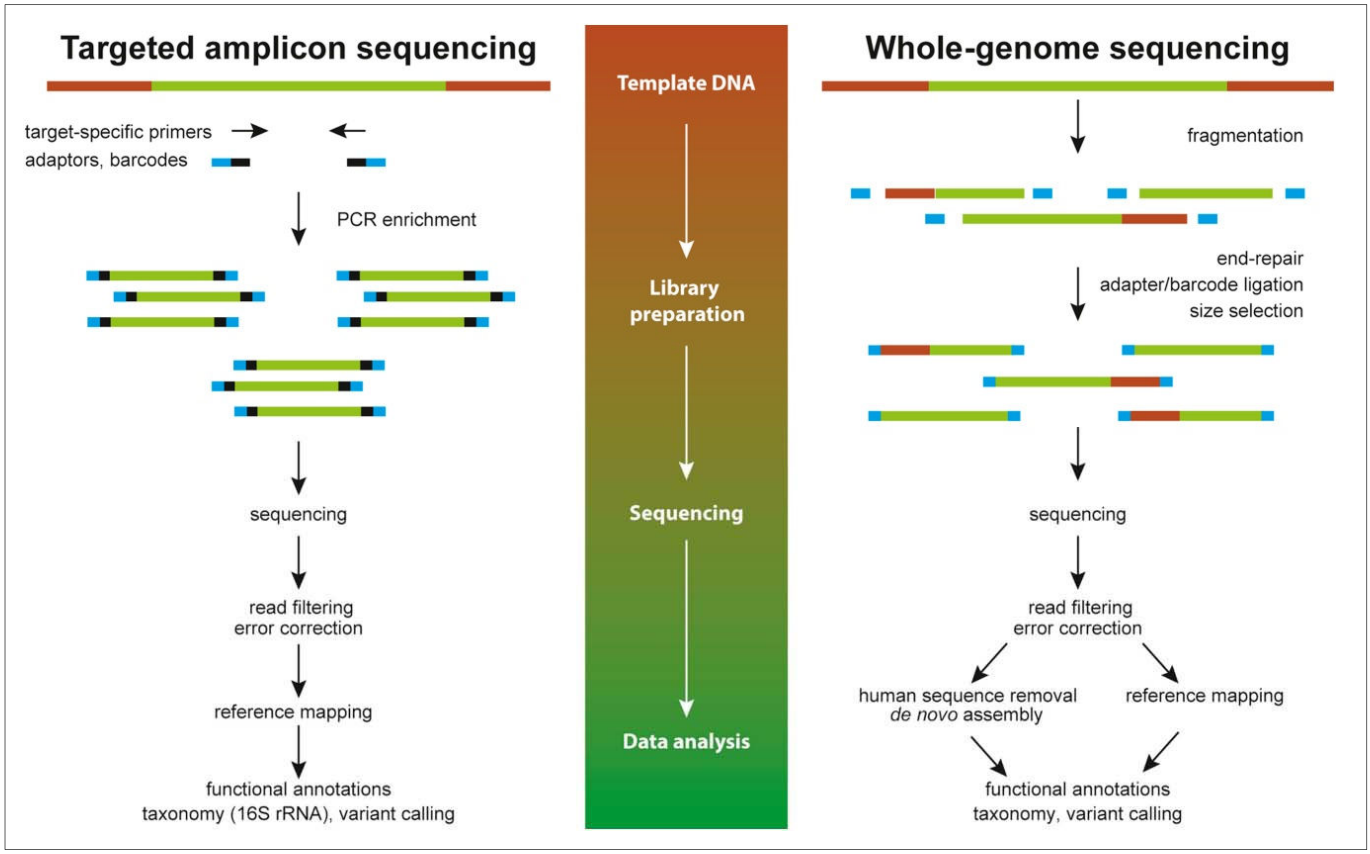


Figure 10. Principales techniques de séquençage à haut débit, d'après [59].

Certaines étapes clés influencent la qualité de préparation et d'analyse d'un métagénome, comme la purification virale, l'amplification ou l'analyse bio-informatique. En effet selon la richesse de l'échantillon et son niveau de contamination en acides nucléiques de l'hôte, il est souvent nécessaire de purifier et d'enrichir les particules virales afin de d'exploiter pleinement la profondeur de séquençage. Il existe de nombreux protocoles de purification, presque autant que de publications portant sur des métagénomes viraux, et au cours de cette thèse j'ai moi-même proposé et développé un protocole expérimental standardisé de purification des particules virales infectieuses d'un échantillon complexe (article n°2). Une autre étape clé est, dans le cas d'un échantillon à faible teneur en contenu génomique, l'amplification préalable du contenu génomique. Cette amplification est dite « séquence-indépendante » (SIA). Il existe différentes techniques de SIA : l'amplification séquence-indépendante utilisant une amorce unique (SISPA), la PCR utilisant des oligonucléotides dégénérés (DOP-PCR), la PCR random ou l'amplification en cercle roulant (RCA, [60]). Ces techniques présentent néanmoins différents biais d'amplification [61-64], résultant en une difficulté à assigner une

abondance en génomes viraux dans l'échantillon originel. Enfin une étape majeure est l'analyse bio-informatique des métagénomes. Cette étape est limitante dans la caractérisation des communautés microbiennes d'un échantillon, car la majorité des séquences obtenues restent non identifiées, du fait d'un manque de séquences homologues dans les bases de données, notamment concernant les viromes à ARN [65].

IV. **Article n°1**: revue "Viral metagenomics on animals as a tool for the detection of zoonoses prior to human infection?"

Viral metagenomics on animals as a tool for the detection of zoonoses prior to human infection?

Sarah Temmam [1], Bernard Davoust[1], Jean-Michel Berenger[1], Didier Raoult [1], and Christelle Desnues [1*]

Published in: Int. J. Mol. Sci. 2014;15(6):10377-97. doi: 10.3390/ijms150610377.

[1] URMITE UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, Marseille, France

***** Author to whom correspondence should be addressed; E-Mail: christelle.desnues@univ-amu.fr; Tel.: +33-491-324-630; Fax: +33-491-387-772.

Int. J. Mol. Sci. **2014**, *15*, 10377-10397; doi:10.3390/ijms150610377

OPEN ACCESS
International Journal of
Molecular Sciences
ISSN 1422-0067
www.mdpi.com/journal/ijms

Review

Viral Metagenomics on Animals as a Tool for the Detection of Zoonoses Prior to Human Infection?

Sarah Temmam, Bernard Davoust, Jean-Michel Berenger, Didier Raoult and Christelle Desnues *

Unité de Recherche sur les Maladies Infectieuses Tropicales Emergentes (URMITE) UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, 13005 Marseille, France; E-Mails: sarah.temmam@gmail.com (S.T.); bernard.davoust@gmail.com (B.D.); jean-michel.berenger@ap-hm.fr (J.-M.B.); didier.raoult@gmail.com (D.R.)

* Author to whom correspondence should be addressed; E-Mail: christelle.desnues@univ-amu.fr; Tel.: +33-491-324-630; Fax: +33-491-387-772.

Received: 12 March 2014; in revised form: 24 May 2014 / Accepted: 28 May 2014 / Published: 10 June 2014

Abstract: Many human viral infections have a zoonotic, *i.e.*, wild or domestic animal, origin. Several zoonotic viruses are transmitted to humans directly via contact with an animal or indirectly via exposure to the urine or feces of infected animals or the bite of a bloodsucking arthropod. If a virus is able to adapt and replicate in its new human host, human-to-human transmissions may occur, possibly resulting in an epidemic, such as the A/H1N1 flu pandemic in 2009. Thus, predicting emerging zoonotic infections is an important challenge for public health officials in the coming decades. The recent development of viral metagenomics, *i.e.*, the characterization of the complete viral diversity isolated from an organism or an environment using high-throughput sequencing technologies, is promising for the surveillance of such diseases and can be accomplished by analyzing the viromes of selected animals and arthropods that are closely in contact with humans. In this review, we summarize our current knowledge of viral diversity within such animals (in particular blood-feeding arthropods, wildlife and domestic animals) using metagenomics and present its possible future application for the surveillance of zoonotic and arboviral diseases.

Keywords: viral metagenomics; hematophagous arthropods; zoonoses; wildlife; domestic animals

1. Introduction

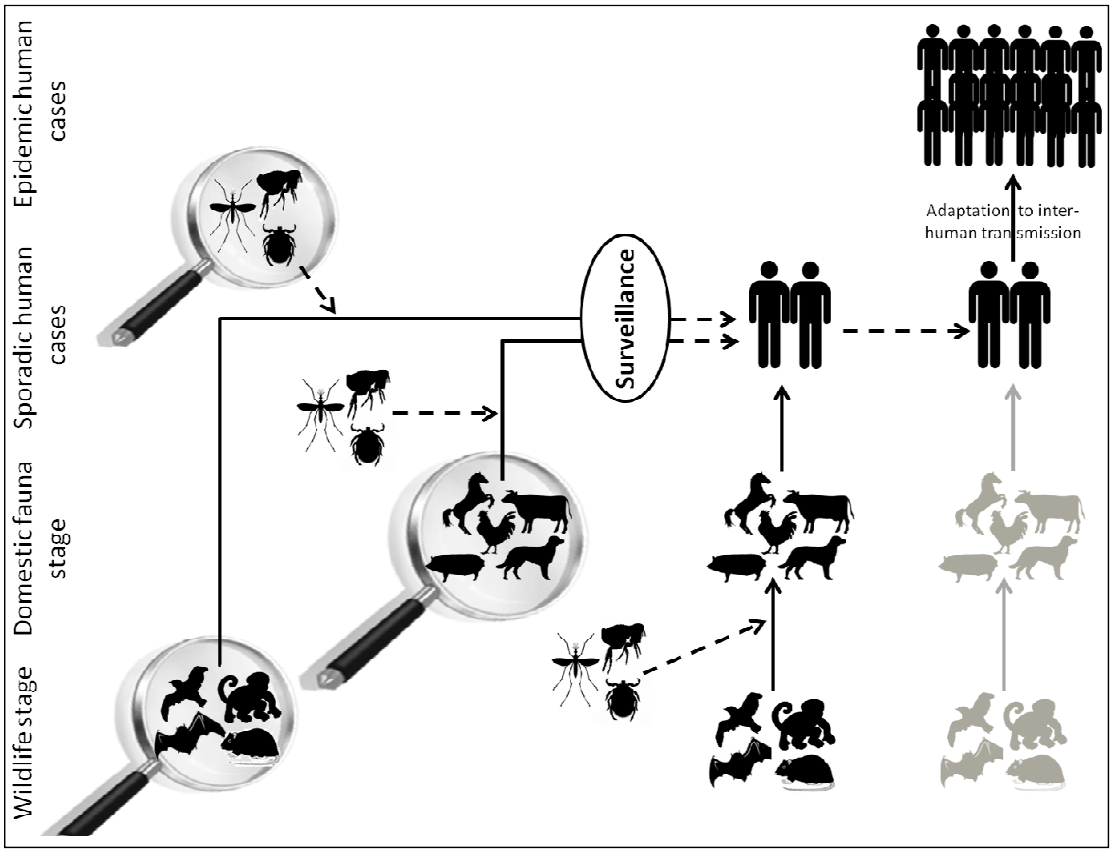
Human microbiologic infections, known as zoonoses, are acquired directly from animals or via arthropods bites and are an increasing public health problem. More than two thirds of emerging human pathogens are of zoonotic origin, and of these, more than 70% originate from wildlife [1,2]. In novel environments, viruses, particularly RNA viruses, can easily cross the species barrier by mutations, recombinations or reassortments of their genetic material, resulting in the capacity to infect novel hosts. Because of their adaptive abilities, RNA viruses represent more than 70% of the viruses that infect humans [3]. When socio-economic and ecologic changes affect their environment, humans may encounter increased contact with emerging viruses that originate in wild or domestic animals.

Wolfe *et al.* in 2007 [4] and Karesh *et al.* in 2012 [5] described different stages in the switch from an animal-specific infectious agent into a human-specific pathogen. The key stage is the transition of a strictly animal-specific infectious agent (originating from wildlife or domestic animals) to exposed human populations, resulting in sporadic human infections (Figure 1). If the pathogen is able to adapt to its human host and acquire the means to accomplish an inter-human transmission, horizontal human-to-human transmission occurs and maintains the viral cycle. Sometimes, an intermediate host, such as a domestic animal, is the link between sylvatic viral circulation and human viral circulation. For example, some human infections originating from bats, such as Nipah, Hendra, SARS and Ebola viral infections, may involve intermediate amplification in hosts such as pigs, horses, civets and primates, respectively [6] (Figure 1). Genetic, biologic, social, political or economic factors may explain a switch in viral host targets. For example, climate changes may influence the geographical repartition of vector arthropods, leading to new areas of the distribution of infectious diseases, like *Aedes albopictus* and Chikungunya infections in the Mediterranean [7]. Morens *et al.* [8] listed different key factors that may contribute to the emergence or re-emergence of infectious diseases, such as microbial adaptation to a new environment, biodiversity loss, ecosystem changes that lead to more frequent contact between wildlife and domestic animals or human populations, human demographics and behavior, economic development and land use, international travel and commerce, *etc.* [9,10]. These patterns of transmission allow identifying different animals to follow in order to monitor the appearance of new or re-emerging infectious agents before its first detection in the human populations. Therefore, hematophagous arthropods, wildlife and domestic animals may serve as targets for zoonotic and arboviral disease surveillance, particularly because sampling procedures and long-term follow-up studies are more easily performed in these hosts than in humans.

Historically, classic viral detection techniques were based on the intracerebral inoculation of suckling mice or viral isolation in culture and the subsequent observation of cytopathic effects on cell lines. Later, immunologic methods, e.g., seroneutralization or hemaglutination, were used to detect viral antigens in various complex samples. These techniques were based on the isolation of viral agents. With the progresses of molecular biology, polymerase chain reaction (PCR)-based methods became the main techniques for virus discovery and allowed the detection of uncultivable viruses [11], but these techniques required prior knowledge of closely related viral genomes. Next-Generation Sequencing (NGS) techniques make it possible to sequence all viral genomes in a given sample without previous knowledge about their nature. These techniques, known as viral metagenomics, have allowed the

discovery of completely new viral species. Because of their low cost, the use of NGS techniques is exponentially increasing.

Figure 1. The origins of zoonotic human infections.



The transmission of infections between humans occurs after a pathogen from a wild or domestic animal contacts with exposed human populations. The human exposures may or may not be mediated by the bite of bloodsucking arthropods. Surveillance programs may target wildlife, domestic animals or arthropods for emerging viruses before their adaptation to human hosts.

2. Viral Metagenomics: A Powerful Technique to Inventory the Viral Diversity among Complex Environments

Viruses can be identified by a wide range of techniques, which are mainly based on comparisons with known viruses. Historic methods include electron microscopy, cell culture, inoculation in suckling mice and serology, but these methods have limitations. For example, many viruses cannot be cultivated, excluding the use of cell line isolation and serologic techniques, and can only be characterized by molecular methods. In 2011, Bexfield summarized the different molecular techniques that identify new viruses such as microarray, subtractive hybridization-based and PCR-based methods [12]. Although these techniques have allowed the discovery of many viruses, the prior knowledge of similar viruses is required. Recent advances in sequence-independent PCR-based methods have overcome this limitation, and Sequence-Independent Single Primer Amplification (SISPA), Degenerate Oligonucleotide Primed PCR (DOP-PCR), random PCR and Rolling Circle Amplification (RCA) methods have emerged [12].

The end result of most of these PCR methods is amplified DNA that requires definitive identification by sequencing.

Novel DNA sequencing techniques, known as "Next-Generation Sequencing" (NGS) techniques, are new tools providing high-throughput sequence data with many possible applications in research and diagnostic settings [13]. With the development of different NGS platforms, it is now possible to sequence all viral genomes in a given sample without previous knowledge about their nature with the use of sequence-independent amplification followed by high-throughput sequencing. This combination of techniques, known as viral metagenomics, allows the discovery of completely new viral species within a complex sample and, due to decreasing costs, are nowadays exponentially increasing.

NGS techniques are able to generate a huge number of sequences, ranging from thousands to millions of reads, in only one reaction. In order to fully benefit from this depth of sequencing to identify infectious agents present in a given environment, host DNA/RNA should previously be removed from samples. Preliminary treatments are therefore required prior to nucleic acid amplification and sequencing, mainly based on nucleases treatments and/or viral purification by ultracentrifugation on sucrose, cesium chloride or glycerol gradients. These strategies are known as "Particle-Associated nucleic acid amplification" [14], *i.e.*, they try to isolate intact (*i.e.*, infectious) viral particles from their environment, protected from the action of nucleases. Subsequent low amount of nucleic acids have required the use of Sequence-Independent Amplifications (SIA) such as SISPA, DOP-PCR, random PCR, RCA [12]. Although these techniques allow generating enough nucleic acid material for sequencing, their main disadvantage remains that they distort quantitative analyzes by introducing bias of amplification in viral diversity studies. As a consequence, quantitative analyses of the composition of resulting viromes may not reflect the reality.

In diagnostic virology, in either human or veterinary medicine, viral metagenomics has allowed the discovery of causative viral agents of disease conditions [13–17]. Virome analyses have also been conducted to describe the baseline viral diversity in healthy human conditions [18], as a prior knowledge before studying the viral flora of pathologic conditions.

In the same way, the use of viral metagenomics as a tool for arboviral and zoonotic disease surveillance requires prior knowledge of the viral diversity associated to hematophagous arthropods and animals in close contact with humans. This review thus summarizes our current knowledge of the diversity of viral communities associated with several arthropods, wildlife and domestic animals and present its potential applications for the surveillance of zoonotic and arboviral diseases.

3. Blood-Feeding Arthropods

Several species of arthropods require a blood meal for their survival, either for their gonotrophic cycle or for feeding. Because humans and arthropods may share a common habitat, arthropods may feed on humans. When an arthropod feeds on a vertebrate host, it injects saliva into the host's blood and, if the micro-organism was previously able to replicate and to migrate into the salivary glands of the arthropod, the micro-organism is concomitantly injected to the vertebrate host's blood when the arthropod feeds. Such arthropods are called vectors. Medical entomology has therefore studied vector-host relationships for the last century. Several human pathogenic micro-organisms, including a large number of viral families, are transmitted via the bite of a bloodsucking arthropod. Reservoirs

may be the hematophagous arthropod itself if viruses are maintained by trans-ovarial and trans-stadial transmission to the progeny, or in several cases, the reservoirs may be of wildlife or domestic animal origin, with arthropods providing the transmission of the virus between the two vertebrate hosts. Table 1a summarizes the principal hematophagous arthropods of medical importance and associated arboviruses (for arthropod-borne viruses) detected by classic methodologies.

Recently, the use of metagenomics has allowed the detection of a large number of known and unknown insect-specific or zoonotic viruses associated with arthropods [19,20]. Broad surveys of viral diversity in arthropods are lacking, although metagenomics is a powerful technique for these studies and may be a promising tool for arbovirus surveillance. Only a few studies, focused on mosquitoes, are available. The major viruses that have been detected may be summarized in 4 categories (Table 1b): insect-specific viruses, plant viruses reflecting the nectar meal of mosquitoes, bacteria-infecting viruses (bacteriophages) and animal viruses possibly reflecting the arthropod blood meal of the vertebrate host.

3.1. Zoonotic Viruses

Ng and collaborators, in 2011, were the first to conduct a wide survey of viral diversity within mosquitoes using metagenomics [21]. Viral reads represented only 1% to 2% of total reads obtained after 454 pyrosequencing, and animal viruses represented not more than 10% of viral reads. Several mosquito species do not have strict host-specific trophic preferences, such as *Culex erythrothorax*, which was studied by Ng, *et al.* [21]. As a consequence, animal viruses detected in mosquitoes possibly reflect the virome of the large variety of vertebrate hosts they fed on (e.g., humans, primates, birds). For example, *Papillomaviridae* detected in mosquitoes have a human origin, while *Circoviridae* have mainly a bird origin [21]. No potential zoonotic viruses were documented in this study.

In 2014, Coffey and collaborators conducted a metagenomic analysis of Australian mosquitoes [22]. They were able to detect many animal viruses belonging to the *Flaviviridae* (Edge Hill virus) and *Reoviridae* (Wallal virus) families, but several viruses described in the study are known to infect marsupials. The authors also detected Ross River virus, an *Alphavirus* belonging to the *Togaviridae* family transmitted by mosquitoes, which is able to infect humans and cause influenza-like illness and/or polyarthritis, mainly in Australia. Coffey *et al.* also reported a novel virus, a dipteran-mammal-associated rhabdovirus called dimarhabdovirus. Viruses belonging to this supergroup are known to replicate both in hematophagous dipterans and vertebrate hosts [23], but no evidence of human infections are reported. Sequences belonging to the *Orthobunyavirus* genus (*Bunyaviridae* family) were also detected [22]. *Bunyaviridae* are single-stranded negative-strand segmented RNA viruses infecting a wide variety of vertebrate and invertebrate hosts. The family comprises 5 genera, and the *Orthobunyavirus* genus is divided into 4 major groups (Bunyamwera, Wyeomyia, Simbu and California encephalitis). Phylogenetic analyses performed on L and M segments of the newly described orthobunyaviruses did not place them into a specific group, nor provided information regarding a vertebrate host origin. However, their identification in anthropophilic mosquitoes coupled with their ability to replicate in mammalian cells may presuppose a human or other mammalian origin.

Although most of the zoonotic viruses transmitted by arthropods are RNA viruses, the RNA virome of arthropods, especially mosquitoes, is poorly described and is a future challenge for arbovirus

surveillance. Indeed, although studying viromes of engorged arthropods does not reflect the viral diversity intrinsically linked to arthropods, it provides a good picture of viruses circulating in a given human or animal population. The "One Health" concept (see below) recognizes that human health is linked to animal health and the environment in which they co-evolve. As a consequence and due to the ease with which they can be studied, arthropods may be used as targets for arboviral infections surveillance programs focusing either on human and veterinary health.

3.2. Insect-Specific Viruses

By sequencing the total DNA, Ng *et al.* [21] detected, 60% to 80% insect-specific viral reads, including a majority of densoviruses, as was also reported by Hall-Mendelin *et al.* [24]. Ma and collaborators [25] were also able to detect densoviruses by sequencing the small RNA molecules produced by the arthropod in response to viral infections. This process involved RNA interference or RNA silencing, resulting in the sequence-specific degradation of viral genomes.

Densoviruses (DNV) belong to the family *Parvoviridae*, sub-family *Densovirinae*. They are small, non-enveloped icosahedral ssDNA viruses, which infect the majority of arthropods (e.g., insect orders *Diptera*, *Hemiptera*, and *Lepidoptera*). DNV seem to be highly host-specific, and because they infect most tissues of their hosts, they are responsible for the death of their arthropod host. They apparently have larvicidal activity, and a few DNV-resistant adults may emerge, resulting in a population that is able to vertically transmit the infection to their descendants. In contrast, it seems that DNV infection of late-stage larvae or young adults results in the establishment of a persistent and vertically transmitted viremia. Thus, these viruses are a promising tool for vector control [26,27].

Other insect-specific viruses were detected by Cook *et al.* [28] using the same small RNA-based metagenomic technique, including *Chronic bee paralysis virus* and viruses belonging to the *Birnaviridae* family (genus *Entomobirnavirus*). The authors also detected sequences related to an insect-specific *Flaviviridae*, but it possibly constituted sequences integrated into the genome of the mosquito (Table 1b).

Table 1. (**a**) Non-exhaustive list of major viruses detected in blood-feeding arthropods by various serological or molecular techniques; (**b**) Examples of viruses detected in mosquitoes by metagenomic studies.

(**a**)

Kingdom	Class	Order	Family	Type Arthropod	Example of Viral Families	Ref.
Arthropoda	*Arachnida*	*Ixodida*	*Ixodidae*	Hard tick	*Flaviviridae* (TBEV, OHFV) *Bunyaviridae* (CCHFV) *Reoviridae* (CTFV)	[29,30]
			Argasidae	Soft tick	*Flaviviridae* (AHFV, SREV, WNV), *Bunyaviridae* (SOLV)	[31–34]
	Insecta	*Anoploura*	*Pediculidae*	Louse	not documented	
		Siphonaptera	*Pulicidae*	Flea	*Flaviviridae* (TBEV)	[35]
		Hemiptera	*Cimicidae*	Bed bug	*Bunyaviridae* (KKV) *Hepadnaviridae* (HBV)	[36–39]
			Reduviidae	Triatoma	not documented	

Table 1a. *Cont.*

Kingdom	Class	Order	Family	Type Arthropod	Example of Viral Families	Ref.
Arthropoda	*Insecta*	*Diptera*	*Simuliidae*	Black fly	not documented	
			Tabanidae	Horse fly	*Bunyaviridae* (LACV, JCV)	[40,41]
			Psychodidae	Sand fly	*Bunyaviridae* (SFNV, TOSV) *Flaviviridae*, *Rhabdoviridae*	[31,42–46]
			Muscidae	Tsetse fly	not documented	
			Culicidae	Mosquito	*Flaviviridae* (WNV, YFV, DENV), *Togaviridae* (CHIKV, ONNV), *Bunyaviridae* (RVFV, NRIV)	[47–50]
			Ceratopogonidae	Biting midge	*Bunyaviridae* (OROV, RVFV, CCHFV)	[51–53]

Abbreviations: Tick-Borne Encephalitis virus (TBEV), Omsk Hemorrhagic Fever virus (OHFV), Colorado Tick Fever virus (CTFV), Alkhurma Hemorrhagic Fever virus (AHFV), Saumarez Reef virus (SREV), West Nile virus (WNV), Soldado virus (SOLV), Kaeng Khoi virus (KKV), Hepatitis B virus (HBV), Lacrosse virus (LACV), Jamestown Canyon virus (JCV), Sandfly Fever Naples virus (SFNV), Toscana virus (TOSV), Yellow Fever virus (YFV), Dengue virus (DENV), Chikungunya virus (CHIKV), O'nyong-nyong virus (ONNV), Rift Valley Fever virus (RVFV), Ngari virus (NRIV), Oropouche virus (OROV), Crimee-Congo Hemorrhagic Fever virus (CCHFV).

(**b**)

Arthropod Species	Type Study	Viral Reads Taxonomic Assignation				Ref.
		Animal Viruses	Insect-Specific Viruses	Plant Viruses	Phages	
Mixed-species female mosquitoes	DNA virome	*Anelloviridae*, *Circoviridae*, *Herpesviridae*, *Poxviridae*, *Papillomaviridae*	*Parvoviridae* (*Densovirinae*), *Poxviridae* (*Entomopoxvirinae*), *Iridoviridae*	*Geminiviridae*, *Nanoviridae*	*Myoviridae*, *Podoviridae*, *Siphoviridae*	[21]
Culex pipiens molestus	Small RNA virome	not documented	*Parvoviridae* (*Densovirinae*)	not documented	not documented	[25]
Aedes sp.	DNA/RNA virome	not documented	*Parvoviridae* (*Densovirinae*)	not documented	*Inoviridae*	[24]
Anopheles sp., *Ochlerotatus* sp.	Small RNA virome	*Reoviridae* (*Orbivirus*) *	*Chronic bee paralysis virus*, *Birnaviridae* (*Entomobirnavirus*), *Flaviviridae* *, *Bunyaviridae* ** (*Phlebovirus*)	*Narnaviridae*, *Partitiviridae* *	not documented	[28]
Anopheles sp., *Culex* sp., *Aedes* sp.	DNA/RNA virome	*Rhabdoviridae*, *Bunyaviridae*, *Flaviviridae*, *Reoviridae*, *Togaviridae*	not documented	not documented	not documented	[22]

* Possibly integrated viral sequences into the genome of the arthropod; ** Classification hypothesized by the authors.

4. Wildlife

In 2004, Bengis *et al.* described two different patterns of transmission of infectious diseases from wildlife to humans [54]. The first pattern is one in which a viral disease of wildlife origin is transmitted rarely to humans, but once viral adaptation to the human host occurs, horizontal human-to-human transmission maintains the viral cycle. A major example of this pattern of transmission is the adaptation of HIV from SIV (Simian Immunodeficiency Virus) [55]. The second pattern involves many animal-to-human transmission events, possibly mediated by arthropods, for which animals are reservoirs and horizontal human-to-human transmissions are rare (Figure 1). A good example is West Nile virus infection, for which the usual viral cycle involves wild birds, mosquitoes, rarely horses and humans, who are accidental hosts [56]. Table 2a summarizes the principal zoonotic viruses associated with wildlife that are able to infect humans.

4.1. Bats

Frugivorous, insectivorous or hematophagous bats worldwide have been studied for their role as reservoirs of infectious agents. Many viruses isolated from bats are able to cross the species barrier and infect humans, regularly causing severe diseases in humans (e.g., SARS, Ebola hemorrhagic fever, Nipah, rabies) (Table 2a). Most metagenomic studies targeting wildlife have been conducted on bats (Table 2b), as Calisher and collaborators reviewed in 2006 [57], Wong and collaborators in 2007 [6], Smith and Wang in 2013 [58] or Luis *et al.* in 2013 [59]. Because "bat science" is a large and well-studied area in infectious diseases, this review will not focus more on this topic.

4.2. Rodents

Because of their close contact with humans, rodents are known reservoirs of pathogens, including many viral families (Table 2a). The major source of human contact with rodent pathogens is the exposure to the urine or feces of infected animals via the environment.

To describe the viral diversity of feces of wild rodents living in contact with humans, Phan and collaborators conducted a metagenomic analysis of 105 fecal specimens from mice, voles and rats [60]. They reported the presence of insect (e.g., *Densovirinae*, *Iridoviridae*) and plant viral sequences (e.g., *Nanoviridae*, *Geminiviridae*) reflecting the diet of rodents (Table 2b). They also detected several mammalian viruses, including the first known mouse sapelovirus and astrovirus, a species-specific mouse papillomavirus and novel picornaviruses possibly forming new genera within the family. Based on phylogenetic and distance-based data, a close relative of the Aichi virus was discovered in the murine stool samples [60]. Aichi virus is a virus belonging to the *Picornaviridae* family that has been identified in human diarrheas but for which the pathogenicity has not clearly been demonstrated. Although the murine Aichi virus shared more than 80% identity with the human Aichi virus, further studies need to be conducted to determine if this new virus is able to infect humans and, because it is excreted in rodents' feces, may represent a potential threat to human health.

Phan *et al.* also noted the presence of plant viruses, such as *Virgaviridae*, in the virome of the rodents' feces [60]. The authors concluded that these viruses reflect the diet of rodents, and usually plant viruses are considered incapable of infecting humans. However, a few studies reported the

presence of plant viral RNA in the human body, including the respiratory system via the use of cigarettes [61] and the gut via the consumption of contaminated food [62] though there is no evidence of a role in human pathologies.

Whether the animal viruses detected in the studies conducted on target animals have the capacity to infect humans is, to our knowledge, unknown, and this capacity needs to be further characterized before developing a metagenomic-target-based tool that is useful for the surveillance of emerging zoonoses.

Table 2. (**a**) Non-exhaustive list of major zoonotic viruses detected in wildlife; (**b**) Examples of viruses detected in bats and rodents by metagenomic studies.

(**a**)

Wildlife	Zoonosis	Virus	Vector-Based Transmission	Domestic Animal Intermediate Host	Ref.
Bat	Nipah/Hendra	*Paramyxoviridae*, *Henipavirus*	No	Pig/horse	[63]
	Ebola hemorrhagic fever	*Filoviridae*, *Ebolavirus*, EBOV	No	No	[64]
	Severe acute respiratory syndrome (SARS)	*Coronaviridae*, *Betacoronavirus*, SARS-CoV	No	Civet, cat	[65]
	Rabies	*Rhabdoviridae*, *Lyssavirus*, RABV	No	Dog	[66]
Rodent	Lymphocytic choriomeningitis	*Arenaviridae*, *Arenavirus*, LCMV	No	No	[67]
	Lassa hemorrhagic fever	*Arenaviridae*, *Arenavirus*, LASV	No	No	[67]
	Pulmonary syndrome and hemorrhagic syndrome	*Bunyaviridae*, *Hantavirus*	No	No	[68,69]
Bird	Japanese encephalitis	*Flaviviridae*, *Flavivirus*, JEV	Yes (mosquitoes)	Swine	[70]
	West Nile	*Flaviviridae*, *Flavivirus*, WNV	Yes (mosquitoes)	Horse	[56]
	Avian influenza	*Orthomyxoviridae*, *Influenzavirus*, A/H5N1, A/H1N1	No	Poultry, swine	[71–73]
Primate	Marburg hemorrhagic fever	*Filoviridae*, *Marburgvirus*, MARV	No	No	[74,75]
	Acquired immunodeficiency syndrome (AIDS)	*Retroviridae*, *Lentivirus*, HIV	No	No	[55]

Abbreviations: Ebola virus (EBOV), SARS-Coronavirus (SARS-CoV), Rabies virus (RABV), Lymphocytic ChorioMeningitis virus (LCMV), Lassa virus (LASV), Japanese Encephalitis virus (JEV), West Nile virus (WNV), Marburg virus (MARV), Human Immunodeficiency Virus (HIV).

Table 2. *Cont.*

(**b**)

Wild Animals	Type Study	Example of the Taxonomic Assignation of Viral Reads				Ref.
		Animal Viruses	Plant/Fungal Viruses	Phages	Insect-Specific Viruses	
Bats	DNA/RNA virome (feces)	*Parvoviridae*, *Circoviridae*, *Picornaviridae*, *Adenoviridae*, *Poxviridae*, *Astroviridae*, *Coronaviridae*	*Luteoviridae*, *Secoviridae*, *Tymoviridae*, *Partitiviridae*	*Microviridae*, *Siphoviridae*	*Dicistroviridae*, *Iflaviridae*, *Tetraviridae*, *Nodaviridae*, *Parvoviridae* (*Densovirinae*)	[76]
	DNA/RNA virome (feces, urine, throat swabs, tissue)	*Coronaviridae*, *Herpesviridae*	*Tymoviridae*	*Podoviridae*	*Iflaviridae*, *Dicistroviridae*	[77]
	DNA/RNA virome (feces, urine, tissue, serum, throat swabs)	*Retroviridae*, *Flaviviridae*, *Caliciviridae*, *Togaviridae*, *Paramyxoviridae*, *Adenoviridae*, *Papillomaviridae*, *Parvoviridae*, *Herpesviridae*, *Hepadnaviridae*	not documented	not documented	not documented	[78,79]
	DNA/RNA virome (feces)	*Papillomaviridae*, *Circoviridae*, *Anelloviridae*	not documented	unclassified	*Parvoviridae* (*Densovirinae*)	[80]
	DNA/RNA virome (feces, throat swabs)	*Adenoviridae*, *Herpesviridae*, *Papillomaviridae*, *Retroviridae*, *Circoviridae*, *Rhabdoviridae*, *Astroviridae*, *Flaviviridae*, *Coronaviridae*, *Picornaviridae*, *Parvoviridae*	*Chrysoviridae*, *Hypoviridae*, *Partitiviridae*, *Totiviridae*	*Inoviridae*, *Microviridae*	*Baculoviridae*, *Iflaviridae*, *Dicistroviridae*, *Tetraviridae*, *Parvoviridae* (*Densovirinae*)	[81]
	DNA/RNA virome (tissue)	*Herpesviridae*, *Papillomaviridae*, *Polyomaviridae*, *Hepadnaviridae*, *Circoviridae*, *Poxviridae*, *Retroviridae*, *Astroviridae*	*Phycodnaviridae*, *Bromoviridae*	*Myoviridae*, *Podoviridae*, *Siphoviridae*	*Baculoviridae*, *Polydnaviridae*, *Parvoviridae* (*Densovirinae*), *Iflaviridae*	[82]
	DNA/RNA virome (urine, throat swabs)	*Herpesviridae*, *Papillomaviridae*, *Adenoviridae*, *Poxviridae*, *Polyomaviridae*, *Retroviridae*, *Parvoviridae*, *Picornaviridae*	not documented	not documented	*Parvoviridae* (*Densovirinae*)	[83]
Rodents	DNA/RNA virome (feces)	*Circoviridae*, *Picobirnaviridae*, *Picornaviridae*, *Astroviridae*, *Parvoviridae*, *Papillomaviridae*, *Adenoviridae*, *Coronaviridae*	*Nanoviridae*, *Geminiviridae*, *Phycodnaviridae*, *Secoviridae*, *Partitiviridae*, *Tymoviridae*, *Alphaflexiviridae*, *Tombusviridae*	unclassified	*Parvoviridae* (*Densovirinae*), *Iridoviridae*, *Polydnaviridae*, *Dicistroviridae*, *Bromoviridae*, *Virgaviridae*	[60]

5. Domestic Animals

Several human infections have their origin in domestic animals. For example, the reservoir of genotype-3 Hepatitis E virus is pigs, but humans may be infected by the consumption of undercooked contaminated meat [84]. Some of these zoonotic viruses may be vector-transmitted to humans, such as the Rift Valley fever virus (Table 3a).

In veterinary medicine, metagenomic studies were conducted chiefly to determine the causal agent of pathologies with unknown etiology [16,17,85,86]. Few studies were conducted to describe the viral flora within domestic animals without a pathologic context. The viral diversity of domestic animals is summarized in Table 3b, which lists the major viruses discovered by veterinary medicine through metagenomics. Analyzing viral circulation within the domestic animals by metagenomics is a promising tool not only for veterinary medicine but also for the surveillance of zoonotic viruses possibly transmissible to humans for whom domestic animals act as reservoirs or intermediate hosts between wildlife and humans. Following the circulation of known and potential emerging viral agents in domestic animals appears to be an important surveillance goal.

5.1. Swine Breeding

Shan *et al.* reported in 2011 the metagenomic analysis of feces from healthy and diarrheic piglets grown in a high-density farm [87]. Viral reads represented 64% to 68% of total reads obtained after 454 sequencing, and RNA viruses accounted for more than 98% of viral reads. Beyond the RNA viral families detected, *i.e.*, *Picornaviridae*, *Astroviridae*, *Caliciviridae*, and *Coronaviridae* families, no zoonotic viruses that could infect humans were detected (Table 3b). In fact, the authors noted the presence of kobuviruses, astroviruses, enteroviruses, sapoviruses, sapeloviruses, coronaviruses, bocaviruses and teschoviruses either in healthy or diarrheic piglets, with only variations in the number of reads between healthy and diarrheic piglets. They concluded that such co-infections, even in healthy animals, may promote recombinations or reassortments between viruses, resulting in the emergence of new viruses, possibly infecting humans.

In 2012, Masembe and collaborators conducted a metagenomic analysis of domestic pig sera as part of a routine general surveillance program for African swine fever [88]. They were able to detect not only strictly swine-specific viruses (such as African swine fever viruses or swine Torque Teno viruses) but also zoonotic arboviruses (Table 3b). In fact, they reported the presence of Ndumu virus, an *Alphavirus* transmitted by mosquitoes, which may infect cattle [89] and humans [90] for whom no symptoms are yet known. Thus, the study of Masembe and collaborators emphasizes the usefulness of using metagenomics on domestic animals as a tool for the surveillance of human-infecting arboviruses.

5.2. Bushmeat and Wild Boars

Illegal bushmeat traffic is a problem for biodiversity conservation and is also a potential threat to human health when contaminated tissues are consumed [91]. Even legal bushmeat is a potential infectious hazard.

Bushpigs are hunted African wild boars which, because of their increasingly exploited habitat, have increased contact with domestically bred pigs and may infect domestic animals living in close contact

with humans. Bushpig meat is also consumed in some African countries. In this context, Blomström and collaborators conducted a metagenomic analysis of bushpig sera collected in Uganda [92]. They detected the presence of sequences related to suid-specific viruses, such as new variants of Porcine Parvovirus 4 and Torque Teno sus viruses, and the presence of a transcriptionally active Porcine Endogenous retrovirus. No zoonotic viruses, possibly infecting humans, were reported.

Reuter *et al.* conducted a similar study on wild boar feces collected in Hungary [93]. They noted the presence of viral reads matching the porcine Kobuvirus, a close relative of the human Aichi virus. Aichi virus and Kobuvirus are both viruses belonging to the *Picornaviridae* family detected in human and swine diarrheas respectively but for which the pathogenicity is not yet clearly demonstrated. Further studies should be conducted to determine whether porcine or wild boar kobuviruses are highly host-specific or if these viruses are able to infect humans.

Table 3. (**a**) Non-exhaustive list of major zoonotic viruses detected in domestic animals; (**b**) Examples of viruses detected in suid species by metagenomic diversity studies.

(**a**)

Domestic Animal	Zoonosis	Virus	Vector-Based Transmission	Ref.
Cats, dogs	Rabies	*Rhabdoviridae*, *Lyssavirus*, RABV	No	[94,95]
Cattle, sheep, goats	Rift Valley fever	*Bunyaviridae*, *Phlebovirus*, RVFV	Yes (mosquitoes)	[96,97]
	Vaccinia	*Poxviridae*, *Orthopoxvirus*, VACV	No	[98]
Pigs	Hepatitis E	*Hepeviridae*, *Hepevirus*, HEV	No	[99]
	Japanese encephalitis	*Flaviviridae*, *Flavivirus*, JEV	Yes (mosquitoes)	[70]
Horses	West Nile	*Flaviviridae*, *Flavivirus*, WNV	Yes (mosquitoes)	[56]
	Hendra	*Paramyxoviridae*, *Henipavirus*, HeV	No	[63]
Poultry	Avian flu	*Orthomyxoviridae*, *Influenzavirus*, A/H5N1	No	[72,73]

Abbreviations: Rabies virus (RABV), Rift Valley Fever virus (RVFV), Vaccinia virus (VACV), Hepatitis E virus (HEV), Japanese Encephalitis virus (JEV), West Nile virus (WNV), Hendra virus (HeV).

(**b**)

Animal Species	Type Studies	Viral Reads Taxonomic Assignation		Ref.
		Animal Viruses *stricto sensu*	Zoonotic Viruses	
Pigs	DNA/RNA virome (serum)	*Asfarviridae*, *Anelloviridae*, *Retroviridae*	*Togaviridae* (*Alphavirus*)	[88]
	DNA/RNA virome (stool)	*Picornaviridae*, *Astroviridae*, *Caliciviridae*, *Coronaviridae*, *Circoviridae*, *Parvoviridae*	not documented	[87]
Bushpigs	DNA/RNA virome (serum)	*Parvoviridae*, *Circoviridae*, *Retroviridae*	not documented	[92]
Wild boars	DNA/RNA virome (feces)	*Picornaviridae*, *Astroviridae*	not documented	[93]

No bacteriophages or plant viruses have been reported yet.

6. Future Perspectives in Metagenomic-Based Surveillance Programs

Viruses are the most abundant biological entities in the environment, including in the human body [18]. Viruses make up over two-thirds of all new human pathogens, a highly significant over-representation given that most current human pathogen species are bacteria, fungi or helminthes [100]. There are 219

viral species (belonging to 23 families) that are known to infect humans, among which more than two-thirds are of zoonotic origin [3].

Rudolf Virchow (1821–1902), a German physician and pathologist said "between animal and human medicine there are no dividing line, nor should there be". Although more than 60% of viruses that infect humans are of zoonotic origin, human and veterinary medicine has each evolved separately until recently. Only recently physicians and researchers working on human infectious diseases have become aware that human interactions with the ecosystem may affect human health. As a consequence, an interdisciplinary approach to health has begun that includes physicians, researchers, veterinarians, epidemiologists, and ecologists. This recent strategy, known as the "One World, One Health" concept [101] seeks to increase communication, collaboration, and cooperation across a wide variety of disciplines, such as human and veterinary medicine, public health, microbiology, and ecology, to attain optimal health for people, animals and the environment in which they evolve.

In this context, zoonotic-borne and arbovirus-borne disease surveillance programs have recently integrated entomology and veterinary medicine. To prevent emerging infectious diseases in humans, surveillance programs now focus on the early detection of new or re-emerging infectious agents in hematophagous arthropods and wild or domestic animals, before viral adaptation to human hosts (Figure 1). Viral metagenomics are well-adapted tools for these surveillance programs because they allow the detection of all viral genomes in a given sample without previous knowledge of their nature.

Because they are easy to sample, arthropods may be used as targets for emerging arbovirus-borne disease surveillance. Recent metagenomic analyses focused on mosquito arthropods have demonstrated the richness of the mosquito virome, including viruses that reflect the nectar or blood meals [19,20] (Table 1b). Because arboviruses are transmitted to vertebrate hosts via the saliva of arthropods, a simple way to determine if emerging viral pathogens may be transmitted to humans is to selectively analyze the virome of the salivary glands of the arthropod, even though dissection is difficult for extremely small arthropods. However, metagenomic studies targeting the entire body of the bloodsucking arthropod not only allow for the description of the viral flora within the arthropod, which highlight the emerging infectious agents or insect-specific viruses as tools for vector population control, but they also allow for the study of interactions between viral and bacterial communities that may result in viral interference (e.g., Wolbachia endosymbiont and Dengue virus interactions [102,103]) This information can lead to the development of new antiviral strategies. Because detecting viruses from the entire arthropod does not conclusively mean there is vector-based transmission of viruses to vertebrates, these studies would therefore require determining whether the virus is able to multiply in the arthropod and to migrate into the salivary glands.

Wild fauna may also be appropriate target animals for emerging zoonoses surveillance. Because of the many restrictions on studying endangered wild animals (such as bats), non-invasive sampling procedures may be used such as collecting urine or feces. Moreover, humans are more frequently in contact with feces or urine of wild animals in their shared environment, rather than with tissues or blood, with the exception of the consumption of bushmeat. As a consequence, most metagenomic studies conducted on wildlife have involved the feces or urine of wild animals (Table 2b) [58,60]. As for arthropods, these studies revealed how diverse and species-specific is the virome, and how unknown viruses have yet to be discovered.

Recent studies searching for the reservoir of Middle-East Respiratory Syndrome-Coronavirus (MERS-CoV) have shown the potential role of camels in the transmission of MERS-CoV to humans [104,105]. Camels are not the usual targets of zoonotic surveillance programs, but these recent examples highlight the interest of focusing future viral metagenomic studies on other animal species interacting with humans if one considers their ability to transmit human infectious agents by crossing the species barriers between animals and humans.

Metagenomics is thus a promising tool for the detection of new viral species that could potentially be a threat for human health. However, it yet suffers several pitfalls when considering new/highly divergent viral genomes. Indeed, the taxonomic assignation of reads generated by NGS techniques is only based on the comparison of sequences or patterns with previously described sequences present in databases. As a consequence, completely new or highly divergent viral sequences might be difficult to identify and subsequently there is a high risk to miss the detection of important viral pathogens. This problem remains the major challenge of metagenomic studies. Future progress in metagenomics should improve *in silico* analyses to overcome or attenuate this problem and would therefore permit to use metagenomics tools for the surveillance of emerging viruses.

Finally, detecting viruses, and especially viral genomes, within a given animal does not provide evidence of the transmissibility of the virus to humans. Determining the viral ability to cross the species barrier and to infect humans is a necessary part of studying viral metagenomics. In 1890, Koch's postulates described 4 criteria to determine the etiology of a pathology, mainly based on the cultivation of infectious agents isolated from diseased organisms [106]. These postulates were recently adapted by Fredericks *et al.* [107] and Mokili *et al.* [108] to molecular and NGS data. Metagenomics is a powerful tool to detect potentially new or re-emerging viruses in complex samples, however, subsequent studies are needed to determine if the viruses that were detected represent a potential threat to human (or animal) health. Defining the causality of a given pathology is a complex task, and the isolation of viral agents via cell culture or intracerebral inoculation of suckling mice remains the gold-standard in conducting studies of the pathogenicity of viruses detected by metagenomics.

Acknowledgments

The authors would like to thank Hervé Pascalis from CRVOI for his helpful reviewing and Noël Tordo from Institut Pasteur for his expertise on rodents and *Rhabdoviridae*. This work was supported by Starting Grant No. 242729 from the European Research Council awarded to C. Desnues.

Author Contributions

S.T. wrote the manuscript, B.D., J.-M.B., D.R. and C.D. reviewed and approved the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Cutler, S.J.; Fooks, A.R.; van der Poel, W.H. Public health threat of new, reemerging, and neglected zoonoses in the industrialized world. *Emerg. Infect. Dis.* **2010**, *16*, 1–7.
2. One Health Initiative. Available online: http://www.onehealthinitiative.com/ (accessed on 18 February 2014).
3. Woolhouse, M.; Scott, F.; Hudson, Z.; Howey, R.; Chase-Topping, M. Human viruses: Discovery and emergence. *Phil. Trans. R. Soc. B* **2012**, *367*, 2864–2871.
4. Wolfe, N.D.; Panosian Dunavan, C.; Diamond, J. Origins of major human infectious diseases. *Nature* **2007**, *447*, 279–283.
5. Karesh, W.B.; Dobson, A.; Lloyd-Smith, J.O.; Lubroth, J.; Dixon, M.A.; Bennett, M.; Aldrich, S.; Harrington, T.; Formenty, P.; Loh, E.H.; *et al*. Ecology of zoonoses: Natural and unnatural histories. *Lancet* **2012**, *380*, 1936–1945.
6. Wong, S.; Lau, S.; Woo, P.; Yuen, K.Y. Bats as a continuing source of emerging infections in humans. *Rev. Med. Virol*. **2007**, *17*, 67–91.
7. Gould, E.A.; Higgs, S. Impact of climate change and other factors on emerging arbovirus diseases. *Trans. R. Soc. Trop. Med. Hyg*. **2009**, *103*, 109–121.
8. Keesing, F.; Belden, L.K.; Daszak, P.; Dobson, A.; Harvell, C.D.; Holt, R.D.; Hudson, P.; Jolles, A.; Jones, K.E.; Mitchell, C.E.; *et al*. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **2010**, *468*, 647–652.
9. Morens, D.M.; Folkers, G.K.; Fauci, A.S. The challenge of emerging and re-emerging infectious diseases. *Nature* **2004**, *438*, 242–249.
10. Jones, K.E.; Patel, N.G.; Levy, M.A.; Storeygard, A.; Balk, D.; Gittleman, J.L.; Daszak, P. Global trends in emerging infectious diseases. *Nature* **2008**, *451*, 990–993.
11. Pozzetto, B. Méthodes de diagnostic en virologie. In *Virologie Médicale*; Mammette, A., Ed.; Presses Universitaires de Lyon: Lyon, France, 2002; pp. 187–210.
12. Bexfield, N.; Kellam, P. Metagenomics and the molecular identification of novel viruses. *Vet. J.* **2011**, *190*, 191–198.
13. Barzon, L.; Lavezzo, E.; Militello, V.; Toppo, S.; Palù, G. Applications of next-generation sequencing technologies to diagnostic virology. *Int. J. Mol. Sci.* **2011**, *12*, 7861–7884.
14. Stang, A.; Korn, K.; Wildner, O.; Uberla, K. Characterization of virus isolates by particle-associated nucleic acid PCR. *J. Clin. Microbiol.* **2005**, *43*, 716–720.
15. Capobianchi, M.R.; Giombini, E.; Rozera, G. Next Generation Sequencing technology in clinical virology. *Clin. Microbiol. Infect*. **2013**, *19*, 15–22.
16. Blomström, A.L. Viral metagenomics as an emerging and powerful tool in veterinary medicine. *Vet. Q*. **2011**, *31*, 107–114.
17. Belák, S.; Karlsson, O.E.; Blomström, A.L.; Berg, M.; Granberg, F. New viruses in veterinary medicine, detected by metagenomic approaches. *Vet. Microbiol*. **2013**, *165*, 95–101.
18. Popgeorgiev, N.; Temmam, S.; Raoult, D.; Desnues, C. Describing the silent human virome with an emphasis on giant viruses. *Intervirology* **2013**, *56*, 395–412.
19. Liu, S.; Vijayendran, D.; Bonning, B.C. Next generation sequencing technologies for insect virus discovery. *Viruses* **2011**, *3*, 1849–1869.

20. Junglen, S.; Drosten, C. Virus discovery and recent insights into virus diversity in arthropods. *Curr. Opin. Microbiol*. **2013**, *16*, 507–513.
21. Ng, T.F.; Willner, D.L.; Lim, Y.W.; Schmieder, R.; Chau, B.; Nilsson, C.; Anthony, S.; Ruan, Y.; Rohwer, F.; Breitbart, M. Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS One* **2011**, *6*, e20579.
22. Coffey, L.L.; Page, B.L.; Greninger, A.L.; Herring, B.L.; Russell, R.C.; Doggett, S.L.; Haniotis, J.; Wang, C.; Deng, X.; Delwart, E.L. Enhanced arbovirus surveillance with deep sequencing: Identification of novel rhabdoviruses and bunyaviruses in Australian mosquitoes. *Virology* **2014**, *448*, 146–158.
23. Bourhy, H.; Cowley, J.A.; Larrous, F.; Holmes, E.C.; Walker, P.J. Phylogenetic relationships among rhabdoviruses inferred using the *L polymerase* gene. *J. Gen. Virol*. **2005**, *86*, 2849–2858.
24. Hall-Mendelin, S.; Allcock, R.; Kresoje, N.; van den Hurk, A.F.; Warrilow, D. Detection of arboviruses and other micro-organisms in experimentally infected mosquitoes using massively parallel sequencing. *PLoS One* **2013**, *8*, e58026.
25. Ma, M.; Huang, Y.; Gong, Z.; Zhuang, L.; Li, C.; Yang, H.; Tong, Y.; Liu, W.; Cao, W. Discovery of DNA viruses in wild-caught mosquitoes using small RNA high throughput sequencing. *PLoS One* **2011**, *6*, e24758.
26. Fédière, G. Epidemiology and pathology of densovirinae. *Contrib. Microbiol*. **2000**, *4*, 1–11.
27. Carlson, J.; Suchman, E.; Buchatsky, L. Densoviruses for control and genetic manipulation of mosquitoes. *Adv. Virus Res*. **2006**, *68*, 361–392.
28. Cook, S.; Chung, B.Y.; Bass, D.; Moureau, G.; Tang, S.; McAlister, E.; Culverwell, C.L.; Glücksman, E.; Wang, H.; Brown, T.D.; *et al*. Novel virus discovery and genome reconstruction from field RNA samples reveals highly divergent viruses in dipteran hosts. *PLoS One* **2013**, *8*, e80720.
29. Estrada-Peña, A.; Jongejan, F. Ticks feeding on humans: A review of records on human-biting Ixodoidea with special reference to pathogen transmission. *Exp. Appl. Acarol*. **1999**, *23*, 685–715.
30. Dobler, G. Zoonotic tick-borne flaviviruses. *Vet. Microbiol*. **2010**, *140*, 221–228.
31. Charrel, R.N.; Fagbo, S.; Moureau, G.; Alqahtani, M.H.; Temmam, S.; de Lamballerie, X. Alkhurma hemorrhagic fever virus in Ornithodoros savignyi ticks. *Emerg. Infect. Dis*. **2007**, *13*, 153–155.
32. Gritsun, T.S.; Nuttall, P.A.; Gould, E.A. Tick-borne flaviviruses. *Adv. Virus Res*. **2003**, *61*, 317–371.
33. Lawrie, C.H.; Uzcátegui, N.Y.; Gould, E.A.; Nuttall, P.A. Ixodid and argasid tick species and West Nile virus. *Emerg. Infect. Dis*. **2004**, *10*, 653–657.
34. Converse, J.D.; Hoogstraal, H.; Moussa, M.I.; Feare, C.J.; Kaiser, M.N. Soldado virus (Hughes group) from *Ornithodoros* (*Alectorobius*) *capensis* (*Ixodoidea*: *Argasidae*) infesting Sooty Tern colonies in the Seychelles, Indian Ocean. *Am. J. Trop. Med. Hyg*. **1975**, *24*, 1010–1018.
35. Sotnikova, A.N.; Soldatov, G.M. Isolation of tick-borne encephalitis viruses from fleas ceratophyllus tamias wagn. *Med. Parazitol.* **1964**, *33*, 622–624.
36. Kolb, A.; Needham, G.R.; Neyman, K.M.; High, W.A. Bedbugs. *Dermatol. Ther*. **2009**, *22*, 347–352.

37. Williams, J.E.; Imlarp, S.; Top, F.H., Jr.; Cavanaugh, D.C.; Russell, P.K. Kaeng Khoi virus from naturally infected bedbugs (cimicidae) and immature free-tailed bats. *Bull. World Health Organ*. **1976**, *53*, 365–369.
38. Silverman, A.L.; Qu, L.H.; Blow, J.; Zitron, I.M.; Gordon, S.C.; Walker, E.D. Assessment of hepatitis B virus DNA and hepatitis C virus RNA in the common bedbug (*Cimex lectularius* L.) and kissing bug (*Rodnius prolixus*). *Am. J. Gastroenterol*. **2001**, *96*, 2194–2198.
39. Usinger, R.L. *Monograph of the Cimicidae*; Entomological Society of America: Annapolis, MD, USA, 1966; pp. 81–167.
40. Wright, R.E.; Anslow, R.O.; Thompsons, W.H.; DeFoliart, G.R.; Seawright, G.; Hanson, R.P. Isolations of lacrosse virus of the California group from Tabanidae in Wisconsin. *Mosquito News* **1970**, *30*, 600–603.
41. DeFoliart, G.R.; Anslow, R.O.; Hanson, R.P.; Morris, C.D.; Papadopoulos, O.; Sather, G.E. Isolation of Jamestown Canyon serotype of California encephalitis virus from naturally infected Aedes mosquitoes and tabanids. *Am. J. Trop. Med. Hyg*. **1969**, *18*, 440–447.
42. Charrel, R.N.; Izri, A.; Temmam, S.; de Lamballerie, X.; Parola, P. Toscana virus RNA in Sergentomyia minuta flies. *Emerg. Infect. Dis*. **2006**, *12*, 1299–1300.
43. Izri, A.; Temmam, S.; Moureau, G.; Hamrioui, B.; de Lamballerie, X.; Charrel, R.N. Sandfly fever Sicilian virus, Algeria. *Emerg. Infect. Dis*. **2008**, *14*, 795–797.
44. Moureau, G.; Ninove, L.; Izri, A.; Cook, S.; de Lamballerie, X.; Charrel, R.N. Flavivirus RNA in phlebotomine sandflies. *Vector Borne Zoonotic Dis*. **2010**, *10*, 195–197.
45. Comer, J.A.; Tesh, R.B. Phlebotomine sand flies as vectors of vesiculoviruses: A review. *Parassitologia* **1991**, *33*, 143–150.
46. Tesh, R.; Saidi, S.; Javadian, E.; Loh, P.; Nadim, A. Isfahan virus, a new vesiculovirus infecting humans, gerbils, and sandflies in Iran. *Am. J. Trop. Med. Hyg*. **1977**, *26*, 299–306.
47. Groseth, A.; Weisend, C.; Ebihara, H. Complete genome sequencing of mosquito and human isolates of Ngari virus. *J. Virol*. **2012**, *86*, 13846–13847.
48. Mwaengo, D.; Lorenzo, G.; Iglesias, J.; Warigia, M.; Sang, R.; Bishop, R.P.; Brun, A. Detection and identification of Rift Valley fever virus in mosquito vectors by quantitative real-time PCR. *Virus Res*. **2012**, *169*, 137–143.
49. Paupy, C.; Kassa Kassa, F.; Caron, M.; Nkoghé, D.; Leroy, E.M. A chikungunya outbreak associated with the vector Aedes albopictus in remote villages of Gabon. *Vector Borne Zoonotic Dis*. **2012**, *12*, 167–169.
50. Calzolari, M.; Bonilauri, P.; Bellini, R.; Albieri, A.; Defilippo, F.; Maioli, G.; Galletti, G.; Gelati, A.; Barbieri, I.; Tamba, M.; *et al*. Evidence of simultaneous circulation of West Nile and Usutu viruses in mosquitoes sampled in Emilia-Romagna region (Italy) in 2009. *PLoS One* **2010**, *5*, e14324.
51. Carpenter, S.; Groschup, M.H.; Garros, C.; Felippe-Bauer, M.L.; Purse, B.V. Culicoides biting midges, arboviruses and public health in Europe. *Antivir. Res*. **2013**, *100*, 102–113.
52. Lee, V.H. Isolation of viruses from field populations of culicoides (*Diptera*: *Ceratopogonidae*) in Nigeria. *J. Med. Entomol*. **1979**, *16*, 76–79.
53. Davies, F.G.; Walker, A.R.; Ochieng, P.; Shaw, T. Arboviruses isolated from culicoides midges in Kenya. *J. Comp. Pathol*. **1979**, *89*, 587–595.

54. Bengis, R.G.; Leighton, F.A.; Fischer, J.R.; Artois, M.; Mörner, T.; Tate, C.M. The role of wildlife in emerging and re-emerging zoonoses. *Rev. Sci. Tech*. **2004**, *23*, 497–511.
55. Hahn, B.H.; Shaw, G.M.; de Cock, K.M.; Sharp, P.M. AIDS as a zoonosis: Scientific and public health implications. *Science* **2000**, *287*, 607–614.
56. Trevejo, R.T.; Eidson, M. Zoonosis update: West Nile virus. *J. Am. Vet. Med. Assoc*. **2008**, *232*, 1302–1309.
57. Calisher, C.H.; Childs, J.E.; Field, H.E.; Holmes, K.V.; Schountz, T. Bats: Important reservoir hosts of emerging viruses. *Clin. Microbiol. Rev*. **2006**, *19*, 531–545.
58. Smith, I.; Wang, L.F. Bats and their virome: An important source of emerging viruses capable of infecting humans. *Curr. Opin. Virol*. **2013**, *3*, 84–91.
59. Luis, A.D.; Hayman, D.T.; O'Shea, T.J.; Cryan, P.M.; Gilbert, A.T.; Pulliam, J.R.; Mills, J.N.; Timonin, M,E.; Willis, C.K.; Cunningham, A.A.; *et al*. A comparison of bats and rodents as reservoirs of zoonotic viruses: Are bats special? *Proc. Biol. Sci*. **2013**, *280*, 20122753.
60. Phan, T.G.; Kapusinszky, B.; Wang, C.; Rose, R.K.; Lipton, H.L.; Delwart, E.L. The fecal viral flora of wild rodents. *PLoS Pathog*. **2011**, *7*, e2218.
61. Balique, F.; Colson, P.; Raoult, D. Tobacco mosaic virus in cigarettes and saliva of smokers. *J. Clin. Virol*. **2012**, *55*, 374–376.
62. Colson, P.; Richet, H.; Desnues, C.; Balique, F.; Moal, V.; Grob, J.J.; Berbis, P.; Lecoq, H.; Harlé, J.R.; Berland, Y.; *et al*. Pepper mild mottle virus, a plant virus associated with specific immune responses, Fever, abdominal pains, and pruritus in humans. *PLoS One* **2010**, *5*, e10041.
63. Middleton, D.J.; Weingartl, H.M. Henipaviruses in their natural animal hosts. *Curr. Top. Microbiol. Immunol*. **2012**, *359*, 105–121.
64. Leroy, E.M.; Kumulungui, B.; Pourrut, X.; Rouquet, P.; Hassanin, A.; Yaba, P.; Délicat, A.; Paweska, J.T.; Gonzalez, J.P; Swanepoel, R. Fruit bats as reservoirs of Ebola virus. *Nature* **2005**, *438*, 575–576.
65. Lau, S.K.; Woo, P.C.; Li, K.S.; Huang, Y.; Tsoi, H.W.; Wong, B.H.; Wong, S.S.; Leung, S.Y.; Chan, K.H.; Yuen, K.Y. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 14040–14045.
66. Johnson, N.; Vos, A.; Freuling, C.; Tordo, N.; Fooks, A.R.; Müller, T. Human rabies due to lyssavirus infection of bat origin. *Vet. Microbiol*. **2010**, *142*, 151–159.
67. Charrel, R.N.; de Lamballerie, X. Zoonotic aspects of arenavirus infections. *Vet. Microbiol*. **2010**, *140*, 213–220.
68. Reusken, C.; Heyman, P. Factors driving hantavirus emergence in Europe. *Curr. Opin. Virol*. **2013**, *3*, 92–99.
69. Zeier, M.; Handermann, M.; Bahr, U.; Rensch, B.; Müller, S.; Kehm, R.; Muranyi, W.; Darai, G. New ecological aspects of hantavirus infection: A change of a paradigm and a challenge of prevention—A review. *Virus Genes* **2005**, *30*, 157–180.
70. Liu, H.; Lu, H.J.; Liu, Z.J.; Jing, J.; Ren, J.Q.; Liu, Y.Y.; Lu, F.; Jin, N.Y. Japanese encephalitis virus in mosquitoes and swine in Yunnan province, China 2009–2010. *Vector Borne Zoonotic Dis*. **2013**, *13*, 41–49.
71. Taubenberger, J.K.; Kash, J.C. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* **2010**, *7*, 440–451.

72. Reperant, L.A.; Kuiken, T.; Osterhaus, A.D. Influenza viruses: From birds to humans. *Hum. Vaccin. Immunother*. **2012**, *8*, 7–16.
73. Ma, W.; Kahn, R.E.; Richt, J.A. The pig as a mixing vessel for influenza viruses: Human and veterinary implications. *J. Mol. Genet. Med*. **2008**, *27*, 158–166.
74. Brauburger, K.; Hume, A.J.; Mühlberger, E.; Olejnik, J. Forty-five years of Marburg virus research. *Viruses* **2012**, *4*, 1878–1927.
75. Leroy, E.M.; Gonzalez, J.P.; Baize, S. Ebola and Marburg haemorrhagic fever viruses: Major scientific advances, but a relatively minor public health threat for Africa. *Clin. Microbiol. Infect*. **2011**, *17*, 964–976.
76. Li, L.; Victoria, J.G.; Wang, C.; Jones, M.; Fellers, G.M.; Kunz, T.H.; Delwart, E. Bat guano virome: Predominance of dietary viruses from insects and plants plus novel mammalian viruses. *J. Virol*. **2010**, *84*, 6955–6965.
77. Donaldson, E.F.; Haskew, A.N.; Gates, J.E.; Huynh, J.; Moore, C.J.; Frieman, M.B. Metagenomic analysis of the viromes of three North American bat species: Viral diversity among different bat species that share a common habitat. *J. Virol*. **2010**, *84*, 13004–13018.
78. Drexler, J.F.; Corman, V.M.; Müller, M.A.; Maganga, G.D.; Vallo, P.; Binger, T.; Gloza-Rausch, F.; Cottontail, V.M.; Rasche, A.; Yordanov, S.; *et al*. Bats host major mammalian paramyxoviruses. *Nat. Commun*. **2012**, *3*, doi:10.1038/ncomms1796.
79. Drexler, J.F.; Geipel, A.; König, A.; Corman, V.M.; van Riel, D.; Leijten, L.M.; Bremer, C.M.; Rasche, A.; Cottontail, V.M.; Maganga, G.D.; *et al*. Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16151–16156.
80. Tse, H.; Tsang, A.K.; Tsoi, H.W.; Leung, A.S.; Ho, C.C.; Lau, S.K.; Woo, P.C.; Yuen, K.Y. Identification of a novel bat papillomavirus by metagenomics. *PLoS One* **2012**, *7*, e43986.
81. Wu, Z.; Ren, X.; Yang, L.; Hu, Y.; Yang, J.; He, G.; Zhang, J.; Dong, J.; Sun, L.; Du, J.; *et al*. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. *J. Virol*. **2012**, *86*, 10999–11012.
82. He, B.; Li, Z.; Yang, F.; Zheng, J.; Feng, Y.; Guo, H.; Li, Y.; Wang, Y.; Su, N.; Zhang, F.; *et al*. Virome profiling of bats from Myanmar by metagenomic analysis of tissue samples reveals more novel Mammalian viruses. *PLoS One* **2013**, *8*, e61950.
83. Baker, K.S.; Leggett, R.M.; Bexfield, N.H.; Alston, M.; Daly, G.; Todd, S.; Tachedjian, M.; Holmes, C.E.; Crameri, S.; Wang, L.F.; *et al*. Metagenomic study of the viruses of African straw-colored fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus. *Virology* **2013**, *441*, 95–106.
84. Meng, X.J. From barnyard to food table: The omnipresence of hepatitis E virus and risk for zoonotic infection and food safety. *Virus Res*. **2011**, *161*, 23–30.
85. Li, L.; Diab, S.; McGraw, S.; Barr, B.; Traslavina, R.; Higgins, R.; Talbot, T.; Blanchard, P.; Rimoldi, G.; Fahsbender, E.; *et al*. Divergent astrovirus associated with neurologic disease in cattle. *Emerg. Infect. Dis*. **2013**, *19*, 1385–1392.
86. Hoffmann, B.; Scheuch, M.; Höper, D.; Jungblut, R.; Holsteg, M.; Schirrmeier, H.; Eschbaumer, M.; Goller, K.V.; Wernike, K.; Fischer, M.; *et al*. Novel orthobunyavirus in Cattle, Europe, 2011. *Emerg. Infect. Dis*. **2012**, *18*, 469–472.

87. Shan, T.; Li, L.; Simmonds, P.; Wang, C.; Moeser, A.; Delwart, E. The fecal virome of pigs on a high-density farm. *J. Virol*. **2011**, *85*, 11697–11708.
88. Masembe, C.; Michuki, G.; Onyango, M.; Rumberia, C.; Norling, M.; Bishop, R.P.; Djikeng, A.; Kemp, S.J.; Orth, A.; Skilton, R.A.; *et al*. Viral metagenomics demonstrates that domestic pigs are a potential reservoir for Ndumu virus. *Virol. J*. **2012**, *9*, doi:10.1186/1743-422X-9-218.
89. Arbovirus Catalog. Available online: https://wwwn.cdc.gov/Arbocat/Default.aspx (accessed on 20 February 2014).
90. Kokernot, R.H.; McIntosh, B.M.; Worth, C.B. Ndumu virus, a hitherto unknown agent, isolated from culicine mosouitoes collected in northern Natal. Union of South Africa. *Am. J. Trop. Med. Hyg*. **1961**, *10*, 383–386.
91. Smith, K.M.; Anthony, S.J.; Switzer, W.M.; Epstein, J.H.; Seimon, T.; Jia, H.; Sanchez, M.D.; Huynh, T.T.; Galland, G.G.; Shapiro, S.E.; *et al*. Zoonotic viruses associated with illegally imported wildlife products. *PLoS One* **2012**, *7*, e29505.
92. Blomström, A.L.; Ståhl, K.; Masembe, C.; Okoth, E.; Okurut, A.R.; Atmnedi, P.; Kemp, S.; Bishop, R.; Belák, S.; Berg, M. Viral metagenomic analysis of bushpigs (*Potamochoerus larvatus*) in Uganda identifies novel variants of Porcine parvovirus 4 and Torque teno sus virus 1 and 2. *Virol. J.* **2012**, *9*, doi:10.1186/1743-422X-9-192.
93. Reuter, G.; Nemes, C.; Boros, A.; Kapusinszky, B.; Delwart, E.; Pankovics, P. Porcine kobuvirus in wild boars (*Sus scrofa*). *Arch. Virol*. **2013**, *158*, 281–282.
94. Yousaf, M.Z.; Qasim, M.; Zia, S.; Khan, M; Ashfaq, U.A.; Khan, S. Rabies molecular virology, diagnosis, prevention and treatment. *Virol. J*. **2012**, *9*, doi:10.1186/1743-422X-9-50.
95. Lackay, S.N.; Kuang, Y.; Fu, Z.F. Rabies in small animals. *Vet. Clin. N. Am.* **2008**, *38*, 851–861.
96. Balenghien, T.; Cardinale, E.; Chevalier, V.; Elissa, N.; Failloux, A.B.; Jean Jose Nipomichene, T.N.; Nicolas, G.; Rakotoharinome, V.M.; Roger, M.; Zumbo, B. Towards a better understanding of Rift Valley fever epidemiology in the south-west of the Indian Ocean. *Vet. Res.* **2013**, *44*, 78.
97. Bird, B.H.; Nichol, S.T. Breaking the chain: Rift Valley fever virus control via livestock vaccination. *Curr. Opin. Virol*. **2012**, *2*, 315–323.
98. Essbauer, S.; Pfeffer, M.; Meyer, H. Zoonotic poxviruses. *Vet. Microbiol*. **2010**, *140*, 229–236.
99. Temmam, S.; Besnard, L.; Andriamandimby, S.F.; Foray, C.; Rasamoelina-Andriamanivo, H.; Héraud, J.M.; Cardinale, E.; Dellagi, K.; Pavio, N.; Pascalis, H.; *et al*. High prevalence of hepatitis E in humans and pigs and evidence of genotype-3 virus in swine, Madagascar. *Am. J. Trop. Med. Hyg*. **2013**, *88*, 329–338.
100. Woolhouse, M.; Gaunt, E. Ecological origins of novel human pathogens. *Crit. Rev. Microbiol*. **2007**, *33*, 231–242.
101. Contributing to One World, One Health: A Strategic Framework for Reducing Risks of Infectious Diseases at the Animal-Human-Ecosystems Interface. Available online: ftp://ftp.fao.org/docrep/fao/011/aj137e/aj137e00.pdf (accessed on 20 February 2014).
102. Moreira, L.A.; Iturbe-Ormaetxe, I.; Jeffery, J.A.; Lu, G.; Pyke, A.T.; Hedges, L.M.; Rocha, B.C.; Hall-Mendelin, S.; Day, A.; Riegler, M.; *et al*. A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and Plasmodium. *Cell* **2009**, *139*, 1268–1278.
103. Bian, G.; Xu, Y.; Lu, P.; Xie, Y.; Xi, Z. The endosymbiotic bacterium Wolbachia induces resistance to dengue virus in *Aedes aegypti*. *PLoS Pathog*. **2010**, *6*, e833.

104. Haagmans, B.L.; Al Dhahiry, S.H.; Reusken, C.B.; Raj, V.S.; Galiano, M.; Myers, R.; Godeke, G.J.; Jonges, M.; Farag, E.; Diab, A.; *et al*. Middle East respiratory syndrome coronavirus in dromedary camels: An outbreak investigation. *Lancet Infect. Dis.* **2014**, *14*, 140–145.
105. Briese, T.; Mishra, N.; Jain, K.; Zalmout, I.S.; Jabado, O.J.; Karesh, W.B.; Daszak, P.; Mohammed, O.B.; Alagaili, A.N.; Lipkin, W.I. Middle east respiratory syndrome coronavirus quasispecies that include homologues of human isolates revealed through whole-genome analysis and virus cultured from dromedary camels in Saudi Arabia. *MBio* **2014**, *5*, e1146-14.
106. Koch, R. Investigations into bacteria: V. The etiology of anthrax, based on the ontogenesis of Bacillus anthracis. *Cohns Beitrage zur Biologie der Pflanzen* **1876**, *2*, 277–310.
107. Fredericks, D.N.; Relman, D.A. Sequence-based identification of microbial pathogens: A reconsideration of Koch's postulates. *Clin. Microbiol. Rev*. **1996**, *9*, 18–33.
108. Mokili, J.L.; Rohwer, F.; Dutilh, B.E. Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol*. **2012**, *2*, 63–77.

V. Objectifs et présentation de la thèse

Dans le contexte des maladies virales émergentes présenté précédemment, et plus précisément des zoonoses virales émergentes, mon travail de thèse a consisté en la description et la caractérisation des communautés virales de différents modèles d'études intervenant dans la transmission zoonotique de virus animaux à l'homme : des arthropodes hématophages étant en contact étroit à la fois avec l'homme et avec l'animal ainsi que des animaux issus de la faune sauvage (singes chassés pour leur viande, rongeurs endémiques) et domestique (bovins).

Etant donné que plus de 70% des virus zoonotiques sont des virus à génome ARN et en l'absence de protocole consensus dans la littérature, la première partie de ce travail a consisté à mettre au point et à valider un protocole de préparation des viromes à ARN qui soit standardisé afin de l'appliquer par la suite à l'ensemble des modèles choisis.

Dans la deuxième partie de cette thèse, je vous présenterai une partie des travaux que j'ai réalisés sur l'étude métagénomique de communautés virales d'arthropodes hématophages collectés au Sénégal lors d'une mission en Novembre 2013. Situé dans le delta du Sine Saloum, les villages ruraux de Dielmo et de Ndiop sont des villages sentinelles de l'Institut Pasteur de Dakar qui y a installé depuis plus de 30 ans deux stations de recherche. Les populations y sont suivies mensuellement pour des programmes de surveillance du paludisme et disposent d'un laboratoire Point-of-Care dans lequel sont réalisés la majorité des tests diagnostics en cas de survenues de pathologies [66]. Cependant, près de la moitié des fièvres développées par les villageois restent sans étiologie connue. Afin de tenter d'expliquer une part de ces fièvres, et dans le but d'identifier de potentiels virus zoonotiques transmissibles par les arthropodes hématophages, j'ai donc caractérisé les communautés virales de différents arthropodes hématophages, à savoir : des tiques molles de rongeurs du genre *Ornithodoros*, des tiques dures de bovins du genre *Rhipicephalus sp.* et *Hyalomma sp.*, des puces de chats et de chiens *Ctenocephalides felis*, des poux de têtes *Pediculus humanus capitis*, des punaises de lits *Cimex hemipterus*, des culicoïdes du genre *Culicoides sp*., et des acariens de rongeurs du genre *Laelaps sp.* Dans la suite de cette thèse, je ne présenterai qu'une partie de ces travaux, portant notamment sur les culicoïdes.

Enfin dans la dernière partie de cette thèse je vous présenterai une étude conduite sur les communautés virales à génomes ADN et ARN de viande de singe originaire de Centrafrique et illégalement importée en France, afin de déterminer le potentiel risque zoonotique d'introduction de pathogènes par ce type de trafic de viande de brousse.

Mise au point d'une méthode de préparation des viromes ARN

Article n°2: "Host-associated metagenomics: a guide to generating infectious RNA viromes."

Host-associated metagenomics: a guide to generating infectious RNA viromes.

Sarah Temmam[1,¶], Sonia Monteil-Bouchard[1,¶], Catherine Robert[1], Hervé Pascalis[2], Caroline Michelle[1], Priscilla Jardot[1], Rémi Charrel[3], Didier Raoult[1], Christelle Desnues[1,*].

Published in: PLoS One. 2015;10(10):e0139810. doi: 0.1371/journal.pone.0139810.

[1] Unité de Recherche sur les Maladies Infectieuses Tropicales Emergentes (URMITE) UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, Marseille, France.

[2] Centre de Recherche et de Veille sur les maladies émergentes dans l'Océan Indien (CRVOI), IRD La Réunion, Plateforme de Recherche CYROI, La Réunion, France.

[3] Emergence des Pathologies Virales (EPV), IRD 190, EHESP, Aix-Marseille Université, Marseille, France.

[¶] These authors have contributed equally to this work.

[*] Corresponding author: christelle.desnues@univ-amu.fr

Préambule à l'article "Host-associated metagenomics: a guide to generating infectious RNA viromes"

La majorité des virus zoonotiques pouvant présenter un risque pour la santé humaine sont des virus à génome ARN [6]. Du fait du faible taux de conservation de leur génome dans leur cycle réplicatif, les virus à ARN sont plus facilement capables de transgresser la barrière d'espèce entre l'animal et l'homme, et ainsi présenter un risque de transmission zoonotique.

Depuis l'avènement des techniques de séquençage à haut débit (NGS), la métagénomique a permis de décrire les communautés microbiennes de différents environnements (écologie microbienne) [67-71] ou d'identifier des micro-organismes responsables de maladies infectieuses inexpliquées, tant en médecine humaine [56,72] qu'en médecine vétérinaire [57,73]. Jusqu'à récemment, les études de viromes ciblaient principalement les communautés de virus à génome ADN. De plus, bien que de nombreux protocoles aient été publiés [74-82], il n'existait pas de méthode standardisée de préparation des viromes, ce qui rendait compliquée la comparaison entre données de séquençage.

Etant donné le contexte d'étude, il a fallu dans un premier temps mettre au point un protocole simple de préparation et d'analyse des communautés virales à génome ARN à partir de matrices complexes. Notre protocole diffère en outre des méthodes classiquement utilisées en métagénomique virale car, contrairement aux méthodes utilisant le chlorure de césium (CsCl), l'utilisation du sucrose préserve l'intégrité et l'infectivité des particules virales, est moins toxique pour les cellules et permet donc des isolements viraux ultérieurs à partir des particules virales purifiées et enrichies. Par ailleurs, notre protocole permet de réduire de façon notable la quantité de séquences d'hôte afin de profiter pleinement de la profondeur de séquençage offerte par le NGS et, ainsi, avoir la capacité de détecter des virus présents en faible quantité dans des échantillons. Nous avons donc dans un premier temps évalué sur un pool représentatif de virus à génome ARN dans des conditions « non biologiques » (surnageants viraux) et « biologiques » (broyats de poux supplémentés en surnageants de virus) différentes étapes de purification virale (filtration, ultracentrifugation, etc.), de reverse transcription et d'amplifications séquence-indépendante (SIA) en matière de (i) maintien de l'infectivité des virions (ii)

rendement de récupération des virions (iii) élimination des acides nucléiques de l'hôte. Une fois le protocole mis au point, nous l'avons validé en séquençant le virome ARN de poux gorgés de sang artificiellement contaminé par des virus à ARN. Le séquençage NGS de ce virome ARN artificiel a non seulement permis de reconstituer le génome complet des virus les plus abondemment absorbés par les poux mais aussi d'observer des différences de couverture de génome des virus à génome ARN segmenté en fonction des méthodes de SIA testées, ce qui reflète les biais d'amplification de ces méthodes.

En conclusion, le protocole développé au cours de ce travail permet non seulement une bonne élimination des acides nucléiques de l'hôte (étape limitante dans ce type d'approches) et l'optimisation des profondeurs de séquençage, mais aussi et surtout l'isolement futur de particules virales, ouvrant ainsi des perspectives importantes de caractérisation virologiques et moléculaires des virus détectés par séquençage.

Ce protocole a été employé par la suite pour l'étude des modèles arthropode et faune sauvage.

PLOS | ONE

RESEARCH ARTICLE

Host-Associated Metagenomics: A Guide to Generating Infectious RNA Viromes

Sarah Temmam[1☯], Sonia Monteil-Bouchard[1☯], Catherine Robert[1], Hervé Pascalis[2], Caroline Michelle[1], Priscilla Jardot[1], Rémi Charrel[3], Didier Raoult[1], Christelle Desnues[1]*

1 Unité de Recherche sur les Maladies Infectieuses Tropicales Emergentes (URMITE) UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, Marseille, France, **2** Centre de Recherche et de Veille sur les maladies émergentes dans l'Océan Indien (CRVOI), IRD La Réunion, Plateforme de Recherche CYROI, La Réunion, France, **3** Emergence des Pathologies Virales (EPV), IRD 190, EHESP, Aix-Marseille Université, Marseille, France

☯ These authors contributed equally to this work.
* christelle.desnues@univ-amu.fr

OPEN ACCESS

Citation: Temmam S, Monteil-Bouchard S, Robert C, Pascalis H, Michelle C, Jardot P, et al. (2015) Host-Associated Metagenomics: A Guide to Generating Infectious RNA Viromes. PLoS ONE 10(10): e0139810. doi:10.1371/journal.pone.0139810

Editor: Patrick Tang, Sidra Medical and Research Center, QATAR

Received: December 15, 2014

Accepted: September 17, 2015

Published: October 2, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was conducted under the frame of an Agence Nationale de la Recherche Young Investigator fellowship (ANR-13-JSV6-0004) awarded to Christelle Desnues. The work of Rémi N. Charrel was conducted under the frame of EurNegVec COST Action TD1303. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Background

Metagenomic analyses have been widely used in the last decade to describe viral communities in various environments or to identify the etiology of human, animal, and plant pathologies. Here, we present a simple and standardized protocol that allows for the purification and sequencing of RNA viromes from complex biological samples with an important reduction of host DNA and RNA contaminants, while preserving the infectivity of viral particles.

Principal Findings

We evaluated different viral purification steps, random reverse transcriptions and sequence-independent amplifications of a pool of representative RNA viruses. Viruses remained infectious after the purification process. We then validated the protocol by sequencing the RNA virome of human body lice engorged *in vitro* with artificially contaminated human blood. The full genomes of the most abundant viruses absorbed by the lice during the blood meal were successfully sequenced. Interestingly, random amplifications differed in the genome coverage of segmented RNA viruses. Moreover, the majority of reads were taxonomically identified, and only 7–15% of all reads were classified as "unknown", depending on the random amplification method.

Conclusion

The protocol reported here could easily be applied to generate RNA viral metagenomes from complex biological samples of different origins. Our protocol allows further virological characterizations of the described viral communities because it preserves the infectivity of viral particles and allows for the isolation of viruses.

Introduction

Viruses are the most ubiquitous and abundant biological entities on Earth [1]. They infect all other biological entities (such as bacteria, archaea, plants, arthropods, and mammals) living in diverse environments (such as soil, water, air, and multi-cellular organisms). Viruses influence other organisms directly by modulating their hosts' survival via host mortality or horizontal gene transfer, or indirectly via the diversion of the host metabolic pathways during viral replication. To study viral diversity within hosts and environments, recent techniques known as viral metagenomics have emerged. Primarily based on Sequence-Independent Amplification (SIA) techniques and followed by Next-Generation Sequencing (NGS) technologies, viral metagenomics allows the description of viral communities within a complex environment without any prior knowledge of their nature. For example, in diagnostic virology, viral metagenomics have been used to identify causative viral agents of disease conditions in human [2–4] and veterinary medicine [5–7], as well as in plant [8,9] and arthropod diseases [10,11]. Virome analyses have also been conducted to describe the baseline viral diversity in healthy human conditions prior to studying the viral flora of pathologic conditions [12]. In viral ecology, metagenomics have been used to describe viral communities of diverse environments, including coastal seawater and sediment, soil, hotsprings, lakes, sewage, and air [13–17].

Viral metagenomic analyses of complex environments usually require pre-treatment steps, such as viral purification and nucleic acid enrichment, before sequencing. Several physical characteristics of viral particles enable viral purification (e.g., capsid durability), but the wide variety of viruses' biological characteristics cause difficulties in developing a standardized protocol compatible with a broad range of particle sizes, shapes, densities, and genome types [18]. Usually, virome preparation is based on dead-end or tangential flow filtration, and nuclease digestion of non-protected viral and host cells. Then, PolyEthylene Glycol (PEG) precipitation or ultracentrifugation is eventually used, followed by nucleic acid extraction [19,20]. Although there are as many protocols to generate viral metagenomes as published metagenomic studies, the vast majority are aimed at purifying viruses from their complex matrices. This strategy, known as "Particle-Associated nucleic acid amplification", is aimed at insolating intact (*i.e.*, infectious) viral particles from their environment, protected from the action of nucleases [21]. Alternative steps exist within this general protocol, depending on the origin of the matrices. For example, marine biological samples such as coral tissue require chloroform homogenization of the matrix before viral purification [22].

During the last decade, several standardized protocols for generating DNA viromes from various environments have been described [18–20], but such standardization has not been reached so far for RNA viruses. Here, we present a simple protocol for the purification and sequencing of RNA viromes from host-associated biological samples of various origins. Our protocol preserves the infectivity of viral particles and allows for further applications. This protocol has been evaluated and validated by sequencing the RNA virome of body lice that were engorged *in vitro* with artificially contaminated human blood. We decided to use artificially engorged body lice as a model for host-associated metagenomics because of the ease of sampling and handling compared to human or animal specimens which require special permissions. Additionally, this system is convenient because arthropods are complex organisms in which viral, bacterial, and parasitic communities coexist.

Materials and Methods

Viral strains

To optimize the purification steps of the protocol, a representative panel of RNA viruses was chosen based on size, density, the presence of an envelope, and genetic composition. The latter

Table 1. Characteristics of reference viruses.

	YF	CPX	CoxB3	H3N2	MS2	T4
Particle size (nm)	50	50–300	30	80–120	26	45–230 x 825
Density (g.cm^{-3})	1.19 (sucrose)	1.17–1.18 (sucrose)	1.33–1.45 (CsCl)	1.19 (sucrose)	1.46 (CsCl)	1.50 (CsCl)
Sensitivity to CsCl [32–35]	yes	yes	no	yes	no	no
Envelope	yes	yes	no	yes	no	no
Capsid	icosahedral	helical	icosahedral	helical	icosahedral	icosahedral with tail
Genome organization	linear ssRNA, positive sense, non-segmented, 10.8 kb	linear ssRNA, ambisense, segmented (N = 2), 11 kb	linear ssRNA, positive sense, non-segmented, 7.4 kb	linear ssRNA, negative sense, segmented (N = 8), 13.5 kb	linear ssRNA, positive sense, non-segmented, 4 kb	linear dsDNA, 169 kb
Viral family	*Flaviviridae*	*Arenaviridae*	*Picornaviridae*	*Orthomyxoviridae*	*Leviviridae*	*Myoviridae*

doi:10.1371/journal.pone.0139810.t001

category included whether the virus had a positive or negative strand RNA genome, and whether it contained a segmented genome. Viruses chosen were: *Yellow fever virus* 17D vaccine strain (YF), *Coxsackievirus B3* strain 2679 (CoxB3), *Influenza A/H3N2* strain Marseille/04046111/2011 (H3N2), *Cupixi arenavirus* (CPX), and MS2 bacteriophage (MS2). Additional DNA bacteriophage (T4 phage) was added to the panel to verify the efficiency of the protocol for DNA viruses and to quantify the remaining contamination of DNA in the RNA fraction of the virome (Table 1).

YF, CPX, and CoxB3 viral strains were propagated in Vero cells while H3N2 was propagated in MDCK cells. After 4 passages, viral supernatants were nuclease-treated with 0.5 U RNAse A (Roche Diagnostics, Meylan, France) and 4 μg (1 U) of DNAse I (Sigma-Aldrich, Saint-Quentin Fallavier, France) per mL of supernatant, then precipitated with 10% PEG 8000 (Sigma-Aldrich, Saint-Quentin Fallavier, France) and 300 mM NaCl (Sigma-Aldrich, Saint-Quentin Fallavier, France) overnight at +4°C. After centrifugation at 12 000 g for 30 min at +4°C, the pellet was resuspended in 2 mL of Phosphate Buffer Saline solution (PBS), aliquoted and stored at -80°C until further use. Viral loads were estimated in Plaque-Forming Units (PFUs).

MS2 and T4 bacteriophages were purchased from the Leibniz Institute DSMZ–German Collection of Microorganisms and Cell Cultures (LGC Standards S.a.r.l., Molsheim, France) and propagated according to the manufacturer's protocol. Bacteriophage suspensions were centrifuged at 1 500 g for 10 min, and then filtrated through 0.45 μm filters (Merck Millipore, Molsheim, France). Viral suspensions were digested with 30 U Turbo DNAse (Life technologies, Saint Aubin, France) and 25 U RNase A (Roche Diagnostics, Meylan, France) in Turbo DNase buffer at 37°C for 1 hour, then precipitated as described above. Viral loads were estimated in PFUs.

Real-time PCR

To evaluate the efficiency of treatments for virome preparation, real-time PCRs were conducted on YF, CPX, CoxB3, H3N2, MS2, and T4 targets using the SuperScript® III Platinum One-Step RT-PCR (Life Technologies, Saint Aubin, France). The QuantiTect SYBR® Green PCR/RT-PCR Kit (Qiagen, Courtaboeuf, France) was used for host-contaminating 18S DNA and RNA. Each of these reagents was used according to the manufacturers' protocols. All quantitative real-time PCR (qPCR) and reverse transcription real-time PCR (qRT-PCR) reactions were performed in a CFX96 thermocycler (Biorad, Marnes-la-Coquette, France). YF, CoxB3,

H3N2, MS2, T4 and 18S primers and probes are published in [23–27] and presented in S1 Table. CPX primers and probes were designed for this study (S1 Table). Additional primers were designed based on H3N2 and YF sequences detected in the metagenomes to verify the absence of cross-contaminations during the construction of the libraries (S1 Table). qPCR were performed using the QuantiTect SYBR® Green PCR Kit (Qiagen, Courtaboeuf, France) and according to the manufacturers' protocol.

Ethical statement

The maintenance of a laboratory colony of *Pediculus humanus corporis* lice has been approved by the Institutional Animal Care and Use Committee of the Faculty of Medicine at Aix-Marseille University, France.

Human blood collected from healthy donors was obtained from the Etablissement Français du Sang (EFS, Marseille, France). In accordance with EFS standardized procedures for blood donation, informed consent was obtained from healthy volunteers, and personal data relative to blood donors were rendered anonymous at the time of blood donation and before blood transfer to our research lab.

Body lice blood meal

Forty body lice were kindly provided by Jean-Michel Berenger (entomologist at "Unité de Recherche sur les Maladies Infectieuses Tropicales Emergentes", CNRS7278, Marseille, France) and divided into two pools. Twenty lice were fed *in vitro* using artificially contaminated hemolyzed human blood, an artificial Parafilm membrane (Sigma-Aldrich, Saint-Quentin Fallavier, France) and the Hemotek system (Hemotek Ltd, Accrington, United Kingdom), an electric heating element which maintains the temperature of the blood meal, as described by Sangaré *et al.*[28]. Briefly, 5.0 x 10^{10} PFU of MS2, 1.5 x 10^{9} PFU of T4, 3.07 x 10^{5} PFU of H3N2, 1.5 x 10^{4} PFU of CoxB3, 1.0 x 10^{3} PFU of YF and 1.19 x 10^{2} PFU of CPX were added to pre-heated hemolyzed human blood to a final volume of 1.5mL, placed on the Hemotek chamber and maintained at 37°C for 30 min. During this time, lice took their blood meal. After 30 min, only 11 lice were engorged. Immediately after external decontamination of the lice, consisting of sequential washes with 10% bleach in sterile water, 70% ethanol and a final wash in sterile water, all 20 lice were stored at -80°C until further analysis. Hereafter these lice will be named "engorged lice".

The second pool of twenty body lice was not fed with blood and was used as a positive control. Briefly, after decontamination, lice were spiked using the same concentrations of each reference virus as the ones used for blood-fed lice, and treated like the engorged ones. Hereafter these lice will be named "spiked lice".

Viral purification and concentration from complex samples and evaluation of the infectivity

Solid samples (*i.e.*, "spiked" and "engorged" lice) were first homogenized in 2 mL of 0.02 μm-filtrated EMEM medium (Life Technologies, Saint Aubin, France) using the TissueLyser homogenizer (Qiagen, Courtaboeuf, France) and two 3 mm tungsten beads at 25 Hz for 2 min. Solid and liquid complex samples were then sequentially centrifuged at +4°C at 300 rpm for 30 min and 10 000 rpm for 15 min. The resulting samples were then filtered through a 0.45 μm filter (Millipore, Molsheim, France) to remove any cellular debris and bacteria.

To eliminate host DNA/RNA and free nucleic acids, 20 U Exonuclease I (New England Biolabs, Évry, France), 25 U Benzonase® (Merck Millipore, Molsheim, France), 25 U RNase A (Roche Diagnostics, Meylan, France), 20 U Turbo DNase (Life Technologies, Saint Aubin,

France) and 10 μL of 10X Turbo DNase buffer were added to the clarified supernatant and incubated at 37°C for 1 hour. A total of 100 μL of the resulting supernatant was harvested to assess the concentration of Virus-Like-Particles (VLP) by fluorescence microscopy, as previously described by Thurber *et al.* [19]. All fluorescence images were acquired with a Leica SP5 inverted confocal microscope with 4 lasers, a 100 X objective and a numerical aperture of 1.4.

The resulting suspension was then deposited onto a discontinuous sucrose gradient consisting of 800 μL of a 0.02 μm-filtered layer of 66% sucrose in EMEM and 2.7 mL of a 0.02 μm-filtered layer of 30% sucrose in PBS, and ultracentrifuged in an MLS50 Beckman-Coulter rotor at 130 000 g for 2 hours at +4°C. The viral fraction was harvested from the interface between the 66% and the 30% sucrose layers using a 23G needle. The same procedure was applied for CsCl gradient ultracentrifugation to compare the maintenance of the infectivity of viral particles with sucrose gradient. Briefly, the viral suspension was deposited onto a CsCl gradient composed of 750 μL of 1.2 g/mL, 1.5 g/mL and 1.7 g/mL CsCl layers and ultracentrifuged in an MLS50 Beckman-Coulter rotor at 130 000 g for 2 hours at +4°C. The viral fraction was harvested between the 1.5 and the 1.2 g/mL CsCl layers using a 23G needle.

An aliquot of 100 μL of sucrose viral fraction of engorged and spiked lice was recovered to assess the VLP concentration by fluorescence microscopy as described above, 2 x 100 μL aliquots were immediately stored at -80°C for further viral isolation, and the resulting supernatant was used to extract nucleic acids.

Vero, MDCK and *E. coli* cells were used to determine the viral load of YF, CPX, CoxB3, H3N2, MS2, and T4 respectively, before and after the purification process by the lysis plaques or by the $TCID_{50}$ methods.

Nucleic acid extraction

Three nucleic acid extraction processes were evaluated: Trizol LS® (Life Technologies, Saint Aubin, France), the QIAmp viral RNA mini kit (Qiagen, Courtaboeuf, France), and the High Pure viral nucleic acid kit (Roche Diagnostics, Meylan, France) according to the manufacturer's protocols. Nucleic acids were eluted in 20 μL for the QIAmp and High Pure kits and 100 μL for the Trizol LS extraction. After Trizol LS extraction, to remove any traces of phenol/chloroform that could interfere with subsequent enzymatic reactions, the RNeasy MinElute Cleanup kit (Qiagen, Courtaboeuf, France) and Agencourt AMPure beads (Beckman-Coulter, Villepinte, France) were used for the RNA and DNA fractions, respectively. Samples were then eluted in a final volume of 20 μL.

DNA fractions were used to assess the remaining host DNA contamination level. For total RNA preservation, 40 U RNase OUT (Life Technologies, Saint Aubin, France) was added.

RNA integrity and quantification

RNA integrity was checked on an RNA6000 Pico chip (Agilent Technologies, Les Ulis, France) according to the manufacturer's protocol, and analyzed on the Agilent 2100 Bioanalyzer.

RNA concentration was estimated with the Quanti-it Ribogreen kit (Life Technologies, Saint Aubin, France) according to the manufacturer's recommendations, and fluorescence was quantified with the Tecan GENios fluorometer.

RNA processing

Total RNA was processed with three different random reverse transcriptions, as previously described by Froussard *et al.* in 1992 [29], Wang *et al.* in 2002 [30], and Victoria *et al.* in 2008 [31]. Briefly, reverse transcription (RT) was conducted on 9 μL of RNA (Trizol LS extraction) or 9 μL of total nucleic acids (QIAamp and High Pure extractions) using the Superscript III

Reverse transcriptase (Life Technologies, Saint Aubin, France) and the tagged-random hexamers described in the Froussard, Wang, and Victoria studies [29–31]. The thermal profile was as follows: 25°C – 5 min, 35°C – 15 min, 55°C – 30 min, and 94°C – 2 min.

Single-stranded DNA (ssDNA) was subsequently used as a template for the Klenow reaction to obtain double-stranded DNA (dsDNA). Briefly, 20 μL of ssDNA was mixed with 8 U of Klenow (Life Technologies, Saint Aubin, France) and 1 μL of 10 mM dNTP to a final volume of 30 μL. Thermal profiles were used as previously described by Froussard [29], Wang [30], and Victoria [31].

The resulting dsDNA was purified twice with the Agencourt AMPure beads (Beckman-Coulter, Villepinte, France), eluted in a final volume of 20 μL, and quantified with the Quanti-it Picogreen reagent (Life Technologies, Saint Aubin, France). Size distribution was then checked on a DNA7500 chip (Agilent Technologies, Les Ulis, France) and analyzed on the Agilent 2100 Bioanalyzer.

RNA Sequence-Independent Amplification

For RNA sequence-independent amplification (SIA), dsDNA generated by random RT-Klenow reactions was used in random PCR. Briefly, 5 μL of dsDNA were mixed with 2.5 U of Long Amp Taq DNA polymerase (New England Biolabs, Évry, France) to a final volume of 25 μL and then randomly amplified according to Froussard [29], Wang [30], and Victoria [31]. Hereafter, the 3 viral metagenomes will be named "Froussard," "Wang", or "Victoria", according to the sequence-independent amplification method used to generate it.

Amplification products were twice purified with Agencourt AMPure beads (Beckman-Coulter, Villepinte, France) according to the manufacturer's protocol and eluted to a final volume of 15 μL. The concentration of dsDNA was estimated with the Quanti-it Picogreen kit (Life Technologies, Saint Aubin, France). Amplified products were analyzed on a DNA7500 chip (Agilent Technologies, Les Ulis, France).

Illumina MiSeq sample preparation and processing

RNA metagenomes of engorged and spiked lice randomly amplified by "Froussard", "Wang", or "Victoria" random PCR were sequenced with the MiSeq Technology using paired-end and barcode strategies according to the Nextera XT library kit in a 2 x 300 bp format (Illumina Inc., San Diego CA 92121, USA). Briefly, cDNA was quantified by Qubit® with the High Sensitivity kit (Life Technologies, Carlsbad, CA, USA) and dilutions were performed to a final quantity of 1 ng of cDNA as the input. The "tagmentation" step fragmented the cDNA, and then limited cycle PCR amplification completed the tag adapters and introduced dual-index barcodes. After purification with AMPure beads (Life Technologies, Carlsbad, CA, USA), the libraries were normalized on specific beads according to the Nextera XT protocol (Illumina Inc., San Diego CA 92121, USA). Normalized libraries were sequenced along with 11 other projects for a total of 18 projects. Automated cluster generation and paired-end sequencing with dual-index reads were performed in a single run of 2 x 300 bp read length.

Sequence processing and virus genome identification

Paired reads were imported into the CLC Genomics Workbench 6.0.1 program (CLC Bio, Aarhus, Denmark) with importing parameters including minimum and maximum distances set at 50 and 400, respectively. Raw Illumina reads were first trimmed according to their quality score (Illumina pipeline 1.8 and later), their length (reads $<$ 50 nt long were discarded) and according to the primers used for random PCR.

Reads were then mapped onto reference genomes using the CLC Genomics Workbench 6.0.1 program (CLC Bio, Aarhus, Denmark) with mapping parameters that included a minimal length fraction of 0.5, a minimal similarity fraction of 0.8, a mismatch cost of 2, and an insertion/deletion cost of 3. The GenBank accession numbers of YF, CPX, CoxB3, H3N2, MS2, and T4 reference genomes used for mapping are JN628279, AY216519, AY896763, CY114421, V00642, and AF158101, respectively.

Un-mapped reads were *de novo* assembled into contigs using the CLC Genomics Workbench 6.0.1 assembler with stringent assembly parameters: minimal length fraction of 0.9, minimal similarity fraction of 0.75, word size of 2, minimal contig length of 200 bp, mismatch cost of 2 and insertion/deletion cost of 3. Contigs and singletons of un-mapped reads were compared to the NCBI nucleotide database using the BlastN program, with a minimum coverage of 50%, a minimum identity of 50%, and an E-value $< 10^{-5}$. Reads and contigs having no significant hits according to the criteria were classified as "unknown". Table 2 presents the data from the Illumina MiSeq sequencing of engorged and spiked lice according to the random PCR method.

Results

Viral particles purification and concentration

When studying viral communities of complex environmental or biological samples, diverse eukaryote, prokaryote, and archaea communities may interfere with the analysis. As a consequence, several pretreatments are required to purify the virome from these contaminants, mainly based on the difference in size and density of viral particles compared to those of eukaryotic and prokaryotic cells. A general overview of the process developed here is presented in Fig 1. To assess the loss of contaminants during the whole purification process, qPCR and qRT-PCR targeting the 18S DNA and RNA, respectively, were performed before and after each step of the process. Similarly, control real-time RT-PCR and PCR targeting the reference RNA and DNA viruses, respectively, were conducted to verify their presence after each step of the purification process.

The first step was a 0.45 μm filtration followed by nuclease digestion. The filtration eliminated 0.6 $\log_{10}$ and 0.4 $\log_{10}$ of 18S DNA and RNA, respectively, and the nuclease treatment eliminated an additional 1.8 $\log_{10}$ and 2.1 $\log_{10}$ of 18S DNA and RNA, respectively. Interestingly, the use of the nuclease cocktail, *i.e.*, Turbo DNAse–RNAse–Benzonase–Exonuclease I, resulted in better elimination of 18S DNA and RNA contamination than the use of Turbo DNAse and RNAse in combination. Indeed, the Ct value of the 18S DNA contamination was 36.38 with the use of Turbo DNAse–RNAse but negative with the use of the nuclease cocktail. Moreover, the 18S RNA contamination was eliminated more effectively with the nuclease cocktail than with the use of the Turbo DNAse–RNAse combination (Ct = 34.85 vs 32.67, respectively). No difference was observed in the 18S contamination when samples were incubated with the nuclease cocktail for between one and two hours (ΔCt between one and two hours of nuclease digestion of 0.48 for 18S DNA and 0.29 for 18S RNA). The reference DNA and RNA viruses were not affected by filtration and nuclease treatments (with ΔCt values ranging from 0.15 to 1.33 depending on the virus).

To increase the elimination of 18S DNA and RNA contamination via ultracentrifugation, two discontinuous sucrose gradient formulations were tested (66%-20% and 66%-30%). For a given centrifugation speed and time, the 66%-30% sucrose gradient was more effective at removing the 18S contamination, with less than one $\log_{10}$ of loss of viral load for the reference DNA and RNA viruses. There was no major difference observed between the two sucrose gradients for the removal of 18S DNA contamination (-3.37 $\log_{10}$ for the 66%-30% gradient and

Table 2. Data of the Illumina MiSeq sequencing of engorged and spiked lice according to the random PCR method.

ENGORGED LICE	"Froussard"	"Wang"	"Victoria"
Total number of reads (R1+R2)	**1,594,552**	**1,673,532**	**1,181,724**
Total mapped reads	721,520	697,950	716,305
Un-mapped reads	873,032	975,582	465,419
Contigs of un-mapped reads	1,262	473	335
Singletons of un-mapped reads	270,157	226,773	130,193
Total assigned reads	**1,360,372**	**763,442**	**1,097,434**
Viral reads:	**728,700**	**749,544**	**738,929**
MS2 reads	617,744	595,402	678,268
CoxB3 reads	107,252	150,502	60,619
H3N2 reads	181	3,552	18
T4 reads	8	0	2
YF reads	0	0	0
CPX reads	0	0	0
Other viral reads:	**3,515**	**88**	**22**
Podoviridae (T7 and T3 phages)	3,477	56	6
Leviviridae (BO1 phage)	1	0	0
Siphoviridae (*Staphylococcus* phage phiETA2)	0	1	0
Retroviridae (Squirrel monkey retrovirus-H)	37	31	11
Mimiviridae (Mimivirus and Mamavirus)	0	0	5
Eukaryote reads	**618,992**	**657,100**	**355,448**
Prokaryote reads	**9,165**	**7,239**	**3,035**
Unknown reads	**234,180**	**910,090**	**84,290**
SPIKED LICE	**"Froussard"**	**"Wang"**	**"Victoria"**
Total number of reads (R1+R2)	**1,460,831**	**1,917,827**	**1,673,347**
Total mapped reads	1,442,668	1,639,590	1,636,226
Un-mapped reads	18,163	278,237	37,121
Contigs of un-mapped reads	381	275	145
Singletons of un-mapped reads	6,208	26,988	873
Total assigned reads	**1,457,577**	**1,629,308**	**1,654,195**
Viral reads:	**1,447,878**	**1,609,176**	**1,644,755**
MS2 reads	1,433,365	1,528,259	1,622,373
CoxB3 reads	6,848	77,161	21,488
H3N2 reads	2,439	3,486	749
YF reads	14	64	4
CPX reads	2	2	2
Other viral reads:	**1,627**	**204**	**139**
Podoviridae (T7 and T3 phages)	899	734	65
Microviridae (phiX174 phage)	12	4	7
Retroviridae (Squirrel monkey retrovirus-H)	713	4,432	67
Mimiviridae (Hirudovirus Sangsue, Samba virus)	2	3	0
Eukaryote reads	**4,742**	**16,543**	**7,849**
Prokaryote reads	**4,957**	**3,589**	**1,591**
Unknown reads	**3,254**	**288,519**	**19,152**

doi:10.1371/journal.pone.0139810.t002

-2.98 $\log_{10}$ for the 66%-20% gradient); however, the difference was important for the 18S RNA contamination, as 0.20 $\log_{10}$ of 18S RNA was removed with the 66%-20% sucrose gradient versus more than 3 $\log_{10}$ of 18S RNA using the 66%-30% sucrose gradient. Additionally, the

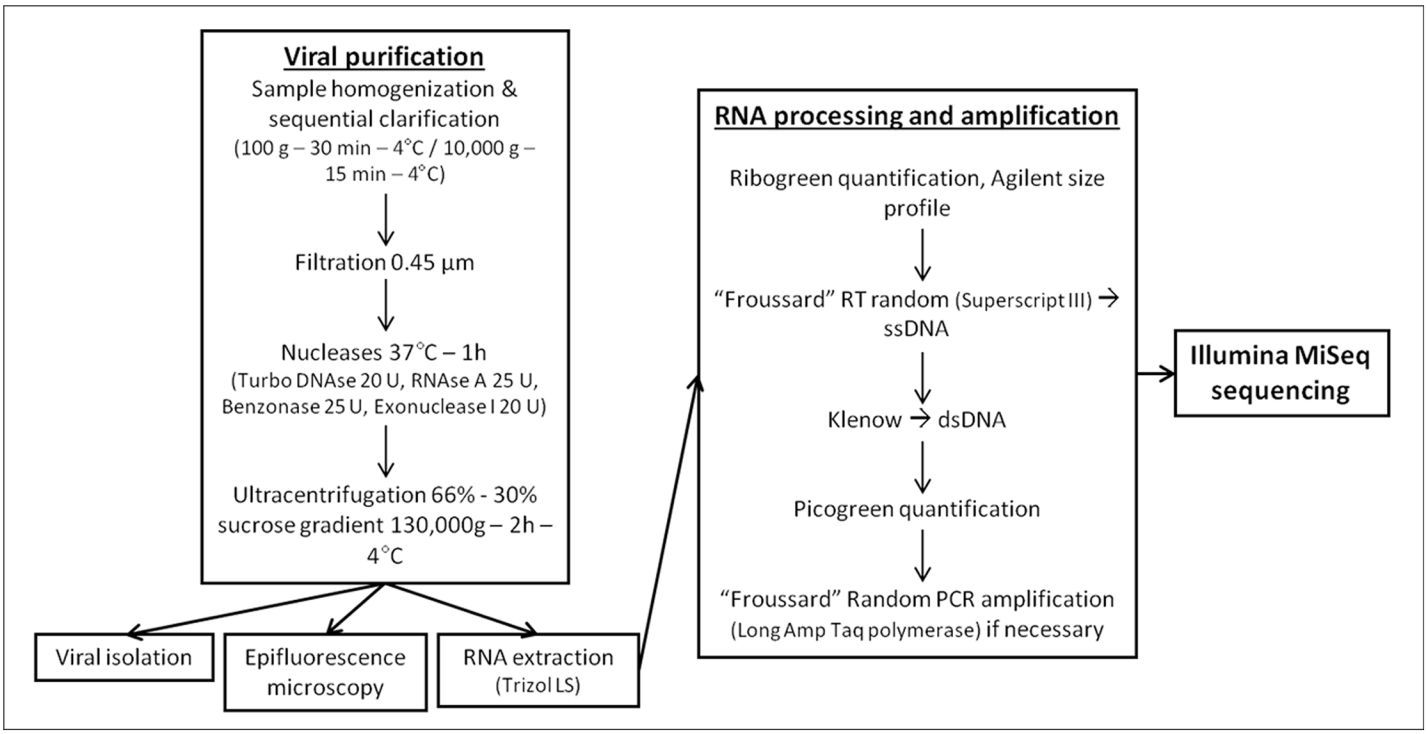


Fig 1. General overview of the protocol.

doi:10.1371/journal.pone.0139810.g001

reference viruses were not affected by the increase in sucrose density (-0.46 $\log_{10}$ for YF, -0.00 $\log_{10}$ for CPX, -0.18 $\log_{10}$ for CoxB3, -0.67 $\log_{10}$ for H3N2, and -0.00 $\log_{10}$ for T4 when increasing the sucrose density from 66%-20% to 66%-30%).

Fig 2A presents the RNA profile of an EMEM sample spiked with the reference DNA and RNA viruses pretreated by a 0.45 µm filtration followed by nuclease digestion and

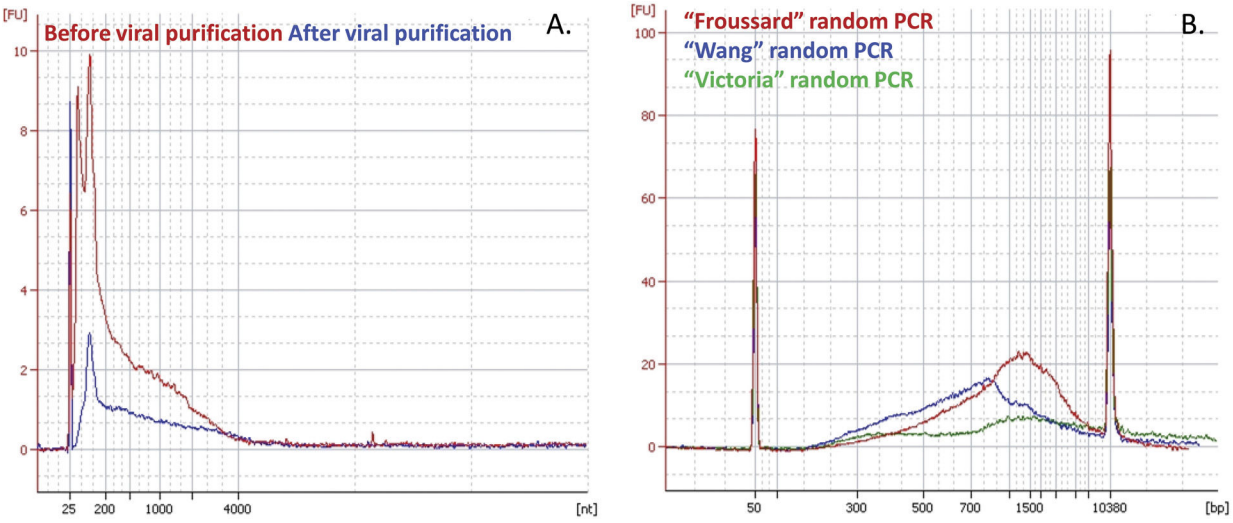


Fig 2. Nucleic acid profiles analyzed on a 2100 Expert Agilent Analyzer. A. RNA profile of an EMEM sample spiked with the reference viruses before (red) and after (blue) viral purification, run on a Pico RNA chip. B. DNA profile of a dsDNA sample (originated from RNA) amplified either with Froussard (red), Wang (blue) or Victoria (green) random PCR, run on a DNA7500 chip.

doi:10.1371/journal.pone.0139810.g002

ultracentrifugation, before and after viral purification. Although a clear decrease in the total amount of RNA was observed, which reflected the loss of 18S RNA, small RNAs (between 100 and 200 nt) were detected. This could suggest the recovery of digested host RNAs, despite the nuclease treatment and ultracentrifugation step (Fig 2A).

Further optimizations, primarily of the ultracentrifugation step, were tested to improve the purification process but did not achieve satisfying results. Briefly, (i) an increase in centrifugation speed from 130 000 g to 160 000 g did not improve the recovery of reference viruses (ΔCt ranging from 0.26 to 0.91, depending on the reference virus), but did result in the recuperation of more contaminating 18S RNA (loss of total 18S RNA estimated at 6.78 Ct when ultracentrifuged at 130 000 g and only at 0.56 Ct when ultracentrifuged at 160 000 g); (ii) the addition of a final pelleting ultracentrifugation step after the discontinuous gradient ultracentrifugation step resulted in an important loss of infectivity of reference viruses and their increased sensitivity to nucleases; (iii) the addition of a final nuclease digestion step after ultracentrifugation, to remove any remaining 18S DNA and RNA contaminant, resulted in a major loss of 18S DNA and RNA contamination (negative PCR for 18S DNA and RNA after a final nuclease treatment but a resulting Ct = 32.91 and Ct = 30.44, respectively, if no final nuclease step was added). However, we observed a concomitant loss of several viral reference genomes (negative PCR for CoxB3 and H3N2 if a final nuclease digestion is performed whereas Ct values of 22.20 and 29.42, respectively, with no final nuclease treatment).

Maintaining viral infectivity

For the ultracentrifugation step, the choice of sucrose gradient instead of cesium chloride (CsCl), usually used to generate viral metagenomes, was motivated by the wish to preserve the infectivity of viral particles. Indeed, the use of CsCl gradients alters the integrity of enveloped virions and affects their infectivity (Table 1) [32–35]. Consequently, we tested the infectivity of recovered viral particles after all purification processes by re-isolating viruses after the purification steps, either in 66%-30% sucrose or in 1.7–1.5–1.2 g/mL CsCl gradients. The resulting cytopathic effects are presented in S1 Fig. After two days of culture, sucrose and CsCl revealed a toxic effect on Vero cells, with higher cytotoxicity for CsCl than for sucrose. Indeed, 1/10 and 1/100 dilutions were necessary to inhibit cell toxicity for sucrose and CsCl, respectively (S1A Fig).

For non-enveloped viruses (*i.e.* MS2 and CoxB3), no difference was observed between the viral titer obtained after either sucrose or CsCl purification, suggesting that the infectivity of virions was preserved. Indeed, for MS2 bacteriophage, we estimated the resulting loss of viral titer after the process to be -0.86 $\log_{10}$ after the sucrose gradient step and -0.80 after the CsCl gradient step; for CoxB3, the resulting loss of viral titer was estimated at -1.12 $\log_{10}$ after the sucrose gradient step and -1.00 after the CsCl gradient step.

Conversely, cytopathic effects due to CPX and YF infections were observed only after sucrose purification (S1B Fig), confirming the alteration of viral structures and the loss of infectivity after CsCl ultracentrifugation and the recovery of infectious viral particles after 66%-30% sucrose gradient.

Nucleic acid extraction

Three nucleic acid extractions processed on an EMEM sample spiked with the reference viruses were evaluated according to the extraction yield and the degradation of RNA: Trizol LS, QIAmp viral RNA mini kit from Qiagen, and High Pure viral nucleic acid kit from Roche. Although no difference was observed in the extraction yield between Qiagen and Trizol LS (13.5 ng/μL and 19.8 ng/μL of total RNA, respectively), Roche extraction did not reach the

same yield (3.1 ng/μL of total RNA). The RNA profile of the 3 extracts presented the same small RNAs as presented in Fig 2A (*i.e.*, those between 100 nt and 200 nt), suggesting a degradation of nucleic acids during the extraction process.

Although real-time PCR targeting the DNA and RNA reference viruses conducted on these 3 extracts revealed an average difference of one $\log_{10}$ (3 Ct) between Qiagen (the best), Trizol LS, and Roche, Trizol LS extraction was chosen because of its capacity to extract large sample volumes ($>$ 1 mL), as recovered after the sucrose gradient ultracentrifugation step. Moreover, the Trizol LS extraction allowed for separate extraction of DNA and RNA, while Qiagen and Roche both extracted total nucleic acids and therefore required a DNAse post-extraction treatment for the RNA virome preparation.

RNA processing and random PCR

Hereafter, the 3 metagenomes are designated as "Froussard", "Wang", or "Victoria", according to the sequence-independent reverse transcription and amplification methods used to generate them.

Total RNA was processed in three different random reverse transcriptions (RT), as previously described by Froussard [29], Wang [30], and Victoria [31]. For each random RT, different random primer concentrations were tested to determine the best concentration for generating cDNA fragments with lengths compatible with Illumina MiSeq requirements. For Froussard RT, we tested 0.001 μg/μL, 0.01 μg/μL, 0.05 μg/μL, 0.10 μg/μL (published concentration), and 0.15 μg/μL of primer. For Wang RT, we tested 0.4 pmol, 4 pmol, 20 pmol, 40 pmol (published quantity), and 60 pmol of primer. For Victoria RT, we tested 1 pmol, 10 pmol, 50 pmol, 100 pmol (published quantity), and 150 pmol of primer. No difference was observed regarding the size of cDNA fragments when increasing or decreasing the quantity of random primer compared with the published primer concentration used for reverse transcription. The effect of different random primer concentrations used in the reverse transcription step on the size of the generated amplicons is presented in S2A–S2F Fig.

For RNA sequence-independent amplification, dsDNA generated by random RT-Klenow reactions was used in random PCR. To minimize the amplification step and the potential resulting bias of sequence representation in the metagenome, three cycling conditions were tested for each random PCR: 10, 20, and 40 cycles. No difference in dsDNA size profile was observed between 20 and 40 PCR cycles, but difference was observed in the yield of amplification. Ten cycles of amplification failed to reach the minimum amount of dsDNA material required by Illumina MiSeq whereas 20 and 40 cycles of amplification produced a sufficient quantity. For example, a cDNA sample quantified at 0.23 ng/μL was amplified at the yield of 0.27 ng/μL, 0.42 ng/μL, and 84.06 ng/μL after 10, 20 and 40 cycles of Froussard random amplification, respectively. Although only 20 cycles were enough to obtain a sufficient quantity of dsDNA for several samples (at least 1 ng, according to Illumina recommendations), we decided to use 40 cycles in all cases to ensure enough sequencing material and to ensure similar sample treatments. Fig 2B presents the DNA profile of a dsDNA sample amplified with Froussard, Wang, or Victoria PCR after 40 cycles of random amplification. A difference in size profiles of dsDNA between the amplification methods should be noted: 500–2 000 bp with a maximum at 800 bp for Wang amplification, and 300–10 000 bp with a maximum at 1 500 bp for Froussard and Victoria random PCRs. As a result, users can adapt random amplifications to their sequencing technology requirements. For example, a shorter elongation step in the Froussard random amplification results in the generation of smaller fragments, which are compatible with Illumina MiSeq or Roche 454 requirements (S2H Fig).

Finally, we compared the effect of the use of different random PCR on the amplicon size profile for a given reference virus. For each tested reference virus, Froussard, Wang or Victoria amplification methods generated different amplicon sizes (S2I–S2L Fig), revealing the amplification bias when SIA is performed.

Protocol validation: artificial arthropod RNA virome

To evaluate the protocol on a complex sample, twenty body lice were fed *in vitro* with human blood supplemented with YF, CPX, CoxB3, H3N2, MS2, and T4 viruses, and then further processed with the methodology for RNA virome preparation and analysis described above. T4 DNA bacteriophage was added to the RNA viruses' panel in order to assess the residual DNA contamination of the RNA fraction. Twenty non-engorged body lice were concomitantly used as a positive control, *i.e*. they were spiked with the same amount of viruses as that used for engorged lice, and further processed and sequenced as previously described. Viral concentrations applied to either fed- or spiked-lice followed a Gaussian distribution that mimics "natural" conditions. This range of concentrations was also used to evaluate the sensitivity of the virome protocol.

We first evaluated by qRT-PCR and qPCR the loss of viral particles during the purification process of the positive control. A difference of less than one $\log_{10}$ was observed for YF (ΔCt = 2.1) and CoxB3 (ΔCt = 2.6), more than one $\log_{10}$ for H3N2 (ΔCt = 4.06), and no difference for MS2 before and after the purification process, suggesting that the process did not affect the viral load of the reference viruses. For T4 DNA phage, a difference of less than one $\log_{10}$ (ΔCt = 1.2) was observed before and after the purification steps. CPX was negative, both before and after the purification process, probably due to a low viral load and a high limit of detection of the qRT-PCR system. We then purified this artificial positive control in the same way as the engorged lice, as described hereafter.

After homogenization of the engorged lice, clarification and filtration through a 0.45 μm filter, the clarified homogenate was treated with a cocktail of DNA and RNA nucleases to remove the majority of host DNA/RNA contaminants. To purify viral particles from their complex environment, the supernatant was ultracentrifuged on a 66%-30% sucrose gradient, and total RNA were extracted from the resulting interface using the Trizol LS® extraction method. To control the recovery of viral particles and the efficiency of viral purification, fluorescence microscopy was conducted on an aliquot of the purified viral fraction (S3 Fig). No particles with a size compatible to that of bacterial and eukaryotic cells were observed, suggesting good recovery of viral particles after treatments.

Quantification estimated 1.3 ng/μL of total RNA in the engorged lice extract. To evaluate the remaining host DNA and RNA contamination, real-time 18S PCR and RT-PCR were conducted on RNA extracts, resulting in a Ct = 28.85 for the DNA 18S PCR and a negative result for the RNA 18S RT-PCR. These results highlight the presence of a residual host DNA contamination in the RNA fraction. Positive qRT-PCR signals were detected for MS2 (Ct = 15.10) and CoxB3 (Ct = 23.02) viruses, but not for YF, CPX, and H3N2 viruses amplified with published primers.

RNA extracted from engorged and spiked lice was further processed with the Froussard, Wang, or Victoria random reverse transcription-Klenow reactions described above. The resulting dsDNA was quantified using the Quanti-it Picogreen, but quantifications were negative. The low amount of nucleic acid material required the use of random PCR. Only 40 cycles of random PCR achieved the quantity of DNA compatible with Illumina MiSeq requirements.

Table 2 presents the data of the Illumina MiSeq sequencing according to the random PCR method for both engorged and spiked lice. After MiSeq sequencing, with a run of 2 x 300 bp

read length, 7.64 Gb of total information was obtained from a 524 K/mm^2 cluster density with 12,380,000 passed filter paired reads (96.1% of clusters passing the quality control filter). Within this pooled run, the index representation for the 6 cDNA samples ranged from 4.8% to 7.8%, corresponding to 590,972 to 969,911 passed filter paired reads.

Bioinformatics analyses of the Froussard, Wang and Victoria engorged lice metagenomes are presented in Fig 3 and in Table 2. The mapping of total reads against reference genomes of YF, CPX, CoxB3, H3N2, MS2, and T4 are presented in Fig 3A. Results of mapping and taxonomic assignation of reads of the positive control (*i.e.* spiked lice) RNA metagenome are presented in Table 2 and S4 Fig. No difference was observed in the genome coverage for MS2 and CoxB3 between the 3 engorged lice and the 3 spiked lice RNA viromes. Coverage was 99.41% for MS2 and 97.45% for CoxB3 in the Froussard engorged metagenomes, 99.30% for MS2 and 100% for CoxB3 in the Wang engorged metagenomes, and 99.41% for MS2 and 98.56% for CoxB3 in the Victoria engorged metagenomes. T4 qPCR was negative and very few reads of T4 DNA bacteriophage were obtained in the RNA Froussard and Victoria metagenomes. However, several reads of T3 DNA bacteriophage were obtained, either for engorged or spiked lice metagenomes (Table 2, Fig 3, and S4 Fig). A recent study has shown that the T4 strain (11303-B4) provided by the ATCC was indeed characterized as a T3 strain after sequencing [36]. Since we spiked the blood with the same strain and since our results from sequencing showed only few reads related to T4 compared to the high abundance of T3 phage reads, we believe that the T3 reads recovered derived from T3 phage spiked in the blood sample. These results thus indicate a remaining viral DNA contamination of the RNA fraction after Trizol LS® extraction (even after a second Trizol/chloroform purification performed in the RNA aqueous phase). Sequencing of H3N2 in the RNA metagenome yielded coverage of 22.21%, 24.21% and 0.40% for Froussard, Wang and Victoria, respectively, after the body lice blood meal following inoculation of 3.07 x 10^5 PFU of H3N2 in the blood meal. Although a negative qRT-PCR result was obtained for H3N2 in the lice extract using primers from the literature, amplification using primers specifically designed on H3N2 reads recovered after sequencing was positive (data not shnown), confirming the presence of H3N2 viruses in the lice. The discrepancies observed between PCR results obtained with published or virome-based primers are likely due to differences in primers sensitivities. Additionally, no reads were obtained for YF and CPX, which were also negative after qRT-PCR (both with published and virome-based primers), most likely due to the low amount of viral particles ingested by the lice during their blood meal (less than 1.0 x 10^3 PFU of YF and 1.19 x 10^2 PFU of CPX added to the blood). Indeed, as described by Cheval *et al.* in 2011 [37] and by Frey *et al.* in 2014 [38], the limit of detection of next-generation sequencing techniques are estimated at 10^3 to 10^4 genome copies per mL.

Sequencing of the spiked lice resulted in the same pattern of viral relative abundances and coverage as the one observed for engorged lice (Table 2, S4 Fig), suggesting that the difference in the observed read abundances, either for the engorged or for the spiked lice, did not result from the purification process and was probably due to intrinsic characteristics of the reference viruses. Interestingly, whereas no reads of YF and CPX were detected in the engorged lice metagenomes, whatever the amplification method, a few reads of YF and CPX were detected in the spiked lice metagenomes (N = 2 to 32 reads for YF, depending on the amplification method, N = 1 read for CPX, whatever the amplification method) (Table 2). YF was positive in qRT-PCR both with the published and virome-based primers after the process, and CPX was negative, suggesting that the viral load of these viruses was probably at the limit of detection of the sequencing method (*i.e.* between 10^3 PFU [as for YF] and 10^2 PFU [as for CPX]).

No difference was observed in the depth of sequencing. The average genome coverage of MS2 and CoxB3 was estimated at 46,035 and 3,707, respectively, for Froussard engorged lice

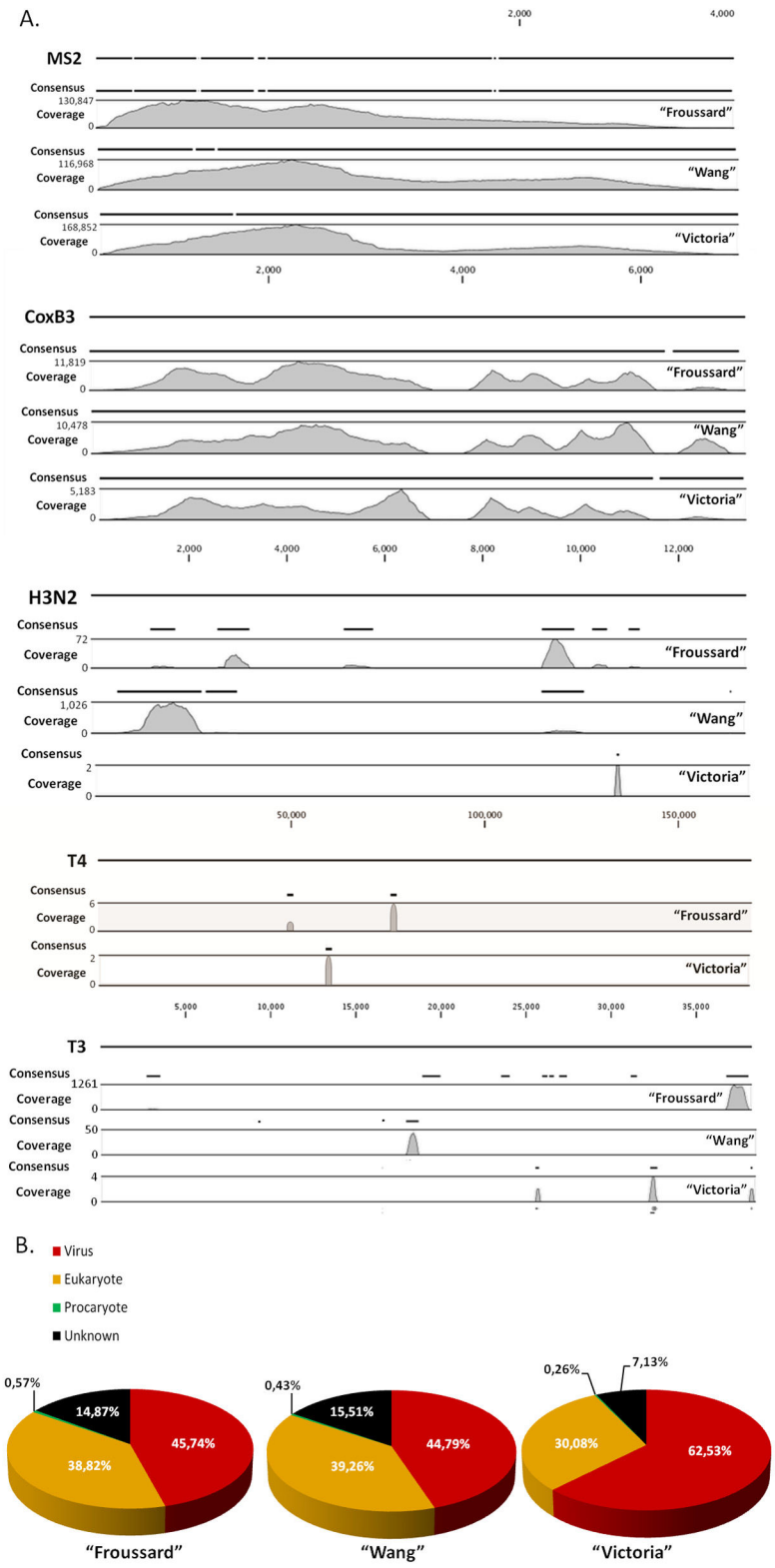


Fig 3. Comparison of the 3 random PCR reactions in engorged lice metagenomes according to: A. reference genome coverage. B. General taxonomic assignment of reads.

doi:10.1371/journal.pone.0139810.g003

metagenomes, 43,559 and 3,495, respectively, for Wang engorged lice metagenomes, and 50,193 and 1,344, respectively, for Victoria engorged lice metagenomes. The main difference between the 3 engorged lice metagenomes was in H3N2 mapping: the Wang random PCR seemed to amplify preferentially large segments (PB2 and PB1 polymerase segments), while the Froussard method seemed to randomly amplify all segments (Fig 3A). Interestingly, this bias of repartition of reads along the whole genome of H3N2 was not observed in the positive control (S4 Fig), suggesting that with larger amounts of viruses, Wang and Victoria SIA are able to randomly amplify each segment of the virus. Nevertheless, in "environmental" conditions (*i.e.* with low viral loads), only the Froussard SIA seemed to be able to randomly amplify each segment of segmented viruses (Fig 3A).

There was no notable difference in the taxonomic assignment of reads when comparing the 3 engorged lice metagenomes (Table 2). Froussard, Wang, and Victoria metagenomes were able to identify 85.13%, 84.49%, and 92.87% of reads, respectively. Of these, viral reads represented 45.74% (721,520 reads) of Froussard, 44.79% (749,587 reads) of Wang, and 62.53% (38,939 reads) of Victoria metagenomes; eukaryote reads represented 38.82%, 39.26%, and 30.08%, respectively, and prokaryote reads (bacteria and archaea) represented only 0.57%, 0.43%, and 0.26% of the respective total reads (Fig 3B). Despite all the pretreatments applied to decrease the host contamination, human and arthropod sequences persisted. Indeed, human contamination of the viromes was estimated at only 0.67%, 0.17%, and 0.044% for Froussard, Wang, and Victoria metagenomes, respectively, but remaining arthropod contamination was estimated at 38.05%, 39.05% and 29.98%.

Discussion

Viral metagenomics has been used worldwide to describe viral communities in various complex environments [13–17] or to identify the etiology of human, animal, and plant pathologies [2–9]. This work resulted in the generation of several matrix-dependent methods to generate viromes. In this study, we aimed at developing a standardized protocol for the purification and sequencing of RNA host-associated viromes from complex biological samples of different origins that would preserve the infectivity of viral particles and allow for further virological characterizations, with an important reduction of host DNA/RNA contaminants in order to fully benefit the depth of sequencing of NGS techniques.

The RNA metagenome preparation protocol is based primarily on the "particle-associated nucleic acid amplification" strategy [21]. That is, it tries to purify intact viral particles from their environment, protected from the action of nucleases used to degrade host-contaminating DNA and RNA. The first step of the protocol is the filtration of the supernatant using 0.45 μm filters. Most previous metagenomic studies using filtration as a purification step used filters with a pore size of 0.22 μm [39,40], however the use of a 0.45 μm filtration step allows the recovery of large DNA viruses whose size would not permit recovery with smaller pore size filters [41].

To eliminate the majority of contaminating host DNAs and RNAs, various strategies are used, such as soft cell lysis, which leaves nuclei intact [42], but the remaining nucleic acids from hosts require nuclease post-treatment. Here we propose the use of a combination of Exonuclease I, Benzonase, RNase A and Turbo DNase to enhance the digestion of host DNA and RNA compared to the use of only RNAse A and Turbo DNAse. Several studies reported that the use of RNAse A in combination with proteinase K could result in the inactivation of some viruses, such as MS2 bacteriophage [43], if no further inactivation of nucleases and proteases is performed. In our protocol, to prevent the inactivation of several particles, which would result in a significant loss of infectivity, the digested supernatant was immediately loaded onto a sucrose gradient to purify viral particles from the action of nucleases.

The next purification step was ultracentrifugation on a discontinuous 66%-30% sucrose gradient. A broad range of methods has been described to purify and concentrate viral particles. These include tangential flow filtration (TFF) [40,44,45], PEG precipitation [18,19,46], cesium chloride (CsCl) gradient ultracentrifugation [22,32,44], and more recently iodixanol (Optiprep^{TM}) density gradient medium. The main disadvantage of CsCl ultracentrifugation is that when analyzing the viral communities of an environment, the lack of prior knowledge of the viruses constituting these communities prevents the harvesting of specific fractions where viruses are located, resulting in recovery of all fractions and thereby diluting the viruses. Iodixanol gradient ultracentrifugation is used either in continuous or in isopycnic gradient to purify viral particles, resulting, as for CsCl, in the necessity to harvest all viral fractions and further diluting the viruses, unless the density of a specific virus is known [47]. Further pelleting by ultracentrifugation would overcome this problem, but we have shown that such a pelleting would damage the integrity of viral particles (recovered viruses were sensitive to nucleases used in a final step, suggesting that the structure of the virion was affected by the two ultracentrifugations). PEG precipitation would also overcome this problem, but would usually result in the precipitation many host DNA, RNA, and protein contaminants.

Recently, a protocol for virome preparation has been published by Kohl *et al.* [48] in which 80%-20% sucrose gradient is used to purify viral particles. The authors used a second ultracentrifugation step to pellet viruses harvested after the gradient ultracentrifugation. In our protocol, we noted that (i) the use of a discontinuous 66%-30% sucrose layer enhanced the purification of viral particles from host nucleic acid contaminants and (ii) a further pelleting ultracentrifugation step followed by nuclease treatments would damage the integrity of virions and increase their sensitivity to nucleases. Here we propose a single ultracentrifugation step based on a discontinuous 66%-30% sucrose gradient. Even if most viruses displayed large differences in their physical-chemical properties, most of them would migrate at the interface between the two sucrose layers. This would result in a single "ring" for recovery, usually < 1 mL and compatible with most nucleic acid extraction protocols.

One key step of metagenome preparation concerns total nucleic acid extraction. Indeed, the assessment of the description of viral diversity of an environment, as well as the genome coverage of viruses, depends on the quality and quantity of extracted nucleic acids. Usually, column-based extraction kits allow for preferential extraction of either DNA or RNA, or both in the same tube, which leads to the splitting in two of the extracts to generate separate DNA and RNA viromes. Recently, column-based kits allowing the extraction of separate DNAs and RNAs from a single sample tube have emerged, but a recent study conducted by Mathieson *et al.* concluded that the quality and integrity of the nucleic acids isolated were compromised using these kits [49]. For this reason, we chose the Trizol LS® system. However, using a DNA bacteriophage as a control of the remaining DNA contamination of the RNA virome, we demonstrated that RNA Trizol LS® extraction allows for the recovery of at least some viral DNA, even if two consecutive extractions of the aqueous phase are performed. Additionally, we selected this reagent for its capacity to extract DNA and RNA separately from large sample volumes (> 1 mL), such as those generated after the recovery of the viral interface resulting from the ultracentrifugation step.

Usually, extraction yields of DNA and RNA obtained after all the viral purification steps are not sufficient for direct sequencing, and they often require a sequence-independent amplification. DNA SIA is biased [50,51] but quite easy, mainly based on the use of phi29 DNA polymerase through Multiple Displacement Amplification (MDA) or Rolling Circle Amplification (RCA) [52]. However, the sequence-independent amplification of RNAs requires several preprocesses to reverse transcript RNA in cDNA before the amplification. For RNA SIA, several amplification methods have been described and compared, including Sequence-Independent

Single-Primer Amplification (SISPA), Degenerate Oligonucleotide-Primed PCR (DOP-PCR), and random PCR [52,53]. Although these techniques allow the generation of enough nucleic acid material for sequencing, their main disadvantage remains that they distort quantitative analyses by introducing bias of amplification in viral diversity studies. A recent study published by Karlsson *et al*. concluded that SIA introduces a strong amplification bias, consisting of inhomogeneous genome coverage and sequence depth [54]. These problems are probably due to the tag sequence, as reported by Rosseel and collaborators [55]. In fact, in the present study similar results were obtained when comparing the genome coverage of reference viruses after Froussard [29], Wang [30], and Victoria [31] random amplifications. These 3 techniques only differ in the tag sequences and the lengths of the 3' random hexamers. Although no notable differences were observed for MS2 and CoxB3 genome coverage, the primary difference between the 3 SIA techniques was for A/H3N2 virus mapping. Indeed, only the Froussard random PCR method performed on engorged lice with artificially contaminated human blood seemed to randomly amplify all segments (Fig 3A). Similarly, other viral reads were detected after Froussard random SIA in higher abundance than the two other random PCRs (Table 2). Additionally, the comparison of the amplicon profiles obtained after the 3 random PCRs resulted for each reference virus in important differences, highlighting the amplification bias (S2I–S2L Fig). Due to the necessity of conducting random amplification after viral purification, quantitative analyses of the composition of resulting viromes may not reflect the initial composition of the viral communities of a given sample, and only qualitative analyses can be conducted.

When characterizing the virome of a biological sample, especially when metagenomics is used in diagnostic virology to determine the etiology of pathology, it is important to isolate the virus. Doing so allows for a clear determination of whether the cultivated virus is able to cause disease in healthy individuals, as described in Koch's postulate in 1876 for bacteria [56] and reviewed by Rivers in 1937 for viruses [57]. Because defining the causality of a given pathology is complex, and because the isolation of viral agents remains the gold standard for conducting studies of the pathogenicity of viruses detected by metagenomics, the need to preserve the infectivity of viral particles during the metagenome analysis process has emerged. By using sucrose instead of CsCl for the gradient preparation, most of the viruses would be purified without compromising the integrity of infectious particles. This can occur with the use of CsCl gradient due to the high osmolarity of CsCl, potentially resulting in the degradation of the structure of several enveloped viruses [32–35] and loss of infectivity, as shown in S1 Fig.

Conclusion

The resulting protocol for host-associated infectious RNA virome preparation is therefore composed of (1) a nuclease digestion of homogenized samples with Turbo DNAse, RNAse A, Benzonase, and Exonuclease I (2) a purification of viral particles on a discontinuous 66%-30% sucrose gradient ultracentrifugation at 130 000 g for 2 hours (3) a Trizol LS® RNA extraction followed by a Turbo DNAse digestion and (4) Froussard-based random RT and PCR.

The protocol reported here could easily be applied to generate RNA viral metagenomes from complex biological samples of different origins, with no loss of viral infectivity and an important elimination of contaminating host DNA and RNA after the process. Moreover, the pipeline described here allows for further virological characterizations of the described viral communities because it preserves the infectivity of viral particles and allows for the isolation of viruses.

Supporting Information

S1 Fig. Infectivity test after the virome process. A. Cytotoxicity of sucrose and CsCl on Vero cells infected with non-purified, sucrose-purified and CsCl purified CPX (day 2 post-infection,

10X objective). B. Cytopathic effects on Vero cells infected with non-purified, sucrose-purified and CsCl purified CPX (day 13 post-infection, 10X objective, dilution 1/100 of the inoculum).
(TIF)

S2 Fig. Nucleic acid profiles analyzed on a 2100 Expert Agilent Analyzer. A: amplification profile using Froussard method with different amount of random primers (red: 0.05 μg/μL / blue: 0.10 μg/μL / green: 0.15 μg/μL). B: amplification profile using Wang method with different amount of random primers (red: 20 pmol / blue: 40 pmol / green: 60 pmol). C: amplification profile using Victoria method with different amount of random primers (red: 50 pmol / blue: 100 pmol / green: 150 pmol). D: amplification profile using Froussard method with different amount of random primers (red: 0.10 μg/μL / blue: 0.01 μg/μL / green: 0.001 μg/μL). E: amplification profile using Wang method with different amount of random primers (red: 40 pmol / blue: 4 pmol / green: 0.4 pmol). F: amplification profile using Victoria method with different amount of random primers (red: 100 pmol / blue: 10 pmol / green: 1 pmol). G: amplification profile with 20 (red) or 40 (blue) cycles of random PCR. H: amplification profile using Froussard method with different elongation durations (red: 1 min / blue: 3 min). I to L: amplification profile of RNA viruses according to Froussard (red), Wang (blue) and Victoria (green) random PCR.
(TIF)

S3 Fig. Fluorescence microscopy of VLP after viral purification and enrichment of body lice. All images were acquired with a Leica SP5 inverted confocal microscope with 4 lasers, a 100X objective and a numerical aperture of 1.4. Scale bar means 30 μm.
(TIF)

S4 Fig. Comparison of the 3 random PCR reactions in spiked lice metagenomes according to the reference genome coverage.
(TIF)

S1 Table. Primers used in this study.
(DOCX)

Acknowledgments

The authors would like to thank Jean-Michel Berenger for his contribution to the lice blood meal, Pascal Weber for his help with the confocal microscope, Ti Thien Nguyen and Noémie Labas for their help in preparing the libraries, and Dr. Nicolás Rascovan for his critical comments on the manuscript. This work was conducted under the frame of an Agence Nationale de la Recherche Young Investigator fellowship (ANR-13-JSV6-0004) awarded to Christelle Desnues. The work of Rémi N. Charrel was conducted under the frame of EurNegVec COST Action TD1303.

Author Contributions

Conceived and designed the experiments: ST SMB CR HP CM DR CD. Performed the experiments: ST SMB CR CM PJ. Analyzed the data: ST SMB CR CM CD. Contributed reagents/materials/analysis tools: RC. Wrote the paper: ST SMB CR HP CM PJ RC DR CD.

References

1. Suttle CA. Viruses in the sea (2005) Nature 437: 356–361. PMID: 16163346
2. Chiu CY. Viral pathogen discovery (2013) Curr Opin Microbiol 16: 468–478. doi: 10.1016/j.mib.2013.05.001 PMID: 23725672

3. Capobianchi MR, Giombini E, Rozera G (2013) Next-generation sequencing technology in clinical virology (2013) Clin Microbiol Infect 19: 15–22. doi: 10.1111/1469-0691.12056 PMID: 23279287
4. Barzon L, Lavezzo E, Militello V, Toppo S, Palù G. Applications of next-generation sequencing technologies to diagnostic virology (2011) Int J Mol Sci 12: 7861–7884. doi: 10.3390/ijms12117861 PMID: 22174638
5. Belák S, Karlsson OE, Blomström AL, Berg M, Granberg F. New viruses in veterinary medicine, detected by metagenomic approaches (2013) Vet Microbiol 165: 95–101. doi: 10.1016/j.vetmic.2013.01.022 PMID: 23428379
6. Blomström AL. Viral metagenomics as an emerging and powerful tool in veterinary medicine (2011) Vet Q. doi: 10.1080/01652176.2011.604971 PMID: 22029881
7. Temmam S, Davoust B, Berenger JM, Raoult D, Desnues C. Viral metagenomics on animals as a tool for the detection of zoonoses prior to human infection? (2014) Int J Mol Sci 15: 10377–10397. doi: 10.3390/ijms150610377 PMID: 24918293
8. Roossinck MJ. Plant virus metagenomics: biodiversity and ecology (2012) Annu Rev Genet 46: 359–369. doi: 10.1146/annurev-genet-110711-155600 PMID: 22934641
9. Ng TF, Duffy S, Polston JE, Bixby E, Vallad GE, Breitbart M. Exploring the diversity of plant DNA viruses and their satellites using vector-enabled metagenomics on whiteflies (2011) PLos One. doi: 10.1371/journal.pone.0019050 PMID: 21544196
10. Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, Moran NA, et al. A metagenomic survey of microbes in honey bee colony collapse disorder (2007) Science. 318: 283–287. PMID: 17823314
11. Granberg F, Vicente-Rubiano M, Rubio-Guerri C, Karlsson OE, Kukielka D, Belák S, et al. Metagenomic detection of viral pathogens in Spanish honeybees: co-infection by Aphid Lethal Paralysis, Israel Acute Paralysis and Lake Sinai Viruses (2013) PLos One. doi: 10.1371/journal.pone.0057459 PMID: 23460860
12. Popgeorgiev N, Temmam S, Raoult D, Desnues C. Describing the silent human virome with an emphasis on giant viruses (2013) Intervirology 56: 395–412. doi: 10.1159/000354561 PMID: 24157886
13. Rosario K, Breitbart M. Exploring the viral world through metagenomics (2011) Curr Opin Virol 1: 289–297. doi: 10.1016/j.coviro.2011.06.004 PMID: 22440785
14. Fancello L, Trape S, Robert C, Boyer M, Popgeorgiev N, Raoult D, et al. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara (2013) ISME J 7: 359–369. doi: 10.1038/ismej.2012.101 PMID: 23038177
15. Whon TW, Kim MS, Roh SW, Shin NR, Lee HW, Bae JW. Metagenomic characterization of airborne viral DNA diversity in the near-surface atmosphere (2012) J Virol 86: 8221–8231. doi: 10.1128/JVI.00293-12 PMID: 22623790
16. Djikeng A, Kuzmickas R, Anderson NG, Spiro DJ. Metagenomic analysis of RNA viruses in a fresh water lake (2009). PLos One. doi: 10.1371/journal.pone.0007264 PMID: 19787045
17. Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, Carlson C, et al. The marine viromes of four oceanic regions (2006) PLos Biol. 4: e368. PMID: 17090214
18. Thurber RV (2011) Methods in Viral Metagenomics. In: de Bruijn FJ, editor. Handbook of Molecular Microbial Ecology, Volume II: Metagenomics in Different Habitats. Wiley-Blackwell. pp 15–24.
19. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F. Laboratory procedures to generate viral metagenomes (2009) Nat Protoc 4: 470–483. doi: 10.1038/nprot.2009.10 PMID: 19300441
20. Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, et al. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery (2014) J Virol Methods 195: 194–204. doi: 10.1016/j.jviromet.2013.08.035 PMID: 24036074
21. Stang A, Korn K, Wildner O, Uberla K. Characterization of virus isolates by particle-associated nucleic acid PCR (2005) J Clin Microbiol 43: 716–720. PMID: 15695669
22. Weynberg KD, Wood-Charlson EM, Suttle CA, van Oppen MJ. Generating viral metagenomes from the coral holobiont (2014) Front Microbiol. doi: 10.3389/fmicb.2014.00206 PMID: 24847321
23. Drosten C, Göttig S, Schilling S, Asper M, Panning M, Schmitz H, et al. Rapid detection and quantification of RNA of Ebola and Marburg viruses, Lassa virus, Crimean-Congo hemorrhagic fever virus, Rift Valley fever virus, dengue virus, and yellow fever virus by real-time reverse transcription-PCR (2002) J Clin Microbiol 40: 2323–2330. PMID: 12089242
24. Watkins-Riedel T, Woegerbauer M, Hollemann D, Hufnagl P. Rapid diagnosis of enterovirus infections by real-time PCR on the LightCycler using the TaqMan format (2002) Diagn Microbiol Infect Dis 42: 99–105. PMID: 11858904

25. van Elden LJ, Nijhuis M, Schipper P, Schuurman R, van Loon AM. Simultaneous detection of influenza viruses A and B using real-time quantitative PCR (2001) J Clin Microbiol 39: 196–200. PMID: 11136770

26. Ninove L, Nougairede A, Gazin C, Thirion L, Delogu I, Zandotti C, et al. RNA and DNA bacteriophages as molecular diagnosis controls in clinical virology: a comprehensive study of more than 45,000 routine PCR tests (2011) PLos One. doi: 10.1371/journal.pone.0016142 PMID: 21347398

27. Breitbart M, Rohwer F. Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing (2005) Biotechniques 39: 729–736. PMID: 16312220

28. Sangaré AK, Boutellis A, Drali R, Audoly G, Weber P, Rolain JM, et al. Doxycycline kills human lice through its activity on their bacterial symbiont (2015) Int J Antimicrob Agents 45: 675–676. doi: 10.1016/j.ijantimicag.2015.02.008 PMID: 25836018

29. Froussard P. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA (1992) Nucleic Acids Res 20: 2900. PMID: 1614887

30. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, et al. Microarray-based detection and genotyping of viral pathogens (2002) Proc Natl Acad Sci U S A 99: 15687–15692. PMID: 12429852

31. Victoria JG, Kapoor A, Dupuis K, Schnurr DP, Delwart EL. Rapid identification of known and new RNA viruses from animal tissues (2008) PLoS Pathog. doi: 10.1371/journal.ppat.1000163 PMID: 18818738

32. Trudel M, Payment P (1989) Purification et analyse de virus par ultracentrifugation. In: Manuel de Techniques Virologiques, Presses de l'Université du Québec.

33. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2011) Virus Taxonomy: ninth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press.

34. Lawrence JE, Steward GF (2010) Purification of viruses by centrifugation. In: Wilhelm SW, Weinbauer MG and Suttle CA, editors. Manual of Aquatic Viral Ecology. ASLO. Pp 166–181.

35. Tidona CA, Darai G (2001). The Springer Index of Viruses. Springer.

36. Kleiner M, Hooper LV, Duerkop BA. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes (2015) BMC Genomics. doi: 10.1186/s12864-014-1207-4 PMID: 25608871

37. Cheval J, Sauvage V, Frageul L, Dacheux L, Guignon G, Dumey N, et al. Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples (2011) J Clin Microbiol. 49: 3268–3275. doi: 10.1128/JCM.00850-11 PMID: 21715589

38. Frey KG, Herrera-Galeano JE, Redden CL, Luu TV, Servetas SL, Mateczun AJ, et al. Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood (2014) BMC Genomics. doi: 10.1186/1471-2164-15-96 PMID: 24495417

39. Phan TG, Vo NP, Bonkoungou IJ, Kapoor A, Barro N, O'Ryan M, et al. Acute diarrhea in West African children: diverse enteric viruses and a novel parvovirus genus (2012) J Virol 86: 11024–11030. PMID: 22855485

40. Ng TF, Marine R, Wang C, Simmonds P, Kapusinszky B, Bodhidatta L, et al. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage (2012) J Virol 86: 12161–12175. doi: 10.1128/JVI.00869-12 PMID: 22933275

41. Colson P, de Lamballerie X, Fournous G, Raoult D. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales (2012) 55: 321–332.

42. Daly GM, Bexfield N, Heaney J, Stubbs S, Mayer AP, Palser A, et al. A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing (2011) PLos One. doi: 10.1371/journal.pone.0028879 PMID: 22216131

43. Yang Y, Griffiths MW. Enzyme treatment reverse transcription-PCR to differentiate infectious and inactivated F-specific RNA phages (2014) Appl Environ Microbiol. 80: 3334–3340. doi: 10.1128/AEM.03964-13 PMID: 24657854

44. Hurwitz BL, Deng L, Poulos BT, Sullivan MB. Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics (2013) Environ Microbiol 15: 1428–1440. doi: 10.1111/j.1462-2920.2012.02836.x PMID: 22845467

45. Sachsenröder J, Twardziok S, Hammerl JA, Janczyk P, Wrede P, Hertwig S, et al. Simultaneous identification of DNA and RNA viruses present in pig faeces using process-controlled deep sequencing (2012) PLos One. doi: 10.1371/journal.pone.0034631 PMID: 22514648

46. Marston DA, McElhinney LM, Ellis RJ, Horton DL, Wise EL, Leech SL, et al. Next generation sequencing of viral RNA genomes (2013) BMC Genomics. doi: 10.1186/1471-2164-14-444 PMID: 23822119

47. Maillard P, Walic M, Meuleman P, Roohvand F, Huby T, Le Goff W, et al. Lipoprotein lipase inhibits hepatitis C virus (HCV) infection by blocking virus cell entry (2011) PLoS One. doi: 10.1371/journal.pone.0026637 PMID: 22039521

48. Kohl C, Brinkmann A, Dabrowski PW, Radonić A, Nitsche A, Kurth A. Protocol for metagenomic virus detection in clinical specimens (2015) Emerg Infect Dis. 21: 48–57. doi: 10.3201/eid2101.140766 PMID: 25532973

49. Mathieson W, Thomas GA. Simultaneously extracting DNA, RNA, and protein using kits: is sample quantity or quality prejudiced? (2013) Anal Biochem 433: 10–18. doi: 10.1016/j.ab.2012.10.006 PMID: 23068038

50. Kim KH, Bae JW. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses (2011) Appl Environ Microbiol 77: 7663–7668. doi: 10.1128/AEM.00289-11 PMID: 21926223

51. Marine R, McCarren C, Vorrasane V, Nasko D, Crowgey E, Polson SW, et al. Caught in the middle with multiple displacement amplification: the myth of pooling for avoiding multiple displacement amplification bias in a metagenome (2014) Microbiome. doi: 10.1186/2049-2618-2-3 PMID: 24475755

52. Berthet N, Reinhardt AK, Leclercq I, van Ooyen S, Batéjat C, Dickinson P, et al. Phi29 polymerase based random amplification of viral RNA as an alternative to random RT-PCR (2008) BMC Mol Biol. doi: 10.1186/1471-2199-9-77 PMID: 18771595

53. Delwart EL. Viral metagenomics (2007) Rev Med Virol 17:115–131. PMID: 17295196

54. Karlsson OE, Belák S, Granberg F. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses (2013) Biosecur Bioterror 11 Suppl 1: 227–234.

55. Rosseel T, Van Borm S, Vandenbussche F, Hoffmann B, van den Berg T, Beer M, et al. The origin of biased sequence depth in sequence-independent nucleic acid amplification and optimization for efficient massive parallel sequencing (2013) PLos One. doi: 10.1371/journal.pone.0076144 PMID: 24086702

56. Koch R. Investigations into bacteria: V. The etiology of anthrax, based on the ontogenesis of *Bacillus anthracis* (1876) Cohns Beitrage zur Biologie der Pflanzen 2: 277–310.

57. Rivers TM. Viruses and Koch's Postulates. (1937) J Bacteriol 33: 1–12. PMID: 16559982

Supporting information

Figure S1. Infectivity test after the virome process. A. Cytotoxicity of sucrose and CsCl on Vero cells infected with non-purified, sucrose-purified and CsCl purified CPX (day 2 post-infection, 10X objective). B. Cytopathic effects on Vero cells infected with non-purified, sucrose-purified and CsCl purified CPX (day 13 post-infection, 10X objective, dilution 1/100 of the inoculum).

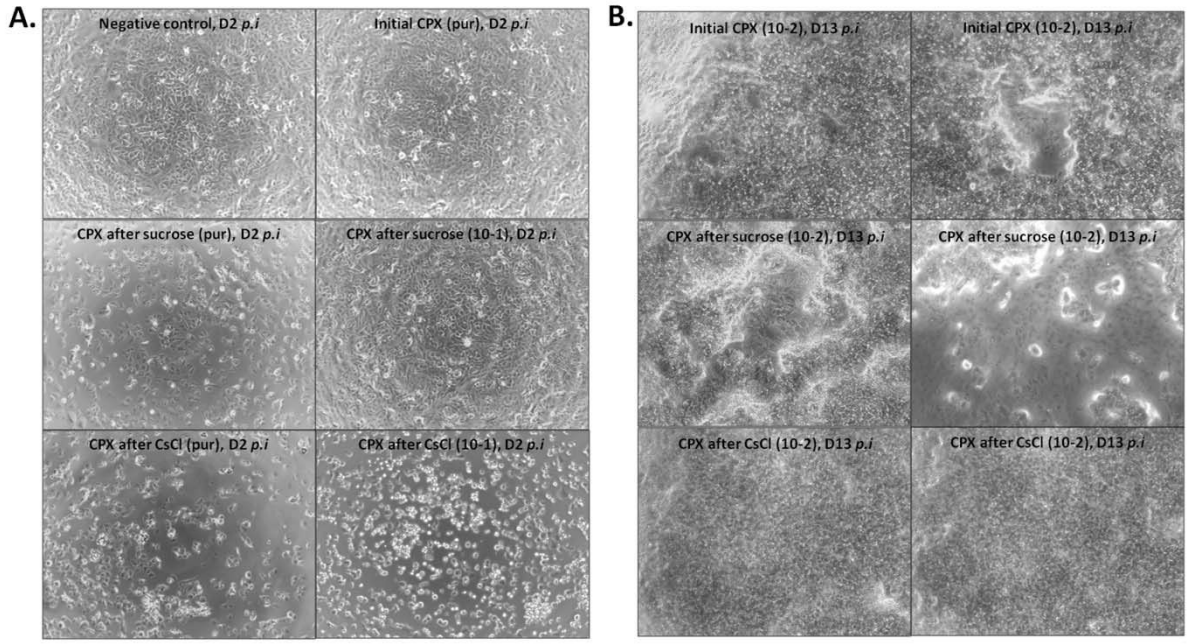


Figure S2. Nucleic acid profiles analyzed on a 2100 Expert Agilent Analyzer. A: amplification profile using Froussard method with different amount of random primers (red: 0.05 µg/µL / blue: 0.10 µg/µL / green: 0.15 µg/µL). B: amplification profile using Wang method with different amount of random primers (red: 20 pmol / blue: 40 pmol / green: 60 pmol). C: amplification profile using Victoria method with different amount of random primers (red: 50 pmol / blue: 100 pmol / green: 150 pmol). D: amplification profile using Froussard method with different amount of random primers (red: 0.10 µg/µL / blue: 0.01 µg/µL / green: 0.001 µg/µL). E: amplification profile using Wang method with different amount of random primers (red: 40 pmol / blue: 4 pmol / green: 0.4 pmol). F: amplification profile using Victoria method with different amount of random primers (red: 100 pmol / blue: 10 pmol / green: 1 pmol). G: amplification profile with 20 (red) or 40 (blue) cycles of random PCR. H: amplification profile using Froussard method with different elongation

durations (red: 1 min / blue: 3 min). I to L: amplification profile of RNA viruses according to Froussard (red), Wang (blue) and Victoria (green) random PCR.

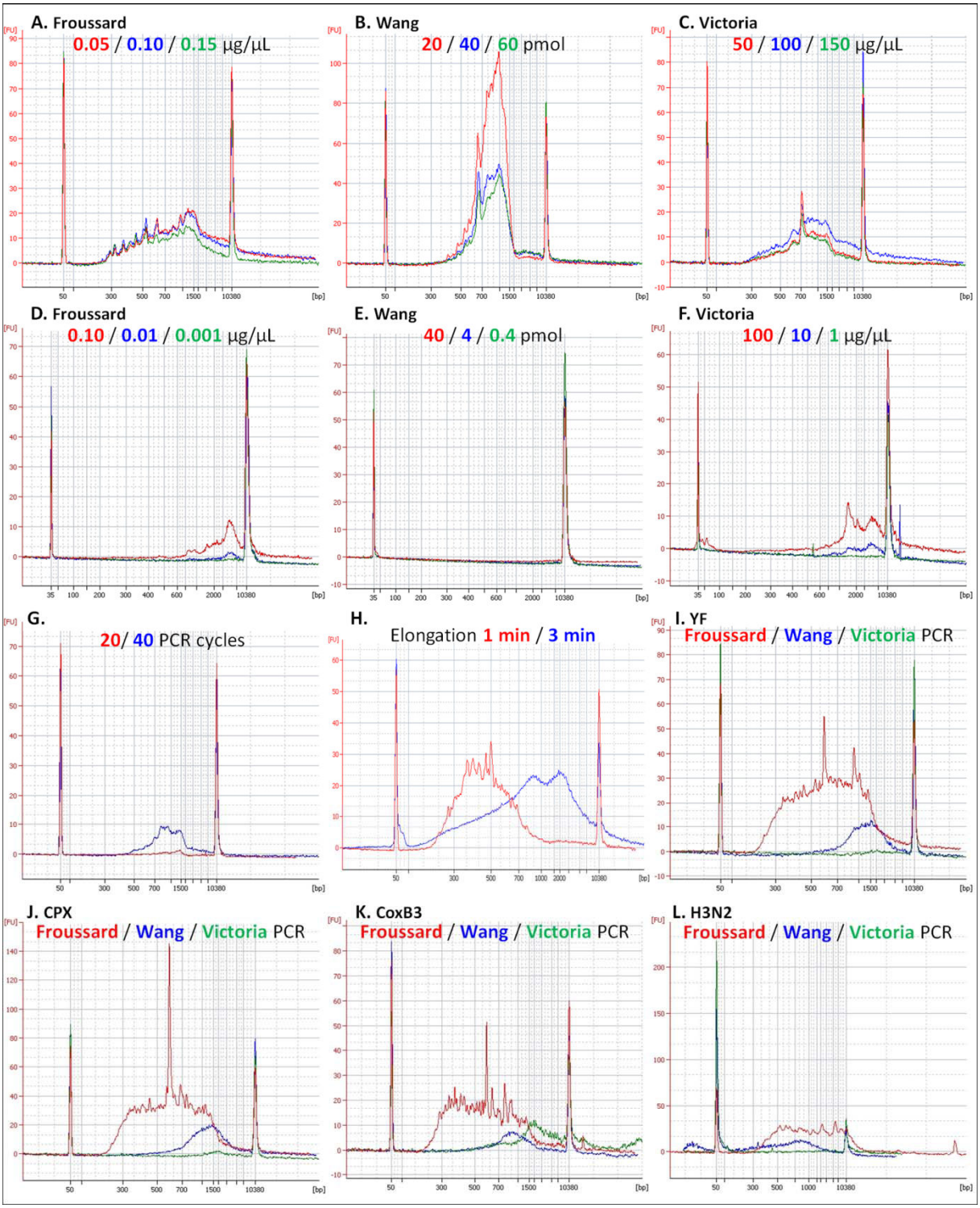


Figure S3. Fluorescence microscopy of VLP after viral purification and enrichment of body lice. All images were acquired with a Leica SP5 inverted confocal microscope with 4 lasers, a 100X objective and a numerical aperture of 1.4. Scale bar means 30 µm.

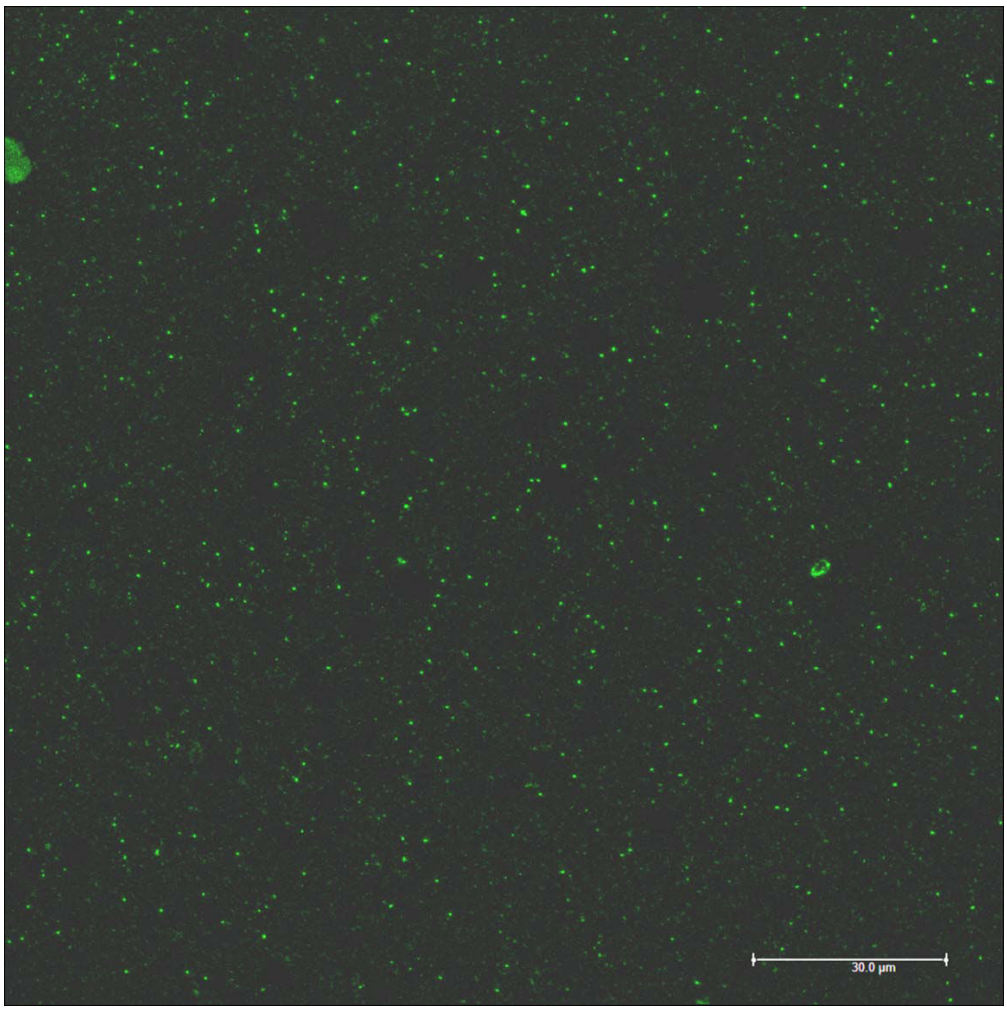


Figure S4. Comparison of the 3 random PCR reactions in spiked lice metagenomes according to the reference genome coverage.

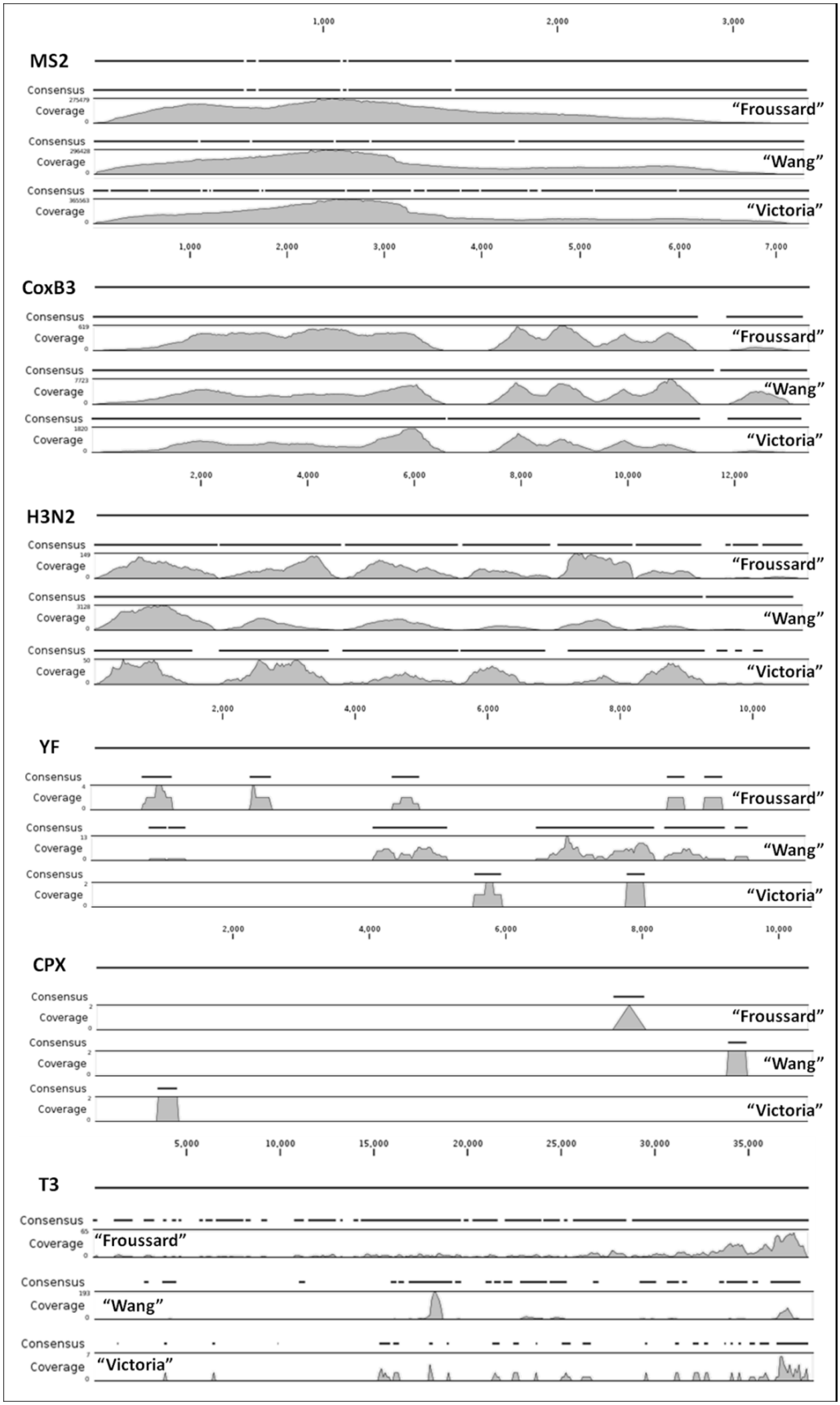


Table S1. Primers used in this study.

Virus	Primer	Sequence (5' → 3')	Reference
YF	forward	AATCGAGTTGCTAGGCAATAAACAC	23
	reverse	TCCCTGAGCTTTACGACCAGA	
CoxB3	forward	CCCCTGAATGCGGCTAATCC	24
	reverse	ATTGTCACCATAAGCAGCCA	
H3N2	forward	CATYCTGTTGTATATGAGGCCCAT	25
	reverse	GGACTGCAGCGTAGACGCTT	
MS2	forward	GTTCCCTACAACGAGCCTAAATTC	26
	reverse	CTCTGAGAGCGGCTCTATTGGT	
T4	forward	CCATCCATAGAGAAAATATCAGAACGA	26
	reverse	CGCTGGGAAAAGAGGAATTATTTA	
CPX	forward	TGGCAGAGAATGGTATAGTAGG	This study
	reverse	GCTTTCCCATATCAGACTTCC	
18S	forward	CCGCAGCTAGGAATAATGGAATAGGAC	27
	reverse	ITTAGCATGCCAGAGTCTCGTTCGT	
YF (virome-based)	forward	TGGTGAAGTTTCATGGGAAGAGG	This study
	reverse	CCAAGATGGAATCAACTTCTTGCC	
H3N2 (virome-based)	forward	TGGATCAAGTGAGAGAAAGTCGG	This study
	reverse	CTCATTTGAGGCAATTTGTACTCC	

Etude du modèle arthropode

I. **Article n°3**: "Viral communities of biting midges reveal novel emerging arboviruses, including novel Thogotovirus species and Rhabdovirus genus."

Viral communities of biting midges reveal novel emerging arboviruses, including novel *Thogotovirus* species and *Rhabdovirus* genus.

Sarah Temmam[1], Sonia Monteil-Bouchard[1], Catherine Robert[1], Jean-Pierre Baudoin[1], Noémie Labas[1], Didier Raoult[1,2], Oleg Mediannikov[1], and Christelle Desnues[1,*].

Under revision in: Viruses

[1] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1095, Aix-Marseille Université, Marseille, France.

[2] Fondation IHU Méditerranée Infection, Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, Centre Hospitalo-Universitaire Timone, Méditerranée Infection, Assistance Publique – Hôpitaux de Marseille, Marseille, France.

[*] Corresponding author: christelle.desnues@univ-amu.fr; tel.: +33(0)491324630; fax: +33(0)491387772.

Préambule à l'article "Viral communities of biting midges reveal novel emerging arboviruses, including novel *Thogotovirus* species and *Rhabdovirus* genus"

Près des deux tiers des pathogènes humains émergents sont originaires de la faune sauvage ou domestique [83], et le mode de transmission de ces virus à l'homme se fait par contact direct ou le plus souvent via un arthropode hématophage. Bien que le rôle des arthropodes soit prépondérant dans la survenue de transmissions zoonotiques, peu d'études ont été conduites sur la composition des communautés virales des arthropodes hématophages, hormis quelques études sur les moustiques [84-88], et plus récemment sur les tiques [89-90]. Ces études ont rapporté la présence de nouveaux virus apparentés à des *bunyavirus*, des *rhabdovirus*, des *reovirus*, des *flavivirus*, des *nairovirus* ou encore des *phlebovirus*, dont certains peuvent constituer un risque d'émergence en population humaine. Aucune autre étude n'a été conduite sur d'autres arthropodes hématophages d'intérêt médical.

Les insectes de la famille des *Ceratopogonidae*, et plus précisément du genre *Culicoides*, sont des vecteurs reconnus de nombreux virus ayant un impact en médecine vétérinaire [91]. Seul le virus *Oropouche*, un *orthobunyavirus*, a été décrit comme étant transmissible à l'homme via la piqûre de culicoïdes, mais sa distribution géographique semble restreinte à l'Amérique du Sud [92] où il cause des syndromes fébriles proches de ceux de la dengue, et pouvant occasionner dans de rares cas des méningites. Au Sénégal, les culicoïdes représentent des nuisances importantes, tant pour l'homme qui est souvent piqué, que pour la santé animale.

Dans ce contexte, nous avons réalisé par séquençage haut débit une analyse de métagénomique des communautés virales de culicoïdes collectés au Sénégal en novembre 2013. Les viromes ARN présentent une grande diversité de virus, infectant à la fois des plantes, des insectes, des bactéries, des amibes et des mammifères. Outre les virus classiquement transmis par les culicoïdes et présents au Sénégal, par exemple les virus de la peste équine ou de la fièvre catarrhale ovine, nous avons pu détecter des séquences apparentées à un nouveau virus du genre *Thogotovirus* ainsi que des séquences d'un nouveau *Rhabdovirus* qui pourrait constituer un nouveau genre viral. En plus de la découverte de nouveaux virus qui pourraient constituer de potentiels risques

d’émergence pour la santé humaine et animale, cette étude illustre l'intérêt de surveiller des populations « sentinelles » telles que les arthropodes vecteurs, qui s’avèrent être de bons modèles d'étude de ce point de vue, puisque nous avons pu montrer la faisabilité de l’identification de nouveaux virus ayant un potentiel zoonotique en amont d'un épisode épidémique qui surviendrait suite à une émergence virale en populations animales ou humaines.

Viral communities of biting midges reveal novel emerging arboviruses, including novel *Thogotovirus* species and *Rhabdovirus* genus.

Sarah Temmam[1], Sonia Monteil-Bouchard[1], Catherine Robert[1], Jean-Pierre Baudoin[1], Masse Sambou[1], Maxence Aubadie-Ladrix[1], Noémie Labas[1], Didier Raoult[1,2], Oleg Mediannikov[1] and Christelle Desnues[1,*].

[1] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1095, Aix-Marseille Université, Marseille, France.

[2] Fondation IHU Méditerranée Infection, Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, Centre Hospitalo-Universitaire Timone, Méditerranée Infection, Assistance Publique – Hôpitaux de Marseille, Marseille, France.

* Corresponding author: christelle.desnues@univ-amu.fr; tel.: +33(0)491324630; fax: +33(0)491387772.

Abstract

More than two thirds of emerging viruses are of zoonotic origin, and among them RNA viruses represent the majority. *Ceratopogonidae* (genus *Culicoides*) are well-known vectors of several viruses responsible for epizooties (bluetongue, epizootic haemorrhagic disease, etc.). They are also vectors of the only known virus infecting humans: the Oropouche virus. Female midges usually feed on a variety of hosts, leading to possible transmission of emerging viruses from animals to humans. In this context and in an attempt to identify potential zoonotic viruses, we report here the analysis of RNA viral communities of Senegalese biting midges using next-generation sequencing techniques.

Sequencing of the RNA virome of three pools of Culicoides revealed the presence of a significant diversity of viruses infecting plants, insects and mammals. Several arboviruses were detected, including a novel *Thogotovirus* species, related but genetically distant from previously described tick-borne thogotoviruses. Novel rhabdoviruses were also detected, possibly constituting a novel *Rhabdoviridae* genus, and putatively restricted to insects. Sequences related to the major viruses transmitted by Culicoides, *i.e.* African horse sickness, bluetongue and epizootic haemorrhagic disease viruses were also detected.

This study highlights the interest of monitoring the emergence and circulation of zoonoses and epizooties using their arthropod vectors.

Keywords

Viral metagenomics, biting midges, zoonoses, epizooties, Thogotovirus, Rhabdovirus.

Introduction

There are more than 200 viral species which are known to be able to infect humans. Since the discovery of the yellow fever virus in 1901, three to four new species have been discovered every year [1]. There is, however, a substantial pool of unknown human viral species which are yet to be discovered and the development and democratisation of Next-Generation Sequencing techniques (NGS) has enabled the identification of many new viruses, for which the potential risk to humans remains mostly unknown. More than two-thirds of viral species infecting humans are of zoonotic origin, and RNA viruses represent more than 70% of these [1,2], resulting in the recent increase in studies of viral communities of wild and domestic animals [3]. However, and despite the fact that haematophagous arthropods usually act as vectors of transmission between animals and humans, few studies have analysed viral communities of arthropods [3]. The studies that have been previously conducted have focused on mosquito viromes [4-8] and have reported the discovery of novel viruses, including bunyaviruses, rhabdoviruses, reoviruses, and flaviviruses. More recently, two studies described the composition of viral communities of hard ticks [9,10] and reported the identification of novel viruses belonging to the *Nairovirus*, *Phlebovirus*, and *Flavivirus* genera, highlighting, as for mosquitoes, potential new zoonotic risks to humans.

Ceratopogonidae, and particularly the genus *Culicoides*, are small (1 to 3 mm) and highly diverse midges, with more than 1,300 species around the world [11,12]. Of these, 96% are haematophagous and only the females require blood meal for egg fertilisation. Biting midges are well-known vectors of several parasites (such as *Mansonella sp*.) [13,14] and viruses infecting animals (*i.e.* bluetongue virus, African horse sickness virus, epizootic haemorrhagic disease virus, Schmallenberg virus, etc.) [15]. The Oropouche virus is the only human virus known to be transmitted by biting midges in Latin and South America [16].

In an attempt to identify potential zoonotic viruses transmitted to humans via Culicoides bites, we report here the first comprehensive analysis of viral communities from

Senegalese biting midges and the identification of several novel arboviruses, including a novel thogotovirus and a novel rhabdovirus.

Materials & Methods

Sample collection

Biting midges were collected using a modified CDC light trap in the villages of Dielmo and Ndiop in the Sine-Saloum region of Senegal, in November 2013. Traps were placed near places where cattle rested and were left overnight. Morphological identification of the arthropods was conducted the following morning. Three types of pools of arthropods were created: STE0043 (more than 200 adult *Culicoides* sp., with no distinction between male and female, or engorged status); STE0044 (N=15 engorged female *Culicoides imicola*) and STE0045 (N=100 non-engorged male and female *Culicoides imicola*).

Virome preparation

The three pools of arthropods were crushed with two 3 mm tungsten beads and a TissueLyser at 25 Hz for two minutes (Qiagen, Courtaboeuf, France). The clarified supernatant was subsequently used as a template for virome preparation, as previously described [17]. Briefly, the clarified supernatant was filtered through a 0.45-µm filter (Millipore, Molsheim, France), and free nucleic acids were digested with a cocktail of nucleases. Finally, the digested supernatant was purified onto a discontinuous 66%-30% sucrose gradient and ultracentrifuged at 130,000 g for two hours at +4°C on a MLS-50 rotor (Beckman-Coulter, Villepinte, France). The viral fraction was harvested at the interphase between the 66% and 30% sucrose layers. Total RNAs were extracted from the purified viral fraction with Trizol LS® reagent (Life Technologies, Saint Aubin, France), according to the manufacturer's recommendations. Random amplification was performed using the Froussard [18] random RT-PCR. and amplification products were purified with Agencourt AMPure Beads (Beckman-Coulter, Villepinte, France) according to the manufacturer's protocol, eluted to a final volume of 15 µL and sequenced using MiSeq Technology using paired-end and barcode strategies according to a Nextera XT library kit in a 2 x 300 bp format (Illumina Inc., San Diego, USA).

Bioinformatic analyses of viromes

Raw reads were imported in pairs into the CLC Genomics Workbench 6.0.1 programme (CLC Bio, Aarhus, Denmark) and trimmed according to their quality score, the

presence of ambiguities, and their length (reads which were shorter than 50 nt were discarded). The pre-processed viral metagenomes are publicly available on the Metavir server (http://metavir-meb.univ-bpclermont.fr) under the 'Arthrovirome' project and on the MG-RAST server (http://metagenomics.anl.gov/) with the identifiers 4604249.3, 4604250.3, and 4604251.3 for the STE0043, STE0044 and STE0045 RNA viromes, respectively.

Cleaned paired reads were assembled into contigs using the CLC Genomics programme and the following parameters: word size of 20 nt, minimum contig length of 200 nt, mismatch cost of 2, insertion/deletion cost of 3, length fraction of 0.5 and similarity fraction of 0.8. Contigs and non-assembled reads were compared to the NCBI nucleotide database using the BlastN algorithm, with a minimum coverage of 50%, minimum identity of 50% and E-value < 10^{-5}. Sequences having no significant hits according to the criteria described above were classified as 'unknown'. Contigs were then compared to the NCBI viral database using the BlastX program with a minimum coverage of 50%, minimum identity of 50% and E-value < 10^{-5}. Finally, to confirm the specificity of the BlastX result, contigs were compared to the NCBI non-redundant nucleotide database using the same criteria. The taxonomic assignation of contigs was conducted by selecting the best BlastX score result between the two Blast run for each contig. Figure S1 presents the pipeline for bioinformatic analyses.

Principal Component Analysis (PCA) was used to compare data in the MG-RAST server [19] with a maximum E-value of 10^{-5}, a minimum identity of 60%, and a minimum alignment length of 15 amino-acids for protein and 15 bp for RNA databases. Data were normalised to values between 0 and 1, and distances were measured using the Bray-Curtis distance matrix.

Phylogenetic analyses

Contigs with a significant hit for viruses were translated and predicted Open Reading Frames (ORFs) were aligned with other amino-acid sequences retrieved from the GenBank database using MUSCLE aligner [20] implemented through MEGA6 [21]. The amino-acid substitutions models that best fitted the data were performed on MEGA6 and were considered for all phylogenetic analyses. The best substitution model was selected using the corrected Akaike information criterion. Phylogenetic trees were constructed using Maximum Likelihood (ML) implemented through the MEGA6 package software, according to the selected substitution model. Nodal support was evaluated using 1,000 bootstrap replicates. Bayesian phylogenetic inference (BI) was carried out using MrBayes [22] with two

independent runs of four incrementally-heated, Metropolis-coupled Markov chain Monte Carlo (MCMC) starting from a random tree. The MCMC were run for 10^6 iterations and associated model parameters were sampled every 500 generations. The initial 2,000 trees in each run were discarded as burning samples and the harmonic mean of the likelihood was calculated by combining the two independent runs.

Molecular evolutionary distances between sequences were calculated using MEGA6 [21]. For analysis of evolutionary distances between thogotoviruses, individual sequences available in GenBank and the *p*-distances algorithm were used. For analysis of molecular evolutionary distances between rhabdoviruses, sequences available in GenBank were grouped according to their recognised or putative genus (defined by phylogenetic analyses) and distances were calculated (i) within genera using the *p*-distance algorithm (ii) between genera using net distance calculations (*i.e.* MEGA6 takes into account the mean distance within genera) and the *p*-distance algorithm.

Transmission electron microscopy (TEM)

Approximately 50 mg of STE0043 arthropod samples were washed in 70% ethanol and crushed in 2 mL of sterile EMEM medium (Life Technologies, Saint Aubin, France). The supernatant was harvested after low speed clarification and subsequently filtered through a 0.8-µm filter (Millipore, Molsheim, France) followed by ultracentrifugation onto a discontinuous 66%-30% sucrose gradient at 130,000 g for two hours at +4°C. The viral fraction was harvested at the interphase between the 66% and 30% sucrose layers and fixed for one hour at +4°C with 2% final glutaraldehyde. The fixed viral fraction was then diluted to a final volume of 4 mL in PBS and directly adsorbed onto formvar carbon films on 400 mesh nickel grids (FCF400-Ni, EMS) by ultracentrifugation at 130,000 g for one hour at +4°C, as previously described [23]. Grids were stained for 10 seconds with 1% molybdate solution in filtered water at room temperature. Electron micrographs were obtained on a Tecnai G2 transmission electron microscope (FEI) operated at 200 keV equipped with a 4096 x 4096 pixel resolution Eagle camera (FEI).

Results

Diversity of viral communities in haematophagous biting midges

RNA viromes of samples STE0043, STE0044 and STE0045 were sequenced using Illumina MiSeq technology. Sequencing statistics are presented in Table 1.

The taxonomic assignment of reads identified only 5 to 25% of sequences which had similarities with known sequences (Figure 1A). Of these, eukaryotes represented the majority of sequences, with 72.52%, 62.10% and 83.95% of total known reads of the STE0043, STE0044 and STE0045 RNA viromes, respectively (Figure 1A). Most eukaryotic reads were assigned to arthropods (> 60% of total eukaryotic reads), and they mainly consisted of arthropod ribosomal RNAs. Bacteria-related sequences ranged from 9% to 37% depending on the sample (Figure 1A).

Virus-related sequences represented 0.73% to 18.48% of total known reads. Of them, plant viruses composed 15.48%, 10.10% and 0.00% of total viral reads for STE0043, STE0044 and STE0045 DNA viromes, respectively (Figure 1B). Insect viruses represented the majority of viral reads, with 55.51%, 76.23% and 33.66% of total viral reads for STE0043, STE0044 and STE0045 RNA viromes, respectively. Several mammalian viruses were detected, such as *Picobirnaviridae*-related viruses, but only in the STE0045 *C. imicola* male and non-engorged female virome, with a global abundance of 33.73% of total viral reads. Several arboviruses (*i.e.* mammalian-infecting viruses transmitted by haematophagous arthropod bites) belonging to the *Reoviridae*, *Orthomyxoviridae* and *Rhabdoviridae* families were identified and represented 28.99%, 4.18% and 19.63% of total viral reads for STE0043, STE0044 and STE0045 RNA viromes, respectively (Figure 1B). Few DNA viruses were also identified in the RNA viromes (bacteriophages and amoeba-infecting giant viruses), possibly due to residual contamination of the RNA fraction by viral DNA (Figure 1B).

Electron microscopy images of the STE0043 *Culicoides sp*. purified viral fraction showed the presence of Virus-Like Particles (VLPs) with various diameters, morphologies, and contrasts (Figure 2). Some VLPs presented a round structure with a distinct envelope, while others appeared with more contrast. The diameters of the particles ranged from 100 nm to 600 nm, compatible with several viruses detected by metagenomic analyses (Figure 2).

Principal Component Analysis (PCA) was used to compare viral communities of biting midges with other haematophagous and non-haematophagous arthropod RNA viromes available in public databases (Figure 3, Table S1). RNA viromes of biting midges clustered together, but the STE0043 *Culicoides sp*. virome was more distant than the STE0044 *C.imicola* engorged female and STE0045 *C. imicola* male and non-engorged female viromes.

In addition, biting midge viromes were closer to field and artificially-infected mosquito metagenomes than to whitefly and butterfly viromes (Figure 3).

Orbiviruses were abundant in Senegalese biting midges

Within the viral reads, *Reoviridae*-related sequences represented 26.34%, 0.045% and 17.80% in STE0043, STE0044 and STE0045 RNA viromes, respectively; with the presence of bluetongue-related sequences in STE0043 (N=3,656 reads) and STE0045 (N=678 reads) viromes while epizootic haemorrhagic disease virus (EHDV) was detected in STE0043 (N=5,454 reads) and STE0044 (N=5 reads) viromes. African horse sickness virus (AHSV) was only detected in the STE0043 *Culicoides sp*. RNA virome (N=1,647 reads).

Various segments of these 10-segmented dsRNA orbiviruses were detected in the metagenomes. For example EHDV-related sequences matched with VP4 protein of segment 4 in the STE0044 *C. imicola* engorged female virome. In the STE0045 sample, all reads matched with segment 8 (NS2 protein) of the bluetongue virus (BTV). Within the STE0043 *Culicoides sp*. virome, sequences related to segments 1-2-3-4-6-7-8 and 9 of AHSV were present, while NS1 (segment 1) and NS3 (segment 10) were not detected. Segments 1-2-3-4-8 and 9 of BTV and segments 1-3-4-6 and 8 of EHDV were detected, with a global coverage of the genome estimated after mapping at 37.27%, 34.58% and 33.16% for AHSV, BTV and EHDV, respectively in the STE0043 metagenome (data not shown).

Novel *Thogotovirus* species

Within the virome of the STE0043 *Culicoides sp*. and STE0045 *C. imicola* male and non-engorged female samples, large contigs of 1,903 nt and 1,217 nt respectively, matched with the viral RNA polymerase PB1 segment of viruses belonging to the genus *Thogotovirus* (family *Orthomyxoviridae*), with a nucleotide identity of 61.26% and 57.61%, respectively. Phylogenetic analyses enabled the identification of a clade formed by the identified thogotovirus-like orthomyxovirus, tentatively named 'Dielmo orthomyxovirus' (DOV), with a high bootstrap value of 99.2 and a high posterior probability of 1 (Figure 4A). The clade formed by DOV, placed at the root of the group formed by viruses belonging to the *Thogotovirus* genus, is supported by high bootstrap value and posterior probability, suggesting that DOV could constitute either a novel species within the *Thogotovirus* genus or a novel genus within the *Orthomyxoviridae* family (Figure 4A). However, analyses of genetic distances between DOV and other orthomyxoviruses supported the classification of DOV among the *Thogotovirus* genus rather than a new genus since it presented similar distances

with other thogotoviruses and distances in the same range as those observed between other thogotoviruses and *Influenzavirus* genus (Figure 4B).

Novel *Rhabdoviridae* genus

Within the virome of the STE0043 *Culicoides sp*. and STE0044 *C. imicola* engorged female samples, large contigs of 1,397 nt and 1,572 nt, respectively, matched with the viral RNA polymerase of North Creek virus (NCV), a novel rhabdovirus detected in Australian mosquito metagenomes [24]. The new Senegalese rhabdovirus, tentatively named 'Dielmo rhabdovirus' (DRV) was distant from North Creek virus, with only 62.61% and 61.06% of nucleotide homologies, respectively. Nucleotide and amino-acid sequences of STE0043 and STE0044 Dielmo rhabdovirus were 100% identical, while they presented a genetic distance from Australian mosquito North Creek virus of 0.352 and 0.377 in nucleotide and amino-acid sequences, respectively.

In order to identify whether DRV could either constitute a novel species or a novel genus within the *Rhabdoviridae* family, we selected GenBank sequences according to the Walker *et al.* dataset [25] in order to clearly identify recognised or putative *Rhabdoviridae* genera (Figure 5). Phylogenetic analysis identified a clade (sub-clade I) formed by biting midge DRV and Australian mosquito NCV, with a high bootstrap value of 99 and a high posterior probability of 1. Beaumont virus, another rhabdovirus identified in Australian mosquito metagenomes [24] and Culex tritaeniorhynchus rhabdovirus (CTR virus), identified in Japanese mosquitoes [26] formed a sub-clade II at the root of sub-clade I (Figure 5, Figure S2). This group, consisting of the two sub-clades, could constitute a novel genus within the *Rhabdoviridae* family (Figure 5, Figure S2). This putative genus was tentatively named *Dielmovirus* genus. Dielmoviruses belong to the Dimarhabdovirus supergroup (dipteran-mammal rhabdoviruses) (Figure S2).

The genetic distances of *Dielmovirus* genus compared to other *Rhabdoviridae* genera, as defined by Walker *et al.* [25], are presented in Figure 6. The mean genetic distance between viruses within the *Dielmovirus* genus is higher than that observed within each recognised or putative genus (Figure 6A), with the exception of Sigmaviruses, supporting the distinction of two sub-clades within the *Dielmovirus* genus: one formed by NCV and DRV, and the other composed of Beaumont and CTR viruses. In addition, the putative *Dielmovirus* genus presented a distribution of distances with other genera in the same range than the global distribution of distances observed between other genera (Figure 6B). Viruses

belonging to the *Dielmovirus* genus diverge by approximately 15% to 26% in the amino-acid sequence of the RNA-dependant RNA polymerase from other *Rhabdoviridae* genera, which is globally observed for all other genera with the exception of the *Lyssavirus*, *Almendravirus*, *Bahiavirus* and *Sawgravirus* genera, which seemed to present a higher genetic distance (Figure 6B). Interestingly, these four recognised and putative genera did not belong to the Dimarhabdovirus supergroup (Figure 6B, Figure S2). The *Sigmavirus* genus presented the lowest distance with *Dielmovirus*, and *Bahiavirus* presented the highest distance, which is consistent with phylogenetic observations.

Other arboviruses detected

Within the virome of the STE0043 *Culicoides sp*. sample, one contig of 609 nt matched with the NS5 segment of Jingmen Tick virus (JTV), a novel chimerical virus isolated in Chinese ticks and composed of four segments: two originating from a flavivirus (NS3 and NS5-like segments) and two with high similarities with *Toxocara canis* nematode cDNA library [27]. The Senegalese biting midge Jingmen Tick-like virus (JTV-like virus) presented a low nucleotide identity of 57.95% with the JTV NS5 segment. Phylogenetic analysis of several representative flaviviruses, JTV and Mogiana tick virus (MTV, another virus isolated in ticks which has similarities with flaviviruses [28]) performed in the NS5 gene revealed that the Senegalese JTV-like virus was located at the root of a clade formed by these new flavi-like viruses with a high bootstrap value of 97 and a high posterior probability of 1. This clade does not belong to the *Flavivirus* genus (posterior probability of 1 for the node defining this clade apart from the *Flavivirus* genus clade), but belongs to the *Flaviviridae* family (Figure 7).

In addition, by re-analysing contigs with low identities and coverage that were previously discarded, we detected one contig which matched the JTV NS3 segment with a homology percentage of 41.3% in nucleotide and an E-value of 10^{-9}, and a contig which matched *Toxocara canis* ANT-5 with an E-value of 10^{-59} and homology of 34.18%.

Presence of endogenous viral elements?

To verify the presence of Endogenous Viral Elements (EVE) within the major detected arboviruses, we screened for the presence of possible integration sites within the viral contigs. Among the 3' portion of the JTV-like viral contig, only 23 nt did not match with a viral sequence but matched with *Ovis canadensis* chromosome 25. We were not able to detect similar sequences in other viral contigs.

In addition, and due to reports of a rhabdoviral EVE in mosquitoes [29], we performed a Bayesian inference phylogenetic analysis of the *Dielmovirus* rhabdovirus genus compared to other rhabdoviruses and *Rhabdoviridae*-related EVEs, which confirmed that Dielmoviruses did not correspond to the previously identified *A. aegyti* RdRP-related EVE (Figure S3).

Finally, the presence of EVEs in the glycoprotein gene of *Orthomyxoviridae* in the genome of *Ixodes scapularis* ticks had previously been reported [30], but not among the PB1 segment of the RdRP detected in our biting midge orthomyxovirus.

Other viruses present in biting midges

Sequencing the viral communities of Senegalese biting midges revealed the presence of viruses infecting a wide variety of hosts, including mammals, insects, plants and bacteria.

Mammalian-infecting viruses were only detected in STE0045 *C. imicola* male and non-engorged female RNA viromes and consisted of 33.73% of total viral reads (Figure 1B). The viral family which was most represented was *Picobirnaviridae* (57.59% of all mammalian viral reads). Interestingly, the *Picobirnaviridae*-related contig matched with a feline picobirnavirus with a nucleotide identity of 53.91%, suggesting the presence of a possible new picobirnavirus either originating from Culicoides or from animals on which arthropods feed.

Insect-specific viruses were also highly abundant in the viromes, representing 55.51%, 76.23% and 33.66% of total viral reads in the STE0043, STE0044 and STE0045 RNA viromes, respectively (Figure 1B). *Iflaviridae* were abundant, but most insect-specific viral reads matched with novel viruses, currently not recognised by the International Committee for Taxonomy of Viruses (ICTV). Indeed, sequences matching the Loreto virus, Negev virus and Negev-like virus 174, Piura virus and Nora virus were retrieved, with low nucleotide identities comprised between 50-56%, 50-60%, 63-74%, 51-63% and 67-69%, respectively.

Plant-infecting viruses belonging to the *Partitiviridae* and *Tymoviridae* families were detected in the metagenomes, mainly in the STE0043 and STE0044 viromes. Several reads related to bacteriophages were also detected in the STE0044 and STE0045 samples, and amoeba-infecting giant viral sequences were identified in the STE0045 virome, probably reflecting a residual contamination of DNA in the RNA preparations (Figure 1B) or the carriage of mRNAs within viral particles.

Discussion

In an attempt to identify potential zoonotic viruses transmitted to humans via Culicoides bites, we report the sequencing of RNA viral communities of Senegalese biting midges. Analysis of the taxonomic assignment of reads revealed a high proportion of unknown sequences. This result, in the same range as those observed in tick [10] and mosquito [8] metagenomes, again reflects the lack of data from RNA viruses in the databases, and highlights the potential pool of unknown viruses yet to be discovered and which could represent future emerging human viruses.

The pattern of composition of RNA viral communities was highly divergent in terms of relative abundance and of the composition of viruses within the three metagenomes, although arthropods were collected at the same place during the same night in the same trap. This suggests that these differences may result from intrinsic characteristics of the insects rather than the environment. However, the three biting midge viromes clustered together in the principal component analysis when compared to other haematophagous and non-haematophagous arthropods, suggesting the presence of a 'core' viral community shared by all biting midges, and 'accessory' viral communities specific to a species, gender or haematophagous status. Indeed, STE0043 *Culicoides sp*. was more distant than the STE0045 pool of *C. imicola* males and non-engorged females and the STE0044 pool of *C. imicola* engorged females, despite the fact that they differ only by arthropod species composition. In addition, biting midge viromes were closer to other haematophagous arthropods than to non-haematophagous arthropods, potentially highlighting the influence of blood meal in the composition of viromes.

Orbivirus-related sequences were the most represented in the viromes. These *Reoviridae* are livestock-restricted viruses which cause severe economic losses: African horse sickness (AHSV) causes malfunctions of the circulatory and respiratory systems leading to the death of equines while bluetongue virus (BTV) and epizootic haemorrhagic disease (EHDV) cause significant decreases in milk production and death in ruminants [15,31-33]. In Europe and Africa, the main vector of AHSV and BTV is *C. imicola,* while EHDV is transmitted by the *C. schultzei* group in Africa [15]. In Senegal, several Culicoides species are present: the *C. imicola*, *C. schultzei*, *C. milnei*, *C. magnus* and *C. fulvithorax* groups [34], which can represent a potential epizootic risk to livestock. In 2007, Senegal reported a significant AHSV epidemic among equines and since then, animals have been vaccinated [35]. BTV also highly

circulates among ruminants, as shown in sero-epidemiological studies [36], although no recent epidemics have been reported. In addition, to our knowledge, no EHD epidemic or study has been reported in Senegal, but the symptoms of BTV or EHDV infections are highly similar, resulting in a possible wrong diagnosis of an etiology to a bluetongue-like pathology [15]. In this study we reported the detection of sequences related to AHSV, BTV and EHDV *Reoviridae* viruses. The STE0043 pool of *Culicoides sp*. presented the majority of *Reoviridae* reads, and within them, AHSV, BTV and EHDV represented 15.3%, 34.0% and 50.7% of total *Reoviridae*-related sequences, respectively. Interestingly, only a few BTV reads were detected in the STE0044 *C. imicola* engorged female and STE0045 *C. imicola* males and non-engorged females despite the fact that *C. imicola* is known to be the main vector of this virus, suggesting that other midge species could be vectors of BTV in Senegal, as demonstrated by the high prevalence of BTV in the STE0043 *Culicoides sp*. pool. In addition, the detection of EHDV-related reads matching nearly all of the viral segments suggests that the virus is probably circulating among the vector populations and may precede the onset of an outbreak. Thus, it highlights the importance of monitoring the emergence of epizooties by studying viral communities of haematophagous arthropods [3].

Arthropod-borne viruses usually originate from wildlife or domestic fauna and can adapt to humans via arthropod vectors [3]. In biting midge RNA viromes, we detected the presence of several reads related to arboviruses, including a novel thogotovirus. Dielmo Orthomyxovirus (DOV) was detected in the pool of *Culicoides sp*. and the pool of *C. imicola* males and non-engorged females. Phylogenetic analyses and calculation of genetic distance resulted in the identification of a new thogotovirus species, distinct from other known thogotoviruses. Thogotoviruses are ssRNA negative-strand segmented viruses belonging to the *Orthomyxoviridae* family. All isolated from hard ticks [37] (with the exception of the Batken virus, which was also isolated from mosquitoes [38]), thogotoviruses are able to infect a wide variety of vertebrate hosts, including birds, rodents, livestock and humans [39-45]. In humans, these viruses cause fever and, in some cases, neurological symptoms such as meningitis or encephalitis [46,47]. Recently, a novel thogotovirus, tentatively named 'Bourbon virus', was responsible for the death of an individual who had previously been bitten by a tick, due to a decrease in blood platelets and white cells but with no neurological symptoms [48]. Unfortunately, no Bourbon viral sequence matching the PB1 segment was available to compare DOV with Bourbon virus in a phylogenetic analysis. However, the

relationships between the newly described DOV and other thogotoviruses lead us to hypothesise that DOV could be capable of infecting vertebrate hosts. Future studies of the tropism of DOV in biting midges, and especially in salivary glands, would help to determine whether DOV constitutes a novel arbovirus.

Rhabdoviridae-related viral sequences were also detected. These sequences clustered together in a monophyletic group with North Creek virus, a virus recently discovered in *Culex sitiens* mosquitoes in Australia [24]. We propose that this sub-clade, in addition to another sub-clade formed by the Beaumont virus [24] and Culex tritaeniorhynchus rhabdovirus [26], form a new genus within the *Rhabdoviridae* family, tentatively named *Dielmovirus*. Dielmoviruses cluster with the *Sigmavirus* genus, within which viruses were only isolated from Drosophila flies. Many rhabdoviruses were previously isolated from biting midges [25]: for example Fukuoka virus (a cattle virus), Vesicular stomatitis New Jersey virus (a cattle virus), Wongabel virus (a seabird virus), Ngaingan virus (a cattle virus), Curionopolis virus (a primate virus) and Tibrogargan virus (a cattle virus). Nearly all of them belong to the 'arbovirus' group, with the exception of the Itacaiunas virus, which is restricted to midges and which form a distinct clade. Dielmoviruses, such as the Sigmaviruses, appear to be restricted to haematophagous (mosquitoes, biting midges) and non-haematophagous (flies) diptera, and phylogenetic analyses revealed that insect-specific rhabdoviruses form distinct monophyletic groups, suggesting that stringent host specificity occurs for these viruses (Figure 5, [25]). In contrast, arbo-rhabdoviruses, possibly due to significant host switching between vertebrate hosts and arthropod vectors, appear to be more diverse. Indeed, higher genetic distances among recognised or putative genera were observed for *Sigmavirus* and *Dielmovirus*, reinforcing the observation that strong host specificity occurs among insect-specific rhabdoviruses. Vasilakis and Tesh recently noted that insect-specific rhabdoviruses, as well as bunyaviruses and flaviviruses, are ancient and probably evolved and diversified in parallel with their insect hosts [49], via vertical transmission or integration within the host genome.

It is well-known that arthropod genomes, as well as vertebrate animals, contain integrated fragments or entire genomes of viral RNA [29,30,50-54]. These regions, called Endogenous Viral Elements (EVE, [30]), can be functional in the genomes of several hosts [54-56] and often derive from ancient viral infections for which the integration was vertically transmitted and evolve in parallel with their eukaryotic host. In our study, we demonstrated

that the newly described *Thogotovirus* species and *Rhabdoviridae* genera did not correspond to previously reported related EVEs [29,30], suggesting that these viruses could constitute novel viral species and genus.

To conclude, this study reports the first description of viral communities of haematophagous arthropods which have an impact upon human and veterinary medicine: the Culicoides. We detected the presence of numerous arboviruses, including a novel *Thogotovirus* species and a novel *Rhabdoviridae* genus which could constitute potential risks for human and animal health. This study thus highlights the importance of monitoring the emergence of epizooties and/or zoonoses using next-generation sequencing techniques for studying viral communities of haematophagous arthropods.

Acknowledgments

The authors would like to thank Dr. Hervé Pascalis for his critical comments on phylogenetic analysis, Dr. Catherine Cêtre-Sossah for her expertise on orbiviruses, and Ti Thien Nguyen for her help in preparing the libraries. This work was partly supported through a SANOFI-PASTEUR international prize awarded to Dr. Christelle Desnues.

Author Contributions

S.T., O.M. and C.D. conceived and designed the experiments; S.T., S.M.B., C.R., J.P.B., and N.L. performed the experiments; S.T., M.S. and M.A.L. collected the samples on the field, S.T. analyzed the data; S.T. wrote the paper; D.R., O.M. and C.D. reviewed the paper.

Conflicts of Interest

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Woolhouse, M.; Scott, F.; Hudson, Z.; Howey, R.; Chase-Topping, M. Human viruses: discovery and emergence. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. **2012**, *367*, 2864-2871.

2. Cutler, S.J.; Fooks, A.R.; van der Poel, W.H. Public health threat of new, reemerging, and neglected zoonoses in the industrialized world. *Emerg. Infect. Dis*. **2010**, *16*, 1-7.
3. Temmam, S.; Davoust, B.; Berenger, J.M.; Raoult, D.; Desnues, C. Viral metagenomics on animals as a tool for the detection of zoonoses prior to human infection? *Int. J. Mol. Sci*. **2014**, *15*, 10377-10397.
4. Bishop-Lilly, K.A.; Turell, M.J.; Willner, K.M.; Butani, A.; Nolan, N.M.; Lentz, S.M.; Akmal, A.; Mateczun, A.; Brahmbhatt, T.N.; Sozhammannan, S.; *et al*. Arbovirus detection in insect vectors by rapid, high-throughput pyrosequencing. *PLoS Negl. Trop. Dis*. **2010**, *4,* doi: 10.1371/journal.pntd.0000878.
5. Dinsdale, E.A.; Edwards, R.A.; Hall, D.; Angly, F.; Breitbart, M.; Brulc, J.M.; Furlan, M.; Desnues, C.; Haynes, M.; Li, L.; *et al*. Functional metagenomic profiling of nine biomes. *Nature* **2008**, *452*, 629-632.
6. Chandler, J.A.; Thongsripong, P.; Green, A.; Kittayapong, P.; Wilcox, B.A.; Schroth, G.P.; Kapan, D.D.; Bennett, S.N. Metagenomic shotgun sequencing of a Bunyavirus in wild-caught *Aedes aegypti* from Thailand informs the evolutionary and genomic history of the Phleboviruses. *Virology* **2014**, *464-465*, 312-319.
7. Chandler, J.A.; Liu, R.M.; Bennett, S.N. RNA shotgun metagenomic sequencing of northern California (USA) mosquitoes uncovers viruses, bacteria, and fungi. *Front. Microbiol*. **2015**, *6*, doi: 10.3389/fmicb.2015.00185.
8. Ng, T.F.; Willner, D.L.; Lim, Y.W.; Schmieder, R.; Chau, B.; Nilsson, C.; Anthony, S.; Ruan, Y.; Rohwer, F.; Breitbart, M. Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS One* **2011**, *6*, doi: 10.1371/journal.pone.0020579.
9. Tokarz, R.; Williams, S.H.; Sameroff, S.; Sanchez Leon, M.; jain, K.; Lipkin, W.I. Virome analysis of *Amblyomma americanum*, *Dermacentor variabilis*, and *Ixodes scapularis* ticks reveals novel highly divergent vertebrate and invertebrate viruses. *J. Virol*. **2014**, *88*, 11480-11492.
10. Xia, H.; Hu, C.; Zhang, D.; Tang, S.; Zhang, Z.; Kou, Z.; Fan, Z.; Bente, D.; Zeng, C. Li, T. Metagenomic profile of the viral communities in Rhipicephalus spp. ticks from Yunnan, China. *PLoS One* **2015**, *10*, doi: 10.1371/journal.pone.0121609.
11. Borkent, A. World species of biting midges (*Diptera*: *Ceratopogonidae*). **2014**, available online: http://wwx.inhs.illinois.edu/files/9913/9144/3328/CeratopogonidaeCatalog.pdf (accessed on 20 October 2015).

12. Borkent, A. The subgeneric classification of species of Culicoides-thoughts and a warning. **2014**, available online: http://wwx.inhs.illinois.edu/files/9613/9136/7590/CulicoidesSubgenera.pdf (accessed on 20 October 2015).
13. Bassene, H.; Sambou, M.; Fenollar, F.; Clarke, S.; Djiba, S.; Mourembou, G.; L Y, A.B.; Raoult, D.; Mediannikov, O. High prevalence of *Mansonella perstans* filariasis in rural Senegal. *Am. J. Trop. Med. Hyg*. **2015**, *93*, 601-606.
14. Agbolade, O.M.; Akinboye, D.O.; Olateju, T.M.; Ayanbiyi, O.A.; Kuloyo, O.O.; Fenuga, O.O. Biting of anthropophilic *Culicoides fulvithorax* (*Diptera*: *Ceratopogonidae*), a vector of *Mansonella perstans* in Nigeria. *Korean J. Parasitol*, **2006**, *44*, 67-72.
15. Mellor, P.S.; Boorman, J.; Baylis, M. Culicoides biting midges: their role as arbovirus vectors. *Annu. Rev. Entomol*. **2000**, *45*, 307-340.
16. Tesh, R.B. The emerging epidemiology of Venezuelan hemorrhagic fever and Oropouche fever in tropical South America. *Ann. N. Y. Acad. Sci*. **1994**, *740*, 129-137.
17. Temmam, S.; Monteil-Bouchard, S.; Robert, C.; Pascalis, H.; Michelle, C.; Jardot, P.; Charrel, R.; Raoult, D.; Desnues, C. Host-Associated Metagenomics: A Guide to Generating Infectious RNA Viromes. *PLoS One* **2015**, *10*, doi: 10.1371/journal.pone.0139810.
18. Froussard, P. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Res*. **1992**, *20*, 2900.
19. Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.M.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A. *et al*. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **2008**, *9*, doi: 10.1186/1471-2105-9-386.
20. Edgar, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **2004**, *5*, doi: 10.1186/1471-2105-5-113.
21. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol*. **2013**, *30*, 2725-2729.
22. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572-1574.

23. Sime-Ngando, T.; Mignot, J.P.; Amblard, C.; Bourdier, G.; Desvilettes, C.; Quiblier-Lloberas, C. Characterization of planktonic virus-like particles in a French mountain lake: methodological aspects and preliminary results. *Annls. Limnol*. **1996**, *32*, 259-263.
24. Coffey, L.L.; Page, B.L.; Greninger, A.L.; Herring, B.L.; Russell, R.C.; Doggett, S.L.; Haniotis, J.; Wang, C.; Deng, X; Delwart, E.L. Enhanced arbovirus surveillance with deep sequencing: Identification of novel rhabdoviruses and bunyaviruses in Australian mosquitoes. *Virology* **2014**, *448*, 146-158.
25. Walker, P.J.; Firth, C.; Widen, S.G.; Blasdell, K.R.; Guzman, H.; Wood, T.G.; Paradkar, P.N.; Holmes, E.C.; Tesh, R.B.; Vasilakis, N. Evolution of genome size and complexity in the *Rhabdoviridae*. *PLoS Pathog*. **2015**, *11*, doi: 10.1371/journal.ppat.1004664.
26. Kuwata, R.; Isawa, H.; Hoshino, K.; Tsuda, Y.; Yanase, T.; Sasaki, T.; Kobayashi, M.; Sawabe, K. RNA splicing in a new rhabdovirus from Culex mosquitoes. *J. Virol*. **2011**, *85*, 6185-6196.
27. Qin, X.C.; Shi, M.; Tian, J.H.; Lin, X.D.; Gao, D.Y.; He, J.R.; Wang, J.B.; Li, C.X. Kang, Y.J.; Yu, B.; et al. A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors. *Proc. Natl. Acad. USA* **2014**, *111*, 6744-6749.
28. Maruyama, S.R.; Castro-Jorge, L.A.; Ribeiro, J.M.; Gardinassi, L.G.; Garcia, G.R.; Brandão, L.G.; Rodrigues, A.R.; Okada, M.I.; Abrão, E.P.; Ferreira, B.R.; *et al*. Characterisation of divergent flavivirus NS3 and NS5 protein sequences detected in Rhipicephalus microplus ticks from Brazil. *Mem. Inst. Oswaldo Cruz* **2014**, *109*, 38-50.
29. Fort, P.; Albertini, A.; Van-Hua, A.; Berthomieu, A.; Roche, S.; Delsuc, F.; Pasteur, N.; Capy, P.; Gaudin, Y.; Weill, M. Fossil rhabdoviral sequences integrated into arthropod genomes: ontogeny, evolution, and potential functionality. *Mol. Biol. Evol*. **2012**, *29*, 381-390.
30. Katzourakis, A.; Gifford, R.J. Endogenous viral elements in animal genomes. *PLoS Genet*. **2010**, *6*, doi: 10.1371/journal.pgen.1001191.
31. Carpenter, S.; Groschup, M.H.; Garros, C.; Felippe-Bauer, M.L.; Purse, B.V. Culicoides biting midges, arboviruses and public health in Europe. *Antiviral Res*. **2013**, *100*, 102-113.
32. Kedmi, M.; Van Straten, M.; Ezra, E.; Galon, N.; Klement, E. Assessment of the productivity effects associated with epizootic hemorrhagic disease in dairy herds. *J. Dairy Sci*. **2010**, *93*, 2486-2495.

33. Nusinovici, S.; Souty, C.; Seegers, H.; Beaudeau, F.; Fourichon, C. Decrease in milk yield associated with exposure to bluetongue virus serotype 8 in cattle herds. *J. Dairy Sci*. **2013**, *96*, 877-888.
34. Sambou, M.; Aubadie-Ladrix, M.; Fenollar, F.; Fall, B.; Bassene, H.; Almeras, L.; Sambe-Ba, B.; Perrot, N.; Chatellier, S.; Faye, N.; *et al*. Comparison of matrix-assisted laser desorption ionization-time of flight mass spectrometry and molecular biology techniques for identification of Culicoides (*Diptera*: *Ceratopogonidae*) biting midges in Senegal. *J. Clin. Microbiol*. **2015**, *53*, 410-418.
35. Diouf, N.D.; Etter, E.; Lo, M.M.; Akakpo, A.J. Outbreaks of African horse sickness in Senegal, and methods of control of the 2007 epidemic. *Vet. Rec*. **2013**, *172*, doi: 10.1136/vr.101083.
36. Lefèvre, P.C.; Calvez, D. La fièvre catarrhale du mouton (bluetongue) en Afrique intertropicale : influence des facteurs écologiques sur la prévalence de l'infection. *Rev. Elev. Med. Vét. Pays Trop*. **1986**, *39*, 263-268.
37. Karabatsos, N. *International catalogue of arboviruses, including certain other viruses of vertebrates*. 3rd ed. San Antonio, Texas: American Society of Tropical Medicine and Hygiene for The Subcommittee on Information Exchange of the American Committee on Arthropod-borne Viruses, USA, 1985.
38. Frese, M.; Weeber, M.; Weber, F.; Speth, V.; Haller, O. Mx1 sensitivity: Batken virus is an orthomyxovirus closely related to Dhori virus. *J. Gen. Virol*. **1997**, *78*, 2453-2458.
39. Hubálek, Z.; Rudolf, I.; Nowotny, N. Arboviruses pathogenic for domestic and wild animals. *Adv. Virus Res*. **2014**, *89*, 201-275.
40. Hubálek, Z.; Rudolf, I. Tick-borne viruses in Europe. *Parasitol. Res*. **2012**, *111*, 9-36.
41. Filipe, A.R.; Calischer, C.H.; Lazuick, J. Antibodies to Congo-Crimean haemorrhagic fever, Dhori, Thogoto and Bhanja viruses in southern Portugal. *Acta Virol*. **1985**, *29*, 324-328.
42. Ogen-Odoi, A.; Miller, B.R.; Happ, C.M.; Maupin, G.O.; Burkot, T.R. Isolation of thogoto virus (*Orthomyxoviridae*) from the banded mongoose, *Mongos mungo* (*Herpestidae*), in Uganda. *Am. J. trop. Med. Hyg*. **1999**, *60*, 439-440.
43. Da Silva, E.V.; Da Rosa, A.P.; Nunes, M.R., Diniz, J.A.; Tesh, R.B.; Cruz, A.C.; Vieira, C.M.; Vasconcelos, P.F. Araguari virus, a new member of the family *Orthomyxoviridae*: serologic, ultrastructural, and molecular characterization. *Am. J. trop. Med. Hyg*. **2005**, *73*, 1050-1058.

44. Lutomiah, J.; Musila, L.; Makio, A.; Ochieng, C.; Koka, H.; Chepkorir, E.; Mutisya, J.; Mulwar, F.; Khamadi, S.; Miller, B.R.; *et al*. Ticks and tick-borne viruses from livestock hosts in arid and semiarid regions of the eastern and northeastern parts of Kenya. *J. Med. Entomol.* **2014**, *51*, 269-277.
45. Briese, T.; Chowdhary, R.; Travassos da Rosa, A.; Hutchison, S.K.; Popov, V.; Street, C.; Tesh, R.B.; Lipkin, W.I. Upolu virus and Aransas Bay virus, two presumptive bunyaviruses, are novel members of the family *Orthomyxoviridae*. *J. Virol*. **2014**, *88*, 5298-5309.
46. Moore, D.L.; Causey, O.R.; Carey, D.E.; Reddy, S.; Cooke, A.R.; Akinkugbe, F.M.; David-West, T.S.; Kemp, G.E. Arthropod-borne viral infections of man in Nigeria, 1964-1970. *Ann. Trop. Med. Parasitol*. **1975**, *69*, 49-64.
47. Butenko, A.M.; Leshchinskaia, E.V.; Semashko, I.V.; Donets, M.A.; Mart'ianova, L.I. Dhori virus--a causative agent of human disease. 5 cases of laboratory infection. *Vopr. Virusol*. **1987**, *32*, 724-729.
48. Kosoy, O.I. Lambert, A.J.; Hawkinson, D.J.; Pastula, D.M.; Goldsmith, C.S.; Hunt, D.C.; Staples, J.E. Novel thogotovirus associated with febrile illness and death, United States, 2014. *Emerg. Infect. Dis*. **2015**, *21*, 760-764.
49. Vasilakis, N.; Tesh; R.B. Insect-specific viruses and their potential impact on arbovirus transmission. *Curr. Opin. Virol*. **2014**, *15*, 69-74.
50. Li, C.X.; Shi, M.; Tian, J.H.; Lin, X.D.; Kang, Y.J.; Chen, L.J.; Qin, X.C.; Xu, J.; Holmes, E.C.; Zhang, Y.Z. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *Elife* **2015**, *4*, doi: 10.7554/eLife.05378.
51. Cui, J.; Holmes, E.C. Endogenous RNA viruses of plants in insect genomes. *Virology* **2012**, *427*, 77-79.
52. Ballinger, M.J.; Bruenn, J.A.; Hay, J.; Czechowski, D.; Taylor, D.J. Discovery and evolution of bunyavirids in arctic phantom midges and ancient bunyavirid-like sequences in insect genomes. *J. Virol*. **2014**, *88*, 8783-8794.
53. Feschotte, C.; Gilbert, C. Endogenous viruses: insights into viral evolution and impact on host biology. *Nat. Rev. Genet*. **2012**, *13*, 283-296.
54. Liu, H.; Fu, Y.; Xie, J.; Cheng, J.; Ghabrial, S.A.; Li, G.; Peng, Y.; Yi, X.; Jiang, D. Widespread endogenization of densoviruses and parvoviruses in animal and human genomes. *J. Virol*. **2011**, *85*, 9863-9876.

55. Boller, K.; Schönfeld, K.; Lischer, S.; Fisher, N.; Hoffmann, A.; Kurth, R.; Tönjes, R.R. Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles. *J. Gen. Virol*. **2008**, *89*, 567-572.
56. Bézier, A.; Louis, F.; Jancek, S.; Periquet, G.; Thézé, J.; Gyapay, G.; Musset, K.; Lesobre, J.; Lenoble, P.; Dupuy, C.; *et al*. Functional endogenous viral elements in the genome of the parasitoid wasp *Cotesia congregata*: insights into the evolutionary dynamics of bracoviruses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2013**, *368*, doi: 10.1098/rstb.2013.0047.
57. Rosario, K.; Capobianco, H.; Ng, T.F.; Breitbart, M.; Polston, J.E. RNA viral metagenome of whiteflies leads to the discovery and characterization of a whitefly-transmitted carlavirus in North America. *PLoS One* **2014**, *9*, doi: 10.1371/journal.pone.0086748.
58. Sparks, M.E.; Gundersen-Rindal, R.E. The *Lymantria dispar* IPLB-Ld652Y cell line transcriptome comprises diverse virus-associated transcripts. *Viruses* **2011**, *3*, 2339-2350.

Tables

Table 1. Virome dataset statistics**.**

	STE0043 ***Culicoides sp.***	**STE0044** ***C. imicola* engorged ♀**	**STE0045** ***C. imicola* non engorged ♂♀**
Raw reads	2,071,144	1,335,388	1,507,966
Cleaned reads *including* :	**2,069,117**	**1,332,764**	**1,505,902**
- Paired reads	2,067,394	1,330,424	1,504,072
- Single reads	1,723	2,340	1,830
Raw read size (nt)	301	301	301
Cleaned read size (nt)	244	244	241
Contigs	**1,849**	**1,139**	**1,134**
Average contig length (nt)	560	536	477
Singletons	48,173	31,804	35,630
MG-RAST no	4604249.3	4604250.3	4604251.3

Figures

Figure 1. Taxonomic assignment of reads **A.** BlastN search against the NCBI nucleotide database (dashes correspond to the arthropod-borne proportion of eukaryotic reads) **B.** Relative abundance of viral families in biting midge metagenomes according to their target hosts (Green: plant viruses, Brown: insect viruses, Grey: bacteriophages, Red: arboviruses, Yellow: mammalian viruses, Blue: amoeba-infecting giant viruses).

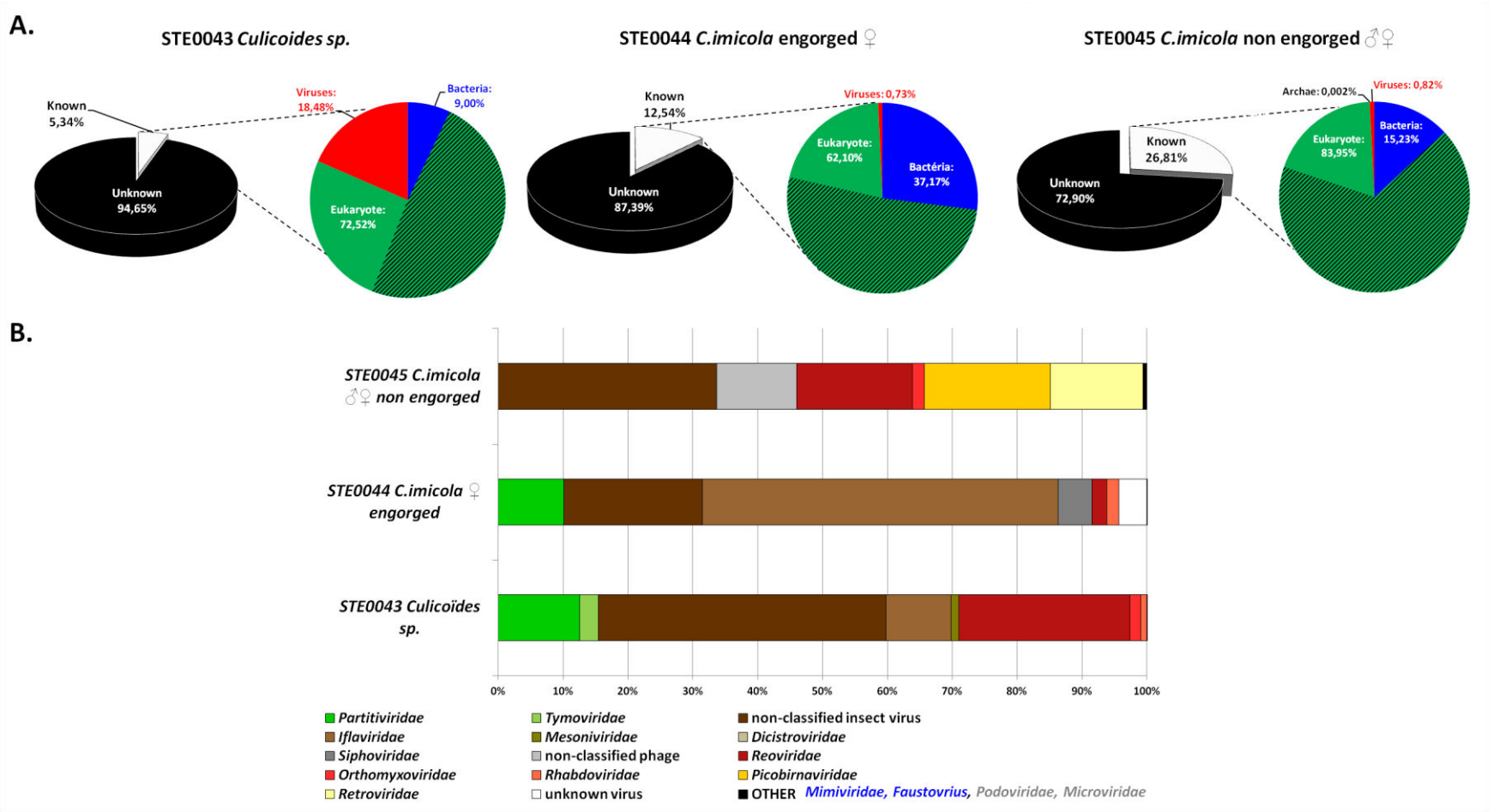


Figure 2. Repertory of transmission electron microscopy images of *Culicoides sp.* viral communities

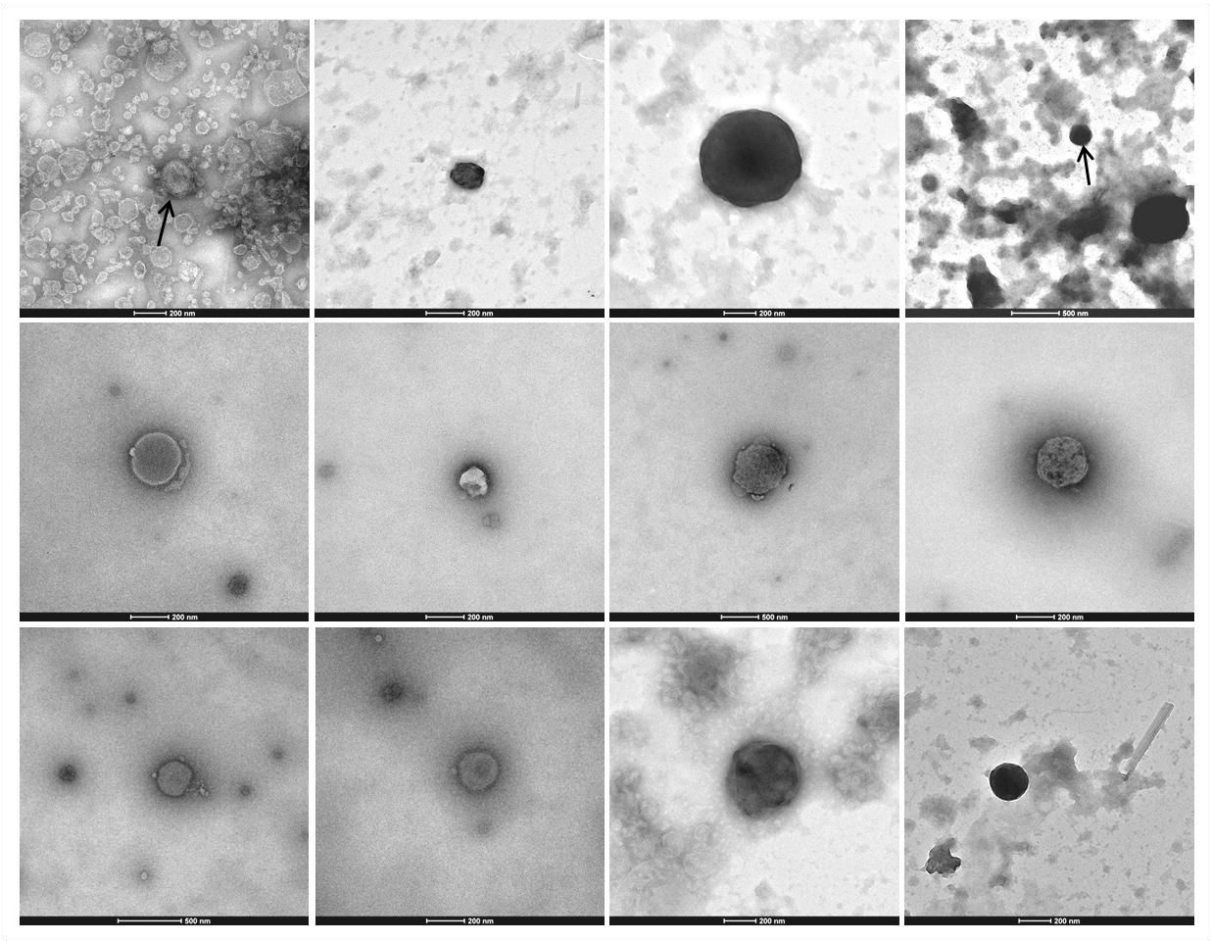


Figure 3. Comparison between viromes of biting midges with available arthropod RNA metagenomes based on a taxonomic classification of reads. Principal Component Analysis (PCA) was used to compare data in MG-RAST server [19] with a maximum E-value of 10^{-5}, a minimum identity of 60%, and a minimum alignment length of 15 amino-acids for protein and 15 bp for RNA databases. Data were normalised to values between 0 and 1 and distances were measured using the Bray-Curtis distance matrix.

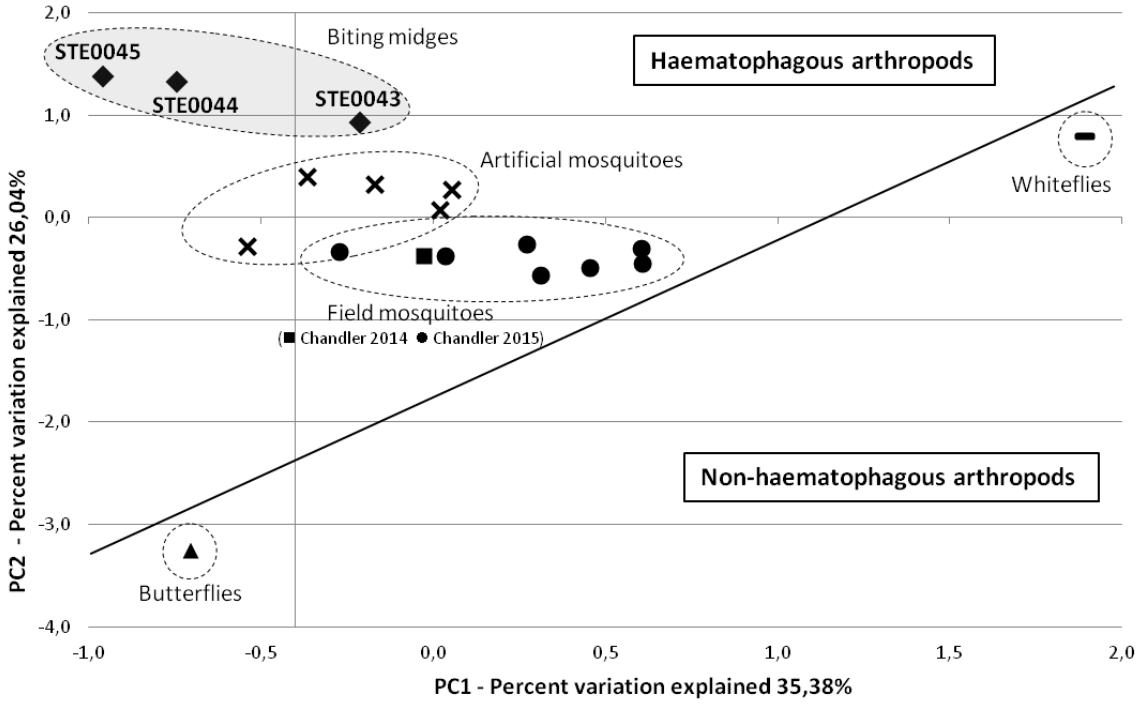
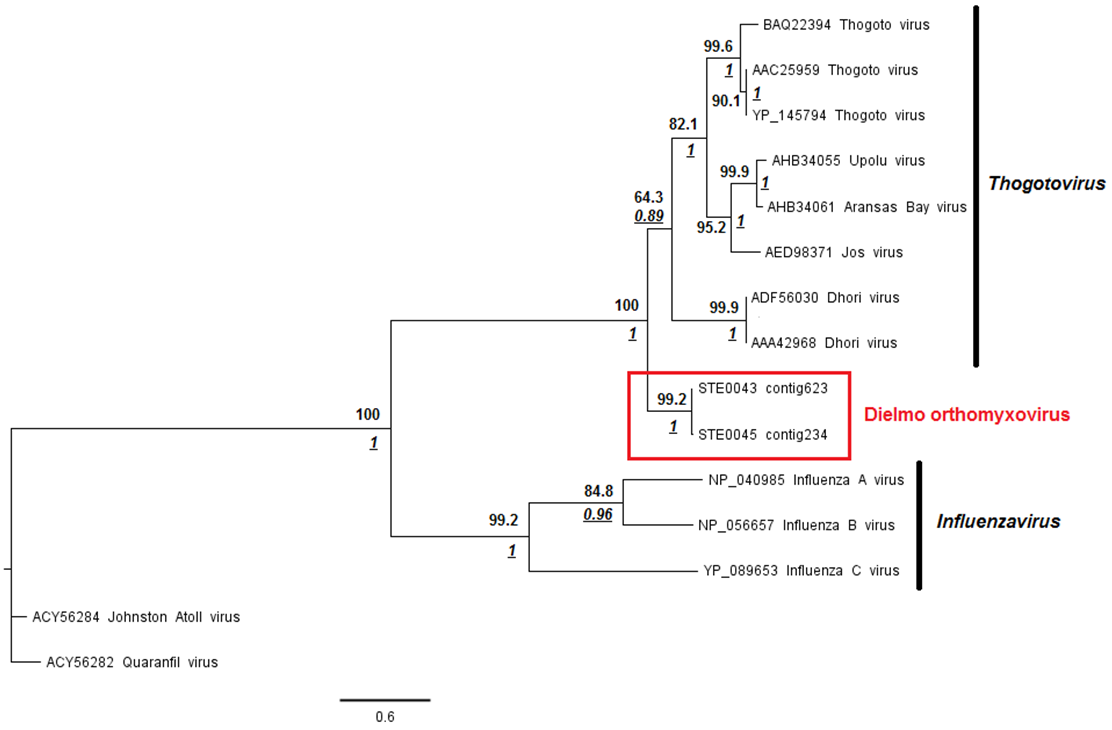


Figure 4. Phylogenetic analyses of Dielmo orthomyxovirus compared to other *Thogotovirus* viruses. **A.** Phylogenetic analysis of a fragment of 358 amino-acids of PB1. ML analysis was used to fix tree topology. ML analysis was performed on 1,000 iterations and bootstrap values are represented in bold. Bayesian posterior probabilities are underlined and represented in italics where nodes coincided with ML. Substitutions models for ML and Bayesian analyses were determined as LG+I+G and rtREV+I+G, respectively. Scale bar indicates the number of amino-acid substitutions per site. **B.** Matrix of genetic distances observed between PB1 amino-acid sequences of Dielmo orthomyxovirus and other representative thogotoviruses. Diversity was calculated by the pairwise-distance algorithm implemented through MEGA [21].

A.



B.

	Thogoto	Thogoto	Thogoto	Jos	Aransas Bay	Upolu	Dhori	Dhori	STE0043	STE0045	Influenza C	Influenza B	Influenza A	Quaranfil
YP145794 Thogoto	-													
AAC25959 Thogoto	0,000	-												
BAQ22394 Thogoto	0,111	0,111	-											
AED98371 Jos	0,248	0,248	0,289	-										
AHB34061 Aransas Bay	0,262	0,262	0,275	0,215	-									
AHB34055 Upolu	0,268	0,268	0,279	0,221	0,070	-								
ADF56030 Dhori	0,379	0,379	0,383	0,403	0,356	0,356	-							
AAA42968 Dhori	0,379	0,379	0,383	0,403	0,356	0,356	0,000	-						
STE0043	**0,379**	**0,379**	**0,376**	**0,389**	**0,399**	**0,389**	**0,383**	**0,383**	-					
STE0045	**0,383**	**0,383**	**0,379**	**0,393**	**0,403**	**0,393**	**0,386**	**0,386**	**0,003**	-				
YP089653 Influenza C	0,705	0,705	0,711	0,711	0,721	0,728	0,718	0,718	**0,725**	**0,725**	-			
NP056657 Influenza B	0,691	0,691	0,718	0,718	0,711	0,721	0,701	0,701	**0,685**	**0,685**	0,594	-		
NP040985 Influenza A	0,695	0,695	0,698	0,718	0,725	0,725	0,695	0,695	**0,681**	**0,681**	0,614	0,413	-	
ACY56282 Quaranfil	0,752	0,752	0,762	0,755	0,748	0,762	0,762	0,762	**0,758**	**0,758**	0,775	0,698	0,745	-
ACY56284 Johnston Atoll	0,762	0,762	0,772	0,762	0,752	0,765	0,768	0,768	**0,772**	**0,772**	0,765	0,708	0,738	0,211

Figure 5. Phylogenetic analysis of *Dielmovirus* genus compared to other *Rhabdoviridae*. Phylogenetic analysis of a fragment of 463 amino-acids of the RNA-dependant RNA polymerase. Bayesian inference (BI) analysis was used to fix tree topology. BI analysis was performed on 1,000,000 iterations and nodes with a posterior probability above 0.80 are represented. ML analysis was performed on 1,000 iterations and nodes above 65 are represented, when nodes coincided with BI. Recognised or a putative genera are defined as described in [25]. Substitutions models for ML and Bayesian analyses were determined as LG+I+G and rtREV+I+G, respectively. Scale bar indicates the number of amino-acid substitutions per site. Cytorhabdoviruses, Novirhabdoviruses and Nucleorhabdoviruses were excluded from the analysis because sequences were too divergent.

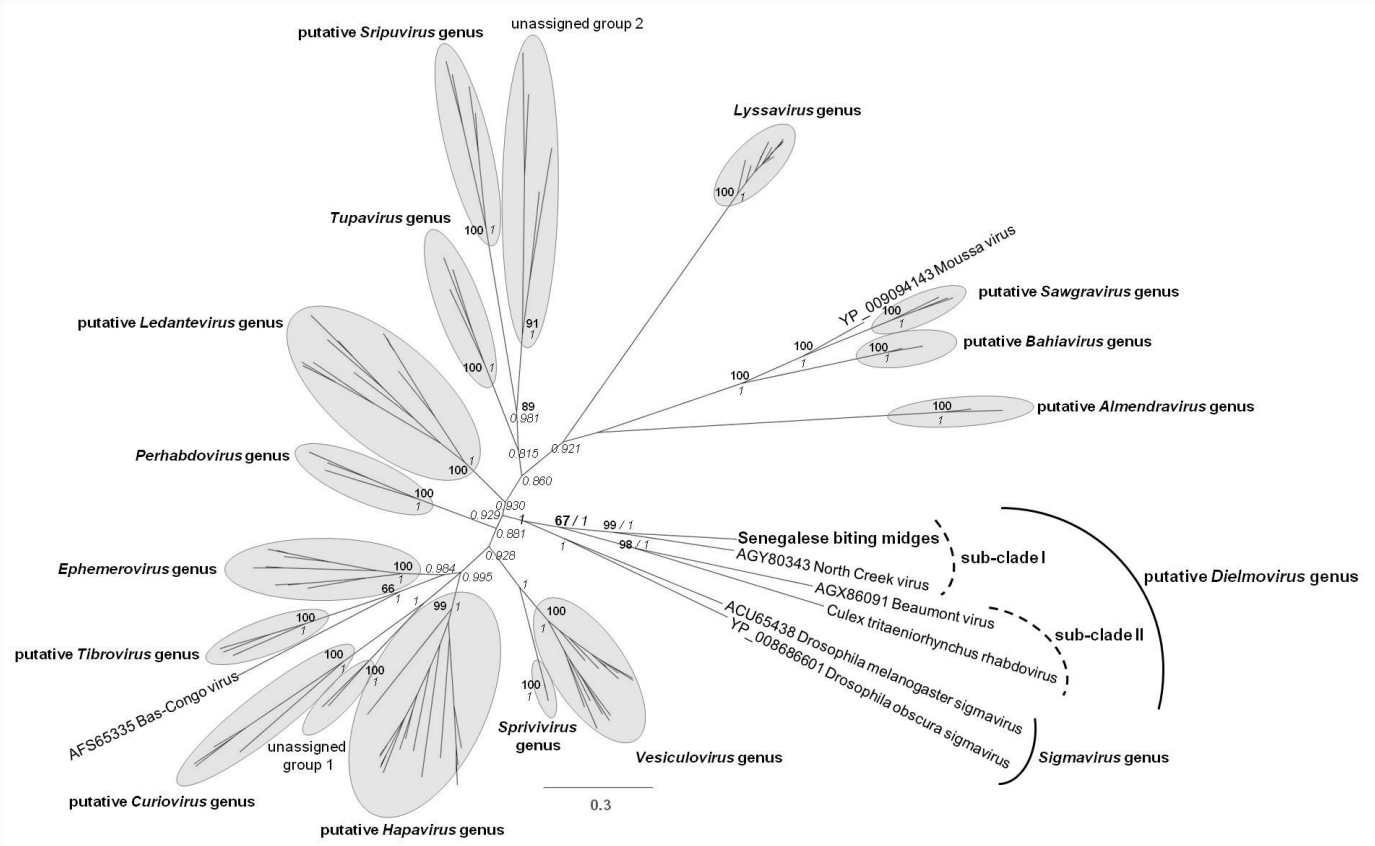


Figure 6. Genetic distances of *Dielmovirus* genus compared to other *Rhabdoviridae*. **A**. Mean distances within recognised and putative *Rhabdoviridae* genera (putative genera, as reported in [25], are indicated by a *). Diversity was calculated by the pairwise-distance algorithm implemented through MEGA6 [21], and 1,000 bootstrap replications. **B**. Distribution of distances between recognised and putative *Rhabdoviridae* genera (putative genera are indicated by a *). Diversity was calculated by the pairwise-distance algorithm implemented through MEGA6 [21].

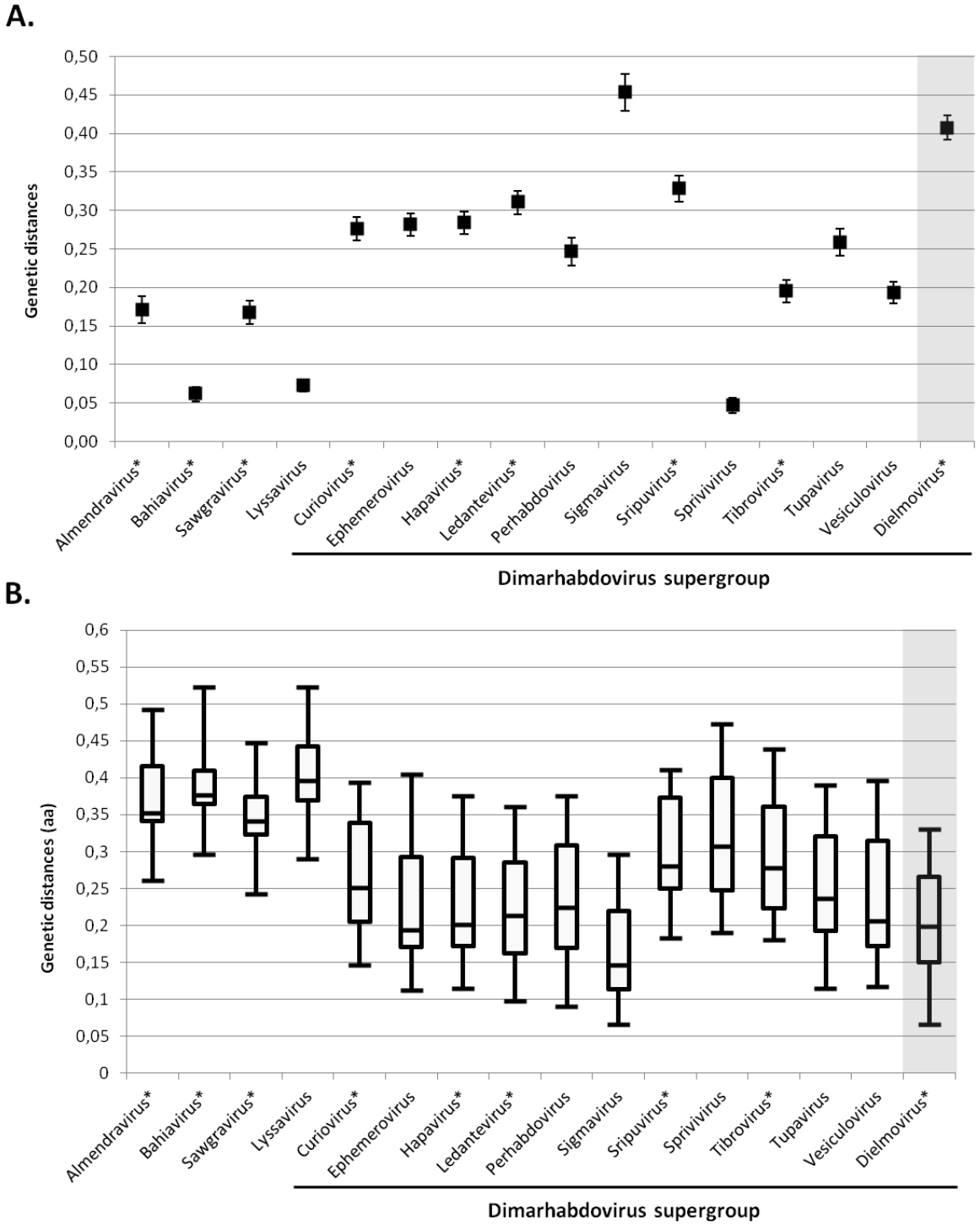
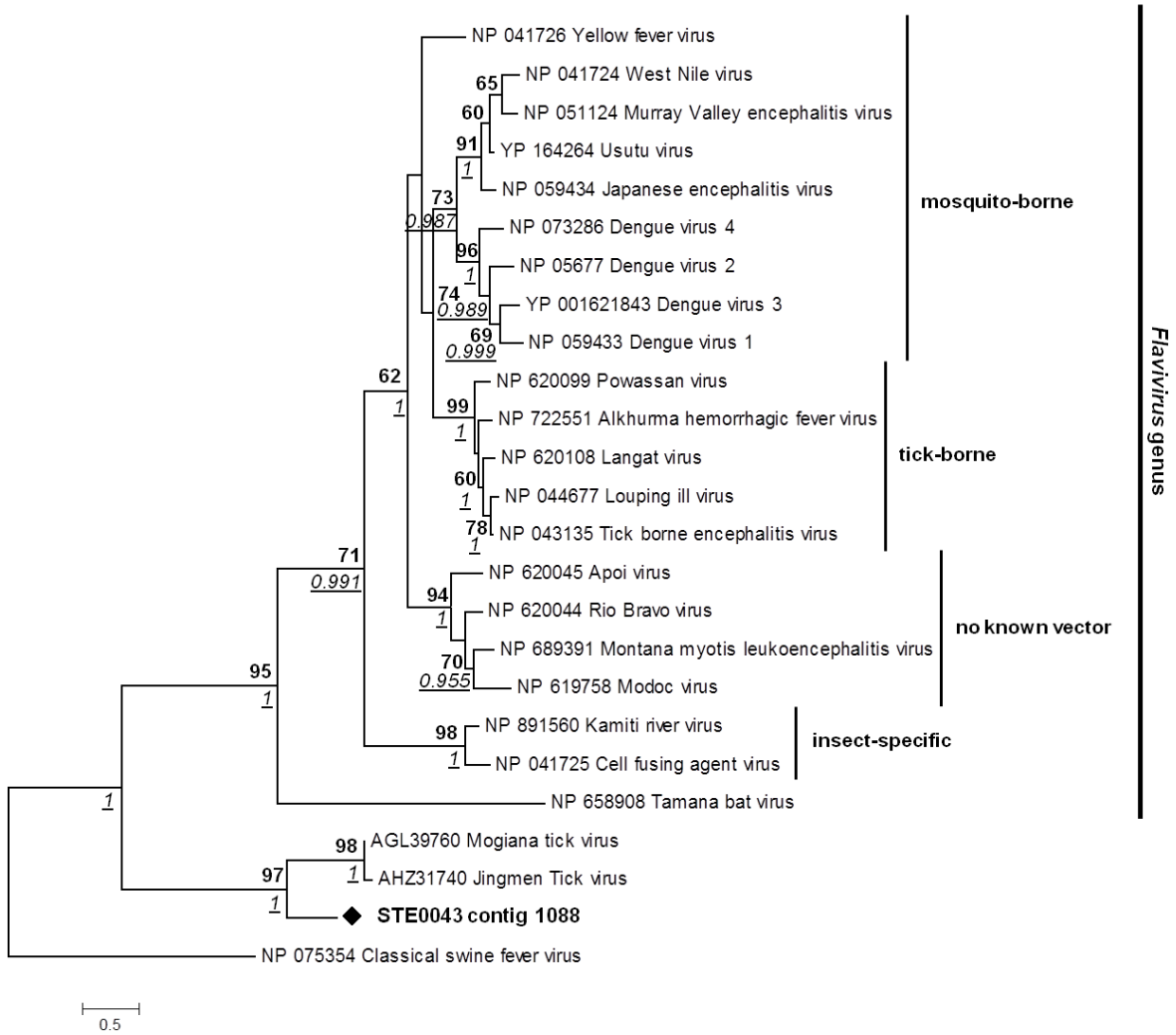


Figure 7. Phylogenetic analysis of Jingmen Tick-like virus. Phylogenetic analysis of a fragment of 319 amino-acids of the NS5 segment. ML analysis was used to fix tree topology. ML analysis was performed on 1000 iterations. Bootstrap values above 60 and posterior probabilities above 0.5 are indicated. Bayesian posterior probabilities are underlined and represented in italics where nodes coincided with ML. Substitution models for ML and Bayesian analyses were determined as LG+G and rtREV+ G, respectively. Scale bar indicates the number of amino-acid substitutions per site.



Supplementary Materials

Figure S1. Pipeline for bioinformatic analyses metagenomes.

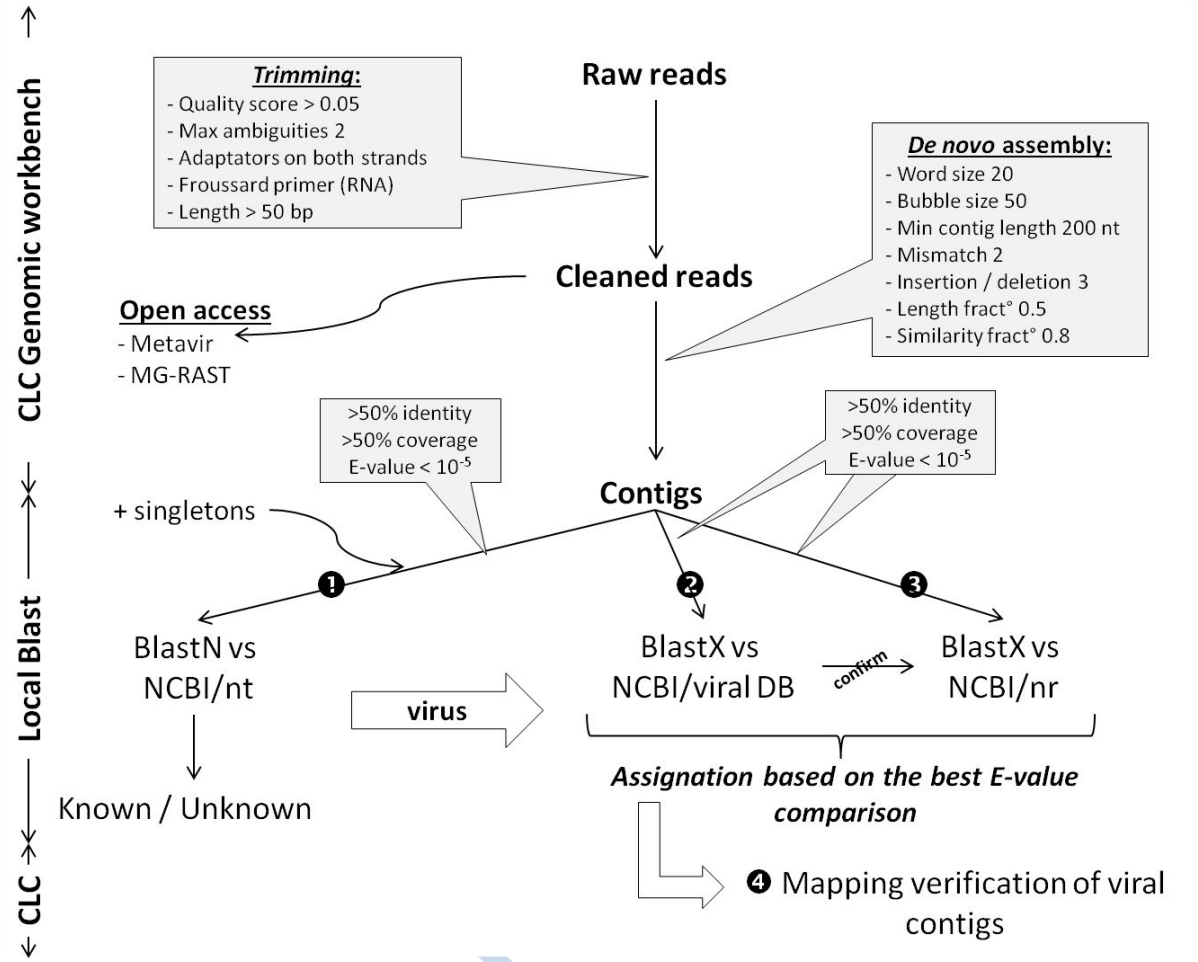


Figure S2. Phylogenetic analysis of *Dielmovirus* genus compared to other *Rhabdoviridae*. Enlarged tree of a fragment of 463 amino-acids of the RNA-dependant RNA polymerase. Bayesian inference (BI) analysis was used to fix tree topology. BI analysis was performed on 1,000,000 iterations and nodes with a posterior probability above 0.80 are represented. ML analysis was performed on 1,000 iterations and nodes above 50 are represented, when nodes coincided with BI. Recognised or putative genera are defined as described in [25]. Substitution models for ML and Bayesian analyses were determined as LG+I+G and rtREV+I+G, respectively. Scale bar indicates the number of amino-acid substitutions per site.

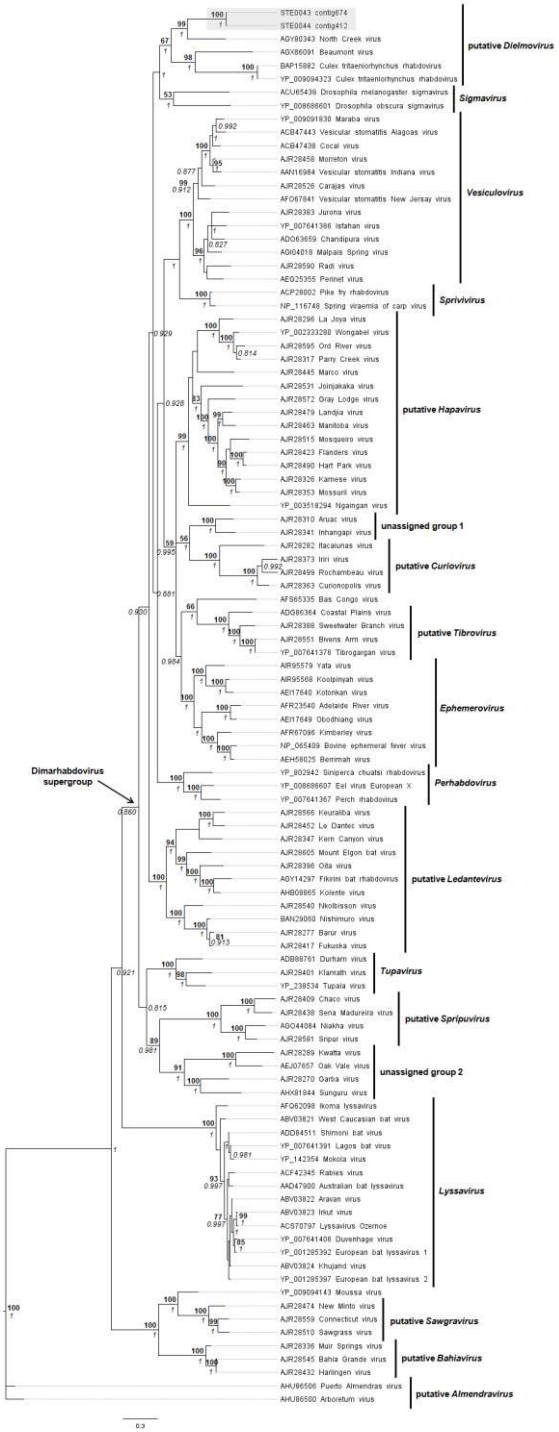


Figure S3. Phylogenetic analysis of *Dielmovirus* genus compared to other *Rhabdoviridae* and the endogenous viral element *A. aegypti*. Phylogenetic analysis of a fragment of 395 amino-acids of the RNA-dependent RNA polymerase. Bayesian inference (BI) analysis was performed on 1,000,000 iterations and nodes with a posterior probability above 0.50 are represented. Recognised or putative genera are defined as described in [25]. Substitution model was determined as rtREV+I+G. Scale bar indicates the number of amino-acid substitutions per site.

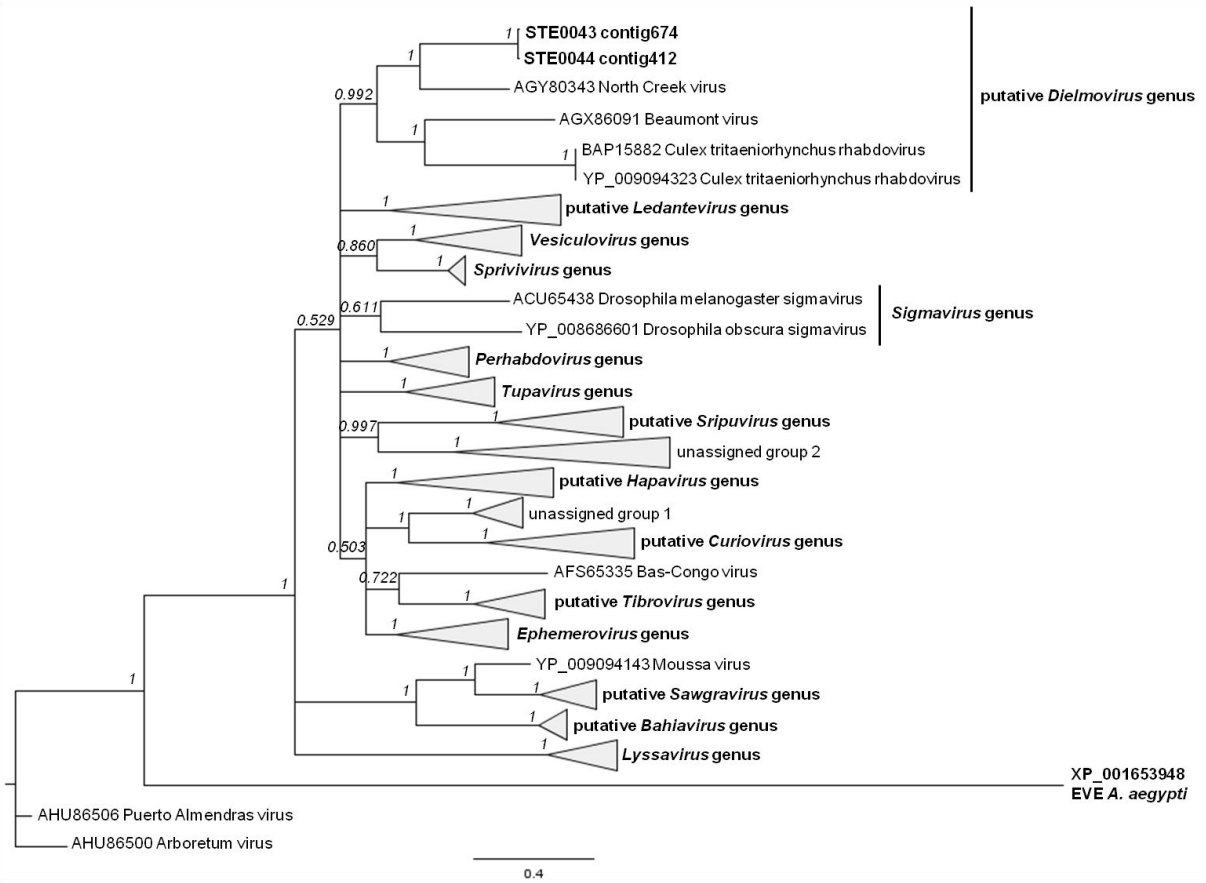


Table S1. Characteristics of metagenomes used for PCA analysis**.**

Arthropod metagenome	Haematophagous	Base no	Reads no	Sequencing method	Remarks	Ref.
Mosquitoes	yes	68,708,092	289,436	Roche 454 FLX	artificially infected with DENV-1	4
		49,582,727	216,164	Roche 454 FLX		
		53,556,733	341,650	Roche 454 FLX		
		78,813,957	390,971	Roche 454 FLX		
		67,224,921	336,822	Roche 454 FLX		
Mosquitoes	yes	16,431,897	89,744	Illumina GA II	-	6
Mosquitoes	yes	20,087,132	29,234	Illumina HiSeq 2000	assembled dataset	7
		32,513,784	44,558	Illumina HiSeq 2000		
		79,323,266	110,242	Illumina HiSeq 2000		
		44,945,652	53,542	Illumina HiSeq 2000		
		20,382,748	29,911	Illumina HiSeq 2000		
		18,912,844	13,577	Illumina HiSeq 2000		
		21,145,145	30,686	Illumina HiSeq 2000		
Whiteflies	no	63,168,431	1,427,809	Illumina GA II	-	57
Butterflies	no	37,301,814	82,099	Sanger	-	58

II. **Article n°4**: "Faustovirus-like asfarvirus in hematophagous biting midges and their vertebrate hosts."

Faustovirus-like asfarvirus in hematophagous biting midges and their vertebrate hosts.

Sarah Temmam[1], Sonia Monteil-Bouchard[1], Masse Sambou[1], Saïd Azza[1], Philippe Decloquement[1], Jacques Bou Khalil[1], Jean-Pierre Baudoin[1], Priscilla Jardot[1], Catherine Robert[1], Bernard La Scola[1,2], Oleg Mediannikov[1], Didier Raoult[1,2] and Christelle Desnues[1,*].

Published in : Frontiers in Microbiology. 6:1406. doi: 10.3389/fmicb.2015.01406

[1] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1095, Aix-Marseille Université, Marseille, France.

[2] Fondation IHU Méditerranée Infection, Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, Centre Hospitalo-Universitaire Timone, Méditerranée Infection, Assistance Publique – Hôpitaux de Marseille, Marseille, France.

[*] Corresponding author: christelle.desnues@univ-amu.fr; tel.: +33(0)491324630; fax: +33(0)491387772.

Préambule à l'article "Faustovirus-like asfarvirus in hematophagous biting midges and their vertebrate hosts"

Le terme « virus géants » désigne un ensemble de virus à génome ADN double brin formant un groupe monophylétique ayant des caractéristiques communes de taille de particule, de taille de génome et de réplication nucleo-cytoplasmique. Préalablement désignés sous le nom de NCLDV (NucleoCytoplasmic Large DNA viruses), ces virus ont été récemment reclassés sous l'ordre proposé des *Megavirales* [93-94] qui comprend des virus de la famille des *Poxviridae*, *Iridoviridae*, *Ascoviridae*, *Phycodnaviridae*, *Asfarviridae*, *Mimiviridae* et *Marseilleviridae* ainsi que les *Pandoravirus* [95] et *Pithovirus* [96] récemment décrits. Parmi ces familles virales, les amibes constituent les hôtes de réplication de prédilection de certaines d'entre elles, comme les *Mimiviridae* et *Marseilleviridae* ainsi que les *Pandoravirus*, les *Pithovirus*, et plus récemment *Faustovirus*, le premier *Asfarviridae* isolé sur amibe [97]. Ces virus associés aux amibes ont été principalement retrouvés dans des prélèvements environnementaux et sont capables d'infecter un grand nombre de protistes, voire même l'homme [98]. L'analyse métagénomique des communautés virales à génome ADN de culicoïdes hématophages collectés au Sénégal nous a permis de mettre en évidence une grande abondance et une importante diversité de séquences apparentées à des génomes de virus géants d'amibes chez des culicoïdes adultes, et principalement de *Faustovirus*, un virus géant de la famille des *Asfarviridae* récemment découvert dans des prélèvements d'égouts et se cultivant sur des amibes du genre *Hartmanella* [97].

Les analyses phylogénétiques du *Faustovirus* identifié dans les culicoïdes du Sénégal ont montré que ce virus est apparenté au *Faustovirus* d'eaux usées, et plus précisément à celui isolé dans les égouts de Dakar, au Sénégal. La microscopie électronique conduite sur les broyats d'arthropodes a présenté des particules dont la taille et la morphologie sont compatibles avec celles du *Faustovirus*. La présence de particules infectieuses de virus d'arthropodes apparenté au Faustovirus a été finalement démontrée en isolant les virions sur l'amibe *Hartmanella vermiformis*, hôte reconnu du *Faustovirus*.

La préférence trophique des arthropodes, déterminée par l'analyse du protéome du pool de *Culicoides sp*., a révélé l'origine du repas de sang des

arthropodes comme étant des bovins, des rongeurs et des humains. Par conséquent nous avons recherché la présence de *Faustovirus* à la fois dans des prélèvements de patients fébriles atteints de fièvre sans étiologie connue, mais aussi dans des sérums de bovins et des organes de rongeurs. Le virus a été détecté dans des prélèvements humains, à la fois chez des personnes fébriles et chez des personnes saines, mais sa présence n'a pas pu être confirmée par amplification sur d'autres gènes viraux. Par contre le virus a été retrouvé et confirmé chez 38% des rongeurs et 14% des bovins testés, suggérant une possible infection des culicoïdes par leur repas de sang sur des hôtes infectés.

Cette étude nous a permis de mettre en évidence le rôle potentiel de la faune sauvage, de la faune domestique et des arthropodes hématophages dans le cycle de réplication et de maintien dans l'environnement de certains virus géants infectant les amibes.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 16 December 2015
doi: 10.3389/fmicb.2015.01406

Faustovirus-Like Asfarvirus in Hematophagous Biting Midges and Their Vertebrate Hosts

*Sarah Temmam[1], Sonia Monteil-Bouchard[1], Masse Sambou[2], Maxence Aubadie-Ladrix[2], Saïd Azza[1], Philippe Decloquement[1], Jacques Y. Bou Khalil[1], Jean-Pierre Baudoin[1], Priscilla Jardot[1], Catherine Robert[1], Bernard La Scola[1,3], Oleg Y. Mediannikov[1], Didier Raoult[1,3] and Christelle Desnues[1]**

[1] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63 Centre National de la Recherche Scientifique 7278 IRD 198 Institut National de la Santé et de la Recherche Médicale U1095, Aix-Marseille Université, Marseille, France, [2] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63 Centre National de la Recherche Scientifique 7278 IRD 198 Institut National de la Santé et de la Recherche Médicale U1095, Aix-Marseille Université, Dakar, Senegal, [3] Fondation IHU Méditerranée Infection, Pôle des Maladies Infectieuses et Tropicales Clinique et Biologique, Fédération de Bactériologie-Hygiène-Virologie, Centre Hospitalo-Universitaire Timone, Méditerranée Infection, Assistance Publique – Hôpitaux de Marseille, Marseille, France

OPEN ACCESS

Edited by:
Gilbert Greub, University of Lausanne, Switzerland

Reviewed by:
Vincent Thomas, BIOASTER, France
Santhosh Chakkaramakkil Verghese, Oregon Health and Science University, USA

***Correspondence:**
Christelle Desnues
christelle.desnues@univ-amu.fr

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *17 September 2015*
Accepted: *25 November 2015*
Published: *16 December 2015*

Citation:
Temmam S, Monteil-Bouchard S, Sambou M, Aubadie-Ladrix M, Azza S, Decloquement P, Bou Khalil JY, Baudoin J-P, Jardot P, Robert C, La Scola B, Mediannikov OY, Raoult D and Desnues C (2015) Faustovirus-Like Asfarvirus in Hematophagous Biting Midges and Their Vertebrate Hosts. Front. Microbiol. 6:1406. doi: 10.3389/fmicb.2015.01406

Faustovirus, a new *Asfarviridae*-related giant virus, was recently isolated in *Vermamoeba vermiformis*, a protist found in sewage water in various geographical locations and occasionally reported in human eye infection cases. As part of a global metagenomic analysis of viral communities existing in biting midges, we report here for the first time the identification and isolation of a Faustovirus-like virus in hematophagous arthropods and its detection in their animal hosts. The DNA virome analysis of three pools of *Culicoides* sp., engorged female *Culicoides imicola* and non-engorged male/female *C. imicola* biting midges collected in Senegal, revealed the presence of amoeba-infecting giant viruses and, among them, a majority of sequences related to Faustovirus. Phylogenetic analyses conducted on several structural genes of Faustovirus confirmed the clustering of the arthropod-borne Faustovirus with sewage-borne Faustoviruses, with a distinct geographical clustering of Senegalese Faustovirus strains. Transmission electron microscopy identified viral particles with morphologies and diameters which were compatible with Faustovirus. The presence of infectious arthropod-borne Faustovirus was finally confirmed by successful isolation on *V. vermiformis* amoeba. Global proteomic analysis of biting midges identified that arthropods' blood meal originating from cattle, rodents and humans. Further screening of cattle sera and rodent tissue resulted in prevalence of Faustovirus being estimated at 38% in rodents and 14% in cattle, suggesting a possible origin of Faustovirus presence in arthropods via the ingestion of contaminated blood meal. Viral loads were the highest in rodents' urine and kidney samples, suggesting a possible excretion of viral particles into the environment. Faustovirus DNA polymerase-related sequences were also detected in more than 9 and 11% of febrile patients and healthy Senegalese human sera, respectively. Our study thus, highlights the need to investigate the role of arthropods, wildlife, and domestic animals in the lifecycle of amoeba-infecting giant viruses and, in particular, the environmental cycle of Faustovirus.

Keywords: biting midges, giant virus, faustovirus, bloodmeal host, environment

INTRODUCTION

Large double-stranded (ds)DNA viruses, also known as "giant viruses," form a monophyletic group consisting of *Poxviridae*, *Iridoviridae*, *Ascoviridae*, *Phycodnaviridae*, *Asfarviridae*, *Mimiviridae,* and *Marseilleviridae* families and are classified under the proposed *Megavirales* order (Colson et al., 2012, 2013). More recently, discovery of *Pandoravirus* and *Pithovirus* genera has been reported (Philippe et al., 2013; Legendre et al., 2014).

Protozoans, and especially amoebas, have been largely used as tools to isolate and cultivate a wide variety of microorganisms, due to their lack of receptor-dependent infection and the ability of some bacteria and viruses to resist phagocytosis and to multiply in these organisms (Greub and Raoult, 2004). So far, giant viruses have been isolated on amoebae from various environments all over the world, mostly from water samples (Pagnier et al., 2013). Recently Faustovirus, a new virus closely related to the *Asfarviridae* family, has been isolated on *Vermamoeba vermiformis* protists in sewage water in various geographical locations (Reteno et al., 2015). *Asfarviridae* are a family of dsDNA viruses consisting of a unique member: the African swine fever virus (ASFV), the only known dsDNA virus transmitted by hematophagous arthropods, i.e., ticks.

Ceratopogonidae, and especially the genus *Culicoides*, are well-known vectors of several parasites (Agbolade et al., 2006; Slama et al., 2014) and viruses (Mellor et al., 2000) infecting animals and human (i.e., Bluetongue virus, African Horse Sickness virus, Epizootic Hemorrhagic Disease virus, and Oropouche virus, the only known human virus transmitted by biting midges). In sub-Saharan countries such as Senegal, biting midges usually feed on livestock but also on humans. Larval stages of *Culicoides* sp. are found in semi-aquatic environments (Harrup et al., 2013), leading to possible contact with amoebae and their associated giant viruses.

In the present study, we report for the first time the detection, isolation, and environmental exploration of Faustovirus in adult *Culicoides* sp. biting midges.

MATERIALS AND METHODS

Sample Collection and Ethical Statement

Arthropods

Biting midges were collected using a modified CDC light trap as previously described (Sambou et al., 2015), in the villages of Dielmo and Ndiop in the Sine-Saloum region of Senegal, in November 2013 (**Figure 1**). Traps were placed near places where cattle rested and were left overnight. Morphological identification of the arthropods was conducted the next morning, as previously described by Sambou et al. (2015). Three types of arthropod pools were created: the STE0043 pool consisted of more than 200 adult *Culicoides* sp., with no distinction between male and female, nor their gorged status; STE0044 and STE0045 pools consisted of 15 engorged female and 100 non-engorged male and female *Culicoides imicola,* respectively. Arthropods were immediately stored in liquid nitrogen directly in the field. All these pools were collected from the same concession in Dielmo during the same night.

Hard ticks collected from cattle were harvested and directly stored in liquid nitrogen in pools according to their animal origin.

Cattle

Cattle sera were collected from animals at the same location as the CDC light traps used to sample biting midges. These were also immediately stored in liquid nitrogen.

Rodents

Rodent trapping was conducted at the same place: traps were left open overnight, and small mammals were sacrificed the next morning by cervical dislocation, according to the guidelines for the handling of wild mammals (Sikes and Gannon, 2011). All animal procedures carried out in this study were approved by the IRD Local Ethics Committee. The spleen, lungs, kidney, liver, brain, bladder, intestine, and serum were collected from trapped animals and directly stored individually in liquid nitrogen. Species identification of trapped small mammals was conducted by sequencing the 18S rRNA, as previously described (Breitbart and Rohwer, 2005).

Water

Drinking water collected from wells in the two rural villages of Dielmo and Ndiop, and water collected from the Néma river in Dielmo were filtered through a 0.80-μm filter, followed by a 0.45-μm filter (Millipore, Molsheim, France). 10% (w/v) of PolyEthylenGlycol (PEG6000, Sigma Aldrich, Saint-Quentin Fallavier, France) and 300 mM NaCl (Sigma Aldrich, Saint-Quentin Fallavier, France) were then added to precipitate viral particles and were incubated overnight at +4°C. After centrifugation at 12,000 g for 20 min, the final pellet was resuspended in 2 mL of 0.02 μm-filtered PBS and stored in aliquots at −80°C.

Human

Human sera were collected through the Point-of-Care (POC) laboratory in Dielmo (Sokhna et al., 2013). One hundred and twelve sera from febrile people with no known etiology and 51 sera from healthy people were collected between November 2013 and June 2014. The National Ethics Committee of Senegal approved the most recent protocol, including the POC laboratory and activities, under the "Avis éthique et scientifique n°00081 du 04 juin 2012."

Sample Processing

Fifty microliters of cattle sera, up to 20 μL of rodent urine and 100 μL of PEGylated water pellets were used to extract total nucleic acids using the High Pure Viral Nucleic Acids kit (Roche Diagnostics, Mannheim, Germany), according to the manufacturer's instructions.

Approximately 0.5 cm^3 of rodent tissues and up to five hard ticks were crushed in pools with two 3-mm tungsten beads and a TissueLyser at 25 Hz for 2 min (Qiagen, Courtaboeuf, France) in 2 mL of sterile EMEM (Life Technologies, Saint Aubin,

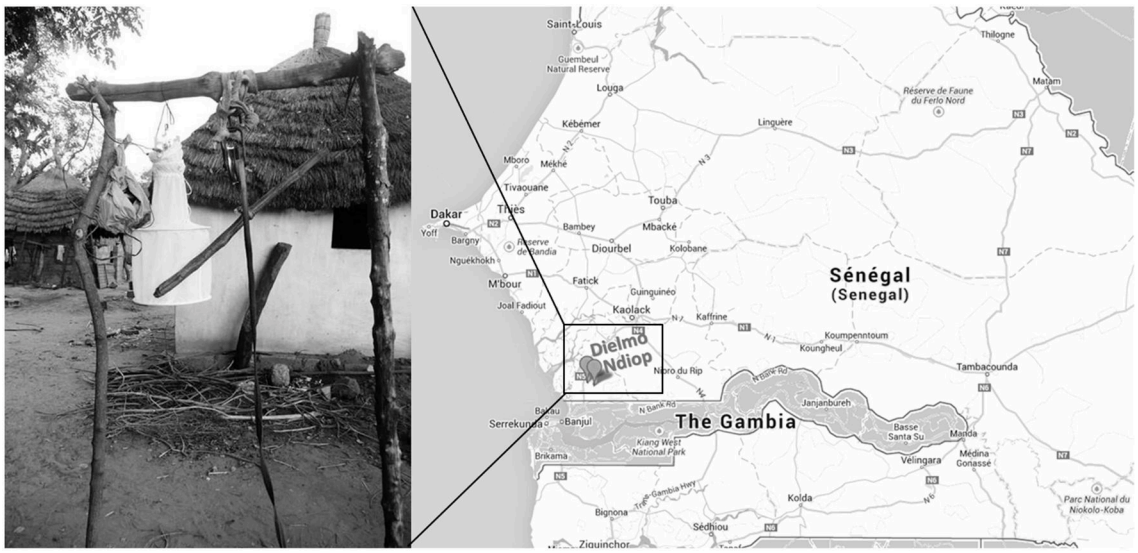


FIGURE 1 | Sampling sites.

France). Two hundred microliters of clarified supernatant was then processed for nucleic acid extraction as for cattle sera.

Total nucleic acids from human sera were previously extracted within the POC laboratory in Dielmo. Briefly, 200 µL of human capillary blood was extracted using the Nucleospin Tissue kit (Macherey-Nagel, Hoerdt, France), according to the manufacturer's recommendations.

Virome Preparation

The three pools of arthropods were crushed with two 3-mm tungsten beads and a TissueLyser at 25 Hz for 2 min (Qiagen, Courtaboeuf, France). The clarified supernatant was subsequently used as a template for virome preparation.

DNA viromes were prepared to purify viral particles from their complex sample. Briefly, the clarified supernatant was filtered through a 0.45-µm filter (Millipore, Molsheim, France), and free nucleic acids were digested with a cocktail of nucleases, i.e., 20U of Turbo DNAse (Life Technologies, Saint Aubin, France), 25U of RNAse A (Roche Diagnostics, Meylan, France), 25U of Benzonase® (Merck Millipore, Molsheim, France) and 20U of Exonuclease I (New England Biolabs, Évry, France), as previously described (Temmam et al., 2015). Finally, the digested supernatant was purified onto a discontinuous 66–30% sucrose gradient and ultracentrifuged at 130,000 g for 2 h at +4°C on a MLS-50 rotor (Beckman-Coulter, Villepinte, France). The viral fraction was harvested at the interphase between the 66 and 30% sucrose layers.

DNA was extracted from the purified viral fraction using Trizol LS® reagent (Life Technologies, Saint Aubin, France), according to the manufacturer's recommendations, and was randomly amplified using a Genomiphi V3 kit (GE Healthcare, Vélizy-Villacoublay, France) in two independent reactions. Amplification products were pooled and purified with Agencourt AMPure Beads (Beckman-Coulter, Villepinte, France) according to the manufacturer's protocol, eluted to a final volume of 15 µL and sequenced using MiSeq Technology with the paired-end and barcode strategies according to a Nextera XT library kit in a 2 × 300 bp format (Illumina Inc., San Diego, USA).

Bioinformatic Analyses of Viromes

Raw reads were imported in pairs into CLC Genomics Workbench 6.0.1 program (CLC Bio, Aarhus, Denmark) and trimmed according to their quality score, the presence of ambiguities, and their length (reads <50 nt long were discarded). The pre-processed viral metagenomes are publicly available on the Metavir server (http://metavir-meb.univ-bpclermont.fr) under the project "Arthrovirome" project and on the MG-RAST server (http://metagenomics.anl.gov/) with the identifiers 4604224.3, 4604225.3, and 4604226.3 for the STE0043, STE0044, and STE0045 DNA viromes, respectively.

Cleaned paired reads were assembled into contigs with the CLC Genomics program using the following parameters: a word size of 20 nt, minimum contig length of 200 nt, mismatch cost of 2, insertion/deletion cost of 3, length fraction of 0.5, and similarity fraction of 0.8. Contigs and non-assembled reads were then compared to the NCBI nucleotide database using the BlastN algorithm, with a minimum coverage of 50%, minimum identity of 50%, and an *E*-value $<10^{-5}$. Sequences with no significant hits according to the criteria described above were classified as "unknown." Contigs were then compared to the *Megavirales* database (Verneau et al. [METADIG: an automated pipeline to search for giant virus-related sequences in metagenomes. *In revision*]) using the BlastX program with a minimum coverage of 50%, minimum identity of 50%, and an *E*-value $<10^{-5}$. To confirm the specificity of the BlastX result, contigs were finally compared to the NCBI RefSeq viral database and to the whole NCBI database under the same criteria. The taxonomic

assignment of contigs was conducted by selecting the best BlastX score result between the three Blasts run for each contig.

Phylogenetic Analyses

Contigs matching with amoeba-infecting giant viruses were extracted and translated using the FragGeneScan tool (Rho et al., 2010), according to the "short" and "complete" parameters. Predicted Open Reading Frames (ORF) were then compared to the *Megavirales* database using the BlastP program to identify *Megavirales* core genes. Phylogenetic analyses were performed on the amino-acid sequences of the RNA diphosphate reductase large sub-unit and the nucleotide sequence of the sub-unit common to RNA polymerase I–II–III, the DNA topoisomerase and the putative helicase C962R of Faustovirus.

Amino-acid and nucleotide sequences were retrieved from the GenBank database and aligned with the MUSCLE aligner (Edgar, 2004) implemented through MEGA6 (Tamura et al., 2013). The DNA/amino-acid substitutions model that best fitted the data were performed on MEGA6 (Tamura et al., 2013) and were considered for all phylogenetic analyses. We selected the best substitution model using the corrected Akaike information criterion. Phylogenetic trees were constructed by Maximum Likelihood (ML) implemented through the MEGA6 package software, according to the selected substitution model. Nodal support was evaluated by 1000 bootstrap replicates.

Detection of Faustovirus in Animal, Human, and Environmental Samples

Quantitative SYBR Green real-time PCR targeting the DNA polymerase of Faustovirus was performed with the Quantitect SYBR Green qPCR kit (Qiagen, Courtaboeuf, France), according to the manufacturer's recommendations, except that 400 nM of forward (5′-CAAAGGCTATTGAGGCGATTTG-3′) and reverse (5′-ATGATTGTGCTGCTAGGATACC-3′) primers were used and mixed with 5 µl of DNA. Annealing temperature was defined as 58°C.

A standard curve was generated after extraction of serial dilutions of a flow cytometer-quantified Faustovirus. Briefly, the quantification was realized using a suspension of concentration-calibrated Cytocount™ fluorescent beads (Dako, Les Ulis, France) and the following formula: (number of counted particles / number of counted beads) × bead concentration (i.e., 1100 beads/µL), as previously reported (Khan et al., 2010). The resulting count was expressed in Virus-Like Particles (VLP) per mL. Dilutions of the quantified virus were performed, and nucleic acids were extracted from each dilution and further used as template for the qPCR standard curve.

Primers targeting Faustovirus RNA polymerase and DNA topoisomerase were designed according to the metagenomes sequences: Fausto_RNApol_F (5′-TACGTCAAGCAGTAG CCAACG-3′), Fausto_RNApol_R (5′-CTACTTGCCG CACAACAGCC-3′), Fausto_DNAtopo_F (5′-CCAGC ACCATATGACACGCG-3′) and Fausto_DNAtopo_R (5′-AATGTATGCGTTCGATTCGCC-3′). PCR targeting Faustovirus RNA polymerase, DNA topoisomerase and capsid (Reteno et al., 2015) were performed using the Hot Star Taq DNA polymerase (Qiagen, Courtaboeuf, France). Annealing temperatures were 57°C, 57°C, and 58°C, respectively.

All PCR products were further analyzed on a 2% agarose gel, and bands of the expected size were extracted from the gel, purified using the QIAex Gel Extraction kit (Qiagen, Courtaboeuf, France) and sequenced with a Big Dye Terminator v1.1 Cycle Sequencing Kit (Life Technologies, Saint Aubin, France) and an ABI 3130 Genetic Analyzer (Life Technologies, Saint Aubin, France).

Total Protein Extraction, Western Blot, and Global Proteomic Analyses

Approximately 50 mg of arthropods from the STE0043 sample were crushed in 300 µL of lysis buffer (Tris-HCl 40 mM pH 7.5, SDS 2% (w/v), DTT 60 mM) with two 3-mm tungsten beads and a TissueLyser at 25 Hz for 2 min (Qiagen, Courtaboeuf, France) before heating at 95°C for 5 min. Proteins from the clarified supernatant were precipitated using the PlusOne 2-D Clean-Up Kit (GE Healthcare, Vélizy-Villacoublay, France). The final pellet was resuspended in 200 µL of solubilization buffer (Urea 8M, Thiourea 2M, 100 mM NaCl, 25 mM Tris, pH 8.2) and dialyzed twice using Slide-ALyzer Dialysis Cassettes 2K MWCO (Pierce Biotechnology, Rockford, USA) against 1 L of 50 mM ammonium bicarbonate pH 7.4, Urea 1M (7 h and overnight). Dialyzed fractions were collected and proteins were quantified by Bradford assay using Coomassie (Biorad, Marnes-la-Coquette, France). The dialyzed fraction was subsequently used as a template for global proteomics and western blot analyses.

Two hundred micrograms of soluble proteins were fractionated on a 12% polyacrylamide gel electrophoresis then revealed by silver staining. Additionally, resolved proteins were transferred onto a nitrocellulose membrane (Trans-blot Transfer Medium, Biorad, Hercules, CA, USA) at 100 V for 1 h using a semi-dry transfer unit (Hoefer TE 77, GE Healthcare, Vélizy-Villacoublay, France). Membranes were then blocked in PBS supplemented with 0.3% Tween-20 and 5% non-fat dried milk (PBS-Tween-Milk) for 90 min, and incubated with mouse polyclonal anti-Faustovirus antibodies (1:1000). The immunoreactive bands were detected using a peroxidase-conjugated goat anti-mouse immunoglobulin (GE Healthcare, Vélizy-Villacoublay, France) diluted to 1:5000 in the blocking buffer for 1 h at room temperature. Three fifteen minutes washes were applied between each step. Immunostained bands were visualized with the chemiluminescence-based kit, as described by the manufacturer (GE Healthcare, Vélizy-Villacoublay, France). The resulting signal was captured by a Fusion FX7 imaging system (Vilber Lourmat, France).

Six pieces corresponding to immunoreactive bands were excised from silver stained gel and subject to mass spectrometry (MS) analysis. Briefly, after several washes, the proteins extracted from the gel were reduced, alkylated and digested with trypsin, as described above. Tryptic peptides were extracted with acetonitrile 100%; the extraction solution was collected and incubated at 45°C to evaporate the acetonitrile and to concentrate it prior to MS analysis. An additional global proteomic analysis was conducted. Briefly, 200 µg of total soluble proteins were

reduced and alkylated with iodoacetamide. Protein digestion was performed by adding 8 μg of sequencing-grade trypsin solution (Promega, Charbonnières, France) to alkylated proteins and incubated overnight at 37°C. The digested sample was then desalted using Pierce Detergent Removal Spin Columns (Thermo Fisher Scientific, Illkirch, France) and analyzed by mass spectrometry, as described hereafter.

A NanoAcquity UPLC System (Waters, Saint-Quentin En Yvelines, France) was coupled with a Synapt-G2 Si HDMS with Traveling-Wave-Ion-Mobility Mass Spectrometry instrument (TWIM-MS; Waters, Saint-Quentin En Yvelines, France). Chromatographic separation was performed on an NanoAcquity UPLC BEH130 C18 column (1.7 μm, 100 μm × 100 mm; Waters, Saint-Quentin En Yvelines, France) preceded by a Symmetry C18 trapping column (5 μm, 180 μm × 20 mm); both were placed in a 40°C oven. The injection volume was set to 2 μL for the digested soluble proteins (200 ng/μL) and 4 μL for gel-extracted proteins. The mobile phase consisted of water (A) and acetonitrile (B) both in 0.1% formic acid. Samples were trapped over 3 min with 99.9% A and 0.1% B. The separation gradient was as follows: 0–100 min, linear from 95% A, 5% B, to 60% A, and 40% B; 100–107 min. Mass spectrometry experiments were performed in positive ion mode and in resolution mode. The settings of the instruments were automatically optimized to obtain the best resolution. The ion source parameters were capillary voltage 3 kV, sampling cone voltage 40 V, ion source temperature 90°C and cone gas flow 50 L/h. Transfer collision low energy was set to 5 V while trap collision low energy was set to 4 V. The high energy ramp was applied from 4 to 5 V for the trap collision and from 19 to 45 V for the transfer collision, enabling fragmentation of the ions after the ion mobility cell and before the time-of-flight (TOF) MS. The instrument was previously calibrated in the mass range of 50–2000 Da using GFP fragments (0.2 pmol/μL). Data were processed using ProteinLynx Global Server version 3.0.1 (Waters, Saint-Quentin En Yvelines, France). Processing parameters were 250 counts for the low energy threshold, 100 counts for the elevated energy threshold and 750 counts for the intensity threshold.

Databases used to compare spectra combined data from Mammalia (2015/Feb/09, Swissprot, 66,370 sequences), Dipteria (2015/Feb/09, Swissprot, 6607 sequences), and giant viruses (2015/Feb/06, TrEMBL and not published giant viruses sequences, 14,866 sequences). An additional database was generated using predicted ORFs generated following the FragGeneScan analysis of giant viral contigs of the three metagenomes. A cut-off was used to remove the matches with only one and two peptides and the option of Merge Data was used with the six gel pieces.

Transmission Electron Microscopy (TEM)

Approximately 50 mg of arthropods from the STE0043 sample were washed in 70% ethanol, as previously described (Slimani et al., 2013) and crushed in 2 mL of sterile EMEM medium (Life Technologies, Saint Aubin, France) using two 3-mm tungsten beads and a TissueLyser at 25 Hz for 2 min (Qiagen, Courtaboeuf, France). The supernatant was harvested after a low speed clarification and subsequently filtered through a 0.8-μm filter (Millipore, Molsheim, France). The resulting supernatant was purified onto a discontinuous 66–30% sucrose gradient and ultracentrifuged at 130,000 g for 2 h at +4°C, as described above.

The viral fraction was harvested at the interphase between the 66 and 30% sucrose layers and fixed for 1 h at +4°C with 2% final glutaraldehyde. The fixed viral fraction was then diluted to a final volume of 4 mL in PBS, and adsorbed directly onto formvar carbon films on 400 mesh nickel grids (FCF400-Ni, EMS) by ultracentrifugation at 130,000 g for 1 h at +4°C, as previously described (Sime-Ngando et al., 1996). Grids were stained for 10 s with 1% molybdate solution in filtered water at room temperature. Electron micrographs were obtained on a Tecnai G2 transmission electron microscope (FEI) operated at 200 keV equipped with a 4096 × 4096 pixel resolution Eagle camera (FEI).

Isolation of Viruses on Amoebae

Approximately 50 mg of arthropods from the STE0043 sample were washed in 70% ethanol, as previously described (Slimani et al., 2013), then with sterile Page's amoebal saline (PAS) solution, and finally crushed in 3 mL of PAS buffer.

V. vermiformis (CDC19 strain) amoebae were used to isolate giant viruses from arthropods, as reported by Pagnier et al. (2013), except that Vancomycin 10 μg/mL, Ciprofloxacine 20 μg/mL, Imipenem/cilastatine 10 μg/mL, Doxycycline 20 μg/Ml, and Voriconazole 20 μg/mL were added to the amoebal suspensions to prevent bacterial and fungal contamination. Briefly, amoebae were cultivated in peptone-yeast extract-glucose (PYG) medium and sub-cultured every 2 days. 5×10^5 amoebae/mL were concomitantly plated in a 12-well microplate with 100 μL of the sample suspension and incubated at 30°C for 3 days. At Day 3 post-infection, the primo-culture was sub-cultured onto a fresh amoebal microplate suspension under the same conditions. The primary cultures and sub-cultures were screened daily for a cytopathogenic effect (CPE) using an inverted microscope and if CPEs were observed, fresh amoeba cells were cytospinned with 100 μL of viral supernatant and further stained with Gimenez and Gram stains, followed by additional Hemacolor staining (Merck, Darmstadt, Germany).

Additionally, 100 μL of positive CPE supernatant was used to extract DNA using the phenol chloroform isoamyl alcohol extraction procedure (Life Technologies, Saint Aubin, France) according to the manufacturer's protocol, and PCRs targeting Faustovirus were conducted, as described above.

RESULTS

Presence and Diversity of Sequences Related to Giant Viruses in the Virome of Biting Midges

The DNA virome of the STE0043, STE0044, and STE0045 samples were sequenced using Illumina MiSeq technology. After trimming, the total number of reads was 1,517,965, 2,163,868, and 2,265,552 reads, respectively (**Table 1**), with 46.29, 76.01, and 48.50% of sequences having homologies after BlastN

TABLE 1 | Sequencing data of the virome datasets.

	STE0043 *Culicoides* sp.	STE0044*C. imicola* ♀engorged	STE0045 *C. imicola*♂♀non-engorged
Raw reads	1,520,202	2,173,228	2,267,752
Cleaned reads	1,517,965	2,163,868	2,265,552
Contigs	19,771	29,995	28,309
Singletons	85,185	134,230	122,805
Average contig length	615 bp	587 bp	581 bp
Total assigned reads:	702,730	1,644,792	1,098,862
eukaryote	576,221	1,457,350	781,498
prokaryote (bacteria + archaea)	122,075	164,982	303,928
virus	4434	22,460	13,436
Total giant viruses:	3465	20,745	4684
Faustovirus	3146	8383	3490
Mimiviridae	317	12,362	1164
Non-classified giant viruses	2	0	30
Other viruses:	969	1715	8752
Nudiviridae	307	0	6237
Poxviridae	182	42	34
Siphoviridae	182	503	808
Non-classified phages	156	277	240
Myoviridae	60	554	785
Papillomaviridae	41	100	2
Iridoviridae	29	0	0
Podoviridae	4	45	29
Phycodnaviridae	2	90	133
Polydnaviridae	2	0	0
Retroviridae	2	92	16
Non-classified plant viruses	2	0	0
Ascoviridae	0	12	2
Herpesviridae	0	0	315
Inoviridae	0	0	149
Iflaviridae	0	0	2

against GenBank nt database, respectively. Eukaryotic sequences represented more than 70% of the total assigned reads in the STE0043, STE0044, and STE0045 metagenomes, mainly identified as human and arthropod reads (**Table 1**).

Among the total assigned reads, 4434 (STE0043), 22,460 (STE0044), and 13,434 (STE0045) sequences were related to viruses (**Table 1**). Most viral sequences were identified as double-stranded DNA viruses, and single-stranded DNA viral sequences were only detected in STE0045 non-engorged male and female *C. imicola* virome. Sequences related to amoeba-infecting giant viruses from the *Mimiviridae* family, Faustovirus and the non-taxonomically classified Pandoravirus represented the majority of viral reads, with more than 78, 92, and 34% of total viral reads, respectively, for STE0043, STE0044, and STE0045 viromes. The most represented virus was Faustovirus, with more than 90, 40, and 74% of total giant viral reads, respectively (**Table 1**).

The presence of Faustovirus in each sample was confirmed by PCR specifically targeting the DNA polymerase, the viral capsid, the RNA polymerase and the DNA topoisomerase genes of the virus. PCR amplification products were obtained for the three metagenomes (see **Supplemental Figure 1** for the capsid amplification) and Faustovirus amplifications were confirmed by sequencing.

The presence of amoeba-infecting giant viral sequences was searched on previously published arthropods metagenomes available in public databases (**Table 2**). Sequences from both hematophagous (mosquitoes, hard ticks, and body lice) and non-hematophagous (termites and whiteflies) arthropods were assembled into contigs and compared to an in-house giant viral database. Mosquito microbiomes showed the presence of *Mimiviridae* and *Pandoraviridae*-related contigs in all of the five studies, although these came from arthropods sampled at different time points in different locations. Hard ticks and experimentally-infected body lice metagenomics revealed also the presence of *Mimiviridae*-related contigs. In contrast, termites and whiteflies metagenomic datasets did not present any amoeba-infecting giant viral contigs. No Faustovirus-related sequences were retrieved in metagenomes, either from hematophagous or non-hematophagous arthropods.

Phylogenetic Analyses of the Faustovirus-Like Virus Identified in the Virome of Biting Midges

Contigs matching with Faustovirus sequences were extracted and translated. Results of the predicted ORFs are presented in **Table 3**. Phylogenetic reconstructions were performed on several conserved genes: the RNA diphosphate reductase large sub-unit and the sub-unit common to RNA polymerase I–II–III that were found in the three biting midges' metagenomes, the DNA topoisomerase only detected on the STE0043 *Culicoides* sp. virome and the putative helicase C962R, both present in the STE0044 *C. imicola* engorged female and the STE0045 *C. imicola* non-engorged male/female viromes.

Phylogenetic analyses performed on the sub-unit common to RNA polymerase I–II–III (**Figure 2A**) and on the RNA diphosphate reductase large sub-unit genes (**Figure 2B**) showed that the biting midges' Faustovirus grouped with other Faustoviruses isolated from French, Senegalese and Lebanese sewage. More precisely, Senegalese biting midges' Faustovirus formed a cluster with Dakar sewage Faustovirus, supported by high bootstrap nodes (100 and 95, respectively). These results were confirmed with phylogenetic analyses performed on the DNA topoisomerase (**Supplemental Figure 2A**) and the putative helicase C962R (**Supplemental Figure 2B**) genes.

Phylogeny performed on the RNA diphosphate reductase large sub-unit gene allowed distinguishing a specific cluster composed of biting midges Faustovirus only within the Senegalese Faustovirus clade, with a high bootstrap value of 78 (**Figure 2B**).

TABLE 2 | Search for the presence of amoeba-infecting giant viral sequences in the metagenomes of other arthropods.

Arthropods	Nb of total reads	Nb of contigs	Sequencing method	Type of metagenome	Amoeba-infecting giant viral contig (nb)	References
Mosquitoes	1,575,043	1964	Roche 454 FLX	RNA shotgun	*Mimiviridae* (1)	Bishop-Lilly et al., 2010
	1,961,290	16,321	Roche 454 GS20	DNA shotgun	*Mimiviridae* (1) *Pandoraviridae* (1)	Dinsdale et al., 2008
	26,403,284	89,744	Illumina GA II	RNA shotgun	*Mimiviridae* (8) *Pandoraviridae* (10)	Chandler et al., 2014
	217,330,434	311,750	Illumina HiSeq 2000	RNA shotgun	*Mimiviridae* (13) *Pandoraviridae* (19)	Chandler et al., 2015
	1,576,489	15,666	Roche 454 GS20	DNA shotgun	*Mimiviridae* (1) *Pandoraviridae* (2)	Ng et al., 2011
Body lice	4,403,873	1733	Illumina MiSeq	RNA shotgun	*Mimiviridae* (5)	Temmam et al., 2015
Whiteflies	1,427,809	193	Illumina GA II	RNA shotgun	0	Rosario et al., 2014
Termites	–	57,641	Sanger	DNA shotgun	0	Warnecke et al., 2007
Hard ticks	350,329	31,881	Roche 454 FLX	DNA shotgun	*Mimiviridae* (1)	Nakao et al., 2013

TABLE 3 | Predicted ORFs for Faustovirus detected in the three metagenomes.

	STE0043		STE0044		STE0045	
	Short	Complete	Short	Complete	Short	Complete
Total contigs/reads	79/3146		137/8383		114/3490	
TOTAL ORFs	93	87	148	145	127	125
Hypothetical protein	83	70	125	122	105	103
62 kDa polyprotein	–	–	1	1	1	1
Ankyrin containing protein	–	–	1	1	1	1
Bacterial MORN repeat-containing protein	–	2	2	2	–	–
BTB/POZ domain-containing protein	–	–	1	1	2	2
BTB/POZ domain-containing protein 9	1	–	1	1	–	–
Deoxyuridine 5′-triphosphate nucleotidohydrolase	–	–	2	2	–	–
DNA topoisomerase small subunit	–	1	–	–	–	–
Glutaredoxin-C3	2	–	–	–	2	2
Helicase	–	–	–	–	1	1
Metallophos_2 containing protein	2	–	–	–	–	–
MORN repeat-containing protein	3	8	2	2	4	4
mRNA-decapping protein g5R	–	1	–	–	–	–
Patatin	–	–	–	–	1	1
Putative ATP-dependent RNA helicase L377	–	1	–	–	–	–
Putative ATP-dependent RNA helicase R563	–	1	–	–	–	–
Putative DNA polymerase family X	–	–	1	1	–	–
Putative DNA-directed RNA polymerase subunit D	–	1	–	–	–	–
Putative helicase C962R	–	–	2	2	2	2
Putative histidinol-phosphate aminotransferase	–	–	–	–	1	1
Putative hydrolase	–	–	1	1	–	–
Putative phosphatidylglycerophosphate synthase	–	–	1	1	–	–
Putative poly(A) polymerase catalytic subunit	–	–	–	–	1	1
Putative T4-like proximal tail fiber			1	1	–	–
Putative UV-damage endonuclease			2	2	–	–
Ribonucleoside-diphosphate reductase large subunit	1	–	2	2	2	2
Ribonucleoside-diphosphate reductase small chain			1	1	–	–
RNA polymerase II subunit Rpb5b	–	1	–	–	–	–
Subunit common to RNA polymerases I, II, and III	1	–	2	2	3	3
Translation initiation factor SUI1	–	1	–	–	–	–
Transcription factor S-II-related protein	–	–	–	–	1	1

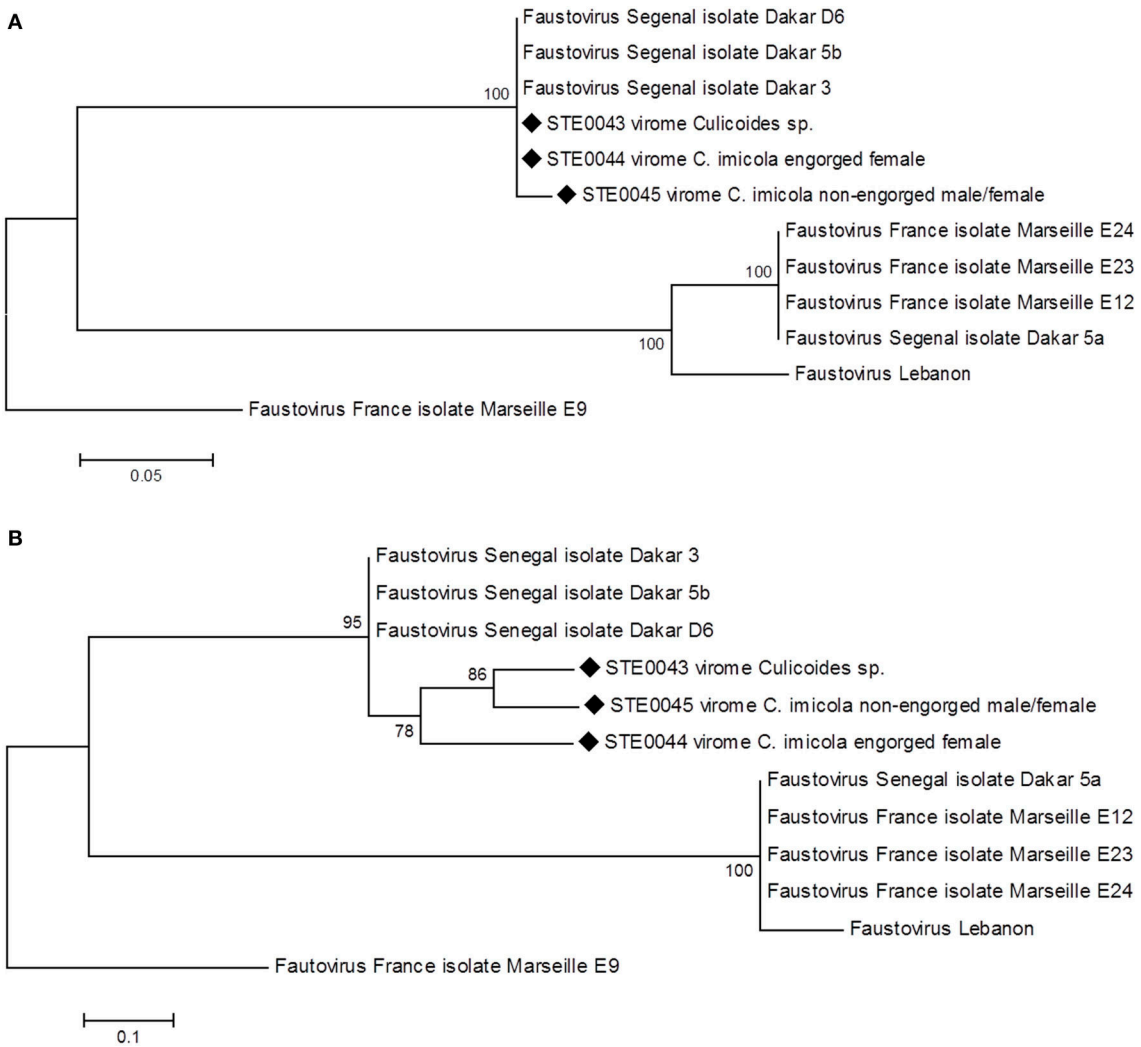


FIGURE 2 | Phylogenetic reconstruction of Faustovirus detected in biting midges' viromes based on the (A) nucleotide sequence of the sub-unit common to RNA polymerase I–II–III (substitution model: Kimura-2-parameters) (B) amino-acid sequence of the RNA diphosphate reductase large sub-unit (substitution model: JTT+G, $G = 3$).

Observation of Virus-Like Particles by Transmission Electron-Microscopy and Isolation of a Faustovirus-Like Virus from Biting Midges

Viral particles purified from biting midges were negatively stained and observed by transmission electron microscopy (**Figure 3**). Virus-like particles were observed with different morphologies and diameters, ranging from 600 (**Figure 3A**) to 200 nm (**Figure 3D**). Some of the observed virions had a diameter (approximately 200 nm) and morphology (icosahedral capsid) compatible with that of Faustovirus (**Figure 3D**).

V. vermiformis protists was used in an attempt to isolate Faustovirus-related viruses detected in STE0043 *Culicoides* sp. virome. One *V. vermiformis* sub-culture was Faustovirus PCR-positive at Day 3 post-infection, and sequencing of a fraction of the capsid gene confirmed that the isolated virus was Faustovirus. However, successful viral production was impaired due to the high bacterial load present in the culture, originating from the arthropods' guts.

Global Proteomics of Biting Midges

Western blot analysis of STE0043 proteins revealed an immunoreactive smear, between 260 and 50 kDa, with anti-Faustovirus polyclonal antibodies (**Figure 4B**). The smear was due to a high load of proteins (200 μg) allowing the detection of viral proteins in very low abundance in these arthropods. As shown in **Figure 4C**, with a loading of 5 μg, no detection was possible in arthropods, although the positive control revealed reactive bands. Within this smear, a putative band at 60 kDa was observed for Faustovirus, possibly corresponding to the viral capsid (arrow of **Figure 4B**). Six pieces of electrophoresis-fractioned proteins among the immunoreactive smear were

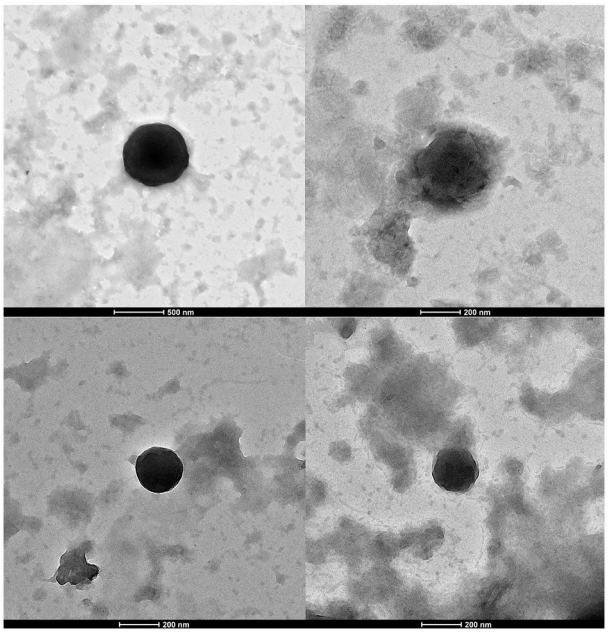


FIGURE 3 | Negative staining electron microscopy imaging of viral particles from *Culicoides* sp. biting midges' samples. Scale bars are indicated under the images.

subsequently extracted from the gel and analyzed by proteomics (**Figure 4A**).

The global proteomic analysis of the six immunoreactive bands identified 4576 different peptides. Many peptides were not identified, due to incomplete databases. The major identified proteins were related to blood tissues: indeed, nearly a quarter of the total identified peptides represented blood proteins components, i.e., serum albumin, hemoglobin and fibrinogen, which represented 5.44, 15.57, and 0.90% of the total identified peptides, respectively. Numerous arthropod-borne peptides (33.77%) were also identified, and more than two thirds of identified peptides were related to the arthropods' major blood meal hosts: *Bovidae* represented 25.00% of the total identified peptides, 15.77% for humans, and 11.42% for *Rodentia*. Less abundantly, 2.95% of peptides were identified as coming from horses, 2.80% from pets (cats and dogs), 1.99% from *Lagomorpha* (mainly rabbits), 1.17% from swine, 1.05% from primates, and 0.87% from *Chiropera*. Interestingly and although no peptide matching with Faustovirus was detected, three peptides matched with the viral proteome of Saudian virus, a new giant virus isolated from sewage, were detected.

Similar results were obtained for the whole proteomic analysis, with slight differences in the proportions of identified peptides, except that four and seven peptides were obtained for Kroon virus and Courdo11 *Mimiviridae* proteins, respectively (Yoosuf et al., 2014; Boratto et al., 2015). In addition, *Mimiviridae*-related sequences were obtained in the metagenomes (data not shown). No match with Faustovirus was obtained.

No hit were obtained when comparing the peptides to the virome-predicted ORFs database.

Detection of Faustovirus in Animals and Environmental Water

The serum from 14 cattle and their associated hard ticks, 13 rodents (four suckling mice and nine *Arvicanthis* sp.), well water from Dielmo and Ndiop, and Néma river water recovered in Dielmo were collected and used to screen for the presence of Faustovirus.

The presence of Faustovirus was detected in five over 13 rodents tested: two were from suckling mice trapped in Ndiop and three were from *Arvicanthis* sp. rodents trapped in Ndiop ($N = 2$) and Dielmo ($N = 1$) respectively. Additionally, Faustovirus was detected in two over 14 cattle sera (one from Ndiop on a healthy cattle and one from Dielmo on a lumpy skin disease-infected veal). Faustovirus was also present in environmental water in both well and river waters from Dielmo and well water from Ndiop.

Two pools of hard ticks (*Boophilus* sp. and *Rhipicephalus evertsi*) collected from the same cattle in Keur Samba Gueye, a village close to Ndiop, were also positive for Faustovirus although the animal serum was negative. One pool of *Amblyomma* sp. hard ticks, collected from cattle in Ndiop, and one pool of *Ornithodoros sonrai* soft ticks, collected from rodents' nests in Keru Serigne Korka (a village located 12 km north-east of Dielmo) were also positive.

These positive detections were all confirmed by sequencing the portion of the DNA polymerase and the capsid genes of Faustovirus. Phylogenetic analyses of rodent-borne, cattle-borne, and water-borne Faustovirus performed on the capsid gene confirmed the relatedness of environmental and mammalian Faustovirus with arthropod-borne Faustovirus, but was not sufficiently discriminant to specifically define clusters of Senegalese Faustovirus (data not shown).

Viral loads of Faustovirus in PCR-positive animals were estimated according to the tissue sample (**Figure 5A**) and the animal species (**Figure 5B**). All harvested organs were positive for Faustovirus, in viral loads ranging from 3.49×10^5 VLP/mL (lung) to 8.01×10^6 VLP/mL (kidney), except for intestine samples which were all negative, even when the extracted DNA was diluted in case of the presence of inhibitors. Interestingly, Faustovirus quantification in kidneys was similar in scale to that in the bladder or urine samples (4.21×10^6 VLP/mL). Faustovirus load in cattle sera was estimated at 5.96×10^6 VLP/mL. Biting midges non-amplified viromes were detected with the highest viral loads, estimated at 2.47×10^7 VLP/mL (**Figure 5A**). No major difference was observed when analyzing the relative abundance of Faustovirus according to animal species (**Figure 5B**). Arvicanthis-positive animals presented the lowest (1.67×10^6 VLP/mL) and biting midges the highest (2.47×10^7 VLP/mL) Faustovirus load.

Detection of Faustovirus in Humans

The serum of 112 febrile patients with no known etiology and 51 healthy people was screened for the presence of Faustovirus with the DNA polymerase targeted system. A total of 11 out of 112 (9.82%) febrile patients and six out of 51 (11.76%) healthy persons were positive. Sequencing of the small PCR product (99 bp) confirmed the positive detection of Faustovirus on 57 bp.

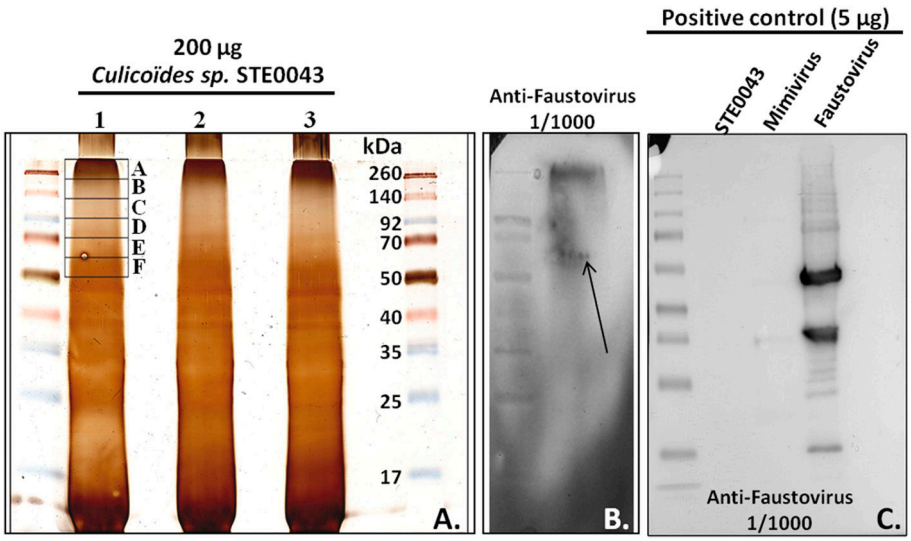


FIGURE 4 | Western blot analysis. (A) Silver staining. **(B)** Staining with anti-Faustovirus polyclonal antibodies on arthropods sample. **(C)** Positive control of Faustovirus western blot with a loading of 5 µg of proteins.

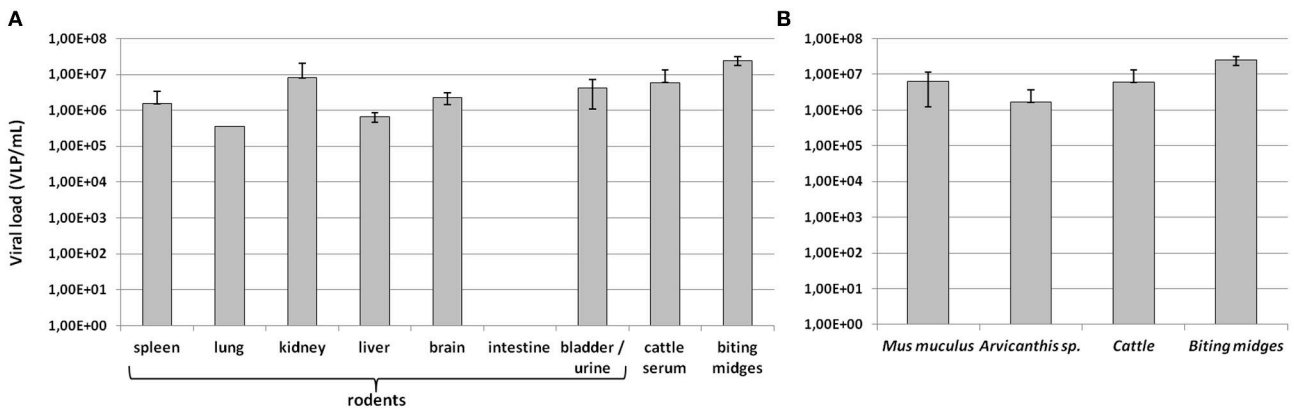


FIGURE 5 | Quantification of Faustovirus viral loads according to (A) tissue samples (B) animal species. Viral loads are expressed in $\log_{10}$ VLP/mL.

Capsid and DNA topoisomerase amplifications of Faustovirus-positive human sera were negative and RNA polymerase amplification of Faustovirus-positive human sera resulted in a non-specific multi-band amplification (data not shown).

DISCUSSION

Amoeba-infecting giant viruses have been isolated in a wide variety of biomes, mostly in environmental (i.e., soil and water) samples (Pagnier et al., 2013). Various amoeba-infecting giant viruses have also been detected in animals, such as in arthropod larvae (Boughalmi et al., 2013a), in the leech *Hirudo medicinalis* (Boughalmi et al., 2013b) or in the sera of cattle and monkeys (Dornas et al., 2014). Recently, the first amoeba-infecting giant virus belonging to the *Asfarviridae* family, Faustovirus, was reported in sewage in various geographical locations (Reteno et al., 2015). The unique other member of the *Asfarviridae* family is the ASFV, a tick-borne virus. In this study we detected and isolated for the first time Faustovirus in adults biting midges and their blood meal-associated mammals.

As part of a global study of viral communities existing in biting midges, pools of *Culicoides* sp., engorged female *C. imicola* and non-engorged male and female *C. imicola* were collected and their corresponding DNA viromes were sequenced. Results revealed the presence of sequences related to giant viruses, mainly Faustovirus with more than 90, 40, and 74% of total giant viral reads, respectively. The presence of an amoeba-infecting *Asfarviridae*-like virus in adult biting midges leads us to question on the mode of contamination of adult biting midges. *Ceratopogonidae* are arthropods with an aquatic and semi-aquatic larval stage, leading to possible contact with amoebae and their associated giant viruses during the larval stage, and a putative trans-stadial transmission of free viral particles or infected amoebae. Moreover, Evans and Schwarz (2011) reported

the infection of adult honeybees by *Malpighamoeba mellificae*, a protozoan developing in the Malpighi tubes of honeybees. We can therefore hypothesize that, adult biting midges could either be infected at the larval or during the adult stage, either with free viral particles or infected amoebae. The mode of contamination of adult biting midges is currently unknown but further studies regarding the presence of Faustovirus in all stages of development of arthropods may help to solve this question.

The engorged female *C. imicola* virome showed the highest abundance of sequences related to Faustovirus, with significant differences with the non-engorged metagenome, suggesting a possible additional viral load of the arthropod via the blood meal of female biting midges. We then searched for the presence of sequences related to Faustovirus and other giant viruses in publicly available arthropods metagenomic datasets. Our results showed that, although in low abundance, giant viral contigs were detected in other arthropods, except for Faustovirus, never detected elsewhere than in the biting midges virome. Mosquitoes and body lice presented similar abundances of *Mimiviridae*- and *Pandoraviridae*-related contigs and hard ticks presented similar abundance of *Mimiviridae*-related contigs, whereas termites and whiteflies present no giant viral sequence. One should note that mosquitoes, body lice and hard ticks are hematophagous arthropods whereas termites and whiteflies are non-hematophagous arthropods, suggesting again the putative role of blood meal in the presence of giant viruses in adult arthropods. Further proteomic analysis of the pool of *Culicoides* sp. revealed the presence of *Bovidae*, *Rodentia,* and human blood-related proteins. As a consequence we subsequently screened for the presence of Faustovirus in human sera, cattle sera and rodent organs, and detected five Faustovirus-positive rodents and two Faustovirus-positive cattle, confirming the possible contamination of female biting midges via their blood meal. Additionally, we detected three Faustovirus-positive cattle-associated engorged hard ticks, again confirming possible infection of arthropods via their blood meal.

Interestingly, we reported high levels of Faustovirus either in rodent tissue or cattle sera (**Figure 5**), and the highest viral loads were found in rodents' kidney and urine samples, suggesting a possible excretion of Faustovirus by rodents in the environment. Finally, the Néma River and the well water from Dielmo and Ndiop were all positive for Faustovirus, suggesting a possible source of contamination of humans and animals via recreational or drinking water. In sub-Saharan countries, such as Senegal, biting midges usually feed on livestock but also on humans, resulting in the vector-borne transmission of pathogens to animals and humans. In this study we reported the detection of Faustovirus in human sera harvested from febrile patients and from healthy people, with no significant difference in the prevalence between the two groups. Although we could not conclude on a putative pathogenic role of Faustovirus, the questions of the mode of infection to humans have to be addressed: is Faustovirus vector-transmitted? And if so, what kind of arthropod can transmit the virus? Or do humans acquire Faustovirus via an environmental source (water, urines of rodents, etc)? The possible reservoir role of rodents in the viral cycle of Faustovirus, both in humans and arthropods, requires further investigations, as for the vector competence of arthropods for Faustovirus.

Faustovirus-related sequences were the most abundant in all viromes. Although no capsid sequence was detected in the metagenomes, we successfully amplified a fragment of the capsid gene and confirmed by sequencing (**Supplemental Figure 1**). Western blot analysis of *Culicoides* sp. proteins using Faustovirus antibodies highlighted a band at the expected size of the capsid protein (**Figure 4B**), and further mass spectrometry sequencing identified giant viral peptides, although not related to Faustovirus. Additionally, the observation of viral particles by transmission electron microscopy with a size and shape compatible with Faustovirus, and further successful isolation of Faustovirus conducted on *V. vermiformis*, a protist commonly found in human environments (Nazar et al., 2012; Coşkun et al., 2013; Niyyati et al., 2014), confirmed the presence of infectious viral particles in the *Culicoides* sp. pool of biting midges. Phylogenetic analyses performed on several core genes revealed that Faustovirus-like viruses detected in the three biting midge viromes branch together in a cluster formed by Dakar 5b, Dakar 3, and Dakar D6 Faustovirus (**Figure 2**), viruses that were previously isolated in sewage from Dakar, Senegal (Reteno et al., 2015). According to the phylogenetic analysis conducted on the capsid gene, rodent-borne, cattle-borne and water-borne Faustovirus clustered together with arthropod-borne Faustovirus. Unfortunately, this portion of the genome of Faustovirus was not sufficiently discriminant to be able to refine the classification of mammalian, arthropod and environmental Faustovirus. Complete full genome sequencing and characterization of these viruses will enable the phylogenetic relationships between arthropod-associated Faustovirus, environmental/mammalian Faustoviruses and human Faustovirus to be refined.

Faustovirus is a recently described giant virus infecting *V. vermiformis* amoebae (Reteno et al., 2015), whose close relative is the ASFV, the only member of the *Asfarviridae* family. *Asfarviridae* are tick-borne dsDNA viruses transmitted by *Ornithodoros* sp. soft ticks and responsible for the African swine fever, a highly contagious and fatal pig infection (Burrage, 2013; Hubálek et al., 2014). Recently, hard ticks have been suspected to be capable of transmitting the virus but without success, although viral DNA was detected up to 8 weeks post-inoculation (de Carvalho Ferreira et al., 2014). In our study we report the detection and isolation of an *Asfarviridae*-like Faustovirus in biting midges, but also in a pool of *O. sonrai* soft ticks and in *Boophilus* sp., *R. evertsi,* and *Amblyomma* sp. pools of hard ticks. Interestingly, *O. sonrai* soft ticks were collected in the dust contained in rodents' nests and were not engorged, whereas hard ticks were collected directly from livestock and were engorged. Additionally, nearly 40% of the rodents tested were Faustovirus-positive, suggesting that *O. sonrai* soft ticks could be a possible vector for Faustovirus, and rodents could be a putative reservoir since no symptoms were observed on the captured rodents and high loads of Faustovirus were detected in their kidneys and urine, resulting in possible excretion of the virus in the environment. The detection of Faustovirus in *Boophilus* sp., *R. evertsi,* and *Amblyomma* sp. hard ticks could reflect the blood

meal of the ticks and the viral persistence of the virus or the viral DNA within the arthropod, as previously reported (de Carvalho Ferreira et al., 2014).

So far, Faustovirus has been only detected in sewage in Marseille, Dakar and in Lebanon and Saudi Arabia (Reteno et al., 2015). In this study, we report for the first time the detection and isolation of Faustovirus in adult biting midges, and the detection of high viral loads of Faustovirus in rodents and cattle. We also reported the detection of Faustovirus in febrile patients and healthy people. This work thus, highlights the need to investigate the role of arthropods and wild or domestic animals on the lifecycle of the *Asfarviridae*-like Faustovirus and, more globally, for the amoeba-infecting giant viruses.

AUTHOR CONTRIBUTIONS

ST, BL, OM, DR, CD designed the experiments. ST, SM, SA, PD, JBK, JB, PJ, CR performed the experiments. ST, MS, MA collected the samples. ST, SA, PD, JBK, JB, OM wrote the article. BL, OM, DR, CD revised the article.

FUNDING

This work was conducted under the frame of the ANR-13-JSV6-0004 awarded to Christelle Desnues.

ACKNOWLEDGMENTS

The authors would like to thank Samia Benamar for providing Faustovirus genome sequences, Ti Thien Nguyen and Noémie Labas for their help in preparing the Illumina libraries, Dr. Nicolás Rascovan for his critical comments and advice on the project, Jean-Michel Bérenger for his help in the morphological identification of ticks, Dr. Georges Diatta for the collection of soft ticks and the veterinarian Babacar Ndao for his help in animal sampling on the field.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.01406

Supplemental Figure 1 | Electrophoresis of PCR targeting the Faustovirus capsid conducted on biting midges viromes.

Supplemental Figure 2 | Phylogenetic reconstruction of Faustovirus detected in biting midges viromes based on the (A) nucleotide sequence of the DNA topoisomerase small sub-unit (substitution model: HKY) (B) nucleotide sequence of the putative helicase C962R (substitution model: HKY).

REFERENCES

Agbolade, O. M., Akinboye, D. O., Olateju, T. M., Ayanbiyi, O. A., Kuloyo, O. O., and Fenuga, O. O. (2006). Biting of anthropophilic Culicoides fulvithorax (Diptera: *Ceratopogonidae*), a vector of *Mansonella perstans* in Nigeria. *Korean J. Parasitol.* 44, 67–72. doi: 10.3347/kjp.2006.44.1.67

Bishop-Lilly, K. A., Turell, M. J., Willner, K. M., Butani, A., Nolan, N. M., Lentz, S. M., et al. (2010). Arbovirus detection in insect vectors by rapid, high-throughput pyrosequencing. *PLoS Negl. Trop. Dis.* 4:e878. doi: 10.1371/journal.pntd.0000878

Boratto, P., Albarnaz, J. D., Almeida, G. M., Botelho, L., Fontes, A. C., Costa, A. O., et al. (2015). *Acanthamoeba polyphaga Mimivirus* prevents amoebal encystment-mediating serine proteinase expression and circumvents cell encystment. *J. Virol.* 89, 2962–2965. doi: 10.1128/JVI.03177-14

Boughalmi, M., Pagnier, I., Aherfi, S., Colson, P., Raoult, D., and La Scola, B. (2013a). First isolation of a Marseillevirus in the *Diptera Syrphidae Eristalis tenax*. *Intervirology* 56, 386–394. doi: 10.1159/000354560

Boughalmi, M., Pagnier, I., Aherfi, S., Colson, P., Raoult, D., and La Scola, B. (2013b). First isolation of a giant virus from wild *Hirudo medicinalis* leech: *Mimiviridae* isolation in *Hirudo medicinalis*. *Viruses* 5, 2920–2930. doi: 10.3390/v5122920

Breitbart, M., and Rohwer, F. (2005). Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. *Biotechniques* 39, 729–736. doi: 10.2144/000112019

Burrage, T. G. (2013). African swine fever virus infection in *Ornithodoros* ticks. *Virus Res.* 173, 131–139. doi: 10.1016/j.virusres.2012.10.010

Chandler, J. A., Liu, R. M., and Bennett, S. N. (2015). RNA shotgun metagenomic sequencing of northern California (USA) mosquitoes uncovers viruses, bacteria, and fungi. *Front. Microbiol.* 6:185. doi: 10.3389/fmicb.2015.00185

Chandler, J. A., Thongsripong, P., Green, A., Kittayapong, P., Wilcox, B. A., Schroth, G. P., et al. (2014). Metagenomic shotgun sequencing of a *Bunyavirus* in wild-caught *Aedes aegypti* from Thailand informs the evolutionary and genomic history of the Phleboviruses. *Virology* 464–465, 312–319. doi: 10.1016/j.virol.2014.06.036

Colson, P., de Lamballerie, X., Fournous, G., Raoult, D. (2012). Reclassification of giant viruses composing a fourth domain of life in the new order *Megavirales*. *Intervirology* 55, 321–332. doi: 10.1159/000336562

Colson, P., de Lamballerie, X., Yutin, N., Asgari, S., Bigot, Y., Bideshi, D. K., et al. (2013). "*Megavirales*," a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch. Virol.* 158, 2517–2521. doi: 10.1007/s00705-013-1768-6

Coşkun, K. A., Ozçelik, S., Tutar, L., Eladi, N., and Tutar, Y. (2013). Isolation and identification of free-living amoebae from tap water in Sivas, Turkey. *Biomed. Res. Int.* 2013:675145. doi: 10.1155/2013/675145

de Carvalho Ferreira, H. C., Tudela Zúquete, S., Wijnveld, M., Weesendorp, E., Jongejan, F., Stegeman, A., et al. (2014). No evidence of African swine fever virus replication in hard ticks. *Ticks Tick Borne Dis.* 5, 582–589. doi: 10.1016/j.ttbdis.2013.12.012

Dinsdale, E. A., Edwards, R. A., Hall, D., Angly, F., Breitbart, M., Brulc, J. M., et al. (2008). Functional metagenomic profiling of nine biomes. *Nature* 452, 629–632. doi: 10.1038/nature06810

Dornas, F. P., Rodrigues, F. P., Boratto, P. V., Silva, L. C., Ferreira, P. C., Bonjardim, C. A., et al. (2014). Mimivirus circulation among wild and domestic mammals, Amazon Region, Brazil. *Emerg. Infect. Dis.* 20, 469–472. doi: 10.3201/eid2003.131050

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Evans, J. D., and Schwarz, R. S. (2011). Bees brought to their knees: microbes affecting honey bee health. *Trends Microbiol.* 19, 614–620. doi: 10.1016/j.tim.2011.09.003

Greub, G., and Raoult, D. (2004). Microorganisms resistant to free-living amoebae. *Clin. Microbiol. Rev.* 17, 413–433. doi: 10.1128/CMR.17.2.413-433.2004

Harrup, L. E., Purse, B. V., Golding, N., Mellor, P. S., and Carpenter, S. (2013). Larval development and emergence sites of farm-associated *Culicoides* in the United Kingdom. *Med. Vet. Entomol.* 27, 441–449. doi: 10.1111/mve.12006

Hubálek, Z., Rudolf, I., and Nowotny, N. (2014). Arboviruses pathogenic for domestic and wild animals. *Adv. Virus Res.* 89, 201–275. doi: 10.1016/B978-0-12-800172-1.00005-7

Khan, M. M., Pyle, B. H., and Camper, A. K. (2010). Specific and rapid enumeration of viable but nonculturable and viable-culturable gram-negative bacteria by using flow cytometry. *Appl. Environ. Microbiol*. 76, 5088–5096. doi: 10.1128/AEM.02932-09

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labedie, K., Adrait, A., et al. (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. U.S.A*. 111, 4274–4279. doi: 10.1073/pnas.1320670111

Mellor, P. S., Boorman, J., and Baylis, M. (2000). Culicoides biting midges: their role as arbovirus vectors. *Annu. Rev. Entomol*. 45, 307–340. doi: 10.1146/annurev.ento.45.1.307

Nakao, R., Abe, T., Nijhof, A. M., Yamamoto, S., Jongejan, F., Ikemura, T., et al. (2013). A novel approach, based on BLSOMs (Batch Learning Self-Organizing Maps), to the microbiome analysis of ticks. *ISME J*. 7, 1003–1015. doi: 10.1038/ismej.2012.171

Nazar, M., Haghighi, A., Taghipour, N., Ortega-Rivas, A., Tahvildar-Biderouni, F., Nazemalhosseini Mojarad, E., et al. (2012). Molecular identification of *Hartmannella vermiformis* and *Vannella persistens* from man-made recreational water environments, Tehran, Iran. *Parasitol. Res*. 111, 835–839. doi: 10.1007/s00436-012-2906-x

Ng, T. F., Willner, D. L., Lim, Y. W., Schmieder, R., Chau, B., Nilsson, C., et al. (2011). Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS ONE* 6:e20579. doi: 10.1371/journal.pone.0020579

Niyyati, M., Rahimi, F., Lasejerdi, Z., and Rezaeian, M. (2014). Potentially pathogenic free-living amoebae in contact lenses of the asymptomatic contact lens wearers. *Iran J. Parasitol*. 9, 14–19.

Pagnier, I., Reteno, D. G., Saadi, H., Boughalmi, M., Gaia, M., Slimani, M., et al. (2013). A decade of improvements in *Mimiviridae* and *Marseilleviridae* isolation from amoeba. *Intervirology* 56, 354–363. doi: 10.1159/000354556

Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., et al. (2013). Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341, 281–286. doi: 10.1126/science.1239181

Reteno, D. G., Benamar, S., Bou Khalil, J., Andreani, J., Armstrong, N., Klose, T., et al. (2015). Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol*. 89, 6585–6594. doi: 10.1128/JVI.00115-15

Rho, M., Tang, H., and Ye, Y. (2010). FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* 38, e191. doi: 10.1093/nar/gkq747

Rosario, K., Capobianco, H., Ng, T. F., Breitbart, M., and Polston, J. E. (2014). RNA viral metagenome of whiteflies leads to the discovery and characterization of a whitefly-transmitted carlavirus in North America. *PLoS ONE* 9:e86748. doi: 10.1371/journal.pone.0086748

Sambou, M., Aubadie-Ladrix, M., Fenollar, F., Fall, B., Bassene, H., Almeras, L., et al. (2015). Comparison of matrix-assisted laser desorption ionization-time of flight mass spectrometry and molecular biology techniques for identification of Culicoides (Diptera: *Ceratopogonidae*) biting midges in Senegal. *J. Clin. Microbiol*. 53, 410–418. doi: 10.1128/JCM.01855-14

Sikes, R. S., and Gannon, W. L. (2011). Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *J. Mammal*. 92, 235–253. doi: 10.1644/10-MAMM-F-355.1

Sime-Ngando, T., Mignot, J. P., Amblard, C., Bourdier, G., Desvilettes, C., Quiblier-Lloberas, C., et al. (1996). Characterization of planktonic virus-like particles in a French mountain lake: methodological aspects and preliminary results. *Ann. Limnol*. 32, 259–263. doi: 10.1051/limn/1996025

Slama, D., Haouas, N., Remadi, L., Mezhoud, H., Babba, H., and Chaker, E. (2014). First detection of *Leishmania infantum* (Kinetoplastida: *Trypanosomatidae*) in *Culicoides* spp. (Diptera: *Ceratopogonidae*). *Parasit. Vectors* 7:51. doi: 10.1186/1756-3305-7-51

Slimani, M., Pagnier, I., Boughalmi, M., Raoult, D., and La Scola, B. (2013). Alcohol disinfection procedure for isolating giant viruses from contaminated samples. *Invervirology* 56, 434–440. doi: 10.1159/000354566

Sokhna, C., Mediannikov, O., Fenollar, F., Bassene, H., Diatta, G., Tall, A., et al. (2013). Point-of-care laboratory of pathogen diagnosis in rural Senegal. *PLoS Negl. Trop. Dis.* 7:e1999. doi: 10.1371/journal.pntd.0001999

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

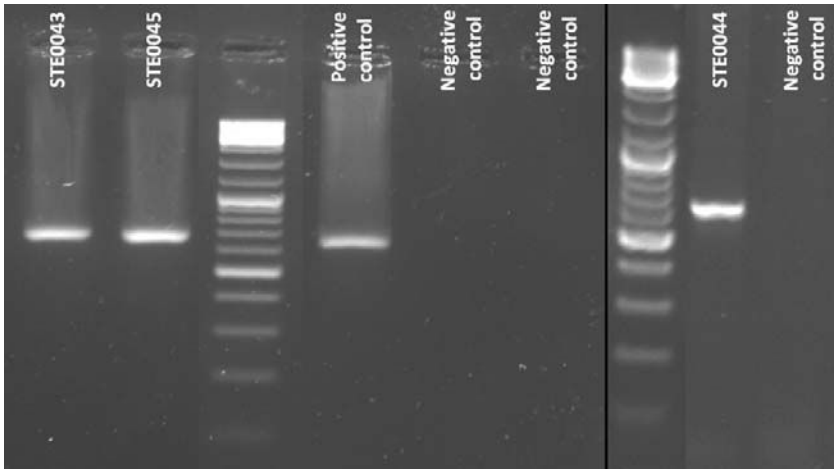
Temmam, S., Monteil-Bouchard, S., Robert, C., Pascalis, H., Michelle, C., Jardot, P., et al. (2015). Host-associated metagenomics: a guide to generating infectious RNA viromes. *PLoS ONE* 10:e0139810. doi: 10.1371/journal.pone.0139810

Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T. H., Stege, J. T., et al. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450, 560–565. doi: 10.1038/nature06269

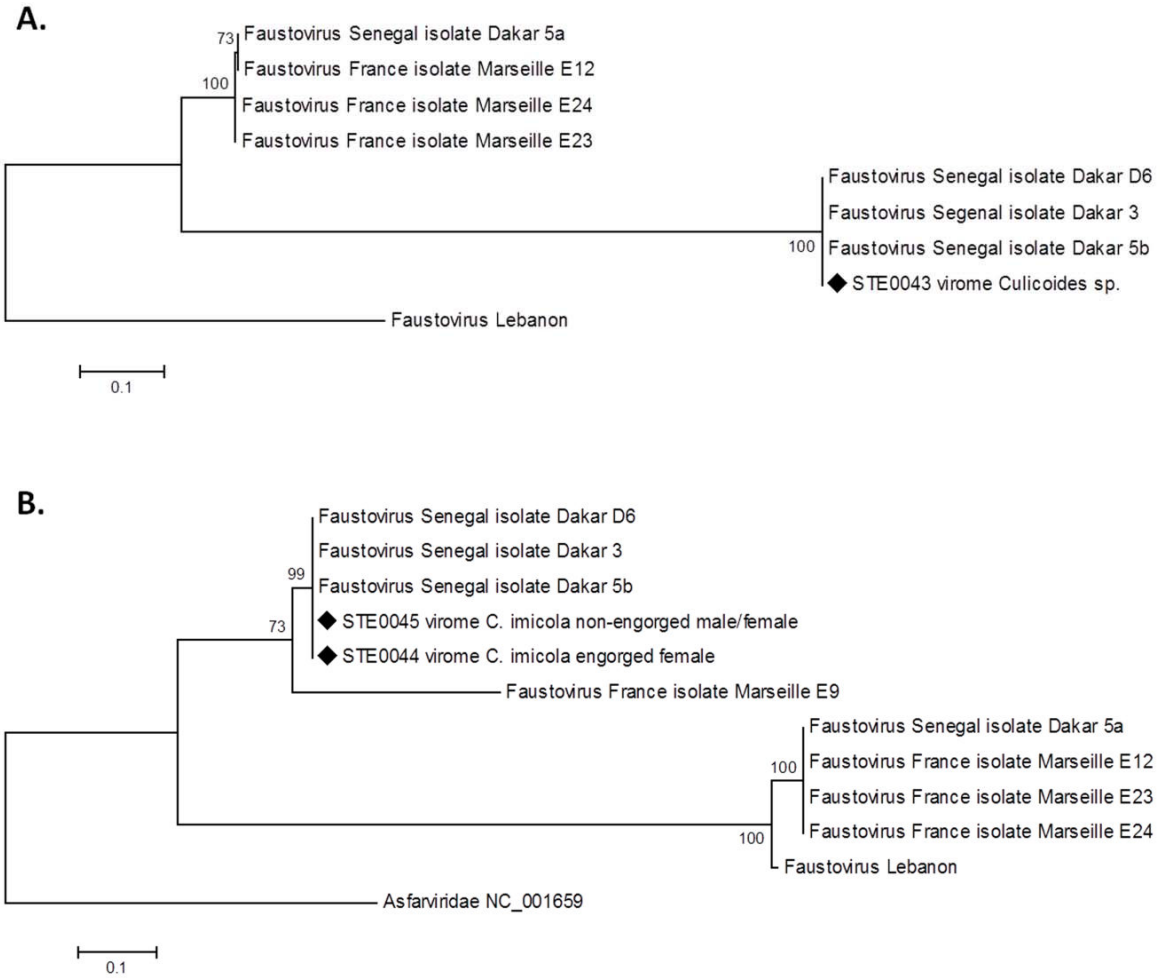
Yoosuf, N., Pagnier, I., Fournous, G., Robert, C., La Scola, B., and Raoult, D. (2014). Complete genome sequence of Courdo11 virus, a member of the family *Mimiviridae*. *Virus Genes* 48, 218–223. doi: 10.1007/s11262-013-1016-x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supporting information



Supplemental Figure 1. Electrophoresis of PCR targeting the Faustovirus capsid conducted on biting midges viromes.



Supplemental Figure 2. Phylogenetic reconstruction of Faustovirus detected in biting midges viromes based on the (A) nucleotide sequence of the DNA topoisomerase small sub-unit (substitution model: HKY) (B) nucleotide sequence of the putative helicase C962R (substitution model: HKY).

Etude du modèle faune sauvage

Article n°5: "Screening for viral pathogens in African simian bushmeat seized at a French airport."

Screening for viral pathogens in African simian bushmeat seized at a French airport

Sarah Temmam[1], Bernard Davoust[1], Anne-Lise Chaber[2,3], Yves Lignereux[4,5], Caroline Michelle[1], Sonia Monteil-Bouchard[1], Didier Raoult[1] and Christelle Desnues[1].

Under revision in: Transboundary and Emerging Diseases

[1] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1095, Aix-Marseille Université, Marseille, France.

[2] Research Unit of Epidemiology and Risk Analysis applied to veterinary sciences (UREAR-ULg), Faculty of Veterinary Medicine, University of Liege, boulevard de Colonster, 42, B-4000 Liège, Belgium.

[3] Wildlife Consultant L.L.C, P.O. Box 13587, Al Ain, United Arab Emirates.

[4] National Veterinary School, INP, 23 chemin des capelles, 31076 Toulouse cedex 03, France.

[5] Natural History Museum, 45 allée Jules-Guesdes, 31000 Toulouse, France.

[*] Corresponding author: christelle.desnues@univ-amu.fr; tel.: +33(0)491324630; fax: +33(0)491387772.

Préambule à l'article "Screening for viral pathogens in African simian bushmeat seized at a French airport"

Dans certains pays d'Afrique, la chasse et la consommation de viande de brousse issue de la faune sauvage sont des pratiques courantes et une source importante de protéines pour certaines populations [99]. La viande de brousse, et notamment la viande de singe, fait aussi partie d'un trafic parallèle illicite très prospère dans les pays occidentaux où elle représente un mets très coûteux et recherché. Il a été estimé par exemple qu'en France, cinq tonnes de denrées interdites originaires d'Afrique sub-saharienne (viande de brousse, viande issue d'animaux d'élevage, poissons, plantes, etc.) entrent illégalement sur le territoire par l'aéroport de Roissy-Charles-de-Gaulle [100]. Ce type de trafic représente donc un danger, à la fois pour la conservation de la biodiversité (notamment lorsque cela concerne des espèces animales en voie de disparition), mais aussi du point de vue sanitaire en terme de santé publique. Il a par exemple été clairement établi que les singes sont des réservoirs ou des hôtes intermédiaires de nombreux virus comme le virus *Marburg* [101-102] ou le *VIH* [103], et ils peuvent donc constituer une source potentielle de virus zoonotiques pour l'homme via la chasse, la manipulation de carcasses fraîches ou la consommation de viande de brousse. Dans cet article, nous avons donc répertorié les virus associés à quatre échantillons de viande de singe interceptée par les douanes de l'aéroport de Toulous-Blagnac afin d'évaluer le risque de transmission de virus zoonotique à l'homme par la manipulation et/ou la consommation de viande de brousse.

Nous avons, dans un premier temps, réalisé une analyse par microscopie en épifluorescence de prélèvements de viande de cercopithèques originaires de Centrafrique qui a montré la présence de particules de taille compatibles avec des particules virales en quantité importante. Puis nous avons séquencé, par des techniques de séquençage à haut débit, les viromes ADN et ARN de ces prélèvements. Nous avons ainsi pu identifier de nombreux bactériophages des familles des *Siphoviridae*, *Myoviridae* et *Podoviridae*, dont certains infectent des bactéries potentiellement pathogènes pour l'homme. Nous n'avons pas détecté de virus zoonotique. Afin d'augmenter la sensibilité de détection, nous avons enfin réalisé des PCR pan-génériques ciblant de nombreux virus zoonotiques, essentiellement à génome ARN, mais sans résultat positif.

En conclusion, et bien que nous n'ayons pas détecté de virus potentiellement pathogènes pour l'homme dans nos échantillons de viande de singe, la présence de nombreuses espèces de bactériophages témoigne de la présence de bactéries dont certaines pourraient être potentiellement pathogènes pour l'homme. Il serait intéressant de conduire une étude similaire sur des prélèvements de viande fraîche car il est possible que les traitements de la viande, à savoir le séchage et le fumage, aient altéré l'intégrité des particules virales.

Screening for viral pathogens in African simian bushmeat seized at a French airport.

Journal:	*Transboundary and Emerging Diseases*
Manuscript ID	TBED-OA-228-15.R3
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	TEMMAM, Sarah; URMITE, Davoust, Bernard; URMITE, Chaber, Anne-Lise; Research Unit of Epidemiology and Risk Analysis applied to veterinary sciences, Faculty of Veterinary Medicine Lignereux, Yves; Natural History Museum, Michelle, Caroline; URMITE, Monteil-Bouchard, Sonia; URMITE, Raoult, Didier; URMITE, Desnues, Christelle; URMITE,
Subject Area:	Conservation, Emerging diseases, Trade/Effects on trade, Transmission, Veterinary epidemiology, Virus, Wildlife, Zoonosis/Zoonotics

Screening for viral pathogens in African simian bushmeat seized at a French airport.

Sarah Temmam[1], Bernard Davoust[1], Anne-Lise Chaber[2,3], Yves Lignereux[4,5], Caroline Michelle[1], Sonia Monteil-Bouchard[1], Didier Raoult[1] and Christelle Desnues[1].

[1] URMITE UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, Faculté de Médecine de la Timone, 27 boulevard Jean Moulin, 13005 Marseille, France.
[2] Research Unit of Epidemiology and Risk Analysis applied to veterinary sciences (UREAR-ULg), Fundamental and Applied Research for Animals & Health (FARAH), Faculty of Veterinary Medicine, University of Liege, boulevard de Colonster, 42, B-4000 Liège, Belgium.
[3] Wildlife Consultant L.L.C, P.O. Box 13587, Al Ain, United Arab Emirates.
[4] National Veterinary School, INP, 23 chemin des capelles, 31076 Toulouse cedex 03, France.
[5] Natural History Museum, 45 allée Jules Guesdes, 31000 Toulouse, France.

Corresponding author: Dr. Christelle Desnues: christelle.desnues@univ-amu.fr; tel.: +33(0)491324630; fax: +33(0)491387772.

Abstract

Illegal bushmeat traffic is an important threat to biodiversity conservation of several endangered species and may contribute to the emergence and spread of infectious diseases in humans. The hunting, manipulation and consumption of wildlife-based products, especially those of primate origin, may be a threat to human health; however, few studies have investigated the role of bushmeat trade and consumption as a potential source of human infections to date. In this study, we report the screening of viral pathogens in African simian game seized by French customs at Toulouse-Blagnac airport. Epifluorescence microscopy revealed the presence of virus-like particles in the samples, and further metagenomic sequencing of the DNA and RNA viromes confirmed the presence of sequences related to the *Siphoviridae*, *Myoviridae* and *Podoviridae* bacteriophage families; some of them infecting bacterial hosts that could be potentially pathogenic for humans. To increase the sensitivity of detection, twelve pan-generic PCRs targeting several viral zoonoses were performed, but no positive signal was detected. A large-scale inventory of bacteria, viruses and parasites is urgently needed to globally assess the risk for human health of the trade, manipulation and consumption of wildlife-related bushmeat.

Keywords

Bushmeat, wild game, virus, zoonoses, epifluorescence microscopy, viral metagenomics, pan-generic PCR.

Introduction

Human zoonoses are directly acquired from animals or indirectly via arthropod bites and are an increasing public health problem. More than two thirds of emerging human pathogens are of zoonotic origin, and of them, more than 70% originate from wildlife (Cutler *et al.*, 2010; One Health Initiative). Hunted wild animals are not only an important source of protein for many poor rural populations (Brashares *et al.*, 2011); they are also an important part of some cultures, for animist practices or to reflect a luxury status in industrialized countries for example. Usually composed of rodent-borne, antelope-borne, primate-borne, and turtle-borne meat, bushmeat is a threat to both biodiversity conservation of several endangered species such as primates and pangolins (Effiom *et al.*, 2013) and public health. The traffic also contributes to decrease the proteins availability in some poor rural regions (Brashares *et al.*, 2011). All industrialized countries are affected by the illegal traffic of bushmeat, as reported by Falk *et al.* in Switzerland (Falk *et al.*, 2013) and by Bair-Brake *et al.* in the United States (Bair-Brake *et al.*, 2014). In 2013 in France, customs reported 647 seizures of endangered species (wildlife and flora), of which Convention on International Trade in Endangered Species (CITES)-listed animals were estimated at 39% of seized bushmeat carcasses (French customs annual report, Chaber *et al.*, 2010). The international illegal traffic of animals has been estimated to total five tons of wild game, livestock meat and fish, carried by passengers arriving from sub-Saharan Africa at Paris Charles de Gaulle Airport (France) (Chaber *et al.*, 2010), of which little is seized and directly destroyed by customs and veterinary services each year. Equivalent traffic probably occur from other countries (Asia or South American continents, for example), but no report was released yet regarding the detailed origins of customs seizures of illegal game entering in France.

In some countries, bats are hunted for their meat, resulting in the infection of humans, as previously reported (Kamins *et al.*, 2014; Pernet *et al.*, 2014). In addition, several known examples of primate hunting have led to the emergence of human epidemics, such as Ebola hemorrhagic fever and AIDS in which the hunting and manipulation of infected

carcasses were shown to be the origins of epidemics (Wolfe *et al.*, 2004). However, few studies have investigated the role of simian bushmeat trade, manipulation and/or consumption as a potential source of human infections. The carriage of pathogens by simian bushmeat, such as parasites (Pourrut *et al.,* 2011), bacteria (Bachand *et al.*, 2012) or viruses (Aghokeng *et al.*, 2010; Smith *et al.*, 2012) were previously reported. For example, in the United States, Smith *et al*. have reported the presence of retroviruses and herpesviruses in non-human primate tissue samples confiscated at different airports (Smith *et al.*, 2012). In France, Chaber *et al*. have reported data suggesting the presence of food-borne pathogens, such as *Listeria monocytogenes*, and carcinogenic concentrations of benzo(a)pyrene in smoked bushmeat confiscated at the Paris Charles de Gaulle and Toulouse-Blagnac airports (Chaber *et al.,* 2015).

With the progress of molecular biology, PCR-based methods, mainly pan-generic PCR, have become the main techniques for virus discovery; however, these techniques require prior knowledge of closely related viral genomes. Next-generation sequencing (NGS) techniques make it possible to sequence all viral genomes in a given sample without any previous knowledge about their nature, using a combination of sequence-independent amplification and high-throughput sequencing. These techniques, which are known as viral metagenomics, are being used exponentially more frequently to characterize the viral diversities of complex environments, such as animal samples (Temmam *et al.*, 2014). Although molecular-biology based methods are not able to discriminate infectious or inactivated viral particles, PCR and NGS-based methods can provide crucial information onto the different viral species that are present in a sample. In this study, we report the first DNA and RNA viral metagenomic analysis and screening of zoonotic viral pathogens in illegally imported African simian bushmeat seized by French customs at Toulouse-Blagnac airport.

Materials and methods

Specimens

A total of 4 specimens of non-human primates originating from the Central African Republic were confiscated by French customs at Toulouse-Blagnac Airport in September 2013. The four animals had been cut into two halves, eviscerated and smoked before importation into France (Figure 1).

Samples were taken from the axillary and popliteal regions, containing lymph nodes and muscle tissues. The tissue samples (samples STE0011 to STE0014) were then directly stored at -20°C by the customs officials until further analyses.

Sample pre-treatment, epifluorescence microscopy and virome preparation

Dilacerated tissues (0.5 cm^3) were rehydrated overnight at +4°C in 2 mL of sterile EMEM (Life Technologies, Saint Aubin, France) and then crushed with two 3-mm tungsten beads and a TissueLyser at 25 Hz for 2 minutes (Qiagen, Courtaboeuf, France). The clarified supernatant was subsequently used as a template for virome preparation.

DNA and RNA viromes of sample n° STE0011 were prepared as previously described (Temmam *et al.*, 2015). Briefly, the clarified supernatant was filtered through a 0.45-µm filter (Millipore, Molsheim, France), and free nucleic acids were digested at 37°C for one hour with the following cocktail of nucleases: 20 U Exonuclease I (New England Biolabs, Évry, France), 25 U Benzonase® (Merck Millipore, Molsheim, France), 25 U RNase A (Roche Diagnostics, Meylan, France), 20 U Turbo DNase (Life Technologies, Saint Aubin, France) and 10 µL of 10X Turbo DNase buffer. The digested supernatant was then deposited onto a discontinuous 66%-30% sucrose gradient and ultracentrifuged in an MLS50 Beckman-Coulter rotor at 130,000 g for 2 hours at +4°C. The viral fraction was harvested at the interphase between the 66% and 30% sucrose layers using a 23G needle.

One hundred microliters of the purified fraction was harvested to assess the Virus-Like Particle (VLP) concentration by fluorescence microscopy, as previously described by Thurber *et al.* (Thurber *et al.*, 2009). All fluorescence images were acquired with a Leica SP5 inverted confocal microscope with 4 lasers, a 100X objective and a numerical aperture of 1.4.

DNA and RNA were extracted from the purified viral fraction with Trizol LS® reagent (Life Technologies, Saint Aubin, France). Two microliters of DNA was randomly amplified in 2 independent reactions using a Genomiphi V3 kit (GE Healthcare, Vélizy-Villacoublay, France) according to the manufacturer's recommendations. Total RNA was processed by random reverse transcription as previously described by Froussard *et al.* (Froussard, 1992) using Superscript III Reverse Transcriptase (Life Technologies, Saint Aubin, France). cDNA was subsequently used as a template for the Klenow reaction and randomly amplified with 2.5 U of Long Amp Taq DNA Polymerase (New England Biolabs, Évry, France) in a final volume of

25 µL, as previously described (*Temmam S. "Host-associated metagenomics: a guide to generating infectious RNA viromes", in revision*).

DNA and RNA amplification products were purified twice with Agencourt AMPure Beads (Beckman-Coulter, Villepinte, France) according to the manufacturer's protocol, eluted to a final volume of 15 µL and sequenced using MiSeq Technology with the paired-end and barcode strategies according to a Nextera XT library kit in a 2 x 300 bp format (Illumina Inc., San Diego, USA).

Bioinformatic analyses of viromes

Raw reads were imported in pairs into CLC Genomics Workbench 6.0.1 program (CLC Bio, Aarhus, Denmark) and trimmed according to their quality score (Illumina pipeline 1.8 and later), the presence of ambiguities (a maximum of 2 ambiguities), length (reads <50 nt long were discarded) and the adaptors and primers used for random PCR. The pre-processed viral metagenomes are publicly available on Metavir server (http://metavir-meb.univ-bpclermont.fr) with the identifiers "STE0011_DNA" and "STE0011_RNA" under the project "Simian_bushmeat" and on MG-RAST server (http://metagenomics.anl.gov/) with the identifiers 4604107.3 and 4604109.3 for the DNA and RNA viromes, respectively.

Cleaned paired reads were assembled into contigs with CLC Genomics program using the following parameters: a word size of 20 nt, minimum contig length of 200 nt, mismatch cost of 2, insertion/deletion cost of 3, length fraction of 0.5 and similarity fraction of 0.8. Contigs and non-assembled reads were then compared to the NCBI nucleotide database using the BlastN algorithm, with a minimum coverage of 50%, minimum identity of 50% and E-value <10-5. Sequences having no significant hits according to the criteria described above were classified as "unknown".

To enhance the detection of viral reads, cleaned paired reads were compared to the NCBI RefSeq viral database using BlastX program with a minimum coverage of 50%, minimum identity of 50% and E-value $<10^{-5}$. Taxonomic assignation of the reads was performed at the family level, and bacterial target hosts of bacteriophages were taxonomically determined at the genus level. Reads taxonomically classified by BlastX were then mapped against reference genomes using CLC Genomics Workbench 6.0.1 program (CLC Bio, Aarhus, Denmark) to verify their correct taxonomic assignation, with the following

parameters: a minimal length fraction of 0.25, minimal similarity fraction of 0.7, mismatch cost of 2 and insertion/deletion cost of 3.

Pan-generic PCRs

Total nucleic acids were extracted from an aliquot of 100 µL of the purified viral fraction with a BioRobot EZ1 and EZ1 Virus Mini Kit (Qiagen, Courtaboeuf, France) in a final volume of 60 µL, and 20 µL was subsequently reverse transcribed into cDNA using SuperScript III Reverse Transcriptase (Life Technologies, Saint Aubin, France) and random hexamers, according to the manufacturer's recommendations.

To control the efficiency of the PCR and the presence of inhibitors originating from the smoked tissue sample extractions, PCR targeting 18S rRNA was performed with 2 µL of cDNA, as previously described (Breitbart *et al.*, 2005).

A total of 12 pan-generic PCRs, nested PCR or real-time PCR were conducted (Supplemental Table) to screen for the presence of several zoonotic viruses possibly infecting humans, including paramyxoviruses (Tong *et al.*, 2008), coronaviruses (de Souza Luna *et al.*, 2007), flaviviruses (Moureau *et al.*, 2007), phleboviruses (Sánchez-Seco *et al.*, 2003), nairoviruses (Rodriguez *et al.*, 1997), poxviruses (Sánchez-Seco *et al.*, 2006), alphaviruses (Sánchez-Seco *et al.*, 2001), hantaviruses (Klempa *et al.*, 2006), orthobunyaviruses (Lambert *et al.*, 2009), arenaviruses (Bowen *et al.*, 1997), filoviruses (Zhai *et al.*, 2007) and herpesviruses (VanDevanter *et al.*, 1996), according to the cited authors' recommendations for primer concentrations and annealing conditions. Briefly, 5 µL of cDNA was used in a final volume of 25 µL for the first round of PCR, along with HotStar Taq DNA Polymerase (Qiagen, Courtaboeuf, France) or a QuantiTect SYBR Green qPCR Kit (Qiagen, Courtaboeuf, France), depending on the pan-generic PCR. A volume of 1 µL of the first-round PCR product was used for nested PCR when needed. PCR products were analyzed on a 2% agarose gel, and bands of the expected size were extracted from the gel, purified with a QIAquick Gel Extraction Kit (Qiagen, Courtaboeuf, France) and sequenced with a Big Dye Terminator v1.1 Cycle Sequencing Kit (Life Technologies, Saint Aubin, France) and an ABI 3130 Genetic Analyzer (Life technologies, Saint Aubin, France).

Results

Specimen collection

Four specimens of non-human primates originating from the Central African Republic were confiscated by French customs at the Toulouse-Blagnac Airport in 2013. Their specific determination was impossible due to their conservation statuses (Figure 1), but they all belonged to the *Cercopithecidae* family.

Detection and quantification of virus-like particles by epifluorescence microscopy

The estimated numbers of VLPs were 8.85×10^6 VLPs/mL, 8.28×10^4 VLPs/mL, 3.72×10^6 VLPs/mL and 1.73×10^7 VLPs/mL for the samples STE0011, STE0012, STE0013 and STE0014, respectively (Figure 2).

DNA and RNA viromes of simian bushmeat sample

The DNA and RNA viromes of sample STE0011 were sequenced with Illumina MiSeq technology. After trimming, the DNA and RNA viromes contained 647,272 and 519,397 paired reads, respectively.

The results of the BlastN search against the nucleotide NCBI database are presented in Figure 3A and 3C, respectively, for the DNA and RNA viromes. The unknown fractions of the DNA and RNA virome datasets represented 25.55% and 6.43%, respectively. Among the known DNA sequences, bacterial sequences represented 74.84%, eukaryotes comprised 25.14%, and viruses included 0.03% of the total assigned reads. Within the known RNA virome, sequences related to bacteria represented the majority of the reads (more than 92% of the total assigned reads), eukaryotic reads comprised 7.23%, and viral reads included 0.64% of the total assigned reads.

The eukaryotic reads were related to humans and primates (95.21% and 63.70% of the total eukaryotic reads for the DNA and RNA metagenomes, respectively) in addition to plants (1.48% and 0.43% of the total eukaryotic reads for the DNA and RNA metagenomes, respectively) and arthropods (0.01% and 11.90% of the total eukaryotic reads for the DNA and RNA metagenomes, respectively). Parasite (mainly nematodes and helminthes) eukaryotic reads were also detected in the DNA and RNA viromes (0.96% and 21.26%, respectively), which probably remained from undigested free DNA. Among them, *Spirometra erinaceieuropaei* platyhelminth, a worm commonly infecting domestic animals that causes several diseases in humans, was quite abundant (N=1,503 reads) in the DNA metagenome. Additionally, *Haemonchus placei*, a nematode mainly infecting cattle in tropical areas, was detected in 558 reads; 125 reads were related to *Wuchereria bancrofti*, a nematode

responsible for lymphatic filariasis in humans, and 5 and 8 reads of the DNA virome were assigned to *Leishmania* and *Plasmodium spp.*, respectively. Finally, we observed the presence of 2,983 sequences representing *Schistosoma rodhaini* in the RNA metagenome, a trematode responsible for infections in small mammals.

The results of the BlastX search against the RefSeq viral database are presented in Figure 3B and 3D for the DNA and RNA viromes, respectively. More than two thirds of the total viral reads of the DNA virome belonged to bacteriophages, the majority of which were from *Siphoviridae* (51.38% of the viral reads) and *Myoviridae* (24.13% of the viral reads). *Podoviridae*, *Microviridae* and non-classified phages were less abundant, comprising only 3.74%, 0.02% and 3.79% of the total viral reads, respectively (Figure 3B). The majority of bacterial genera infected by *Myoviridae* belonged to the *Bacillus sp.* genus (46.31%), followed by *Lactobacillus sp.* (12.09%), *Enterococcus sp.* (9.12%), *Cronobacter sp.* (7.83%) and *Enterobacteria sp.* (6.89%) (Supplemental Figure 1). In addition, more than 70% of the bacterial genera infected by *Podoviridae* were *Planktothrix sp.* (71.90%) and *Cellulophaga sp.* (16.01%). Most of the *Siphoviridae* bacterial hosts belonged to the *Enterococcus sp.* (44.46%), *Staphylococcus sp.* (15.51%), *Lactococcus sp.* (14.87%) and *Listeria sp.* (7.76%) genera (Supplemental Figure 1).

Sequences belonging to the *Phycodnaviridae* (N=651 reads), *Iridoviridae* (N=29 reads), *Nudiviridae* (N=21 reads) and *Baculoviridae* (N=4 reads) families (Figure 3B) were also detected. *Iridoviridae*, *Nudiviridae* and *Baculoviridae* are insect-infecting viruses, whereas *Phycodnaviridae* are large viruses infecting algae. Several reads belonging to other *Megavirales*, an order of large viruses that infect eukaryotic hosts (Colson *et al.*, 2013), were present in the DNA virome, belonging to the *Mimiviridae* family (N=503 reads), the non-classified Faustovirus (N=982 reads) and Pandoravirus (N=89 reads) (Figure 3B). Further verification of the taxonomic assignation of reads in the *Poxviridae* (N=258 reads), *Ascoviridae* (N=61 reads), *Herpesviridae* (N=49 reads) and *Papillomaviridae* (N=1 read) families (Figure 3B) by mapping against reference genomes showed that they corresponded to repeated patterns and thus could not be confidently attributed to these families.

The sequences of the RNA virome were related to *Siphoviridae*, *Myoviridae* and *Podoviridae* bacteriophages or *Mimiviridae*, *Phycodnaviridae* and *Ascoviridae*. These sequences probably reflected the DNA remaining in the RNA fraction during Trizol LS®

extraction (Figure 3D). Similar to the DNA virome, the presence of *Poxviridae* was verified by mapping the reads against reference genomes, which did not result in good taxonomic assignation. Reads belonging to the *Retroviridae* and *Anelloviridae* families may reflect the presence of residual simian genomic material and blood, respectively. Only two reads of RNA viruses belonging to the *Tymoviridae* family were detected in the RNA virome. *Tymoviridae* are non-enveloped plant-infecting RNA viruses. The presence of such viruses may be linked with the diets of the primates, *i.e.* the plants, fruits or insects that they ingest (Figure 3D).

Pan-generic PCR screening of zoonotic viruses

Several studies have shown that the NGS sequencing of viromes is a less sensitive technique than PCR (Frey *et al.*, 2014; Cheval *et al.*, 2011). Moreover, due to the limited depth of sequencing and the high abundance of bacteriophages detected by metagenomics, the presence of rare viral species, especially zoonotic viruses, may have been missed. Thus, we used pan-generic PCR to screen the four bushmeat samples for the presence of most human viral zoonoses, including paramyxoviruses, coronaviruses, flaviviruses, phleboviruses, nairoviruses, alphaviruses, hantaviruses, orthobunyaviruses, arenaviruses, filoviruses, poxviruses and herpesviruses; however, none of the 4 specimens tested positive using these 12 pan-generic PCRs.

The presence of inhibitors and the efficiencies of the nucleic acid extractions from the smoked tissue samples were controlled by performing PCR targeting 18S rRNA. The four samples tested positive (Supplemental Figure 2), indicating the good efficiency of the amplification of DNA originating from smoked tissues.

Discussion

Quantifying the global wildlife trade is hard since it ranges from live to dead animals, from local barter to major international routes, and is almost always conducted illegally or through informal networks (Karesh *et al.*, 2005). However, illegal bushmeat traffic may contribute to the emergence and spread of infectious diseases in humans and need to be addressed. The origin of AIDS epidemics (Gao *et al.*, 1999), case reports of hepatitis E originating from wild boars (Vasickova *et al.*, 2007; Meng *et al.*, 2009), brucellosis (CDC, 2009), Ebola previous epidemics (Pourrut *et al.*, 2005), and trichinellosis (Roy *et al.*, 2003) are some examples of food-borne illnesses acquired after hunting of wild animals (Karesh *et*

al., 2012). Indeed, some human food-borne pathologies may either be caused by the consumption or during the preparation (*i.e.* butchering, cutting or washing) of meat originating from infected animals. In France, important quantities of meat enter illegally each year but only few are seized by customs (Chaber *et al.*, 2010) because these are out of the priorities of customs officers. Moreover, since 2014 Ebola epidemics and the potential risks of contamination for customs workers, the difficulties to sample bushmeat for research purposes before their immediate destruction have increased. In this context, it is important to obtain the necessary support from relevant authorities to seize and analyze wild game before their destruction. A transparent partnership between customs and health authorities will provide the best opportunity for improving the effectiveness of efforts to control the risk of international bushmeat trade.

In this study, we were able to obtain four non-human primate samples that were first screened for the presence of virus-like particles using epifluorescence microscopy. Estimated viral loads were high, ranging from 10^4 to 10^7 VLPs/mL depending on the sample, but these may have been overestimated due to the use of a non-specific fluorochrome which can also stain membrane-derived vesicles or cell debris (Forterre *et al.*, 2013). Observations confirmed the absence of bacterial contamination in the treated samples.

We thus decided to go further and sequence the purified viral fraction of one bushmeat sample. Viral metagenomic analyses performed on the simian game showed that a large majority of the sequences were related to bacteriophages belonging to the *Siphoviridae*, *Myoviridae*, and *Podoviridae* families. Regardless of infectivity, the bacterial hosts of these viruses mainly belong to the *Firmicutes* and *Proteobacteria* phyla, which include environmental bacteria in addition to potential human pathogenic species, such as some of the *Staphylococcus sp.*, *Listeria sp.*, and *Enterococcus sp*. For example, *Cronobacter spp.* (formerly *Enterobacter sakazakii*) includes a group of Gram-negative bacteria ubiquitously found in the environment. *Cronobacter spp.* are particularly resistant to osmotic stress and elevated temperatures, explaining their ability to survive in a wide variety of dried foods, such as infant formulas (Edelson-Mammel *et al.*, 2005; Estuningsih *et al.*, 2006). They are emerging foodborne pathogens (Estuningsih *et al.*, 2006) that may be responsible for life-threatening infections in newborns, immunocompromised adults and the elderly (Mullane *et al.*, 2007; Bowen *et al.*, 2006; Gosney *et al.*, 2006; Than *et al.*, 2007). The

presence of *Cronobacter* and other pathogenic bacteria, and the risk of human infection caused by bushmeat manipulation and consumption have to be addressed. The source of bacterial contamination of the bushmeat, due for example to non-hygienically handlings (during hunting, carcass evisceration, smoking treatments, or during unsafe transportation conditions), or due to the presence of several bacteria in the original tissue samples, is actually unknown. But whatever the origin of micro-organisms, their presence results in the same potential risk of transmission to humans. A recent study by Chaber *et al*. reported the detection of viable aerobic bacteria above levels considered safe for human consumption, and unsafe levels of carcinogens in fish (Chaber *et al.*, 2015). An exhaustive inventory of the bacterial communities present in bushmeat samples, for example by high-throughput sequencing of the 16S rRNA gene and by isolation, would help to clarify and identify the potential risks of introducing emerging bacterial pathogens by the illegal import of bushmeat.

~~In addition, in order to propose measures for consumers that wild game is well handled and cooked regarding viral, bacterial and parasitic biohazard, the effects of transformation treatments (such as smoking) on the external and internal contamination of meat, have also to be addressed.~~

Recent studies have demonstrated that metagenomic analyses are less sensitive than PCR-based methods for detecting viruses present in low abundance. Indeed, the detection limit for next-generation sequencing techniques is estimated to be 10^3 to 10^4 genomic copies (Frey *et al.*, 2014; Cheval *et al.*, 2011), while PCR-based methods, especially those based on a nested format, can detect up to one genomic copy. Moreover, analysis of the rarefaction curve of bushmeat metagenomes revealed a plateau (Supplemental Figure 3), indicating that full sequence diversity and characterization were not achieved. Thus, we screened for the presence of several zoonotic viruses by pan-generic PCR. No human viral pathogens were detected using the 12 pan-generic PCRs targeting the most common viral zoonoses, although most PCR were in nested format and presented low limits of detection. In addition, the control PCR targeting the 18S rRNA gene demonstrated that the negative results were not due to the presence of inhibitors and/or degraded nucleic acids after smoking treatments. These findings could suggest that either the twelve targeted viral genera were not present in the studied populations at the time of sampling or that the viral particles and their genomic

content were degraded during the smoking and drying of the bushmeat, resulting in their increased sensitivity to nucleases used to purify viromes. In addition, the small sampling size could explain this negative result, especially if a low prevalence of the targeted viruses occurs.

To our knowledge, no outbreaks directly linked to the consumption of contaminated ~~wild game~~bushmeat (*e.g.* of primate, rodent or antelope origins) have been officially reported in France, but outbreaks may have been underestimated as the illegal origin of the products may prevent diseased consumers to declare consumption. Although smoking and thoroughly cooking the meat may reduce the risk of consumption, human contamination may still occur during handling infected meat that escaped sanitary controls. This risk is even higher for people manipulating and consuming fresh carcasses (Subramanian *et al*., 2012; Paige *et al*., 2014; Wolfe *et al*., 2005) which is of great concerns as international traffic of bushmeat also frequently involves fresh products. Large-scale studies targeting the bacterial, viral and parasitic levels from wild-caught game of different geographical and/or animal origins and, whenever possible, using large cohorts are urgently needed to identify risks for human health. These studies should be accompanied with extensive diffusion of information to exposed populations (regarding viral, bacterial and parasitic biohazard) about precautions that have to be taken while manipulating/consuming bushmeat. In addition, the effects of transformation treatments (such as smoking) on the external and internal contamination of meat have also to be addressed.

Formatted: Font: Italic

Acknowledgments

The authors thank Dr. Sonia Kaeuffer from the DDSV (France) for taking care of the material, David Danède from DREAL Midi-Pyrénées (France) for his help in obtaining permits of sampling protected animals seized at the international Toulouse-Blagnac Airport, and Pascal Weber for his help with confocal microscopy. This work was partially supported by ANR-13-JSV6-0004.

References

Aghokeng, A.F., A. Ayouba, E. Mpoudi-Ngole, S. Loul, F. Liegeois, E. Delaporte, and M. Peeters, 2010: Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions. Infect. Genet. Evol. 10, 386-96.

Bachand, N., A. Ravel, R. Onanga, J. Arsenault, and J.P. Gonzalez, 2012: Public health significance of zoonotic bacterial pathogens from bushmeat sold in urban markets of Gabon, Central Africa. J. Wildl. Dis. 48, 785-9.

Bair-Brake, H., T. Bell, A. Higgins, N. Bailey, M. Duda, S. Shapiro, H.E. Eves, N. Marano, and G. Galland, 2014: Is that a rodent in your luggage? A mixed method approach to describe bushmeat importation into the United States. Zoonoses Public Health. 61, 97-104.

Bowen, A.B., and C.R. Braden, 2006: Invasive Enterobacter sakazakii disease in infants. Emerg. Infect. Dis. 12, 1185-1189.

Bowen, M.D., C.J. Peters, and S.T. Nichol, 1997: Phylogenetic analysis of the Arenaviridae: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. Mol. Phylogenet. Evol. 8, 301-16.

Brashares, J.S, C.D. Golden, K.Z. Weinbaum, C.B. Barrett, and G.V. Okello, 2011: Economic and geographic drivers of wildlife consumption in rural Africa. Proc. Natl. Acad. Sci. USA. 108, 13931-6.

Breitbart, M., and F. Rohwer, 2005: Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. Biotechniques. 39, 729-736.

Chaber, A.L., S. Allebone-Webb, Y. Lignereux, A. Cunningham, and J.M. Rowcliffe, 2010: The scale of illegal meat importation from Africa to Europe via Paris. Conservation Letters. 00, 1-7.

Chaber, A.L., and A. Cunningham, 2015: Public Health Risks from Illegally Imported African Bushmeat and Smoked Fish: Public Health Risks from African Bushmeat and Smoked Fish. EcoHealth, in press.

Centers for Disease Control and Prevention (CDC), 2009: *Brucella suis* infection associated with feral swine hunting - three states, 2007-2008. MMWR Morb Mortal Wkly Rep. 58, 618-621.

Cheval, J., V. Sauvage, L. Frageul, L. Dacheux, G. Guignon, N. Dumey, K. Pariente, C. Rousseaux, F. Dorange, N. Berthet, S. Brisse, I. Moszer, H. Bourhy, J.C. Manuguerra, M. Lecuit, A. Burguiere, V. Caro, and M. Eloit, 2011: Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. J. Clin. Microbiol. 49, 3268-3275.

Colson, P., X. de Lamballerie, N. Yutin, S. Asgari, Y. Bigot, D.K. Bideshi, X.W. Cheng, B.A. Federici, J.L. Van Etten, E.V. Koonin, B. La Scola, and D. Raoult, 2013: "Megavirales", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. Arch. Virol. 158, 2517-2521.

Cutler, S.J., A.R. Fooks, and W.H. van der Poel, 2010: Public health threat of new, reemerging, and neglected zoonoses in the industrialized world. Emerg. Infect. Dis. 16, 1-7.

de Souza Luna, L.K., V. Heiser, N. Regamey, M. Panning, J.F. Drexler, S. Mulangu, L. Poon, S. Baumgarte, B.J. Haijema, L. Kaiser, and C. Drosten, 2007: Generic detection of coronaviruses and differentiation at the prototype strain level by reverse transcription-PCR and nonfluorescent low-density microarray. J. Clin. Microbiol. 45, 1049-52.

Edelson-Mammel, S.G., M.K. Porteous, and R.L. Buchanan, 2005: Survival of Enterobacter sakazakii in a dehydrated powdered infant formula. J. Food Prot. 68, 1900-1902.

Effiom, E.O., G. Nuñez-Iturri, H.G. Smith, U. Ottosson, and O. Olsson, 2013: Bushmeat hunting changes regeneration of African rainforests. Proc. Biol. Sci. 280, 20130246.

Estuningsih, S., C. Kress, A.A. Hassan, O. Akineden, E. Schneider, and E. Usleber, 2006: Enterobacteriaceae in dehydrated powdered infant formula manufactured in Indonesia and Malaysia. J. Food Prot. 69, 3013-3017.

Falk, H., S. Dürr, R. Hauser, K. Wood, B. Tenger, M. Lörtscher, and G. Schüpbach-Regula, 2013: Illegal import of bushmeat and other meat products into Switzerland on commercial passenger flights. Rev. Sci. Tech. 32, 727-739.

Forterre, P., N. Soler, M. Krupovic, E. Marguet, and H.W. Ackermann, 2013: Fake virus particles generated by fluorescence microscopy. Trends Microbiol, 21, 1-5.

French customs annual report, available on 2015.09.29 at : http://www.douane.gouv.fr/Portals/0/fichiers/datadouane/publication-douane/bilans-resultats/resultats-2013.pdf

Gao, F., E. Bailes, D.L. Robertson, Y. Chen, C.M. Rodenburg, S.F. Michael, L.B. Cummins, L.O. Arthur, M. Peeters, G.M. Shaw, P.M. Sharp, and B.H. Hahn, 1999: Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. Nature. 397, 436-441.

Frey, K.G., J.E. Herrera-Galeano, C.L. Redden, T.V. Luu, S.L. Servetas, A.J. Mateczun, V.P. Mokashi, and K.A. Bishop-Lilly, 2014: Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. BMC Genomics. doi: 10.1186/1471-2164-15-96.

Froussard, P. 1992: A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. Nucleic Acids Res. 20:2900.

Gosney, M.A., M.V. Martin, A.E. Wright, and M. Gallagher, 2006: Enterobacter sakazakii in the mouths of stroke patients and its association with aspiration pneumonia. Eur. J. Intern. Med. 17, 185-188.

Healy, B., S. Cooney, S. O'Brien, C. Iversen, P. Whyte, J. Nally, J.J. Callanan, and S. Fanning, 2010: Cronobacter (Enterobacter sakazakii): an opportunistic foodborne pathogen. Foodborne Pathog. Dis. 7, 339-350.

Kamins, A.O., J.M. Rowcliffe, Y. Ntiamoa-Baidu, A.A. Cunningham, J.L. Wood, and O. Restif, 2014: Characteristics and risk perceptions of Ghanaians potentially exposed to bat-borne zoonoses through bushmeat. EcoHealth. doi:10.1007/s10393-014-0977-0.

Karesh, W.B., R.A. Cook, E.L. Bennett, and J. Newcomb, 2005: Wildlife trade and global disease emergence. Emerg. Infect. Dis. 11, 1000-1002.

Karesh, W.B., E. Loh, and C. Machalaba, 2012: Food safety: a view from the wild side. In "Improving food safety through a One Health approach: workshop summary", National Academies Press (US).

Klempa, B., E. Fichet-Calvet, E. Lecompte, B. Auste, V. Aniskin, H. Meisel, C. Denys, L. Koivogui, J. ter Meulen, and D.H. Krüger, 2006: Hantavirus in African wood mouse, Guinea. Emerg. Infect. Dis. 12, 838-40.

Lambert, A.J., and R.S. Lanciotti, 2009: Consensus amplification and novel multiplex sequencing method for S segment species identification of 47 viruses of the *Orthobunyavirus*, *Phlebovirus*, and *Nairovirus* genera of the family *Bunyaviridae*. J. Clin. Microbiol. 47, 2398-404.

Meng, X.J., D.S. Lindsay, and N. Sriranganathan, 2009: Wild boars as sources for infectious diseases in livestock and humans. Philos Trans R Soc Lond B Biol Sci., 364, 2697-2707.

Moureau, G., S. Temmam, J.P. Gonzalez, R.N. Charrel, G. Grard, and X. de Lamballerie, 2007: A real-time RT-PCR method for the universal detection and identification of flaviviruses. Vector Borne Zoonotic Dis. 7, 467-77.

Mullane, N.R., C. Iversen, B. Healy, C. Walsh, P. Whyte, P.G. Wall, T. Quinn, and S. Fanning, 2007: Enterobacter sakazakii, an emerging bacterial pathogen with implications for infant health. Minerva Pediatr. 59, 137-148.

One Health Initiative. Available at: http://www.onehealthinitiative.com/

Paige, S.B., S.D. Frost, M.A. Gibson, J.H. Jones, A. Shankar, W.M. Switzer, N. Ting, and T.L. Goldberg, 2014: Beyond bushmeat: animal contact, injury, and zoonotic disease risk in western Uganda. EcoHealth. doi: 10.1007/s10393-014-0942-y.

Pernet, O., B.S. Schneider, S.M. Beaty, M. LeBreton, T.E. Yun, A. Park, T.T. Zachariah, T.A. Bowden, P. Hitchens, C.M. Ramirez, P. Daszak, J. Mazet, A.N. Freiberg, N.D. Wolfe, and B. Lee, 2014: Evidence for henipavirus spillover into human populations in Africa. Nat. Commun. 5, 5342.

Pourrut, X., B. Kumulungui, T. Wittmann, G. Moussavou, A. Delicat, P. Yaba, D. Nkoghe, J.P. Gonzalez, and E.M. Leroy, 2005: The natural history of Ebola virus in Africa. Microbes and Infection. 7, 1005-1014.

Pourrut, X., J.L. Diffo, R.M. Somo, C.F. Bilong Bilong, E. Delaporte, M. LeBreton, and J.P. Gonzalez, 2011: Prevalence of gastrointestinal parasites in primate bushmeat and pets in Cameroon. Vet. Parasitol. 175, 187-91.

Rodriguez, L.L., G.O. Maupin, T.G. Ksiazek, P.E. Rollin, A.S. Khan, T.F. Schwarz, R.S. Lofts, J.F. Smith, A.M. Noor, C.J. Peters, and S.T. Nichol, 1997: Molecular investigation of a multisource outbreak of Crimean-Congo hemorrhagic fever in the United Arab Emirates. Am. J. Trop. Med. Hyg. 57, 512-8.

Roux, S., M. Faubladier, A. Mahul, N. Paulhe, A. Bernard, D. Debroas, and F. Emault, 2011: Metavir: a web server dedicated to virome analysis. Bioinformatics, 27, 3074-3075.

Roy, S.L., A.S. Lopez, and P.M. Schantz, 2003: Trichinellosis surveillance--United States, 1997-2001. MMWR Surveill Summ. 52, 1-8.

Sánchez-Seco, M.P., D. Rosario, E. Quiroz, G. Guzmán, and A. Tenorio, 2001: A generic nested-RT-PCR followed by sequencing for detection and identification of members of the Alphavirus genus. J. Virol. Methods. 95, 153-61.

Sánchez-Seco, M.P., J.M. Echevarría, L. Hernández, D. Estévez, J.M. Navarro-Marí, and A. Tenorio, 2003: Detection and identification of Toscana and other phleboviruses by RT-nested-PCR assays with degenerated primers. J. Med. Virol. 71, 140-9.

Sánchez-Seco, M.P., L. Hernández, J.M. Eiros, A. Negredo, G. Fedele, and A. Tenorio, 2006: Detection and identification of orthopoxviruses using a generic nested PCR followed by sequencing. Br. J. Biomed. Sci. 63, 79-85.

Smith, K.M., S.J. Anthony, W.M. Switzer, J.H. Epstein, T. Seimon, H. Jia, M.D. Sanchez, T.T. Huynh, G.G. Galland, S.E. Shapiro, J.M. Sleeman, D. McAloose, M. Stuchin, G. Amato, S.O. Kolokotronis, W.I. Lipkin, W.B. Karesh, P. Daszak, and N. Marano, 2012: Zoonotic viruses associated with illegally imported wildlife products. PLos One. doi:10.1371/journal.pone.0029505.

Subramanian, M. 2012: Zoonotic disease risk and the bushmeat trade: assessing awareness among hunters and traders in Sierra Leone. EcoHealth 9, 471-482.

Temmam, S., B. Davoust, J.M. Berenger, D. Raoult, and C. Desnues, 2014: Viral metagenomics on animals as a tool for the detection of zoonoses prior to human infection? Int. J. Mol. Sci. 15, 10377-97.

Temmam, S., S. Monteil-Bouchard, C. Robert, H. Pascalis, C. Michelle, P. Jardot, R. Charrel, D. Raoult, and C. Desnues, 2015: Host-associated metagenomics: a guide to generating infectious RNA viromes. PLos One. *In production*.

Than, H.A., and T. Tang, 2007: Enterobacter sakazakii bacteraemia with multiple splenic abscesses in a 75-year-old woman: a case report. Age Ageing. 36, 595-596.

Thurber, R.V., M. Haynes, M. Breitbart, L. Wegley, and F. Rohwer, 2009: Laboratory procedures to generate viral metagenomes. Nat. Protoc. 4, 470-83.

Tong, S., S.W. Chern, Y. Li, M.A. Pallansch, and L.J. Anderson, 2008: Sensitive and broadly reactive reverse transcription-PCR assays to detect novel paramyxoviruses. J. Clin. Microbiol. 46, 2652-8.

Vasickova, P., I. Psikal, P. Chalupa, M. Holub, R. Svoboda, and I. Pavlik, 2007: Hepatitis E virus: A review. Veterinarni Medicina. 52, 365–384.

VanDevanter, D.R., P. Warrener, L. Bennett, E.R. Schultz, S. Coulter, R.L. Garber, and T.M. Rose, 1996: Detection and analysis of diverse herpesviral species by consensus primer PCR. J. Clin. Microbiol. 34, 1666-71.

Wolfe, N.D., P. Daszak, A.M. Kilpatrick, and D.S. Burke, 2005: Bushmeat hunting, deforestation, and prediction of zoonotic disease emergence. Emerg. Infect. Dis. 11, 1822-1827.

Wolfe, N.D., T.A. Prosser, J.K. Carr, U. Tamoufe, E. Mpoudi-Ngole, J.N. Torimiro, M. LeBreton, F.E. McCutchan, D.L. Birx, and D.S. Burke, 2004: Exposure to nonhuman primates in rural Cameroon. Emerg. Infect. Dis. 10, 2094-2099.

Woolhouse, M., F. Scott, Z. Hudson, R. Howey, and M. Chase-Topping, 2012: Human viruses: discovery and emergence. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 367, 2864-71.

Zhai, J., G. Palacios, J.S. Towner, O. Jabado, V. Kapoor, M. Venter, A. Grolla, T. Briese, J. Paweska, R. Swanepoel, H. Feldmann, S.T. Nichol, and W.I. Lipkin, 2007: Rapid molecular strategy for filovirus detection and characterization. J. Clin. Microbiol. 45, 224-6.

Figure legends

Figure 1. Primate bushmeat specimens seized at the Toulouse-Blagnac Airport.

Figure 2. Fluorescence microscopy of VLPs in the bushmeat sample n° STE0014. All images were acquired with a Leica SP5 inverted confocal microscope with 4 lasers, a 100X objective and a numerical aperture of 1.4. The scale bar represents 20 µm.

Figure 3. Taxonomic assignation of reads according to a BlastN search against the NCBI nucleotide database (A and C for the DNA and RNA viromes, respectively) and according to a BlastX search against the viral RefSeq NCBI database (B and D for the DNA and RNA viromes, respectively).

Supplemental Figure 1. Bacterial genera infected by bacteriophages detected in the virome dataset.

Supplemental Figure 2. Agarose gel electrophoresis of the 18S rRNA PCR performed on the four bushmeat samples.

Supplemental Figure 3. Rarefaction curve of DNA and RNA metagenomes according to a similarity cut-off of 75% for clustering, generated through METAVIR software (Roux et al., 2011).

1 **Tables**

2

3 **Supplemental Table.** Primers used in this study.

Pan-generic PCR	Primers	1st round PCR	2nd round PCR	Primers concentration (nM)	Annealing	Bibliographie
Flavivirus	PF1S PF2R-bis	SYBR Green	N/A	550	50°C	Moureau *et al.*, 2007
Paramyxovirus	PAR-F1 PAR-F2 PAR-R	Hot Star Taq	Hot Star Taq	400	48°C	Tong *et al.*, 2008
	RES-MOR-HEN-F1 RES-MOR-HEN-F2 RES-MOR-HEN-R	Hot Star Taq	Hot Star Taq	400	48°C	
Coronavirus	PC2S2-primerA PC2S2-primerB PC2As1-primerA PC2As1-primerB PC2As1-primerC PCS-primerA PCS-primerB PCNAs	Hot Star Taq	SYBR Green	F1 : 200, R1 : 900 F2 : 80, R2 : 400	I : 42°C II: 60°C	de Souza Luna *et al.*, 2007
Phlebovirus	NPHLEBO1+ NPHLEBO1- NPhlebo2+ NPhlebo2-	Hot Star Taq	Hot Star Taq	400	60°C	Sánchez-Seco *et al.*, 2003
Nairovirus	CCHF-F2 CCHF-R3 CCHF-F3 CCHF-R2	Hot Star Taq	SYBR Green	200	60°C	Rodriguez *et al.*, 1997
Poxvirus	PoxS361 PoxAS989 PoxS457 PoxAS927	Hot Star Taq	SYBR Green	800	55°C	Sánchez-Seco *et al.*, 2006
Orthobunyavirus	Cal/Bwa-F Cal/Bwa-R Bun-F Bun-R Weo-F Weo-R Oro-F Oro-R	Hot Star Taq	N/A	1000	55°C	Lambert *et al.*, 2009
Hantavirus	HAN-L-F1 HAN-L-R1 HAN-L-F2 HAN-L-R2	Hot Star Taq	Hot Star Taq	400	I : 48°C II : 50°C	Klempa *et al.*, 2006
Alphavirus	Alpha1+ Alpha1- Alpha2+	Hot Star Taq	SYBR Green	400	52°C	Sánchez-Seco *et al.*, 2001

Arenavirus	Alpha2-ARE-1010C ARE-NW1696R ARE-OW1696R	Hot Star Taq	N/A	1000	45°C	Bowen *et al.*, 1997
Herpesvirus	Herpes-DFA Herpes-ILK Herpes-KG1 Herpes-TGV Herpes-IYG	Hot Star Taq	SYBR Green	1000	46°C	VanDevanter *et al.*, 1996
Filovirus	Filo-A-F Filo-A-R Filo-B-F Filo-B-R Filo-C-F Filo-C-R Filo-D-F Filo-D-R	Hot Star Taq	N/A	500	58°C	Zhai *et al.*, 2007

1

Figure 1. Primate bushmeat specimens seized at the Toulouse-Blagnac Airport.

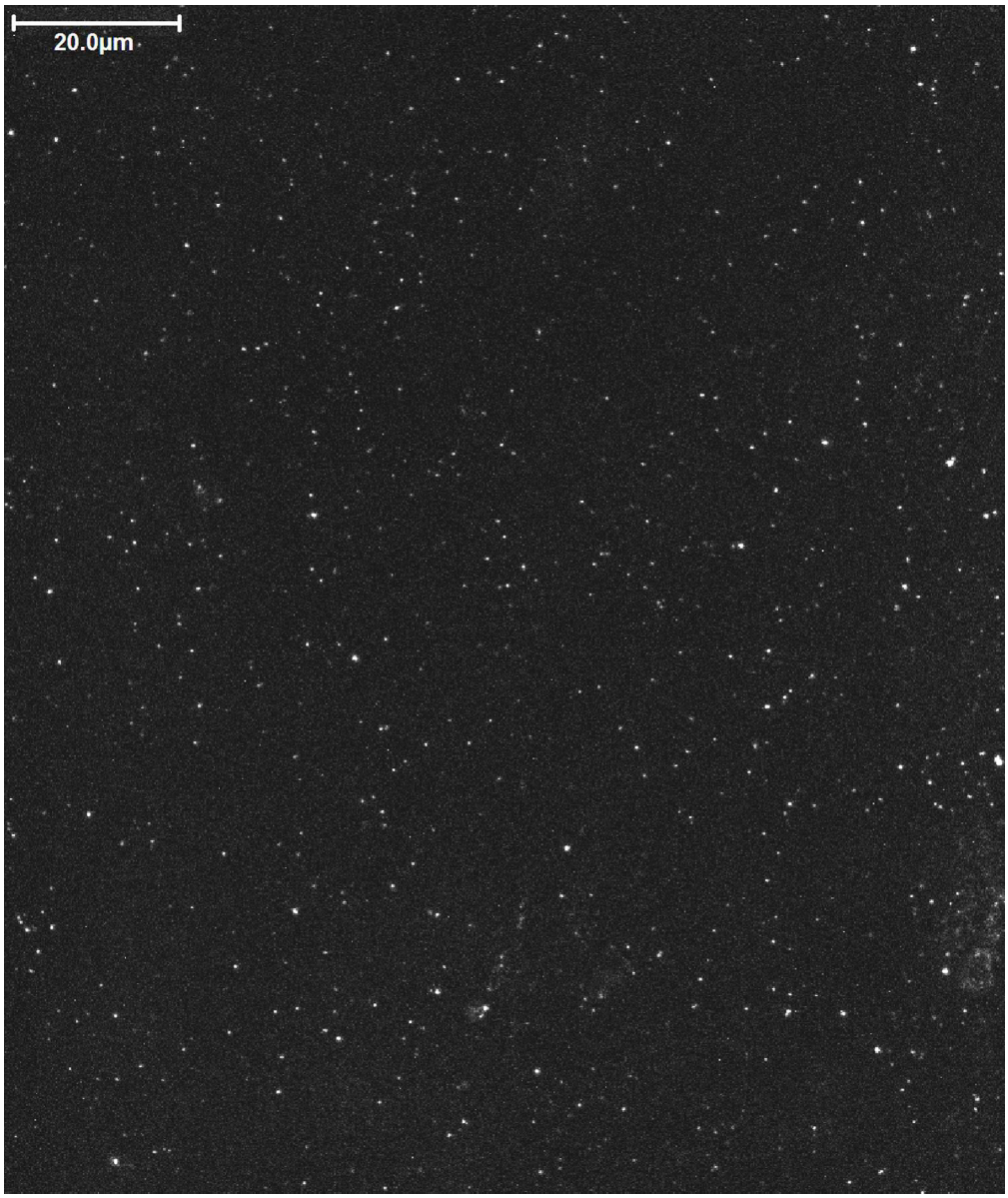


Figure 2. Fluorescence microscopy of VLPs in the bushmeat sample n° STE0014. All images were acquired with a Leica SP5 inverted confocal microscope with 4 lasers, a 100X objective and a numerical aperture of 1.4. The scale bar represents 20 µm.
313x371mm (300 x 300 DPI)

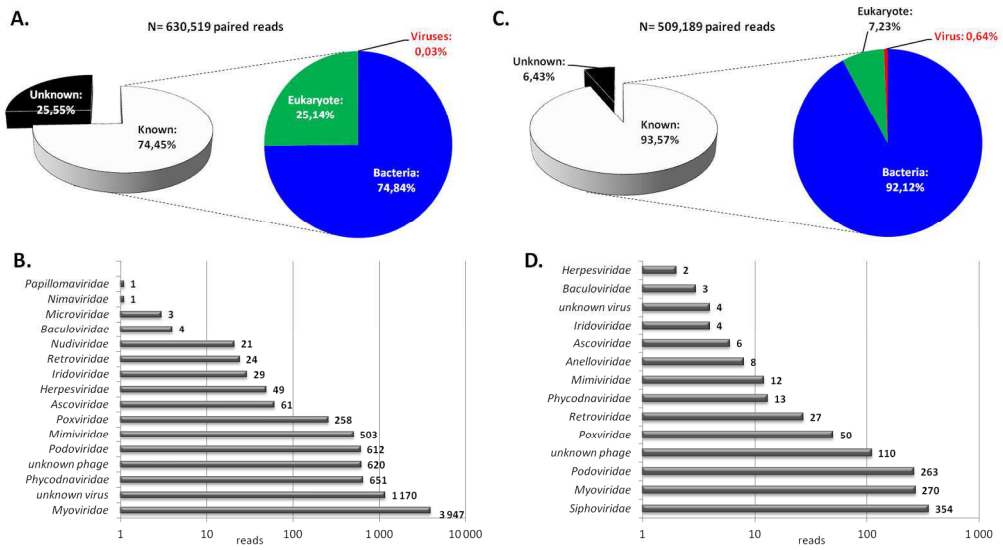
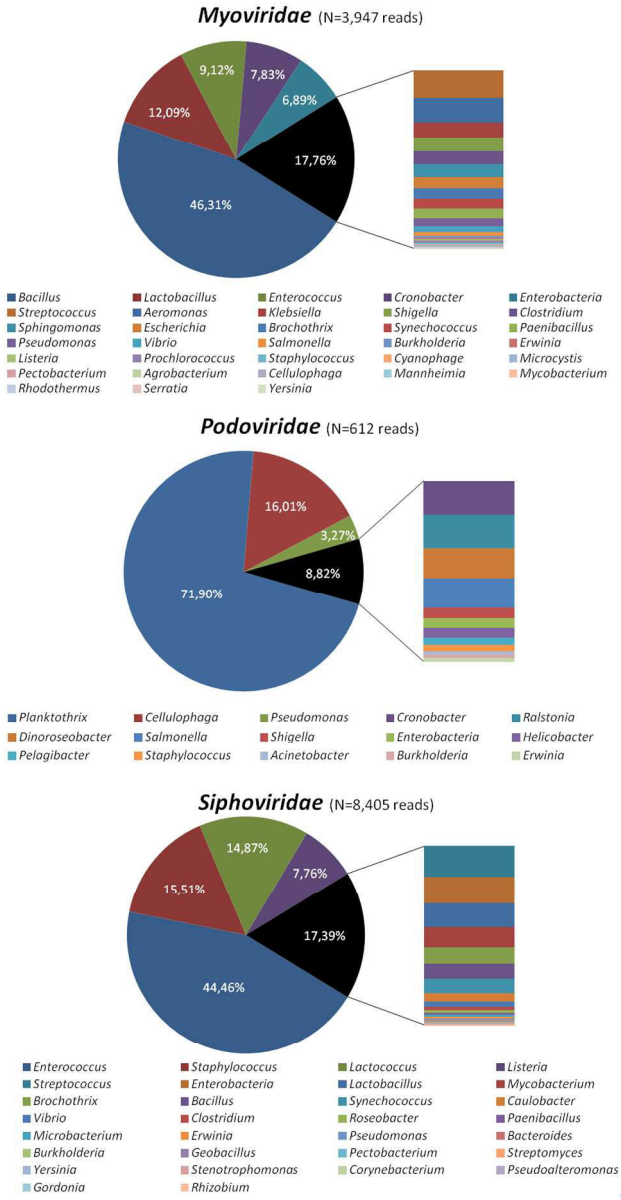
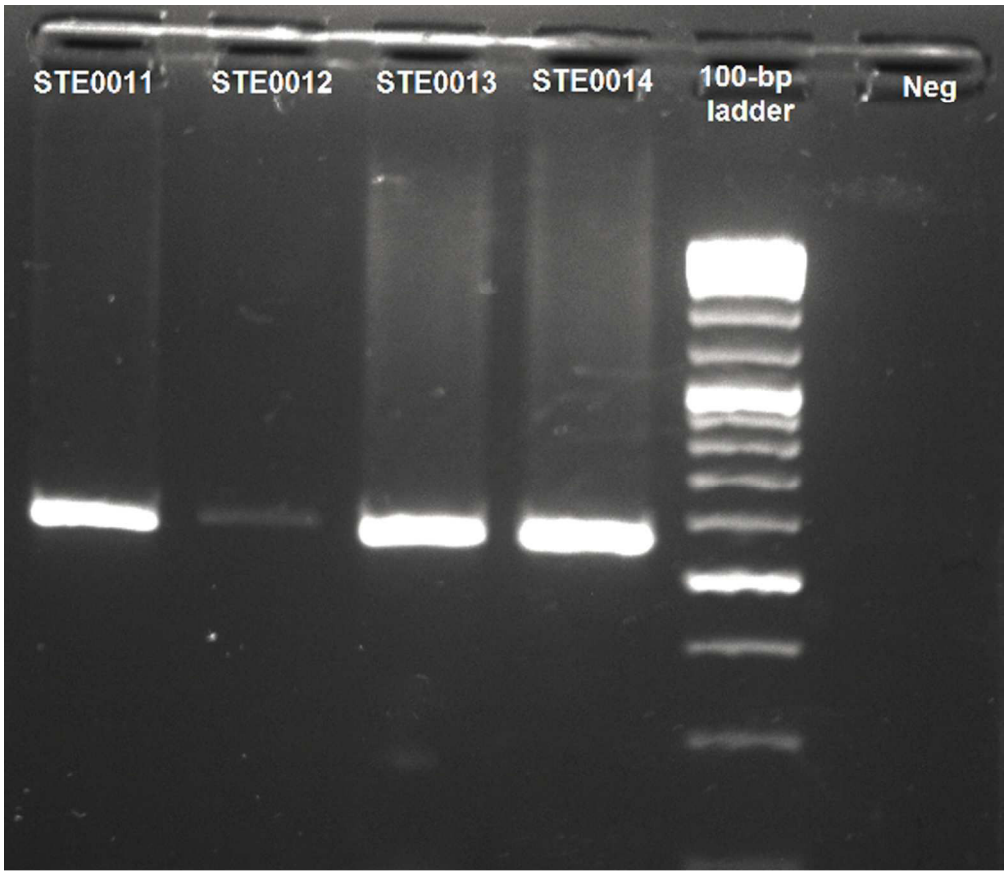
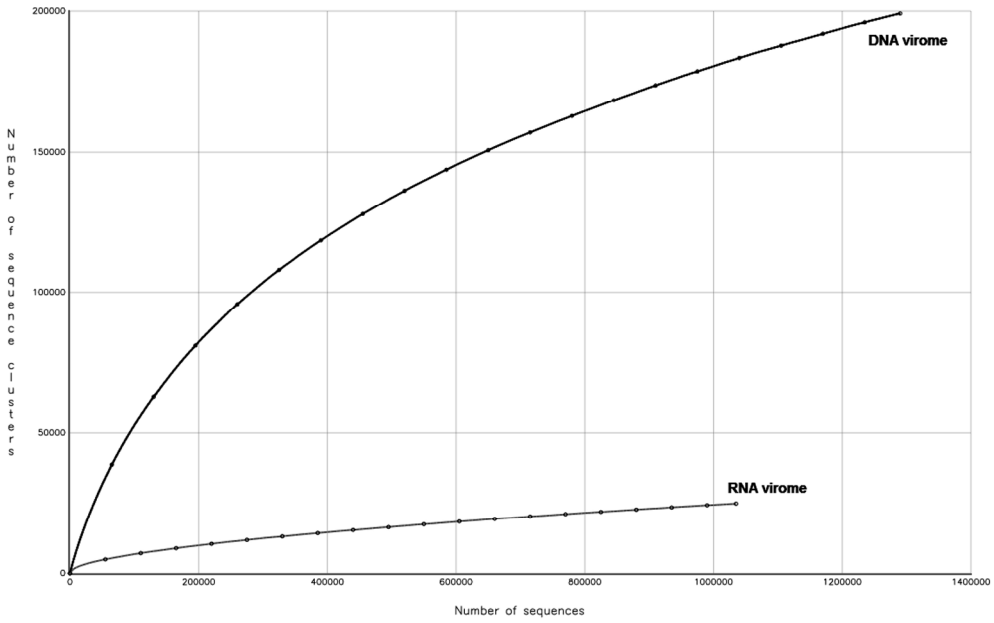


Figure 3. Taxonomic assignation of reads according to a BlastN search against the NCBI nucleotide database (A and C for the DNA and RNA viromes, respectively) and according to a BlastX search against the viral RefSeq NCBI database (B and D for the DNA and RNA viromes, respectively).


Myoviridae (N=3,947 reads)
9,12%
7,83%
6,89%
12,09%
17,76%
46,31%
Bacillus
Lactobacillus
Enterococcus
Cronobacter
Enterobacteria
Streptococcus
Aeromonas
Klebsiella
Shigella
Clostridium
Sphingomonas
Escherichia
Brochothrix
Synechococcus
Paenibacillus
Pseudomonas
Vibrio
Salmonella
Burkholderia
Erwinia
Listeria
Prochlorococcus
Staphylococcus
Cyanophage
Microcystis
Pectobacterium
Agrobacterium
Cellulophaga
Mannheimia
Mycobacterium
Rhodothermus
Serratia
Yersinia
Podoviridae (N=612 reads)
16,01%
3,27%
8,82%
71,90%
Planktothrix
Cellulophaga
Pseudomonas
Cronobacter
Ralstonia
Dinoroseobacter
Salmonella
Shigella
Enterobacteria
Helicobacter
Pelagibacter
Staphylococcus
Acinetobacter
Burkholderia
Erwinia
Siphoviridae (N=8,405 reads)
14,87%
7,76%
15,51%
17,39%
44,46%
Enterococcus
Staphylococcus
Lactococcus
Listeria
Streptococcus
Enterobacteria
Lactobacillus
Mycobacterium
Brochothrix
Bacillus
Synechococcus
Caulobacter
Vibrio
Clostridium
Roseobacter
Paenibacillus
Microbacterium
Erwinia
Pseudomonas
Bacteroides
Burkholderia
Geobacillus
Pectobacterium
Streptomyces
Yersinia
Stenotrophomonas
Corynebacterium
Pseudoalteromonas
Gordonia
Rhizobium



155x134mm (300 x 300 DPI)



305x190mm (300 x 300 DPI)

Pan-generic PCR	Primers	1st round PCR	2nd round PCR	Primers concentration (nM)	Annealing	Bibliographie
Flavivirus	PF1S PF2R-bis	SYBR Green	N/A	550	50°C	Moureau *et al.*, 2007
Paramyxovirus	PAR-F1 PAR-F2 PAR-R	Hot Star Taq	Hot Star Taq	400	48°C	Tong *et al.*, 2008
	RES-MOR-HEN-F1 RES-MOR-HEN-F2 RES-MOR-HEN-R	Hot Star Taq	Hot Star Taq	400	48°C	
Coronavirus	PC2S2-primerA PC2S2-primerB PC2As1-primerA PC2As1-primerB PC2As1-primerC PCS-primerA PCS-primerB PCNAs	Hot Star Taq	SYBR Green	F1 : 200, R1 : 900 F2 : 80, R2 : 400	I : 42°C II: 60°C	de Souza Luna *et al.*, 2007
Phlebovirus	NPHLEBO1+ NPHLEBO1- NPhlebo2+ NPhlebo2-	Hot Star Taq	Hot Star Taq	400	60°C	Sánchez-Seco *et al.*, 2003
Nairovirus	CCHF-F2 CCHF-R3 CCHF-F3 CCHF-R2	Hot Star Taq	SYBR Green	200	60°C	Rodriguez *et al.*, 1997
Poxvirus	PoxS361 PoxAS989 PoxS457 PoxAS927	Hot Star Taq	SYBR Green	800	55°C	Sánchez-Seco *et al.*, 2006
Orthobunyavirus	Cal/Bwa-F Cal/Bwa-R Bun-F Bun-R Weo-F Weo-R Oro-F Oro-R	Hot Star Taq	N/A	1000	55°C	Lambert *et al.*, 2009
Hantavirus	HAN-L-F1 HAN-L-R1 HAN-L-F2 HAN-L-R2	Hot Star Taq	Hot Star Taq	400	I : 48°C II : 50°C	Klempa *et al.*, 2006
Alphavirus	Alpha1+ Alpha1- Alpha2+ Alpha2-	Hot Star Taq	SYBR Green	400	52°C	Sánchez-Seco *et al.*, 2001
Arenavirus	ARE-1010C ARE-NW1696R ARE-OW1696R	Hot Star Taq	N/A	1000	45°C	Bowen *et al.*, 1997
Herpesvirus	Herpes-DFA Herpes-ILK	Hot Star Taq	SYBR Green	1000	46°C	VanDevanter *et al.*, 1996

	Herpes-KG1 Herpes-TGV Herpes-IYG					
Filovirus	Filo-A-F Filo-A-R Filo-B-F Filo-B-R Filo-C-F Filo-C-R Filo-D-F Filo-D-R	Hot Star Taq	N/A	500	58°C	Zhai *et al.*, 2007

Conclusion et perspectives

Dans le cadre de cette thèse, j'ai proposé et mis en place un projet de recherche afin d'étudier le phénomène d'émergence de zoonoses en caractérisant les communautés virales de différents acteurs : les arthropodes vecteurs et les faunes sauvage et domestique. Parmi les modèles choisis se retrouvaient les puces, les poux, les tiques, les punaises, les acariens et les culicoïdes en ce qui concerne les arthropodes vecteurs, et les primates (viande de brousse), les bovins et les rongeurs en ce qui concerne la faune sauvage et domestique. Ce manuscrit s'est focalisé uniquement sur les résultats des analyses des communautés virales de culicoïdes et de viande de singe, mais les analyses des autres modèles sont en cours. De plus, ce projet a permis d'initier de nombreuses collaborations scientifiques (internes et extérieures au laboratoire, notamment avec des vétérinaires) qui permettront, à terme, de poursuivre ce travail de surveillance des zoonoses virales émergentes via le suivi d'animaux et d'arthropodes « sentinelles ».

L'émergence est un phénomène complexe, influencé par de nombreux facteurs et paramètres biotiques (liés notamment à la pression de l'homme sur la faune) et abiotiques (influence de l'environnement) ou encore par des facteurs intrinsèques à l'agent infectieux (comme par exemple la capacité des virus à ARN à muter facilement). Depuis de nombreuses décénnies, néanmoins, c'est l'action de l'homme qui est le plus souvent à l'origine de l'émergence de nombreuses maladies infectieuses, la plupart étant par ailleurs d'origine zoonotique. Jusqu'à très récemment, les études sur les agents infectieux émergents ne se focalisaient que sur un type de médecine, humaine ou vétérinaire, sans tenir compte des interactions de l'homme avec son environnement. Depuis peu, un nouveau concept dit de « santé globale » est apparu. Ce concept, appelé "One world, one medicine, one health" ou plus simplement "One Health", tend à replacer l'homme dans son environnement en concevant des programmes à l'interface homme-animal-environnement et combinant différentes disciplines telles que la santé humaine, la santé animale, l'écologie, la climatologie, l'entomologie ou encore l'épidémiologie, la

sociologie, et l'économie afin de comprendre de façon globale, et à terme de pouvoir un jour prévenir, les futures émergences [104-108]. Bien que le concept One Health soit apparu dès les années 1855 avec les travaux du docteur Rudolf Virchow, puis repris en 1984 par Calvin Schwabe, vétérinaire, ce n'est que récemment (en 2004), suite à l'épidémie de SARS, que les instances internationales telles que l'OMS ont réellement pris conscience de la nécessité d'avoir des approches intégrées sur l'étude des maladies infectieuses émergentes lors d'une réunion définissant les « 12 principes de Manhattan » fondateurs du mouvement One Health [109]. Depuis, la majorité des études ciblant les zoonoses ont intégré des composantes « animales » en plus de celles ciblant les populations humaines [110]. Dans l'exemple récent du MERS-CoV, le réservoir animal du virus a été activement recherché dès l'apparition des premiers cas humains, sans certitude à ce jour du rôle de réservoir des chauves-souris et/ou des chameaux [111-112]. La lutte contre la grippe aviaire de type A/H5N1 passe désormais par une approche intégrée combinant par exemple surveillance des populations aviaires, arrêt de la mixité des espèces et vaccination [113-114]. Ces études *One Health* ont permis de démontrer la faisabilité et l'efficacité de la trans-disciplinarité dans la surveillance des émergences de maladies infectieuses. C'est dans ce contexte que j'ai tenté de mener à bien ce projet de thèse. Dès les captures d'arthropodes, la composante *One Health* a été intégrée en réalisant en parallèlle un échantillonnage des animaux environnants (bovins, rongeurs et animaux domestiques entre autres). Ces échantillonnages ont notamment permis la mise en évidence de la présence de *Faustovirus* dans la faune sauvage et domestique associée aux culicoïdes (article n°4). La présence de *Faustovirus* a également pu être démontrée dans l'eau de la rivière environnant le village de Dielmo ainsi que dans les deux puits alimentant en eau les villages de Dielmo et Ndiop. Des séquences virales ont également été retrouvées dans des prélèvements sanguins de villageois. Bien que la source de contamination par *Faustovirus* des populations humaines ne soit à ce jour pas encore connue et reste à confirmer et à investiguer, cette étude souligne et renforce l'idée que ce type d'approche globale est nécessaire à l'investigation épidémiologique d'un virus en population humaine.

Dans des études de surveillance et de criblage à grande échelle de pathogènes viraux connus ou émergents, il existe de nombreuses techniques cellulaires, moléculaires, microscopiques, ou encore immunologiques permettant d'identifier des virus plus ou moins connus. Cependant, la métagénomique est très vite apparue comme l'outil de choix pour de telles investigations, du fait d'une puissance et d'une profondeur de séquençage sans précédent, d'un coût qui tend à diminuer et de sa capacité à identifier des virus complètement nouveaux sans aucune connaissance préalable. Ainsi, le séquençage des communautés virales des culicoïdes a permis de mettre en évidence un certain nombre de virus nouveaux pouvant potentiellement présenter un risque pour la santé humaine (article n°3). La détermination de leur potentiel infectieux pour l'homme et leur éventuelle pathogénicité pour ce dernier reste cependant difficile à déterminer. La simple identification par NGS de séquences virales constitue la première étape mais n'est pas suffisant. Ainsi, de plus en plus d'études de métagénomique conduites afin d'identifier l'agent causal d'une maladie infectieuse ne se contentent plus de la simple identification d'un virus dans un prélèvement clinique mais nécessitent désormais la preuve de la causalité des symptômes par le virus [53,115]. Répondant au nom de « postulat de Koch » [116], un virus est considéré comme responsable d'une maladie s'il répond aux critères suivants : (1) il doit être retrouvé dans tous les patients malades mais pas chez des personnes saines (2) il doit pouvoir être isolé de l'organisme malade et cultivé *in vitro* (3) il doit pouvoir recréer des symptômes identiques chez un organisme sain (par exemple un modèle animal) inoculé avec le virus cultivé *in vitro* (4) il doit pouvoir être ré-isolé de l'organisme sain préalablement inoculé et être identique au virus originel. Cependant, les virus identifiés par métagénomique ont souvent du mal à répondre à ces critères, d'où la nécessité de tenir compte de nouveaux critères, mieux adaptés à ce type d'études [53-54,115,117]. Ainsi Lipkin proposa en 2013 trois niveaux dans la preuve de causalité [118] : le niveau 1 correspondrait à une relation causale possible entre un agent infectieux et une pathologie dans laquelle le virus suspecté serait retrouvé chez un certain nombre de malades (par des techniques de culture, de microscopie électronique, ou encore de détection d'acides nucléiques viraux ou de protéines virales). Le niveau 2 correspondrait à une relation causale probable entre l'agent infectieux et la pathologie dans laquelle le virus suspecté serait

retrouvé chez un grand nombre de malades souffrant de la même pathologie et chez qui le virus (ou un de ces composants génomique ou protéique) serait retrouvé en grande quantité ou chez qui on retrouverait des preuves immunitaires d'infection (ex : séroconversion). Enfin le niveau 3 correspondrait à une relation causale certaine entre un agent infectieux et une pathologie dans laquelle le virus suspecté répondrait parfaitement aux critères du postulat de Koch originel ou pour qui l'utilisation de traitements préventifs ou curatifs ciblant ce virus tendrait à atténuer ou à éliminer les symptômes de la maladie. Quelle que soit la définition choisie (postulat de Koch originel ou récent), dans la majorité des cas, il est indispensable de parvenir à isoler le virus suspecté afin, par la suite, de réaliser des études complémentaires en population humaine (épidémiologie par sérologie par exemple) ou animale ou encore afin de caractériser le virus (séquençage complet de son génome, détermination de son protéome, etc.). Il est aisé pour les études de métagénomique virale d'atteindre le niveau 1 des critères de Lipkin, mais cela ne démontre en rien la preuve de causalité entre la présence d'un virus et une pathologie. Par ailleurs, dans la plupart des cas ainsi que dans l'étude de métagénomique conduite sur les culicoïdes (article n°3), les séquences virales identifiées par séquençage ne sont pas systématiquement reliées à l'isolement effectif des particules virales correspondantes. Dans le cadre de cette étude, la mise au point d'un protocole de purification des communautés virales qui maintenait l'intégrité et l'infectivité des virions (article n°2) a été motivée par la volonté de pouvoir *in fine* isoler les virus identifiés lors du séquençage. Malheureusement ces isolements ont échoué. Néanmoins d'autres essais sur un plus grand nombre de lignées cellulaires et en utilisant l'inoculation intracérébrale de souriceaux nouveaux-nés permettraient de (1) confirmer la présence de particules virales infectieuses dans les échantillons d'arthropodes et de lever le doute, par exemple, sur la présence d'éléments viraux endogènes dans les génomes de culicoïdes (article n°3) ; et (2) pouvoir conduire des enquêtes épidémiologiques en population humaine afin de déterminer si les virus identifiés dans les arthropodes sont capables d'infecter l'homme et s'ils causent ou non une pathologie chez ces derniers en criblant par exemple la collection de sérums que possède le laboratoire Point-of-Care implanté dans les villages de Dielmo et Ndiop [66].

Il existe de nombreuses zoonoses causées par un arbovirus. Les arbovirus sont des virus transmis activement par la piqûre d'un arthropode hématophage, c'est-à-dire que dans l'arthropode, une fois le repas de sang pris sur un hôte infecté (le plus souvent d'origine animale), le virus est capable de se multiplier et de gagner les glandes salivaires de l'arthropode. Ainsi, la prochaine piqûre sur un hôte sain résultera en la transmission du virus à ce dernier. Il existe cependant un mode de transmission dit « mécanique », où le virus est transmis à un nouvel hôte par simple souillure des pièces buccales de l'arthropode [47]. La simple détection d'un virus dans un arthropode hématophage ne signifie donc pas forcément que celui-ci est un arbovirus et qu'il puisse être transmis de façon vectorielle. On appelle compétence vectorielle l'aptitude d'un arthropode à ingérer, répliquer et transmettre un agent infectieux, à la différence de la capacité vectorielle qui correspond à l'aptitude de l'arthropode à transmettre un agent infectieux en fonction de sa bio-écologie et des conditions environnementales. La capacité vectorielle est donc la résultante de la compétence vectorielle et de la bio-écologie du vecteur (abondance, longévité, préférences trophiques, etc.). Afin de démontrer le rôle vectoriel d'un arthropode hématophage, l'isolement du virus dans les glandes salivaires de ce dernier serait un premier pas. L'étude des viromes de culicoïdes ayant permis la mise en évidence de deux nouveaux virus potentiellement intéressants quant à leur possible effet sur la santé humaine (le nouveau thogotovirus et le nouveau rhabdovirus, cf article n°3), il sera intéressant de poursuivre la caractérisation de ces virus en réalisant des expériences de compétence vectorielle en laboratoire, ce qui indiquerait une possibilité de ces arthropodes à vectoriser ces nouveaux virus. Par ailleurs, afin de vérifier si, dans la nature, les culicoïdes seraient effectivement capables de transmettre de façon active le thogotovirus et le rhabdovirus à des hôtes vertébrés, des campagnes de suivi des populations de « vecteurs » seront mises en place sur une année afin de tenter de corréler l'abondance des arthropodes avec celle des virus. En effet la majorité des zoonoses transmises par des arthropodes hématophages sont soumises à des phénomènes de saisonnalité, liés à la distribution temporelle des populations de vecteurs [119]. Par ailleurs, il a été démontré que la préférence trophique, c'est-à-dire la préférence de repas de sang d'un arthropode hématophage, pouvait varier en fonction de la disponibilité d'un hôte vertébré [48] et que ce changement, ou « host

switching », était souvent à l'origine de l'émergence d'une zoonose en population humaine. Dans l'étude portant sur les communautés virales des culicoïdes, il a été démontré que les culicoïdes se gorgeaient à la fois sur des bovins, des rongeurs et des hommes (article n°4), ce qui laisse supposer une possibilité d'émergence d'un virus zoonotique en population humaine. Il serait donc intéressant d'étudier la capacité vectorielle de ces populations de vecteurs et notamment de voir si le pic d'abondance des vecteurs est lié à un pic d'abondance des virus dans ces derniers et une recrudescence de pathologies humaines et/ou animales, ou encore d'étudier les préférences trophiques de ces arthropodes en fonction des saisons.

La capacité vectorielle d'un arthropode à transmettre un agent infectieux dépend de nombreux facteurs, et notamment des interactions existant entre les communautés microbiennes présentes au sein de l'arthropode. Il a ainsi pu être démontré que la présence concomitante d'une infection par la bactérie Wolbachia modifierait la capacité du moustique à transmettre certains arbovirus [120,121]. A l'inverse, l'infection par le virus *Chikungunya* semble moduler la composition des communautés bactériennes des moustiques [122]. Dans l'analyse des viromes de culicoïdes, nous avons pu mettre en évidence la présence de nombreuses séquences reliées à des bactéries et à des parasites (article n°3). Outre le fait que les culicoïdes puissent également transmettre activement des bactéries et des parasites, comme cela a déjà été démontré pour les parasites du genre *Mansonella* [123,124], ces micro-organismes peuvent aussi influencer la capacité vectorielle en interférant de façon positive ou négative avec les différentes communautés virales présentes. A ce jour, aucune étude de criblage à grande échelle ciblant les bactéries et les parasites des culicoïdes n'a été conduite. Le protocole mis en place dans cette thèse ne permettait pas de caractériser les communautés bactériennes et parasitaires présentes, mais cette étude paraît indispensable pour essayer de caractériser les interactions possibles entre les différentes communautés microbiennes. Par ailleurs, il a été suggéré que la présence de virus dits « insectes-spécifiques » pouvait également interférer, de façon positive ou négative, avec la capacité de transmission d'un arbovirus [125]. Dans les viromes de culicoïdes ont été retrouvés des séquences plus ou moins apparentées à des virus insecte-spécifique tels que les virus *Loreto*, *Negev*, *Piura* et *Nora* préalablement

détectés chez des moustiques, des phlébotomes ou des drosophiles (article n°3). L'implication de ces virus dans la capacité vectorielle des culicoïdes serait également à évaluer afin de proposer des moyens de contrôle des populations de vecteurs.

Dans le contexte de l'émergence des zoonoses, la plupart des études tentent d'identifier les futurs pathogènes responsables de maladies. Mais outre le fait qu'une pathologie puisse être directement liée à la présence d'un agent infectieux au sein d'un organisme, ce qui est démontré dans les postulats de Koch originel ou moderne, un déséquilibre des interactions entre communautés virales, bactériennes, fongiques ou parasitaires présentes dans l'organisme pourrait également conduire à une pathologie. Ainsi, si l'on considère l'organisme humain comme un écosystème à part entière, l'infection par l'homme d'un virus extérieur (d'origine animale ou arthropode par exemple, et a priori non pathogène pour ce dernier) pourrait déséquilibrer la symbiose existant entre les virus commensaux de l'homme et l'homme (article n°6). Développé par Mirko Grmek à la fin des années 1960 et adapté récemment aux communautés microbiennes par Jean-Paul Gonzalez [126], le concept de pathocénose désigne l'état d'équilibre des communautés microbiennes à un moment donné et dans un écosystème donné. Ainsi, la perturbation de cet équilibre pathocénotique pourrait conduire à une pathologie. Dans l'étude de métagénomique virale de viande de brousse (article n°5) ainsi que dans l'étude des communautés virales des culicoïdes, nous avons détecté de nombreux bactériophages qui, bien qu'étant non pathogènes pour l'homme, pourraient poser un problème de déséquilibre des communautés bactériennes présentes dans le corps humain et potentiellement résulter en une pathologie. Il a par exemple été démontré que des interactions mutualistes entre bactériophages et hôte humain interviennent dans le système digestif de ce dernier en contrôlant les populations bactériennes présentes. Il existe *a contrario* des interactions négatives pour l'homme, comme dans le cas de transfert de gènes de résistance aux antibiotiques de bactériophages vers les populations bactériennes [127]. Il est donc nécessaire de concevoir l'étude de l'émergence de zoonoses, non pas seulement comme la détection d'agents infectieux pathogènes pour l'homme, mais de façon plus globale à l'échelle du microbiome tout entier. Ce type d'études ne pourra être

envisagé que par l'utilisation des outils de métagénomiques, les plus exhaustifs existant à ce jour.

Peter Daszak, responsable de l'Alliance EcoHealth, disait « le Saint Graal pour la recherche sur les maladies infectieuses émergentes est le développement d'une stratégie valide de prédiction des futures zoonoses émergentes » [128]. Murphy commentait dès 1998 la difficulté de telles prédictions : « en général, il n'y a aucun moyen de prédire quand et où émergera le prochain agent zoonotique et quelle sera son importance au regard de la santé humaine » [129]. Cependant, depuis la prise de conscience du risque croissant d'infections zoonotiques capables d'impacter l'homme, il semblerait que la communauté scientifique cherche à découvrir les nouveaux pathogènes responsables de pandémies telles que le VIH ou le SARS, même si cela constitue un challenge très important, voire insurmontable. Le premier verrou est la faible connaissance de la diversité microbienne, et en particulier virale, que nous possédons. Daszak [128] a en effet estimé que si l'on considère les 50 000 espèces de vertébrés connues (qui représentent le plus grand réservoir de virus zoonotiques) et que si chacune d'entre elles porte une vingtaine de virus inconnus et endémiques (les chauves-souris en comptent plus de 20 000), on atteindrait alors une diversité virale globale de l'ordre du million de virus pouvant potentiellement infecter l'homme, parmi lesquels seuls environ 2000 sont connus, soit 0.2% ! L'avènement des techniques de métagénomique tend à minimiser cette méconnaissance, mais l'utilisation de tels outils pose de nouveaux problèmes, liés notamment aux analyses bio-informatiques de telles données, qui restent à ce jour, outre leur coût, le principal frein à l'utilisation massive de la métagénomique [65]. Par ailleurs et comme précisé plus haut, la détection de séquences virales chez l'animal par métagénomique ne prédit en rien la possibilité de ces virus à franchir la barrière d'espèces séparant l'animal de l'homme. Néanmoins il apparaitrait, si l'on observe les virus ayant émergé depuis les années 1980, que les nouveaux virus émergents auraient les caractéristiques suivantes [7] :

- ce serait des virus à génome ARN (du fait de leur plus grande plasticité du génome que les virus à ADN)
- ils proviendraient d'un réservoir animal

- ils seraient capables d'infecter un grand nombre d'hôtes différents (phénomène de « host switching » qui confère au virus une plus grande capacité d'adaptation à de nouvelles espèces, y compris à l'homme)
- ils auraient une capacité plus ou moins grande à se transmettre entre hommes. En effet, bien que cette capacité soit limitée dans les exemples récents du MERS-CoV ou du SARS, la pandémie de VIH et l'épidémie en cours d'Ebola sont de bons exemples d'une adaptation à l'espèce humaine réussie pour un virus.

Malgré ces connaissances, la route est encore longue avant que l'on arrive à prévoir quelle sera la nouvelle épidémie ou pandémie virale du prochain siècle. Depuis quelques années, les médias grands publics se sont emparés de ce type d'études afin d'alerter, souvent à des fins sensationnelles, les populations sur les potentiels risques zoonotiques, sans forcément relativiser ces découvertes, ce qui a eu pour conséquence de souvent dramatiser des risques qui n'existent pas forcément. L'exemple le plus marquant serait sans doute celui de la pandémie de grippe A/H1N1 (dite grippe porcine) en 2009 qui a été déclarée comme étant la première pandémie du XXI^{ème} siècle par l'OMS et qui, en France, et bien que relativement bénigne, a eu pour conséquences d'une part l'achat massif de doses inutilisées de vaccins au nom du principe de précaution, mais aussi des conséquences économiques importantes, notamment dans les élevages porcins et aviaires, dues à la peur des consommateurs, souvent entretenue par les médias.

La surveillance de l'émergence de zoonoses est donc bien un enjeu majeur de santé publique, mais elle est à tempérer au regard de tous les biais liés à l'utilisation de techniques, jusqu'alors imparfaites, de découvertes d'agents infectieux.

Bibliographie

1. Morse SS. Factors in the emergence of infectious diseases. *Emerg. Infet. Dis*. 1995; 1:7-15.
2. Morse SS. Emerging infections: condemned to repeat? Chapter 5: Infectious disease emergence: Past, present and Future. In: Microbial Evolution and Co-Adaptation, A Tribute to the Life and Scientific Legacies of Joshua Lederberg: Workshop Summary. *Washington (DC): National Academies Press (US)*; 2009.
3. Chua KB, Gubler DJ. Perspectives of public health laboratories in emerging infectious diseases. *Emerg. Microbes Infect*. 2013; doi: 10.1038/emi.2013.34.
4. Morens DM, Folkers GK, Fauci AS. Emerging infections: a perpetual challenge. *Lancet Infect. Dis*. 2008; 8:710-719.
5. Tompkins DM, Carver S, Jones ME, Krkošek M, Skerratt LF. Emerging infectious diseases of wildlife: a critical perspective. *Trends Parasitol*. 2015; 31: 149-159.
6. Woolhouse M, Scott F, Hudson Z, Howey R, Chase-Topping M. Human viruses: discovery and emergence. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 2012; 367: 2864-2871.
7. Woolhouse M, Gaunt E. Ecological origins of novel human pathogens. Chapter 5: Infectious disease emergence: Past, present and Future. In: Microbial Evolution and Co-Adaptation, A Tribute to the Life and Scientific Legacies of Joshua Lederberg: Workshop Summary. *Washington (DC): National Academies Press (US)*; 2009.
8. Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. *Nature*. 2004; 430: 242-249.
9. Lankau EW, Cohen NJ, Jentes ES, Adams LE, Bell TR, Blanton JD, Buttke D, Galland GG, Maxted AM, Tack DM, Waterman SH, Rupprecht CE, Marano N. Prevention and control of rabies in an age of global travel: a review of travel- and trade-associated rabies events—United States, 1986-2012. *Zoonoses Public Health*. 2014; 61: 305-316.

10. Schaffner F, Medlock JM, Van Bortel W. Public health significance of invasive mosquitoes in Europe. *Clin. Microbiol. Infet.* 2013; 19: 685-692.
11. Smith KM, Anthony SJ, Switzer WM, Epstein JH, Seimon T, Jia H, Sanchez MD, Huynh TT, Galland GG, Shapiro SE, Sleeman JM, McAloose D, Stuchin M, Amato G, Kolokotronis SO, Lipkin WI, Karesh WB, Daszak P, Marano N. Zoonotic viruses associated with illegally imported wildlife products. *PLoS One*. 2012; doi: 10.1371/journal.pone.0029505.
12. Liu Q, Cao L, Zhu XQ. Major emerging and re-emerging zoonoses in China: a matter of global health and socioeconomic development of 1.3 billion. *Int. J. Infect. Dis*. 2014; 25: 65-72.
13. McFarlane RA, Sleigh AC, McMichael AJ; Land-use change and emerging infectious disease on an island continent. *Int. J. Environ. Res. Public Health*. 2013; 10: 2699-2719.
14. Christiansen-Jucht C, Erguler K, Shek CY, Basáñez MG, Parham PE. Modelling *Anopheles gambiae* s.s. population dynamics with temperature- and age-dependent survival. *Int. J. Environ. Res. Public Health*. 2015; 12: 5975-6005.
15. Murray KA, Daszak P. Human ecology in pathogenic landscapes: two hypotheses on how land use change drives viral emergence. *Curr. Opin. Virol*. 2013; 3: 79-83.
16. Fonseca JD, Knight GM, McHugh TD. The complex evolution of antibiotic resistance in *Mycobacterium tuberculosis*. *Int. J. Infect. Dis*. 2015; 32: 94-100.
17. Sallie R. Replicative homeostasis II: influence of polymerase fidelity on RNA virus quasispecies biology: implications for immune recognition, viral autoimmunity and other "virus receptor" diseases. *Virol. J*. 2005; 2: 70.
18. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev*. 2012; 76: 159-216.
19. Lauring AS, Andino R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog*. 2010; doi: 10.1371/journal.ppat.1001005.
20. Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team, Dawood FS, Jain S, Finelli L, Shaw MW, Lindstrom S, Garten RJ, Gubareva LV, Xu X, Bridges CB, Uveki TM. Emergence of a novel swine-origin *Influenza A* (H1N1) virus in humans. *N. Engl. J. Med*. 2009; 360: 2605-2615.

21. Woolhouse ME, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. *Emerg. Infect. Dis*. 2005; 11: 1842-1847.
22. Bianchine PJ, Russo TA. The role of epidemic infectious diseases in the discovery of America. *Allergy Proc*. 1992; 13: 225-232.
23. Patterson KB, Runge T. Smallpox and the Native American. *Am. J. Med. Sci*. 2002; 323: 216-222.
24. Guerra F. The European-American exchange. *Hist. Philos. Life Sci*. 1993; 15: 313-327.
25. McMichael AJ. Environmental and social influences on emerging infectious diseases: past, present and future. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 2004; 359: 1049-1058.
26. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*. 1999; 397: 436-441.
27. Rewar S, Mirdha D. Transmission of ebola virus disease: an overview. *Ann. Glob. Health*. 2014; 80: 444-451.
28. Weyer J, Grobbelaar A, Blumberg L. Ebola virus disease: history, epidemiology and outbreaks. *Curr. Infect. Dis. Rep*. 2015; 17: 480.
29. Marí Saéz A, Weiss S, Nowak K, Lapeyre V, Zimmermann F, Düx A, Kühl HS, Kaba M, Regnaut S, Merkel K, Sachse A, Thiesen U, Villányi L, Boesch C, Dabrowski PW, Radonić A, Nitsche A, Leendertz SA, Petterson S, Becker S, Krähling V, Couacy-Hymann E, Akoua-Koffi C, Weber N, Schaade L, Fahr J, Borchert M, Gogarten JF, Calvignac-Spencer S, Leendertz FH. Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol. Med*. 2014; 7: 17-23.
30. Carroll MW, Matthews DA, Hiscox JA, Elmore MJ, Pollakis G, Rambaut A, Hewson R, García-Dorival I, Bore JA, Koundouno R, Abdellati S, Afrough B, Aiyepada J, Akhilomen P, Asogun D, Atkinson B, Badusche M, Bah A, Bate S, Baumann J, Becker D, Becker-Ziaja B, Bocquin A, Borremans B, Bosworth A, Boettcher JP, Cannas A, Carletti F, Castilletti C, Clark S, Colavita F, Diederich S, Donatus A, Duraffour S, Ehichioya D, Ellerbrok H, Fernandez-Garcia MD, Fizet A, Fleischmann E, Gryseels S, Hermelink A, Hinzmann J, Hopf-Guevara U, Ighodalo Y, Jameson L, Kelterbaum A, Kis Z, Kloth S, Kohl C, Korva M, Kraus A, Kuisma E, Kurth A, Liedigk B, Logue CH, Lüdtke A,

Maes P, McCowen J, Mély S, Mertens M, Meschi S, Meyer B, Michel J, Molkenthin P, Muñoz-Fontela C, Muth D, Newman EN, Ngabo D, Oestereich L, Okosun J, Olokor T, Omiunu R, Omomoh E, Pallasch E, Pályi B, Portmann J, Pottage T, Pratt C, Priesnitz S, Quartu S, Rappe J, Repits J, Richter M, Rudolf M, Sachse A, Schmidt KM, Schudt G, Strecker T, Thom R, Thomas S, Tobin E, Tolley H, Trautner J, Vermoesen T, Vitoriano I, Wagner M, Wolff S, Yue C, Capobianchi MR, Kretschmer B, Hall Y, Kenny JG, Rickett NY, Dudas G, Coltart CE, Kerber R, Steer D, Wright C, Senyah F, Keita S, Drury P, Diallo B, de Clerck H, Van Herp M, Sprecher A, Traore A, Diakite M, Konde MK, Koivogui L, Magassouba N, Avšič-Županc T, Nitsche A, Strasser M, Ippolito G, Becker S, Stoecker K, Gabriel M, Raoul H, Di Caro A, Wölfel R, Formenty P, Günther S. Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. *Nature*. 2015; 524: 97-101.

31. Hubálek Z. Emerging human infectious diseases: anthroponoses, zoonoses, and sapronoses. *Emerg. Infect. Dis*. 2003; 9: 403-404.
32. WHO update report. Available at http://www.who.int/emergencies/mers-cov/en/ on 11.09.2015.
33. Wang Y, Liu D, Shi W, Lu R, wang W, Zhao Y, Deng Y, Zhou W, Ren H, Wu J, Wang Y, Wu G, Gao GF, Tan W. Origin and possible genetic recombination of the Middle East Respiratory Syndrome Coronavirus from the first imported case in China: phylogenetics and coalescence analysis. *MBio*. 2015; doi: 10.1128/mBio.01280-15.
34. Choi JY. An outbreak of Middle East Respiratory Syndrome Coronavirus infection in South Korea, 2015. *Yonsei Med. J*. 2015; 56: 1174-1176.
35. Cowling BJ, Park M, Fang VJ, Wu P, Leung GM, Wu JT. Preliminary epidemiological assessment of MERS-CoV outbreak in South Korea, May to June 2015. *Euro Surveill*. 20: 7-13.
36. WHO update report. Available at http://apps.who.int/ebola/current-situation/ebola-situation-report-30-september-2015 on 30.09.2015.
37. Kosoy OI, Lambert AJ, Hawkinson, DJ, Pastula DM, Goldsmith CS, Hunt DC, Staples JE. Movel thogotovirus associated with febrile illness and death, United States, 2014. *Emerg. Infect. Dis*. 2015; 21: 760-764.
38. Mouinga-Ondémé A, Kazanji M. Simian foamy virus in non-human primates and cross-species transmission to humans in Gabon: an emerging zoonotic disease in central Africa? *Viruses*. 2013; 5: 1536-1552.

39. Pfau CJ. Arenaviruses (Chapter 57). In Baron S, editor. Medical Microbiology, 4th edition. *Galveston (TX): University of Texas Medical Branch at Galveston*; 1996.
40. Hardestam J, Karlsson M, Falk KI, Olsson G, Klingström J, Lundkvist A. Puumala hantavirus excretion kinetics in bank voles (*Myodes glareolus*). *Emerg. Infect. Dis*. 2008; 14: 1209-1215.
41. Taylor LH, Latham SM, Woolhouse ME. Risk factors for human disease emergence. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 2001; 356: 983-989.
42. Geering WA, Davies FG, Martin V. Manuel FAO de santé animale : préparation des plans d'intervention contre la fièvre de la vallée du rift. Chapitre 2 : nature de la maladie. 2003.
43. Mackenzie JS, Jeggo M. Reservoirs and vectors of emerging viruses. *Curr. Opin. Virol.* 2013; 3: 170-179.
44. Mandl JN, Ahmed R, Barreiro LB, Daszak P, Epstein JH, Virgin HW, Feinberg MB. Reservoir host immune responses to emerging zoonotic viruses. *Cell*. 2015; 160: 20-35.
45. Plowright RK, Eby P, Hudson PJ, Smith IL, Westcott D, Bryden WL, Middleton D, Reid PA, McFarlane RA, Martin G, Tabor GM, Skerratt LF, Anderson DL, Crameri G, Quammen D, Jordan D, Freeman P, Wang LF, Epstein JH, Marsh GA, Kung NY, McCallum H. Ecological dynamics of emerging bat virus spillover. *Proc. Biol. Sci*. 2015; 282: doi: 10.1098/rspb.2014.2124.
46. Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, Wong BH, Wong SS, Leung SY, Chan KH, Yuen KY. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. USA*. 2005; 102: 14040-14045.
47. Drucker M, Then C. Transmission activation in non-circulative virus transmission: a general concept? *Curr. Opin. Virol*. 2015; 15: 63-68.
48. Liang G, Gao X, Gould EA. Factors responsible for the emergence of arboviruses; strategies, challenges and limitations for their control. *Emerg. Microbes Infect*. 2015; doi: 10.1038/emi.2015.18.
49. Devaux CA. Emerging and re-emerging viruses: a global challenge illustrated by chikungunya virus outbreaks. *World J. Virol.* 2012; 1: 11-22.

50. Cleaveland S, Laurenson MK, Taylor LH. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 2001; 356: 991-999.
51. Woolhouse ME, Howey R, Gaunt E, Reilly L, Chase-Topping M, Savill N. Temporal trends in the discovery of human viruses. *Proc. Biol. Sci*. 2008; 275: 2111-2115.
52. Pozzetto B. Méthodes de diagnostic en virologie. In : Virologie médicale. *Presses Universitaires de Lyon, A. Mammette* ; 2002.
53. Sridhar S, To KK, Chan JF, Lau SK, Woo PC, Yuen KY. A systematic approach to novel virus discovery in emerging infectious disease outbreaks. *J. Mol. Diagn*. 2015; 17: 230-241.
54. Mokili JL, Rohwer F, Dutilh BE. Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol*. 2012; 2: 63-77.
55. Chiu CY. Viral pathogen discovery. *Curr. Opin. Microbiol*. 2013; 16: 468-478.
56. Barzon L, Lavezzo E, Militello V, Toppo S, Palù G. Applications of next-generation sequencing technologies to diagnostic virology. *Int. J. Mol. Sci*. 2011; 12: 7861-7884.
57. Belák S, Karlsson OE, Blomström AL, Berg M, Granberg F. New viruses in veterinary medicine, detected by metagenomic approaches. *Vet Microbiol*. 2013; 165: 95-101.
58. Culley AI, Lang AS, Suttle CA. High diversity of unknown picorna-like viruses in the sea. *Nature*. 2003; 424: 1054-1057.
59. Lefterova MI, Suarez CJ, Banaei N, Pinsky BA. Next-generation sequencing for infectious disease diagnosis and management: a report of the Association for Molecular Pathology. *J. Mol. Diagn.* 2015; doi: 10.1016/j.jmoldx.2015.07.004.
60. Bexfield N, Kellam P. Metagenomics and the molecular identification of novel viruses. *Vet. J.* 2011; 190: 191-198.
61. Kim KH, Bae JW. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Appl. Environ. Microbiol*. 2011; 77: 7663-7668.
62. Marine R, McCarren C, Vorrasane V, Nasko D, Crowgey E, Polson SW, Wommack KE. Caught in the middle with multiple displacement amplification: the myth of pooling for avoiding multiple displacement

amplification bias in a metagenome. *Microbiome*. 2014; doi: 10.1186/2049-2618-2-3.

63. Karlsson OE, Belák S, Granberg F. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. *Biosecur. Bioterror*. 2013; 1: 227-234.
64. Rosseel T, Van Borm S, Vandenbussche F, Hoffmann B, van den Berg T, Beer M, Höper D. The origin of biased sequence depth in sequence-independent nucleic acid amplification and optimization for efficient massive parallel sequencing. *PLos One*. 2013; doi: 10.1371/journal.pone.0076144.
65. Scholz MB, Lo CC, Chain PS. Next-generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. *Curr. Opin. Biotechnol*. 2012; 23: 9-15.
66. Sokhna C, Mediannikov O, Fenollar F, Bassene H, Diatta G, Tall A, Trape JF, Drancourt M, Raoult D. Point-of-care laboratory of pathogen diagnosis in rural Senegal. *PLoS Negl. Trop. Dis.* 2013; doi: 10.1371/journal.pntd.0001999.
67. Rosario K, Breitbart M. Exploring the viral world through metagenomics. *Curr. Opin. Virol*. 2011; 1: 289-297.
68. Fancello L, Trape S, Robert C, Boyer M, Popgeorgiev N, Raoult D, Desnues C. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. *ISME J.* 2013; 7: 359-369.
69. Whon TW, Kim MS, Roh SW, Shin NR, Lee HW, Bae JW. Metagenomic characterization of airborne viral DNA diversity in the near-surface atmosphere. *J. Virol*. 2012; 86: 8221-8231.
70. Djikeng A, Kuzmickas R, Anderson NG, Spiro DJ. Metagenomic analysis of RNA viruses in a fresh water lake. *PLoS One*. 2009; doi: 10.1371/journal.pone.0007264.
71. Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, Carlson C, Chan AM, Haynes M, Kelley S, Liu H, Mahaffy JM, Mueller JE, Nulton J, Olson R, Parsons R, Rayhawk S, Suttle CA, Rohwer F. The marine viromes of four oceanic regions. *PLoS Biol*. 2006; doi: 10.1371/journal.pbio.0040368.
72. Capobianchi MR, Giombini E, Rozera G. Next-generation sequencing technology in clinical virology. *Clin. Microbiol. Infect*. 2013; 19: 15-22.

73. Blomström AL. Viral metagenomics as an emerging and powerful tool in veterinary medicine. *Vet. Q.* 2011; 31: 107-114.
74. Thurber RV. Methods in viral metagenomics. In: Handbook of Molecular Microbial Ecology, Volume II: Metagenomics in Different Habitats. *FJ de Brujin, Wiley-Blackwell*; 2011.
75. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwser F. Laboratory procedures to generate viral metagenomes. *Nat. Protoc*. 2009; 4: 470-483.
76. Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, Moore NE, Ren X, Huang QS, Carter PE, Peacey M. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J. Virol. Methods*. 2014; 195: 194-204.
77. Breitbart M, Rohwer F. Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. *Biotechniques*. 2005; 39: 729-736.
78. Kleiner M, Hooper LV, Duerkop BA. Evaluation of methods to purify virus-like particles for metagenomic sequencing of intestinal viromes. *BMC Genomics*. 2015; doi: 10.1186/s12864-014-1207-4.
79. Cheval J, Sauvage V, Frangeul L, Dacheux L, Guigon G, Dumey N, Pariente K, Rousseaux C, Dorange F, Berthet N, Brisse S, Moszer I, Bourhy H, Manuguerra JC, Lecuit M, Burguiere A, Caro V, Eloit M. Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. *J. Clin. Microbiol*. 2011; 49: 3268-3275.
80. Daly GM, Bexfield N, Heaney J, Stubbs S, Mayer AP, Palser A, Kellam P, Drou N, Caccamo M, Tiley L, Alexander GJ, Bernal W, Heeney JL. A viral discovery methodology for clinical biopsy samples utilizing massively parallel next generation sequencing. *PLoS One*. 2011; doi: 10.1371/journal.pone.0028879.
81. Marston DA, McElhinney LM, Ellis RJ, Horton DL, Wise EL, Leech SL, David D, de Lamballerie X, Fooks AR. Next generation sequencing of viral RNA genomes. *BMC Genomics*. 2013; doi: 10.1186/1471-2164-14-444.
82. Kohl C, Brinkmann A, Dabrowski PW, Radonić A, Nitsche A, Kurth A. Protocol for metagenomic virus detection in clinical specimens. *Emerg. Infect. Dis*. 2015; 21: 48-57.

83. Cutler SJ, Fooks AR, van der Poel WH. Public health threat of new, reemerging, and neglected zoonoses in the industrialized world. *Emerg. Infect. Dis*. 2010; 16: 1-7.
84. Bishop-Lilly KA, Turell MJ, Willner KM, Butani A, Nolan NM, Lentz SM, Akmal A, Mateczun A, Brahmbhatt TN, Sozhamannan S, Whitehouse CA, Read TD. Arbovirus detection in insect vectors by rapid, high-throughput pyrosequencing. *PLoS Negl. Trop. Dis*. 2010; doi: 10.1371/journal.pntd.0000878.
85. Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, Moran MA, Nelson KE, Nilsson C, Olson R, Paul J, Brito BR, Ruan Y, Swan BK, Stevens R, Valentine DL, Thurber RV, Wegley L, White BA, Rohwer F. Functional metagenomic profiling of nine biomes. *Nature*. 2008; 452: 629-632.
86. Chandler JA, Thongsripong P, Green A, Kittayapong P, Wilcox BA, Schroth GP, Kapan DD, Bennett SN. Metagenomic shotgun sequencing of a bunyavirus in wild-caught *Aedes aegypti* from Thailand informs the evolutionary and genomic history of the phleboviruses. *Virology*. 2014; 464-465: 312-319.
87. Chandler JA, Liu RM, Bennett SN. RNA shotgun metagenomic sequencing of northern California (USA) mosquitoes uncovers viruses, bacteria, and fungi. *Front. Microbiol.* 2015; doi: 10.3389/fmicb.2015.00185.
88. Ng TF, Willner DL, Lim YW, Schmieder R, Chau B, Nilsson C, Anthony S, Ruan Y, Rohwer F, Breitbart M. Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS One*. 2011; doi: 10.1371/journal.pone.0020579.
89. Tokarz R, Williams SH, Sameroff S, Sanchez Leon M, Jain K, Lipkin WI. Virome analysis of *Amblyomma americanum*, *Dermacentor variabilis*, and *Ixodes scapularis* ticks reveals novel highly divergent vertebrate and invertebrate viruses. *J. Virol*. 2014; 88: 11480-11492.
90. Xia H, Hu C, Zhang D, Tang S, Zhang Z, Kou Z, Fan Z, bente D, Zeng C, Li T. Metagenomic profile of the viral communities in *Rhipicephalus spp*. ticks from Yunnan, China. *PLoS One.* 2015; doi: 10.1371/journal.pone.0121609.
91. Mellor PS, Boorman J, Baylis M. Culicoides biting midges: their role as arbovirus vectors. *Annu. Rev. Entomol*. 2000; 45: 307-340.

92. Tesh RB. The emerging epidemiology of Venezuelan hemorrhagic fever and Oropouche fever in tropical South America. *Ann. N.Y. Acad. Sci*. 1994; 740: 129-137.
93. 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.
94. Colson P, de Lamballerie X, Yutin N, Asgari S, Bigot Y, Bideshi DK, Cheng XW, Federici BA, Van Etten JL, Koonin EV, La Scola B, Raoult D. "*Megavirales*", a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch. Virol*. 2013; 158: 2517-2521.
95. Philippe N, Legendre M, Doutre G, Couté Y, Poirot O, Lescot M, Arslan D, Seltzer V, Bertaux L, Bruley C, Garin J, Claverie JM, Abergel C. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science*. 2013; 341: 281-286.
96. Legendre M, Bartoli J, Shmakova L, jeudy S, Labadie K, Adrait A, Lescot M, Poirot O, Bertaux L, Bruley C, Couté Y, Rivkina E, Abergel C, Claverie JM. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. USA.* 2014; 111: 4274-4279.
97. Reneto DG, benamar S, Khalil JB, Andreani J, Armstrong N, Klose T, Rossmann M, Colson P, Raoult D, La Scola B. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J. Virol*. 2015; 89: 6585-6594.
98. Pagnier I, Reteno DG, Saadi H, Boughalmi M, Gaia M, Slimani M, Ngounga T, Bekliz M, Colson P, Raoult D, La Scola B. A decade of improvements in *Mimiviridae* and *Marseilleviridae* isolation from amoeba. *Intervirology*. 2013; 56: 354-363.
99. Brashares JS, Golden CD, Weinbaum KZ, Barrett CB, Okello GV. Economic and geographic drivers of wildlife consumption in rural Africa. *Proc. Natl. Acad. Sci. USA*. 2011; 108: 13931-13936.
100. Chaber AL, Allebone-Webb S, Lignereux Y, Cunningham A, Rowcliffe JM, The scale of illegal meat importation from Africa to Europe via Paris. *Conservation Letters*. 2010; 00: 1-7.

101. Leroy EM, Gonzalez JP, Baize S. Ebola and Marburg haemorrhagic fever viruses: major scientific advances, but a relatively minor public health threat for Africa. *Clin. Microbiol. Infect*. 2011; 17: 964-976.
102. Brauburger K, Hume AJ, Mühlberger E, Olejnik J. Forty-five years of Marburg virus research. *Viruses*. 2012; 4: 1878-1927.
103. Hahn BH, Shaw GM, De Cock KM, Sharp PM. AIDS as a zoonosis: scientific and public health implications. *Science*. 2000; 287: 607-614.
104. Rabinowitz P, Conti L. One Health and emerging infectious diseases: clinical perspectives. *Curr. Top. Microbiol. Immunol*. 2013; 365: 17-29.
105. Preston ND, Daszak P, Colwell RR. The human environment interface: applying ecosystem concepts to health. *Curr. Top. Microbiol. Immunol*. 2013; 365: 83-100.
106. Mackey TK, Liang BA, Cuomo R, Hafen R, Brouwer KC, Lee DE. Emerging and reemerging neglected tropical diseases: a review of key characteristics, risk factors, and the policy and innovation environment. *Clin. Microbiol. Rev*. 2014; 27: 949-979.
107. McCloskey B, Dar O, Zumla A, heymann DL. Emerging infectious diseases and pandemic potential: status quo and reducing risk of global spread. *Lancet Infect. Dis*. 2014; 14: 1001-1010.
108. Daszak P, Epstein JH, Kilpatrick AM, Aguirre AA, Karesh WB, Cunningham AA. Collaborative research approaches to the role of wildlife in zoonotic disease emergence. *Curr. Top. Microbiol. Immunol*. 2007; 315: 463-475.
109. Conference Summary "One World, One Health: Building interdisciplinary bridges to health in a globalized world" 29th September 2004, The Rockefeller University. Available at http://www.oneworldonehealth.org/sept2004/owoh_sept04.html.
110. Rabinowitz PM, Kock R, Kachani M, Kunkel R, Thomas J, Gilbert J, Wallace R, Blackmore C, Wong D, Karesh W, Natterson B, Dugas R, Rubin C, Stone Mountain One Health Proof of Concept Working Group. Toward proof of concept of a one health approach to disease prediction and control. *Emerg. Infect. Dis.* 2013; doi: 10.3201/eid1912.130265.
111. Haagmans BL, Al Dhahiry SH, Reusken CB, Raj VS, Galiano M, Myers R, Godeke GJ, Jonges M, Farag E, Diab A, Ghobashy H, Alhajri F, Al-Thani M, Al-Marri SA, Al Romaihi HE, Al Khal A, Bermingham A, Osterhaus AD, AlHajri MM, Koopmans MP. Middle East respiratory syndrome coronavirus

in dromedary camels: an outbreak investigation. *Lancet Infect. Dis*. 2014; 14: 140-145.
112. Milne-Price S, Miazgowicz KL, Munster VJ. The emergence of the Middle East respiratory syndrome coronavirus. *Pathog. Dis*. 2014; 71: 121-136.
113. Guan Y, Chen H, Li K, Riley S, Leung G, Webster R, Peiris J, Yuen K. A model to control the epidemic of H5N1 influenza at the source. *BMC Infect. Dis*. 2007; doi: 10.1186/1471-2334-7-132.
114. Longini IM Jr, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings DA, Halloran ME. Containing pandemic influenza at the source. *Science*. 2005; 309: 1083-1087.
115. Lipkin WI, Anthony SJ. Virus hunting. *Virology*. 2015; 479-480: 194-199.
116. Koch R. Investigations into bacteria: V. The etiology of anthrax, based on the ontogenesis of *Bacillus anthracis*. *Cohns Beitrage zur Biologie der Pflanzen*. 1876; 2: 277-310.
117. Fredericks DN, Relman DA. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clin. Microbiol. Rev*. 1996; 9: 18-33.
118. Lipkin WI. The changing face of pathogen discovery and surveillance. *Nat. Rev. Microbiol*. 2013; 11: 133-141.
119. Geoghegan JL, Walker PJ, Duchemin JB, Jeanne I, Holmes EC. Seasonal drivers of the epidemiology of arthropod-borne viruses in Australia. *PLoS Negl. Trop. Dis*. 2014; doi: 10.1371/journal.pntd.0003325.
120. Lambrechts L, Ferguson NM, Harris E, Holmes EC, McGraw EA, O'Neill SL, Ooi EE, Ritchie SA, Ryan PA, Scott TW, Simmons CP, Weaver SC. Assessing the epidemiological effect of wolbachia for dengue control. *Lancet Infect. Dis*. 2015; 15: 862-866.
121. Jeffries CL, Walter T. The Potential Use of Wolbachia-Based Mosquito Biocontrol Strategies for Japanese Encephalitis. *PLoS Negl. Trop. Dis*. 2015; doi: 10.1371/journal.pntd.0003576.
122. Zouache K, Michelland RJ, Failloux AB, Grundmann GL, Mavingui P. Chikungunya virus impacts the diversity of symbiotic bacteria in mosquito vector. *Mol. Ecol*. 2012; 21: 2297-2309.
123. Bassene H, Sambou M, fenollar F, Clarke S, Djiba S, Mourembou G, L Y AB, Raoult D, Mediannikov O. High Prevalence of *Mansonella perstans* Filariasis in Rural Senegal. *Am. J. Trop. Med. Hyg*. 2015; 93: 601-606.

124. Agbolade OM, Akinboye DO, Olateju TM, Ayanbiyi OA, Kuloyo OO, Fenuga OO. Biting of anthropophilic *Culicoides fulvithorax* (*Diptera*: *Ceratopogonidae*), a vector of *Mansonella perstans* in Nigeria. *Korean J. Parasitol*. 2006; 44: 67-72.
125. Vasilakis N, Tesh RB. Insect-specific viruses and their potential impact on arbovirus transmission. *Curr. Opin. Virol*. 2015; 15: 69-74.
126. Gonzalez JP, Guiserix M, Sauvage F, Guitton JS, Vidal P, Bahi-Jaber N, Louzir H, Pontier D. Pathocenosis: a holistic approach to disease ecology. *Ecohealth*. 2010; 7: 237-241.
127. Willner D, Furlan M, Schmieder R, Grasis JA, Pride DT, Relman DA, Angly FE, McDole T, Mariella RP Jr, Rohwer F, Haynes M. Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity. *Proc. Natl. Acad. Sci. USA*. 2011; 108: 4547-4553.
128. Daszak P. Can we predict future trends in disease emergence? Chapter 5: Infectious disease emergence: Past, present and Future. In: Microbial Evolution and Co-Adaptation, A Tribute to the Life and Scientific Legacies of Joshua Lederberg: Workshop Summary. *Washington (DC): National Academies Press (US)*; 2009.
129. Murphy FA. Emerging zoonoses. *Emerg. Infect. Dis*. 1998; 4: 429-435.

Annexes

Article n°6: revue "Describing the silent human virome with an emphasis on giant viruses."

Describing the silent human virome with an emphasis on giant viruses

Nikolay Popgeorgiev[1*], **Sarah Temmam**[1*], Didier Raoult[1] and Christelle Desnues[1,a]

Published in: Intervirology. 2013;56(6):395-412. doi: 10.1159/000354561.

[1] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1095, Aix-Marseille Université, Marseille, France.

[a] Corresponding author: christelle.desnues@univ-amu.fr; tel.: +33(0)491324630; fax: +33(0)491387772.

*** These authors contributed equally to this work.**

Dans cet article j'ai pu collaborer avec le Dr. Popgeorgiev sur la réalisation d'une revue sur le virome humain en conditions non pathologiques en apportant ma contribution et mes connaissances sur les virus à génome ARN.

Intervirology

Intervirology 2013;56:395–412
DOI: 10.1159/000354561

Published online: October 17, 2013

Describing the Silent Human Virome with an Emphasis on Giant Viruses

Nikolay Popgeorgiev[a] Sarah Temmam[a] Didier Raoult[a, b] Christelle Desnues[a]

[a]URMITE UM63, CNRS 7278, IRD 198, INSERM 1095, Aix-Marseille Université, Marseille, France;
[b]Special Infectious Agents Unit, King Abdulaziz University, Jeddah, Saudi Arabia

Key Words
Papillomaviruses · Viral flora · Megavirome · Humans

Abstract
Viruses are the most abundant obligate intracellular entities in our body. Until recently, they were only considered to be pathogens that caused a broad array of pathologies, ranging from mild disease to deaths in the most severe cases. However, recent advances in unbiased mass sequencing techniques as well as increasing epidemiological evidence have indicated that the human body is home to diverse viral species under non-pathological conditions. Despite these studies, the description of the presumably healthy viral flora, i.e. the normal human virome, is still in its infancy regarding viral composition and dynamics. This review summarizes our current knowledge of the human virome under non-pathological conditions.

Humans and Viruses: An 'I Love You… Me Neither' Story

Since their discovery more than 100 years ago, viruses have been commonly described as obligate intracellular pathogens. Historically, the first studied virus was the one causing rabies, by Louis Pasteur. However, it was the Russian biologist, Dmitri Ivanovsky, and the Dutch botanist, Martinus Willem Beijerinckwent, who first isolated a tobacco-infecting microbe that caused tobacco mosaic disease. Ivanovsky demonstrated that crushed, infected tobacco leaf extracts remained infectious even after Chamberland filtration, which normally retains bacteria. He suggested that the infection might be caused by a bacterial toxin. However, Beijerinck went one step further, concluding that this new pathogen required living plants to replicate and multiply [1]. Subsequent studies showed that viruses infect all domains of life, including bacteria, archaea and eukaryotes, and are found in all ecological niches [2]. This pleiotropic distribution on our planet allows viruses to play the role of 'natural motors' that drive global energy and nutrient cycling [3, 4]. Until very recently, human viruses were considered only pathogens that were capable of causing human pandemics and a wide range of diseases that in some cases lead to a fatal outcome. With the development of new sequencing technologies (see the following section), which have allowed the analysis of the global viral population (DNA and RNA) in humans, known as the human virome, completely new human-associated viruses have emerged [5, 6]. However, the majority of these high-throughput sequencing techniques were performed with the use of filters with pore sizes in the range of 0.2–0.45 μm, which

N.P. and S.T. contributed equally to this work.

KARGER

E-Mail karger@karger.com
www.karger.com/int

Christelle Desnues
Unité de Recherche sur les Maladies Infectieuses et Tropicales Émergentes
URMITE CNRS-IRD UMR 7278, Faculté de Médecine, Aix-Marseille Université
27, boulevard Jean-Moulin, FR–13385 Marseille Cedex 05 (France)
E-Mail christelle.desnues@univ-amu.fr

filter larger viruses (see the section entitled 'The human megavirome'), resulting in a technical bias of the human virome. In this context, it became rapidly clear that viral richness and diversity in the human body under non-pathological conditions were widely underestimated. As an example, a rough estimation based on bacteria-infecting viruses (bacteriophages) indicates that there are 100 times more viruses than eukaryotic cells in our body [2, 7]. Human-associated viruses control the microbial diversity of the human gut and skin [8, 9]. Viruses affect the very foundation of our nature, our genome. Reminiscences of ancestral human-viral cohabitation are imprinted in our genome with approximately 100,000 known endogenous viral fragments, representing approximately 8% of our genome [10]. Finally, endogenous viral proteins have been associated with important physiological functions, such as mammal placental morphogenesis [11, 12].

In the present review, we briefly present the evolution of the virological techniques employed in the discovery of human-associated viruses. We then explore existing knowledge of the viral diversity found in human physiological systems under non-pathological conditions. Finally, we discuss the consequences of this human-virus cohabitation.

'Tracking the Small Guys': Tools for Viral Detection in Humans

Describing the human viral flora requires the right molecular and cellular tools. Historically, classical virology techniques were based on viral isolation from cells and the subsequent observation of cytopathic effects on cell lines or the intracerebral inoculation of suckling mice. Immunological methods, such as seroneutralization or hemagglutination, were then used to detect viral antigens. These techniques were largely used for the isolation of new pathogenic viruses that could be cultivated [13]. With progress in the field of molecular biology, PCR-based methods became the main techniques for viral detection from diverse environmental and clinical samples [14]. However, the identification of new or highly divergent viruses that could not be cultivated remained challenging. The development of next-generation sequencing techniques made it possible to sequence all viral genomes in a given sample without previous assumptions about their nature. These techniques, known as viral metagenomics, allowed the discovery of completely new viral species. Currently, the majority of viral metagenomics studies have been performed with DNA viruses [15–17]. To our knowledge, the overrepresentation of metagenomic studies performed on DNA viruses compared with RNA viruses is mainly due to technical limitations [18]. In the near future, advances in methodology will certainly enable routine implementation of RNA viral metagenomics studies in humans.

Exploring the Viral Flora in Humans

Digestive Tract

The most extensively studied part of the human body with respect to normal viral communities is the human gastrointestinal tract. The study of this system provides several practical advantages; it represents a non-invasive and easy sampling site as well as provides a sufficient amount of material, thereby allowing for the analysis of the viral composition and dynamics in the gut during a normal life. The first large-scale survey of the human gut virome was performed by Rohwer and colleagues [17] 10 years ago. Using partial shotgun sequencing on viral isolates obtained from healthy feces, they detected the presence of bacteriophages that were mainly related to the Siphoviridae family with an estimated 1,200 genotypes. Interestingly, the majority of detected sequences were unclassified, suggesting that the human gut virome was far more complex than expected. The same group undertook a more detailed study of the composition of DNA viruses from the feces of a healthy 1-week-old infant [19]. The results revealed a viral community with extremely low diversity, with an estimated 8 viral genomes corresponding to Podo-, Sipho- and Myo-virus DNA phages. Interestingly, the overall viral community in the human gut proved to be highly dynamic, changing dramatically between 1 and 2 weeks of age. A more detailed analysis of the infant gut was undertaken by Gordon et al. [16], who performed a comparative study of the viruses present in the fecal microbiota of monozygotic twins and their mothers. Interestingly, they found a high prevalence (>75%) of eukaryotic viral genomes in the gut virome, consisting of sequences related to Herpesviridae, Tymoviridae, Reoviridae and Poxviridae. The majority of bacteriophages and prophages were double-stranded DNA (dsDNA) phages and mostly members of the order Caudovirales. Notably, interindividual viral composition was highly divergent between monozygotic twins, whereas the intraindividual viral flora varied little over a year. All studies agreed that phage communities in the human gut played a critical

role in the control of the bacterial population. However, deciphering the phage-bacteria-human interactome has only recently begun to emerge. For instance, the viral metagenomics analysis of the oral cavity of healthy individuals performed by Willner et al. [20] showed that phages represent an important reservoir for bacterial virulence genes; thus, phages play a dual role in which they control the bacterial population but also contribute to bacterial pathogenicity and resistance via horizontal gene transfer.

A continually increasing number of eukaryotic single-stranded DNA (ssDNA) viruses in healthy human stool samples has also been identified through high-throughput sequencing or by PCR-based methods [21]. Interesting examples of ssDNA viruses are those from the Circoviridae family. For example, Li et al. [22] found new cycloviruses and circoviruses in human stool samples from Pakistan, Nigeria, Tunisia, and the USA. Another gyrovirus, the Chicken anemia virus, which is an important avian pathogen, was found with a high prevalence (25%) in the feces of Chilean children, suggesting a possible cross-species transmission from farm animals to humans [22–24].

Persistent viral shedding of dsDNA viruses of the Polyomaviridae family from the gastrointestinal tract has been reported in several studies. PCR-based detection of the BK, JC and SV40 viruses were identified in healthy children and adults. Viral detection was more frequent in stool samples from children compared with adults. These findings support the hypothesis that the gastrointestinal tract may be a site of Polyomavirus persistence with a possible fecal-oral route of viral transmission [25].

Multiple RNA viruses, generally considered as human pathogens, have also been detected in the normal gut viral flora. PCR-based or metagenomic analyses on 'healthy' human feces revealed the presence of several eukaryotic viral families, such as Astroviridae [26, 27], Caliciviridae [28, 29], Picornaviridae, Reoviridae and Picobirnaviridae, as well as plant viral families, such as Virgaviridae. Picornaviridae is the largest (+) ssRNA viral family with more than 12 recognized genera. Viruses belonging to this family have relatively strict host specificity but can infect a wide range of animals, including humans. Cellular tropism ranges from the gut to the central nervous and respiratory systems. In the gut viral flora, *Enterovirus* (Poliovirus, Echovirus, Coxsackievirus), *Kobuvirus* (Aichi virus), *Parechovirus* and *Cardiovirus* (Saffold virus) [30] have mainly been found, even in a non-pathological context as demonstrated by Kapusinszky [31]. Human Enterovirus type C has also been identified among healthy children [32, 33]. Human Cosavirus (for the common stool-associated Picornavirus) and human Salivirus (for the stool Aichi-like virus), which are not yet recognized as new species, have been reported in several studies in stool samples from healthy children [5, 34–38]; however, an understanding of their pathogenicity is lacking because they can also be present in cases of gastroenteritis.

Reoviridae and Picobirnaviridae are two dsRNA virus families responsible for gastroenteritis, but both may be present in apparently healthy humans. For example, rotaviruses (Reoviridae, *Rotavirus* genus) are a major cause of mortality in children under the age of 5 in developing countries, but some genotypes, such as G10P strains, have frequently been associated with asymptomatic neonatal infections in India [39]. The authors reported no significant differences in the sequences obtained from strains infecting symptomatic and asymptomatic neonates, suggesting that host-specific or environmental factors may contribute to the pathogenicity of a virus in a given population. Similar findings concerning Picobirnaviridae were reviewed by Ganesh [40] in 2012. These interesting findings suggest that frequent enteric infections with diverse enteric viruses occur during early childhood and less frequently in adults without clinical symptoms, indicating a change in the virome based on the age and environment of individuals.

Zhang et al. [41] performed the first metagenomic study on the RNA viral community in human feces. They found that the fecal flora was mainly composed of plant-infecting RNA viruses, specifically Pepper mild mottle virus and Tobacco mosaic virus. Plant viruses are generally considered incapable of infecting humans. However, a few studies have reported the presence of plant viral RNA in the human body, including the respiratory system via cigarette use [42] and the gut via contaminated food consumption [43]. Colson et al. [43] noted a higher prevalence of Pepper mild mottle virus in the stools of adults but not children, possibly due to a difference in their diet. In fact, the presence of plant viruses in humans may not represent an infection of the human body but may be due instead to a passive mechanism, such as the ingestion of contaminated food products, suggesting a role of mammals, including humans, as vectors for plant viruses.

The presence of plant viruses in the human gut highlights the fact that the virome may vary between individuals based on diet as demonstrated for bacteria [44]. The virome of the gut may also depend on environmental factors, such as geography, eating habits or ethnic differences, resulting in interindividual variability.

Blood

The human blood and derived products represent a constant need for blood transfusions and medical treatment. However, the blood also represents an important viral reservoir, and some viruses may be pathogenic. Thus, describing the viral flora in the blood has direct consequences for public health. An increasing body of evidence argues that in apparently healthy individuals, the blood is not sterile and may contain many viral species. The majority of the 'normal' blood viral flora is composed of ssDNA viruses of the Anelloviridae family with Torque teno viruses (TTVs) being the most commonly detected. TTVs are small non-enveloped viruses with icosahedral symmetry that have high genetic diversity. Indeed, the first genus of Anelloviridae, *Alphatorquevirus*, contains 29 TTV species. Initially detected in a Japanese patient with posttransfusion hepatitis [45], TTVs are now considered commensal with a worldwide distribution [46–48]. Although replicative forms of TTV DNA have been detected in peripheral blood mononuclear cells [49], viral loads higher than those in the blood have been identified in the bone marrow, lung, spleen and liver [50]. Thus, it is tempting to speculate that the human blood may play a double role in TTV, both in viral replication and viral dissemination. Several studies have proposed that the main routes for TTV spread are via blood transfusion, oral transmission and sexual contact [48, 51, 52]. Mother-to-child transmission of TTV has also been reported [53]. These multiple routes of dissemination may contribute to the pandemic nature of TTV infection.

Another frequently detected ssDNA virus family is the Parvoviridae family. Parvoviruses are small non-enveloped viruses with icosahedral symmetry and are approximately 18–26 nm in diameter. Human Parvovirus (PARV)4 was originally detected in the plasma of a person at risk for infection with HIV through intravenous drug use [54]. However, frequent detection of PARV4 and PARV5 in the plasma of apparently healthy blood donors as well as in symptomatic individuals has been reported [55]. In some parts of the world, including sub-Saharan Africa, PARV4 seropositivity is frequently detected with high prevalence in the population [56]. Although infections with PARV4 are not accompanied by long-term viremia, viral DNA sequences can likely be detected in tissues for a long time after exposure [57–59], thereby encompassing a form of latency or persistence that is shared with other human PARV, e.g. human PARV B19 and adeno-associated viruses [60–62].

Eukaryotic dsDNA viruses have also been detected in blood donors. Egli et al. [63] reported the prevalence of the BK and JC polyomaviruses by testing the blood of 400 donors. Interestingly, they found significant differences between the BK and JC viruses with respect to virus-host interaction and epidemiology. Moreover, lymphotropic Polyomavirus and human Bocavirus (HBoV) have also been frequently found in the peripheral blood of immunocompromised and apparently healthy subjects [64, 65].

An increasing number of studies have reported the emergence of new retroviral infections in primate hunters in Africa. Viruses from Retroviridae, such as Simian foamy virus, Spumaretrovirus or Human T-lymphotropic virus 3/4, are naturally acquired by apparently healthy individuals in central Africa after hunting and the butchering of infected meat [66, 67]. Moreover, zoonotic retroviruses are frequently detected in the blood of research workers in zoos [68–70]. Although the viruses are found in apparently healthy individuals, the long-term consequences of these viral infections must be evaluated. Indeed, it is possible that in the case of persons with immune disorders, these viruses may contribute to the development of chronic pathologies.

RNA viruses are also part of the viral flora in the blood, but they are mainly pathogenic, and in such cases they represent the viremic phase of infection. Only a few examples of circulating 'asymptomatic' RNA viruses have been reported, but their pathogenicity is not understood. Recently, several arthropod-borne viruses (arboviruses) belonging to the Flaviviridae family, such as Dengue virus, have been detected in the blood of apparently healthy individuals [71]; however, Dengue virus infections can cause undifferentiated fevers and even deaths in some cases. In 2001, Sonoda and Nakayama [72] described circulating Measles virus in peripheral blood mononuclear cells from healthy children exposed to an environment in which measles was circulating. The Measles virus belongs to the Paramyxoviridae family (*Morbilivirus* genus) and is a major cause of child death in non-vaccinated populations. The authors found a high prevalence of Measles virus (23.4%) in exposed populations, but no detection of viral RNA was observed in unexposed children, suggesting an asymptomatic circulation of the virus.

Respiratory Tract

The respiratory tract is a major gateway of infections for the human body, mainly due to environmental exposure. We distinguish upper respiratory tract infections, which refer to infections of the nasopharynx, larynx, tonsils, sinuses and ears, from lower respiratory tract infections, which refer to infections of the trachea, bronchi and alveoli. The frequency of symptomatic viral respiratory

tract infections is higher in young children compared with adults. Although many viruses are responsible for pathologies of the respiratory system (including human rhinoviruses, hRVs, respiratory syncytial virus, influenza and coronaviruses), a number of viruses may be found without any pathological context. In 2009, Willner et al. [73] compared the DNA virome of the upper respiratory tract in people with or without cystic fibrosis to determine whether there was a core respiratory tract virome in non-diseased individuals. In comparison with other viromes, the authors found that the respiratory tract virome had low species richness, most likely due to physical and biological barriers. Although more than 90% of the sequences were unknown, the authors reported the presence of a core set of 19 bacteriophage genomes in the sputum of healthy individuals, reflecting the airborne contamination of each individual. For example, *Streptococcus* phage Cp-1, *Haemophilus influenza* phage HP-1 and *Brucella melitensis* 16 M BrucI prophage were detected along with a random distribution of other phage genotypes. The composition of this phage community may reflect a specific environment, and we can assume that interindividual variability may be due to a difference in environmental exposure. Indeed, some organs, such as the respiratory tract, having frequent contact with the environment, are exposed to different viral communities. In contrast, in cystic fibrosis metagenomes, the pathology appears to favor a phage composition. The study revealed the presence of a core of 20 eukaryotic DNA viral genomes in healthy individuals, mainly composed of adenoviruses, herpesviruses and human papillomaviruses (HPVs). The authors suggested that eukaryotic viral communities in apparently healthy individuals likely represent transient infections that are rapidly cleared by immune cells or viral particles that are removed from the airway via mucociliary clearance.

A metagenomic study conducted in 2012 by Wylie et al. [74] on young children with or without unexplained fever revealed the presence of DNA viruses, including human Parvoviridae viruses (*Dependovirus* and *Bocavirus* genera), in the nasal swabs of healthy children. HBoV is the fourth most common virus found in respiratory samples and may be found in healthy subjects [75], but at a lower frequency than it is found in diseases. HBoV may persist in the respiratory tract for a longer period of time than other respiratory agents, resulting in detection of low levels of HBoV [6]. The role of HBoV as a pathogen remains unclear, but the replication mode of this virus, i.e. with the need of 'helper viruses' (e.g. adenoviruses or herpesviruses), may associate it with respiratory tract diseases [76]. In their metagenomic study, Wylie et al. [74] reported the presence of human adenoviruses in the nasal swabs of healthy children. Adenoviridae (*Mastadenovirus* genus) viruses are classified into 7 subgroups (A–G) with 55 known serotypes. These viruses usually cause asymptomatic or mild disease in humans, but occasionally some specific subtypes (mainly types 3 and 7) cause severe syndromes, including neurological disorders or deaths in immunocompromised populations or children. In 2011, Heydari et al. [77] reported a case of fatal infection due to the combination of HBoV and human Adenovirus in a previously healthy child. Although a single infection by one of these 2 viruses mainly remains asymptomatic, coinfection with both HBoV and human Adenovirus may result in lethal disease, suggesting that interactions between viruses of the viral communities can lead to pathology.

hRVs are small, non-enveloped, positive ssRNA viruses belonging to the Picornaviridae family (*Enterovirus* genus). They comprise 3 major genotypes (hRV-A, B and C) that cause a wide range of respiratory illnesses, from mild common colds to serious lower respiratory tract infections [76]. hRVs are also frequently found in asymptomatic children and adults. In 2006, Winther et al. [78] conducted a prospective cohort study of 15 children aged 1–9 years over a 9- to 12-month period. They found a high hRV presence (21%) in the nasal swabs of young children without any reported symptoms. Viral shedding began several days prior to the onset of symptoms and several days after symptoms occurred. They also noted that the maximum duration of viral presence was relatively short (1–3 weeks). Longer hRV presence may be due to reinfection with a new hRV genotype as reported by Van der Zalm et al. [79]. In 2012, Annamalay et al. [80] conducted a similar study on a prospective cohort of 95 children in Australia. No significant difference was observed in the hRV-A prevalence among children with or without symptoms (i.e. a blocked or runny nose).

Wylie et al. [74] revealed the presence of paramyxoviruses (e.g. Paramyxoviridae, *Respirovirus* and *Pneumovirus* genera) in the nasal swabs of apparently healthy individuals. They also reported the presence of *Influenzavirus A*, *Parechovirus* and *Coronavirus* in nasopharyngeal swabs, similar to that reported by Van der Bergh et al. [81]. Wylie et al. [74] reported a difference in the abundance of viral sequences with febrile children exhibiting 1.5-fold more viral sequences than samples from afebrile children. They also reported a difference in the diversity of the viral genera present in the samples with

a lower diversity found in apparently healthy children. However, no causal relationship between a specific virus and the pathology was found. These observations support the hypothesis that pathology may be due to an imbalance of the microbial communities present in the human body.

Due to the non-invasive nature of the sampling, mainly viromes of the upper respiratory tract of apparently healthy people have been assessed. The viral composition of the lower respiratory tract has been studied using bronchoalveolar lavage samples. One recent study on bronchoalveolar lavage samples from intensive care unit patients identified the presence of viruses from Herpesviridae, Paramyxoviridae and Picornaviridae families [82]. Notably, these viruses were found not only in pneumonia patients, but also in control subjects without pneumonia illness. Thus, additional studies are needed to assess the viral composition of this part of the respiratory system.

Teguments

The human teguments comprise the skin, hair and nails, and play a major role as a barrier protecting the human body from the outside environment. They also represent a complex ecosystem harboring diverse bacterial, fungal and viral species. High-throughput sequencing data on the viral flora of the skin have just begun to be generated. Using Illumina technology, Foulongne et al. [15] detected a high diversity of prokaryotic and eukaryotic viral species in DNA extracts from healthy skin swabs. The most abundant were eukaryotic DNA viruses, such as ssDNA viruses of the Circoviridae family as well as dsDNA viruses of the Polyomaviridae and Papillomaviridae families. Members of Circoviridae (*Gyrovirus* genus) have been previously reported in the human skin of 4% of healthy persons [83]. Sauvage et al. [83] identified a new virus, the human Gyrovirus, in a skin swab sample from an apparently healthy donor. The host range and infection cycle of human Gyrovirus remains unknown. Other ssDNA viruses from the Parvoviridae family were also found in non-diseased human skin. Although initially reported as the etiological agent of erythema infectiosum, PARV B19 is commonly harbored in apparently healthy human skin. Bonvicini et al. [84] found the prevalence of B19 to be 25% in apparently healthy skin biopsies. Interestingly, the group found that young subjects had a significantly higher rate of B19 viremia compared with adults, suggesting that long-term viral persistence may be the common outcome after primary infection.

Polyomaviruses are also common skin viruses. They have a circular dsDNA genome of approximately 5,000 bp that is surrounded by a non-enveloped icosahedral capsid. Polyomaviruses were first described in 1953 in mice, but since then these viruses have been detected in other vertebrate species, including humans. In humans, a new Polyomavirus, Merkel cell Polyomavirus (MCPyV), was recently identified [85, 86]. The presence of MCPyV in human skin has been associated with an aggressive form of skin cancer, Merkel cell carcinoma (MCC). MCPyV infections are found in 80% of MCCs. However, MCPyV and two newly identified polyomaviruses, HPyV6 and HPyV7, are also frequently shed from apparently healthy human skin [15, 87]. In the case of MCC, the accumulation of deleterious mutations in the MCPyV genome, including the viral T antigen gene, render the virus non-infectious. Thus, the oncogenic role of MCPyV does not necessary reflect its lifestyle but rather the consequence of deleterious viral mutations. Other dsDNA viruses that are associated with neoplastic development have also been identified in healthy skin. Detection of α- and β-HPVs as well as human Herpesvirus (HHV)7 has been reported recently in skin biopsies [88, 89]. HHV7 was initially isolated from CD4+ T cells obtained from peripheral blood lymphocytes of an apparently healthy individual [90] and was later associated with primary cutaneous T cell lymphomas (CTCLs). However, the low prevalence of HHV7 in CTCL as well its presence in healthy skin biopsies suggests that HHV7 may not be the primary cause of CTCL [89, 91].

Bacteria-infecting viruses are also frequently found in the human skin and most likely play an important role in controlling the bacterial population. Using viral metagenomics, viruses belonging to the Myoviridae, Siphoviridae, Microviridae, Podoviridae and Inoviridae families were identified, and viruses from the Siphoviridae and Microviridae families were the most abundant. One common phage genera present in healthy human skin consisted of bacteriophages infecting *Propionibacterium acnes* (Siphoviridae family). The *P. acnes* bacterium represents a dominant member of the skin microflora and has also been implicated in the pathogenesis of acne. Multiple *P. acnes* bacteriophages isolated from the sebaceous follicles of healthy skin donors have recently been characterized [9]. Interestingly, these phages showed reduced genetic variability with a broad range of infecting bacterial strains, suggesting the existence of evolutionary constraints that preserve the homogeneity of the phage population.

Nervous System

Little information is available concerning the viral flora in the human nervous (central and peripheral) system in apparently healthy conditions. Examples of neurotropic human viruses are the Herpes simplex virus (HSV)1 and HSV2, which belong to the Herpesviridae family. These viruses have a dsDNA genome located within an icosapentahedral capsid surrounded by an amorphous protein-like material (known as the tegument), which is in turn encapsulated by an envelope consisting of polyamines, lipids and glycoproteins [76]. Genetically, HSV1 and HSV2 are closely related, sharing approximately 70% homology. During primary infection, the virus enters the nerve endings at the peripheral mucocutaneous region. The viral capsid is brought via fast axonal transport into the neuronal cell body of the dorsal root ganglia or the trigeminal ganglia. The viral DNA enters the nucleus of the neuron where it enters a latent state [92]. Notably during this period, two latency-associated transcripts are expressed [93]. Latency-associated transcripts have been shown to have antiapoptotic activity, thereby sustaining the survival of neurons. This activity illustrates the virus-to-host adaption and the benefit of a latent persistence in the nervous system. Although HSV1 and HSV2 are associated with clinical complications, the majority of the infections remain asymptomatic for years or even decades. Indeed, under immunocompetent conditions, the reactivated infection usually remains confined to the vicinity of a single dorsal root ganglion. It has been estimated that asymptomatic reactivation of HSV1 may exceed clinical recrudescence, and asymptomatic HSV2 shedding can occur in more than two-thirds of seropositive individuals [94, 95].

Another interesting example of a neurotropic virus is the Borna disease virus (BDV), which is part of the Bornaviridae family. BDV is an 80- to 100-nm enveloped virion, containing an 8.9-kb (–) ssRNA genome that replicates in the cell nucleus [96, 97]. In vitro BDV induces non-cytopathic chronic infections in neurons [98]. BDV infection was first identified in horses, and natural infections with BDV were subsequently detected in other vertebrates, including humans [99]. In this context, BDV was suggested as a causative agent of diverse human psychiatric disorders [100–102]. Despite these findings, the seroprevalence of the virus in healthy control groups makes the causal relationship between BDV infection and brain disorders hardly verifiable [103]. Recently, endogenous BDV sequences homologous to the viral nucleoprotein were detected in several mammalian species, including humans, suggesting an ancient cohabitation with a BDV ancestor [104, 105]. Overall, further efforts, especially using a viral metagenomics approach, should be put into the study of the viral diversity of the human nervous system.

Genito-Urinary Tract

The viral flora of the genito-urinary tract has been mainly studied in pathological situations, and gaps in the knowledge of the viral flora in apparently healthy conditions need to be filled. Asymptomatic shedding from the genito-urinary tract was reported mainly for dsDNA eukaryotic viruses of the Adenoviridae, Herpesviridae, Papillomaviridae and Polyomaviridae families with the exception of ssDNA viruses of the Anelloviridae family [83, 106–111]. In the case of polyomaviruses, it appears that viral excretion was correlated with the host immune status. Indeed, Csoma et al. [112] detected KI virus and WU virus in the urine of renal transplants but not in the control groups. Moreover, immunosuppression due to pregnancy led to a higher prevalence of BK virus in urine samples in pregnant women compared to non-pregnant women [113].

Multiple herpesviruses were also frequently detected in the genito-urinary tract, especially in the semen of apparently healthy donors. In this case it appears that some herpesviruses, such as human Herpesvirus 6 A/B or the Cytomegalovirus, were able to attach to the sperm head with an intact acrosome [108, 113]. Thus, given the potential risk some herpesviruses may represent to the newborns, additional research is required to evaluate the impact of this asymptomatic shedding from herpesvirus-positive donor semen.

Broad Distribution and Impact of Papillomaviruses in the Human Body

When examining the repartition of viruses according to their distribution in the human body (fig. 1), one can note that DNA viruses of Herpesviridae, Papillomaviridae, Polyomaviridae and Anelloviridae families are present both in the respiratory tract, the gut, the skin, the blood and the genito-urinary tract. One hypothesis may be related to the viral-host adaptation process. For sustained infection, viruses need to have wide range of body repartition allowing them to proliferate efficiently.

Papillomaviruses represent good examples of pleiotropic human viruses in the human body. Papillomaviruses are 55- to 60-nm non-enveloped DNA viruses composed of a single, circular dsDNA molecule. This viral family

consists of more than 120 different HPV types, about 40 of which are sexually transmitted HPVs and a dozen have been identified as the causative agents of cervical, anal, vaginal and penile cancer [114]. HPVs are present in more than 99% of cervical cancers, and HPV type 16 (HPV-16) and HPV-18 are the predominant causes of infection in these cases [115]. These two HPV types are indeed associated with 70% of all cervical cancers with predominance of HPV-16 accounting for about 50% of cases [116]. More recently, papillomaviruses were linked to head and neck malignancies as well. In these cases, the primary causes for these carcinomas were attributed to alcohol and tobacco consumption. However, the number of respiratory and digestive tract cancers caused by HPV infections is constantly increasing [117–119]. Indeed patients with HPV-positive carcinoma are generally younger adults and not alcohol and tobacco users. These carcinomas are mainly localized in the oropharynx and in particular at the tonsils. HPV is found with a prevalence of 40–90% of the oropharynx cancers, depending on the geographical distribution [120–122].

HPVs have cellular tropism for the stratified squamous epithelia. Although the exact mechanism of Papillomavirus tumorigenesis is not fully elucidated it is generally accepted that this effect is mediated through E6, E7 viral proteins which control cell death and proliferation [123–125]. Despite the oncogenic properties of these viruses, the majority of HPV infections remain asymptomatic, and they are cleared by most people without medical consequences. Indeed, the clearance of HPV 18 months postinfection in the male population is 100%, whereas in females it is 97%, suggesting that in the case of an immunocompetent host, HPV infection manifests as a transient phenomenon [126, 127]. The significance of their presence in an apparently healthy context remains unknown.

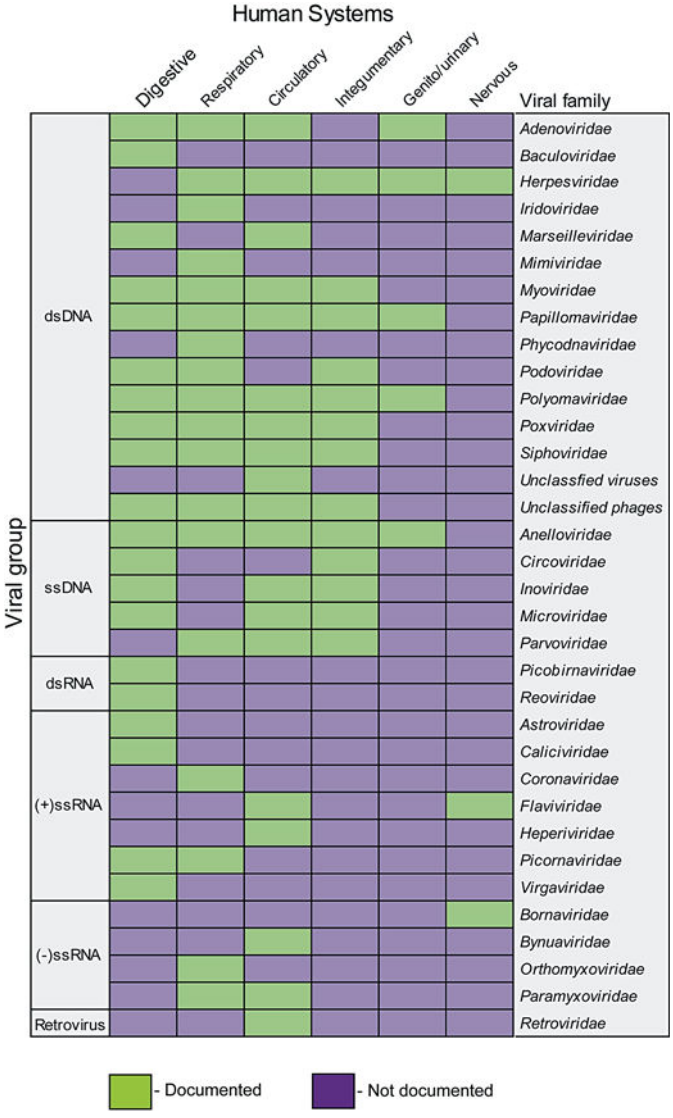


Fig. 1. Description of the viral composition in the human body. Table summarizing the viral families documented (in green) or not documented (in violet) in each human system.

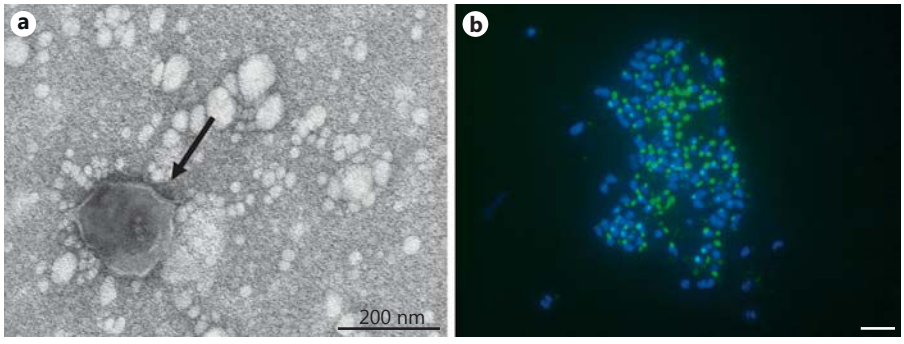


Fig. 2. Detection of GBM. **a** Negative staining of a Marseillevirus-like particle (arrow) present in the virus-purified fraction of serum from blood donor No. 27725. **b** Epifluorescent microscopy images from fluorescent in situ hybridization of GBM in serum from blood donor No. 27725. The DNA probe was synthesized using the Marseillevirus genomic region, orf 152–153, and is marked in green; nuclear staining with DAPI dye is in blue. Scale bar = 10 μm.

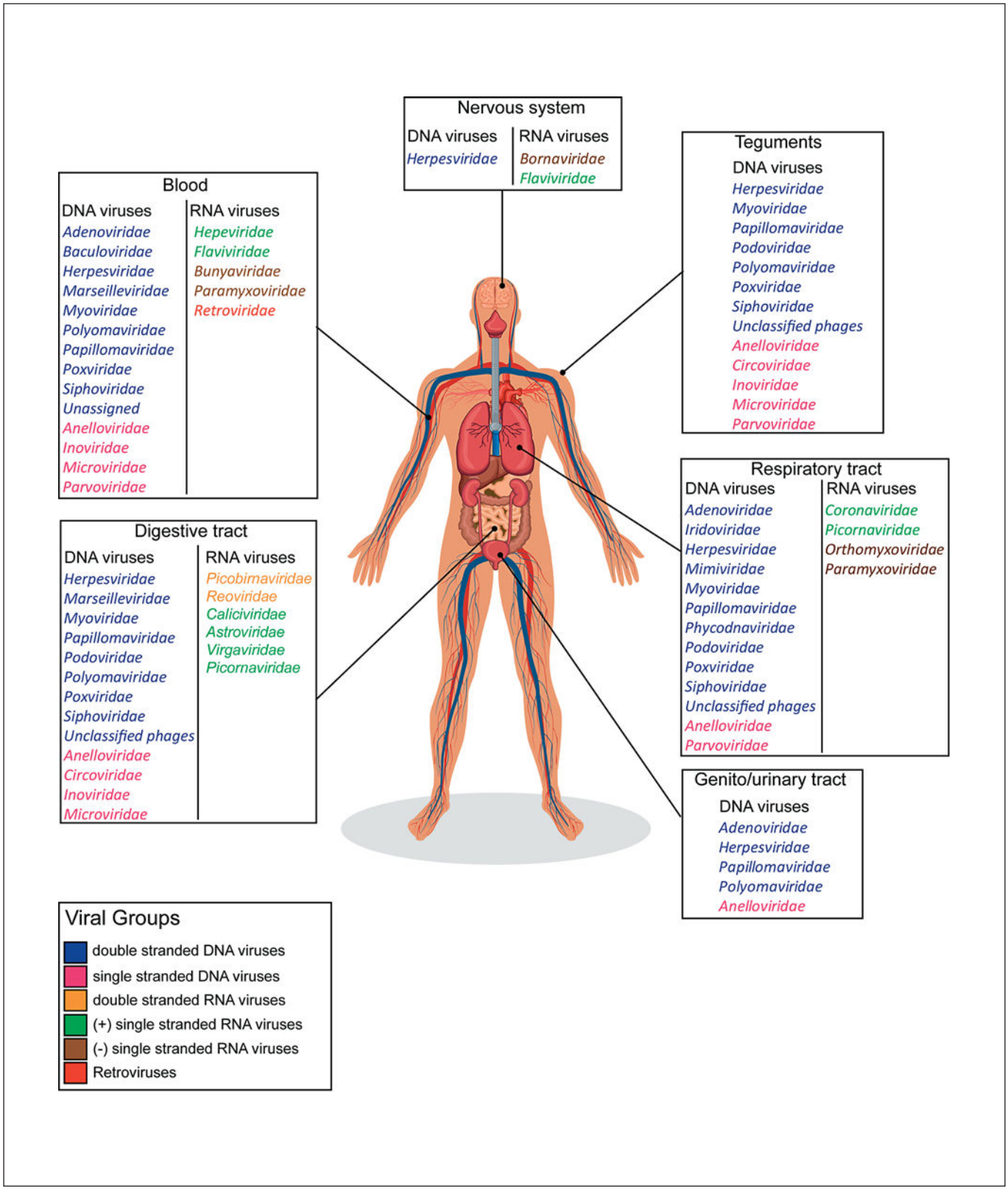


Fig. 3. The human virome in non-pathogenic conditions: distribution of the viral families found in the major human systems. Each viral group is represented with a unique color.

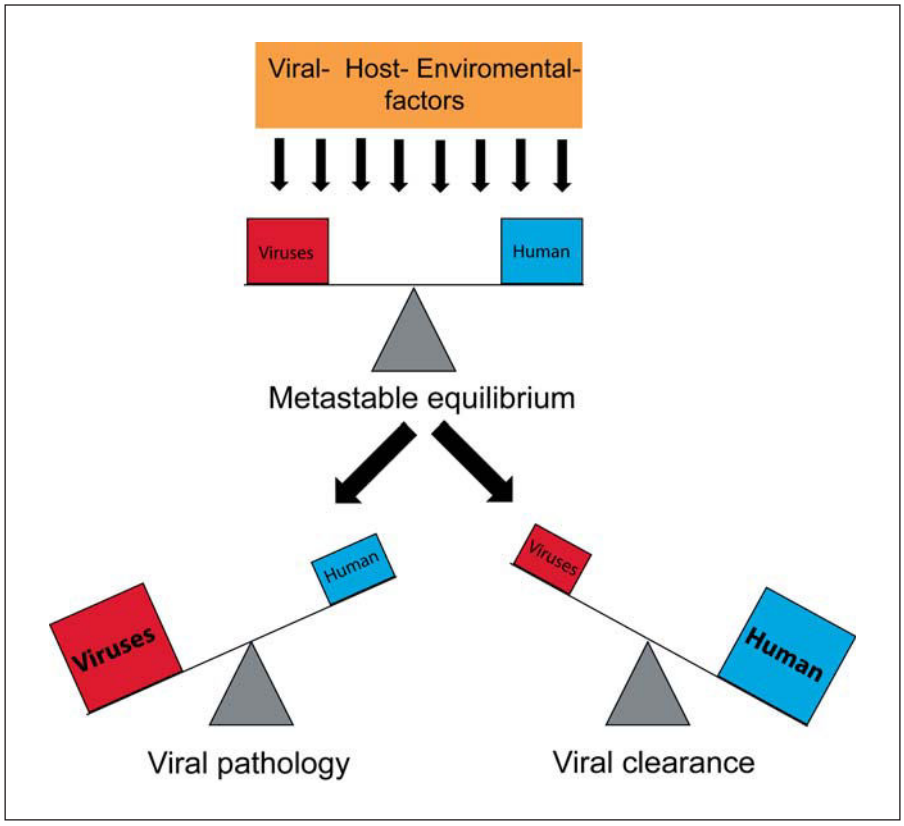


Fig. 4. Human virus metastable equilibrium in non-pathogenic conditions. Schematic representation of the steady state of the human virome in non-pathogenic conditions as regulated by three major factors (virus, host and environment). The disequilibrium of this metastable system leads either to viral spreading or to viral clearance.

The Human Megavirome

dsDNA viruses with large genomes (also known as giant viruses) represent a monophyletic group of viruses classified under the order of Megavirales [128]. Giant viruses are divided into seven viral families, including Poxviridae, Iridoviridae, Ascoviridae, Mimiviridae, Phycodnaviridae, Asfaviridae and the recently described Marseilleviridae [128, 129]. These viruses infect a wide range of eukaryotes, including phagocytic protists and humans [130]. In humans, members of only two of the families, Poxviridae and Mimiviridae, have been linked to disease [131–133]. With next-generation sequencing technologies, an accumulating body of evidence indicates the presence of these viruses in non-pathological conditions. For instance, a metagenomics study carried out by Willner et al. [73] detected multiple DNA sequences related to Poxviridae, Iridoviridae, Mimiviridae and Phycodnaviridae. Moreover, several studies have identified the presence of giant viruses in the human gut in both adults and babies [16, 19, 134]. Breitbart et al. [19] detected sequences homologous to Lymphocystis disease virus (Iridoviridae), a fish-infecting pathogen, whereas Gordon et al. [16] detected previously uncharacterized Pox-related viral sequences in the infant gut.

Recently, a new giant virus closely related to Marseilleviridae, Senegalvirus, was recovered from a stool sample of a 20-year-old Senegal man [134]. Senegalvirus was detected by ultradeep sequencing and was isolated using an amoebal coculture. The Senegalvirus dsDNA genome is approximately 373 kbp in length, making this genome the largest among marseilleviruses. In the same stool, sequences related to the giant Mimivirus were also found [135].

Another virus closely related to the Marseilleviridae family was recently identified in human blood. This new virus, Giant Blood Marseillevirus (GBM), has an estimated 357-kbp dsDNA genome surrounded by a 200-nm capsid (fig. 2). The GBM virus was initially isolated from a blood transfusion pocket using a 0.45-μm filter coupled with high-throughput sequencing from a 32-year-old healthy female donor [136]. Further testing identified concomitant elevated IgG levels and viral DNA in some blood donors, suggesting the persistence of the GBM virus in the blood. Interestingly, GBM was found to infect and replicate in human T cells, but not in amoebas.

Table 1. Summary of the viral families, genera and, in some cases, species found in each human system

Source	Viral group	Viral family	Viral genera/species	Reference
Digestive tract	dsDNA	Adenoviridae	*Enteric adenovirus 40, 41*	139
		Iridoviridae	*Lymphocystis disease virus*	19
		Myoviridae	*phiBCD7, Bacillus phage G, phiP-SSM4*	17, 19
		Podoviridae	*Enterobacteria phage P22, phage T3*	17, 19
		Siphoviridae	*Listeria phage A118, phiE125 Lactococcus phage bIL285, phiCP39-O, Clostridium phage phiCP39-O, Mycobacterium phage Athena, phage PA6, phage SM*	16, 17, 19, 140
		Unclassified phages	*Halophage eHP-10*	17, 19
		Papillomaviridae	*Human papillomavirus 6, 18, 66*	141
		Polyomaviridae	*BK virus, JC virus, SV40 virus, Human polyomavirus 9, 12*	25, 142, 143
		Herpesviridae	*Epstein-Barr virus, Human cytomegalovirus*	16
		Poxviridae	ND	16
		Marseilleviridae	*Senegalvirus*	134
	ssDNA	Anelloviridae	TTV	144, 145
		Circoviridae	*Chicken anemia virus*	22 – 24
		Microviridae	*Chlamydia phage 1,3,4, Bdellovibrio phage phiMH2K, Chlamydia phage CPG1, Spiroplasma phage 4, Chlamydia phage CPAR39*	140
	dsRNA	Picobirnaviridae	*Human picobirnavirus*	31, 41, 146
		Reoviridae	*Human rotavirus*	39
	(+) ssRNA	Caliciviridae	*Norwalk virus*	28, 29, 147
		Astroviridae	*Human astrovirus*	26, 27
		Virgaviridae	*Pepper mild mottle virus, Tobacco mosaic virus*	41, 42
		Picornaviridae	*Human cosavirus, Human klassevirus/salivirus, Aichi virus, Human enterovirus, Human parechovirus, Saffold cardiovirus, Human echovirus, Human coxsackievirus, Human poliovirus*	5, 31, 36, 37, 32 – 35, 140
	(–) ssRNA	Not documented		
	Retroviruses	Not documented		
Respiratory tract	dsDNA	Adenoviridae	*Human adenovirus, Bovine adenovirus A*	73, 74, 77, 81
		Iridoviridae	*Aedes taeniorhynchus iridescent virus*	73
		Herpesviridae	HHV 1, 2, *Bovine herpesvirus 5, Cercopithecine herpesvirus 1, 2, 9, Suid herpesvirus 1*	73
		Mimiviridae	*Acanthamoeba polyphaga mimivirus*	73
		Myoviridae	*Haemophilus phage HP1, Aeromonas hydrophila phi Aeh1, Aeromonas phi 31, Escherichia coli phi CP073-4 prophage, Lactobacillus plantarum phi LP65, Mycobacterium phi Bxz1, Pseudomonas phi KZ, Staphylococcus phi Twort, Vibrio parahaemolyticus phi KVP40*	73
		Papillomaviridae	*HPV*	73
		Phycodnaviridae	*Chlorella virus ATCV-1, Chlorella virus FR483, Ectocarpus siliculosus virus 1, Paramecium bursaria Chlorella virus AR158*	73
		Podoviridae	*Streptococcus phage Cp-1, Brucella melitensis 16M BrucI*	73
		Polyomaviridae	*KI virus, WU virus*	81
		Poxviridae	*Amsacta moorei entomopoxvirus 'L', Melanoplus sanguinipes entomopoxvirus, Taterapox virus*	73
		Siphoviridae	*Bacillus subtilis phi SPBc2, Bacillus subtilis phi 105*	73
		Unclassified phages	*Bacillus cereus phage phBC6A51, Escherichia coli phi CP4-6 prophage, Escherichia coli phi QIN prophage, Escherichia coli phi Sp18 prophage, Mycobacterium phage CJW 1, Shigella flexneri phi Flex4 prophage, Xylella fastidiosa phi Xpd5*	73
	ssDNA	Anelloviridae	TTV, TTV-midi	110
		Parvoviridae	*Adeno-associated virus, Human bocavirus*	6, 74, 75, 77
	dsRNA	Not documented		
	(+) ssRNA	Coronaviridae	*Human coronavirus OC43, NL63, HKU, 229E*	74, 81
		Picornaviridae	*Human rhinovirus, Human parechovirus*	74, 78 – 81

Table 1 (continued)

Source	Viral group	Viral family	Viral genera/species	Reference
Respiratory tract	(–) ssRNA	Orthomyxoviridae	*Influenzavirus A*	74, 81
		Paramyxoviridae	*Human parainfluenzavirus 1–4, Human respiratory syncytial virus, Human metapneumovirus*	74, 148
	Retroviruses	Not documented		
Blood	dsDNA	Adenoviridae	*Adenovirus-36, Human adenovirus*	106, 149
		Baculoviridae	*Spodoptera litura nucleopolyhedrovirus*	47
		Herpesviridae	HHV 3-like, *Suid herpesvirus 1*-like, *Human cytomegalovirus*, *Epstein-Barr virus*, HHV 6B, HHV 7	47, 90, 150
		Marseilleviridae	*Giant Blood Marseillevirus*	136
		Myoviridae	*Streptococcus pneumoniae bacteriophage EJ-1*-like	47
		Polyomaviridae	*Lymphotropic polyomavirus, BK virus, JC virus, KI virus, WU virus, Human polyomavirus 9*	63, 64, 143, 151
		Papillomaviridae	*α-, β-, γ-HPVs*	152
		Poxviridae	*Cowpox virus*-like	47
		Siphoviridae	*Clostridium perfringens bacteriophage Φ3626, Methanobacterium phage psiM2*-like, *Enterobacteria phage λ*	47, 136
		Unassigned	*Nidivirus* (*Heliothis zea virus 1*-like)	47
	ssDNA	Anelloviridae	TTV, TTV-midi, TTV-mini SEN virus, unclassified anelloviruses	47, 136, 153 – 155
		Inoviridae	*Ralstonia phage RSM 1, 3*	136
		Microviridae	*Chlamydia phage φCPAR39*-like	47
		Parvoviridae	*Human bocavirus*, PARV 4, 5, *Adeno-associated virus*	55, 65, 156
	dsRNA	Not documented		
	(+) ssRNA	Hepeviridae	*Hepatitis E virus*	157 – 159
		Flaviviridae	*Dengue virus, Usutu virus, GB virus C*	71, 154, 160 – 162
	(–) ssRNA	Bunyaviridae	*Toscana virus, Puumala hantavirus, Dobrava hantavirus*	163, 164
		Paramyxoviridae	*Measles virus*	72
	Retroviruses	Retroviridae	*Simian foamy virus, Spumaretrovirus, Human T-lymphotropic virus 3, 4, Human endogenous retrovirus H*	47, 66 – 69
Teguments	dsDNA	Herpesviridae	HHV 7	89
		Myoviridae	ND	15
		Papillomaviridae	α-, β-, γ- and unclassified *HPVs*	15, 88
		Podoviridae	ND	15
		Polyomaviridae	MCPyV, *Human polyomavirus 6, 7, 9, Human polyomavirus 9*-like	15, 165
		Poxviridae	*Vaccinia virus*	166
		Siphoviridae	*Propionibacterium phage P100A, P100D, 100.1, 101A, P105*	15
		Unclassified phages	ND	15
	ssDNA	Anelloviridae	TTV	167
		Circoviridae	*Cyclovirus NG2*-like, *Human gyrovirus, Circovirus*-like *CB-A, RW-E*	15, 83
		Inoviridae	ND	15
		Microviridae	ND	15
		Parvoviridae	*Human PARV B19, Human PARV4*	168
	dsRNA	Not documented		
	(+) ssRNA	Not documented		
	(–) ssRNA	Not documented		
	Retroviruses	Not documented		
Genito-urinary tract	dsDNA	Adenoviridae	*Human adenovirus 11, 21, 34, 35*	106
		Herpesviridae	*Human cytomegalovirus, Herpes simplex virus 1, 2, Epstein-Barr virus,* HHV 6A/B, HHV 7, HHV 8	107 108, 113, 169
		Papillomaviridae	*HPV16, 18*	107 170
		Polyomaviridae	*BK virus, JC virus, Human polyomavirus 9*	83, 109, 143

Table 1 (continued)

Source	Viral group	Viral family	Viral genera/species	Reference
Genito-urinary tract	ssDNA	Anelloviridae	TTV, TTV-midi	110, 111
	dsRNA	Not documented		
	(+) ssRNA	Not documented		
	(–) ssRNA	Not documented		
	Retroviruses	Not documented		
Nervous system	dsDNA	Herpesviridae	*Human cytomegalovirus, Herpes simplex virus 1, 2,* HHV 6A/B, HHV 8	171, 172
	ssDNA	Not documented		
	dsRNA	Not documented		
	(+) ssRNA	Flaviviridae	*Dengue virus*	173
	(–) ssRNA	Bornaviridae	BDV	102 103, 174
	Retroviruses	Not documented		

ND = No data available.

In the environment, the majority of Marseillevirus-related viruses have been isolated from aquatic and soil environments, suggesting the possibility of a common infectious route in humans [129, 137, 138]. Although they are found in non-pathological conditions, the consequences of long-term viral persistence should be further evaluated.

Conclusion

The Human-Virus Interactome Goes Beyond Simple Parasitism

Viruses and humans coexist and are constantly interacting. Historically, viruses have been classified as strict intracellular pathogens. However, with the development of new technologies for viral detection, it has become clear that their presence within the healthy human body goes beyond simple parasitism (fig. 1, 3; table 1). The role of the majority of eukaryotic viruses in the healthy human body remains unclear. Although the long-term consequences of viral presence in terms of pathological conditions should be evaluated, it is possible that such viruses may be considered commensals. In other cases, it is not a single virus that is pathogenic for humans but the coinfection with different viruses. The combination of HBoV and Adenovirus represents a good example of such coinfection [77]. The presence of viruses in the human body without any pathological context could also be beneficial for the body or for the human microbial flora. An example of symbiosis between viruses and the human host is the phage communities of the human gut, and these communities may play an important role in the control of the bacterial population. Conversely, a negative interaction (negative for humans) is that phages may represent an important reservoir for bacterial resistance genes and may contribute to bacterial pathogenicity via horizontal gene transfer [20]. As a result, the boundaries between mutualistic and pathogenic viruses remain elusive and are most likely highly dynamic throughout life [2].

The human-virus interactome should be considered as a complex web of interactions, defined by multiple factors. These factors can be classified into three categories: virus-specific (e.g. viral genotype, replication mode, host range, abundance), host-specific (e.g. genetic background, age, immune system) and environment-specific (e.g. geographic location, demographic distribution, animal proximity). In the case of the human virome under healthy conditions, the weight of each factor lays at a metastable equilibrium point, allowing viruses and humans to coexist naturally (fig. 4). A change in one of these parameters could lead to the development of disease conditions or the clearance of the virus from the body. From a medical point of view, a new paradigm is thus emerging; if we define an illness as a disruption of the normal 'healthy' virome, then the restoration of this equilibrium should be the goal of medical treatment, not the elimination of all non-human microorganisms.

Acknowledgements

This work was supported by the Starting Grant No. 242729 from the European Research Council awarded to C. Desnues.

Disclosure Statement

The authors declare that they have no conflict of interest.

References

1 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.
2 Haynes M, Rohwer F: The human virome: Metagenomics of the Human Body, 2011, pp 63–77.
3 Suttle CA: Marine viruses – major players in the global ecosystem. Nat Rev Microbiol 2007;5:801–812.
4 Fuhrman JA: Marine viruses and their biogeochemical and ecological effects. Nature 1999;399:541–548.
5 Greninger AL, Runckel C, Chiu CY, Haggerty T, Parsonnet J, Ganem D, DeRisi JL: The complete genome of klassevirus – a novel picornavirus in pediatric stool. Virol J 2009;6:82.
6 Allander T: Human bocavirus. J Clin Virol 2008;41:29–33.
7 Weinbauer MG: Ecology of prokaryotic viruses. FEMS Microbiol Rev 2004;28:127–181.
8 Stern A, Mick E, Tirosh I, Sagy O, Sorek R: CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. Genome Res 2012;22:1985–1994.
9 Marinelli LJ, Fitz-Gibbon S, Hayes C, Bowman C, Inkeles M, Loncaric A, Russell DA, Jacobs-Sera D, Cokus S, Pellegrini M, Kim J, Miller JF, Hatfull GF, Modlin RL: *Propionibacterium acnes* bacteriophages display limited genetic diversity and broad killing activity against bacterial skin isolates. mBio 2012;3:e00279-12.
10 Belshaw R, Pereira V, Katzourakis A, Talbot G, Paces J, Burt A, Tristem M: Long-term reinfection of the human genome by endogenous retroviruses. Proc Natl Acad Sci USA 2004;101:4894–4899.
11 Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC Jr., McCoy JM: Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 2000;403:785–789.
12 Blaise S, de Parseval N, Benit L, Heidmann T: Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc Natl Acad Sci USA 2003;100:13013–13018.
13 Specter S: Clinical Virology Manual, ed 2. New York, Elsevier, 1992.
14 Ratcliff RM, Chang G, Kok T, Sloots TP: Molecular diagnosis of medical viruses. Curr Issues Mol Biol 2007;9:87–102.
15 Foulongne V, Sauvage V, Hebert C, Dereure O, Cheval J, Gouilh MA, Pariente K, Segondy M, Burguiere A, Manuguerra JC, Caro V, Eloit M: Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing. PloS One 2012;7:e38499.
16 Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, Gordon JI: Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 2010;466:334–338.
17 Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, Rohwer F: Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol 2003;185:6220–6223.
18 Delwart EL: Viral metagenomics. Rev Med Virol 2007;17:115–131.
19 Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, Mahaffy JM, Mueller J, Nulton J, Rayhawk S, Rodriguez-Brito B, Salamon P, Rohwer F: Viral diversity and dynamics in an infant gut. Res Microbiol 2008;159:367–373.
20 Willner D, Furlan M, Schmieder R, Grasis JA, Pride DT, Relman DA, Angly FE, McDole T, Mariella RP Jr, Rohwer F, Haynes M: Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity. Proc Natl Acad Sci USA 2011;108(suppl 1):4547–4553.
21 Delwart E, Li L: Rapidly expanding genetic diversity and host range of the circoviridae viral family and other rep encoding small circular ssdna genomes. Virus Res 2012;164:114–121.
22 Li L, Kapoor A, Slikas B, Bamidele OS, Wang C, Shaukat S, Masroor MA, Wilson ML, Ndjango JB, Peeters M, Gross-Camp ND, Muller MN, Hahn BH, Wolfe ND, Triki H, Bartkus J, Zaidi SZ, Delwart E: Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. J Virol 2010;84:1674–1682.
23 Phan TG, Li L, O'Ryan MG, Cortes H, Mamani N, Bonkoungou IJ, Wang C, Leutenegger CM, Delwart E: A third gyrovirus species in human faeces. J Gen Virol 2012;93:1356–1361.
24 Li L, Shan T, Soji OB, Alam MM, Kunz TH, Zaidi SZ, Delwart E: Possible cross-species transmission of circoviruses and cycloviruses among farm animals. J Gen Virol 2011;92:768–772.
25 Vanchiere JA, Abudayyeh S, Copeland CM, Lu LB, Graham DY, Butel JS: Polyomavirus shedding in the stool of healthy adults. J Clin Microbiol 2009;47:2388–2391.
26 Gabbay YB, Luz CR, Costa IV, Cavalcante-Pepino EL, Sousa MS, Oliveira KK, Wanzeller AL, Mascarenhas JD, Leite JP, Linhares AC: Prevalence and genetic diversity of astroviruses in children with and without diarrhea in Sao Luis, Maranhao, Brazil. Mem Inst Oswaldo Cruz 2005;100:709–714.
27 Mendez-Toss M, Griffin DD, Calva J, Contreras JF, Puerto FI, Mota F, Guiscafre H, Cedillo R, Munoz O, Herrera I, Lopez S, Arias CF: Prevalence and genetic diversity of human astroviruses in Mexican children with symptomatic and asymptomatic infections. J Clin Microbiol 2004;42:151–157.
28 Barreira DM, Ferreira MS, Fumian TM, Checon R, de Sadovsky AD, Leite JP, Miagostovich MP, Spano LC: Viral load and genotypes of noroviruses in symptomatic and asymptomatic children in southeastern Brazil. J Clin Virol 2010;47:60–64.
29 Ayukekbong J, Lindh M, Nenonen N, Tah F, Nkuo-Akenji T, Bergstrom T: Enteric viruses in healthy children in Cameroon: viral load and genotyping of norovirus strains. J Med Virol 2011;83:2135–2142.
30 Himeda T, Ohara Y: Saffold virus, a novel human cardiovirus with unknown pathogenicity. J Virol 2012;86:1292–1296.
31 Kapusinszky B, Minor P, Delwart E: Nearly constant shedding of diverse enteric viruses by two healthy infants. J Clin Microbiol 2012;50:3427–3434.
32 Sadeuh-Mba SA, Bessaud M, Massenet D, Joffret ML, Endegue MC, Njouom R, Reynes JM, Rousset D, Delpeyroux F: High frequency and diversity of species C enteroviruses in Cameroon and neighboring countries. J Clin Microbiol 2013;51:759–770.
33 Rakoto-Andrianarivelo M, Guillot S, Iber J, Balanant J, Blondel B, Riquet F, Martin J, Kew O, Randriamanalina B, Razafinimpiasa L, Rousset D, Delpeyroux F: Co-circulation and evolution of polioviruses and species C enteroviruses in a district of Madagascar. PLoS Pathog 2007;3:e191.
34 Kapoor A, Victoria J, Simmonds P, Slikas E, Chieochansin T, Naeem A, Shaukat S, Sharif S, Alam MM, Angez M, Wang C, Shafer RW, Zaidi S, Delwart E: A highly prevalent and genetically diversified *Picornaviridae* genus in South Asian children. Proc Natl Acad Sci USA 2008;105:20482–20487.

35 Dai XQ, Hua XG, Shan TL, Delwart E, Zhao W: Human cosavirus infections in children in China. J Clin Virol 2010;48:228–229.
36 Li L, Victoria J, Kapoor A, Blinkova O, Wang C, Babrzadeh F, Mason CJ, Pandey P, Triki H, Bahri O, Oderinde BS, Baba MM, Bukbuk DN, Besser JM, Bartkus JM, Delwart EL: A novel picornavirus associated with gastroenteritis. J Virol 2009;83:12002–12006.
37 Shan T, Wang C, Cui L, Yu Y, Delwart E, Zhao W, Zhu C, Lan D, Dai X, Hua X: Picornavirus salivirus/klassevirus in children with diarrhea, China. Emerg Infect Dis 2010;16:1303–1305.
38 Stocker A, Souza BF, Ribeiro TC, Netto EM, Araujo LO, Correa JI, Almeida PS, de Mattos AP, Ribeiro Hda C Jr., Pedral-Sampaio DB, Drosten C, Drexler JF: Cosavirus infection in persons with and without gastroenteritis, Brazil. Emerg Infect Dis 2012;18:656–659.
39 Iturriza Gomara M, Kang G, Mammen A, Jana AK, Abraham M, Desselberger U, Brown D, Gray J: Characterization of G10P[11] rotaviruses causing acute gastroenteritis in neonates and infants in vellore, india. J Clin Microbiol 2004;42:2541–2547.
40 Ganesh B, Banyai K, Martella V, Jakab F, Masachessi G, Kobayashi N: Picobirnavirus infections: viral persistence and zoonotic potential. Rev Med Virol 2012;22:245–256.
41 Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SW, Hibberd ML, Liu ET, Rohwer F, Ruan Y: RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol 2006;4:e3.
42 Balique F, Colson P, Raoult D: Tobacco mosaic virus in cigarettes and saliva of smokers. J Clin Virol 2012;55:374–376.
43 Colson P, Richet H, Desnues C, Balique F, Moal V, Grob JJ, Berbis P, Lecoq H, Harle JR, Berland Y, Raoult D: Pepper mild mottle virus, a plant virus associated with specific immune responses, fever, abdominal pains, and pruritus in humans. PloS One 2010;5:e10041.
44 Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI: A core gut microbiome in obese and lean twins. Nature 2009;457:480–484.
45 Nishizawa T, Okamoto H, Konishi K, Yoshizawa H, Miyakawa Y, Mayumi M: A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. Biochem Biophys Res Commun 1997;241:92–97.
46 Virgin HW, Wherry EJ, Ahmed R: Redefining chronic viral infection. Cell 2009;138:30–50.
47 Breitbart M, Rohwer F: Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. BioTechniques 2005;39:729–736.
48 Biagini P, Gallian P, Cantaloube JF, De Micco P, de Lamballerie X: Presence of TT virus in French blood donors and intravenous drug users. J Hepatol 1998;29:684–685.
49 Okamura A, Yoshioka M, Kubota M, Kikuta H, Ishiko H, Kobayashi K: Detection of a novel DNA virus (TTV) sequence in peripheral blood mononuclear cells. J Med Virol 1999;58:174–177.
50 Okamoto H, Nishizawa T, Takahashi M, Asabe S, Tsuda F, Yoshikawa A: Heterogeneous distribution of TT virus of distinct genotypes in multiple tissues from infected humans. Virology 2001;288:358–368.
51 Zheng MY, Lin Y, Li DJ, Ruan HB, Chen Y, Wu TT: TTV and HPV co-infection in cervical smears of patients with cervical lesions in littoral of zhejiang province (in Chinese). Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi 2010;24:110–112.
52 Biagini P, Gallian P, Touinssi M, Cantaloube JF, Zapitelli JP, de Lamballerie X, de Micco P: High prevalence of TT virus infection in French blood donors revealed by the use of three PCR systems. Transfusion 2000;40:590–595.
53 Bagaglio S, Sitia G, Prati D, Cella D, Hasson H, Novati R, Lazzarin A, Morsica G: Mother-to-child transmission of TT virus: sequence analysis of non-coding region of TT virus in infected mother-infant pairs. Arch Virol 2002;147:803–812.
54 Jones MS, Kapoor A, Lukashov VV, Simmonds P, Hecht F, Delwart E: New DNA viruses identified in patients with acute viral infection syndrome. J Virol 2005;79:8230–8236.
55 Fryer JF, Delwart E, Hecht FM, Bernardin F, Jones MS, Shah N, Baylis SA: Frequent detection of the parvoviruses, parv4 and parv5, in plasma from blood donors and symptomatic individuals. Transfusion 2007;47:1054–1061.
56 Sharp CP, LeBreton M, Kantola K, Nana A, Diffo Jle D, Djoko CF, Tamoufe U, Kiyang JA, Babila TG, Ngole EM, Pybus OG, Delwart E, Delaporte E, Peeters M, Soderlund-Venermo M, Hedman K, Wolfe ND, Simmonds P: Widespread infection with homologues of human parvoviruses b19, parv4, and human bocavirus of chimpanzees and gorillas in the wild. J Virol 2010;84:10289–10296.
57 Longhi E, Bestetti G, Acquaviva V, Foschi A, Piolini R, Meroni L, Magni C, Antinori S, Parravicini C, Corbellino M: Human parvovirus 4 in the bone marrow of Italian patients with aids. AIDS 2007;21:1481–1483.
58 Manning A, Willey SJ, Bell JE, Simmonds P: Comparison of tissue distribution, persistence, and molecular epidemiology of parvovirus B19 and novel human parvoviruses PARV4 and human bocavirus. J Infect Dis 2007;195:1345–1352.
59 Schneider B, Fryer JF, Reber U, Fischer HP, Tolba RH, Baylis SA, Eis-Hubinger AM: Persistence of novel human parvovirus PARV4 in liver tissue of adults. J Med Virol 2008;80:345–351.
60 Isa A, Kasprowicz V, Norbeck O, Loughry A, Jeffery K, Broliden K, Klenerman P, Tolfvenstam T, Bowness P: Prolonged activation of virus-specific CD8+ T cells after acute b19 infection. PLoS Med 2005;2:e343.
61 Norja P, Hokynar K, Aaltonen LM, Chen R, Ranki A, Partio EK, Kiviluoto O, Davidkin I, Leivo T, Eis-Hubinger AM, Schneider B, Fischer HP, Tolba R, Vapalahti O, Vaheri A, Soderlund-Venermo M, Hedman K: Bioportfolio: lifelong persistence of variant and prototypic erythrovirus DNA genomes in human tissue. Proc Natl Acad Sci USA 2006;103:7450–7453.
62 Soderlund-Venermo M, Hokynar K, Nieminen J, Rautakorpi H, Hedman K: Persistence of human parvovirus B19 in human tissues. Pathol Biol 2002;50:307–316.
63 Egli A, Infanti L, Dumoulin A, Buser A, Samaridis J, Stebler C, Gosert R, Hirsch HH: Prevalence of polyomavirus BK and JC infection and replication in 400 healthy blood donors. J Infect Dis 2009;199:837–846.
64 Delbue S, Tremolada S, Elia F, Carloni C, Amico S, Tavazzi E, Marchioni E, Novati S, Maserati R, Ferrante P: Lymphotropic polyomavirus is detected in peripheral blood from immunocompromised and healthy subjects. J Clin Virol 2010;47:156–160.
65 Bonvicini F, Manaresi E, Gentilomi GA, Di Furio F, Zerbini M, Musiani M, Gallinella G: Evidence of human bocavirus viremia in healthy blood donors. Diagn Microbiol Infect Dis 2011;71:460–462.
66 Wolfe ND, Switzer WM, Carr JK, Bhullar VB, Shanmugam V, Tamoufe U, Prosser AT, Torimiro JN, Wright A, Mpoudi-Ngole E, McCutchan FE, Birx DL, Folks TM, Burke DS, Heneine W: Naturally acquired simian retrovirus infections in central African hunters. Lancet 2004;363:932–937.
67 Zheng H, Wolfe ND, Sintasath DM, Tamoufe U, Lebreton M, Djoko CF, Diffo Jle D, Pike BL, Heneine W, Switzer WM: Emergence of a novel and highly divergent HTLV-3 in a primate hunter in Cameroon. Virology 2010;401:137–145.
68 Heneine W, Switzer WM, Sandstrom P, Brown J, Vedapuri S, Schable CA, Khan AS, Lerche NW, Schweizer M, Neumann-Haefelin D, Chapman LE, Folks TM: Identification of a human population infected with simian foamy viruses. Nat Med 1998;4:403–407.
69 Lerche NW, Switzer WM, Yee JL, Shanmugam V, Rosenthal AN, Chapman LE, Folks TM, Heneine W: Evidence of infection with simian type D retrovirus in persons occupationally exposed to nonhuman primates. J Virol 2001;75:1783–1789.
70 Sandstrom PA, Phan KO, Switzer WM, Fredeking T, Chapman L, Heneine W, Folks TM: Simian foamy virus infection among zoo keepers. Lancet 2000;355:551–552.
71 Dias LL, Amarilla AA, Poloni TR, Covas DT, Aquino VH, Figueiredo LT: Detection of dengue virus in sera of Brazilian blood donors. Transfusion 2012;52:1667–1671.
72 Sonoda S, Nakayama T: Detection of measles virus genome in lymphocytes from asymptomatic healthy children. J Med Virol 2001;65:381–387.

73 Willner D, Furlan M, Haynes M, Schmieder R, Angly FE, Silva J, Tammadoni S, Nosrat B, Conrad D, Rohwer F: Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. PloS One 2009;4:e7370.
74 Wylie KM, Mihindukulasuriya KA, Sodergren E, Weinstock GM, Storch GA: Sequence analysis of the human virome in febrile and afebrile children. PloS One 2012;7:e27735.
75 Fry AM, Lu X, Chittaganpitch M, Peret T, Fischer J, Dowell SF, Anderson LJ, Erdman D, Olsen SJ: Human bocavirus: a novel parvovirus epidemiologically associated with pneumonia requiring hospitalization in Thailand. J Infect Dis 2007;195:1038–1045.
76 King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds): Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. 2011, Academic Press, London.
77 Heydari H, Mamishi S, Khotaei GT, Moradi S: Fatal type 7 adenovirus associated with human bocavirus infection in a healthy child. J Med Virol 2011;83:1762–1763.
78 Winther B, Hayden FG, Hendley JO: Picornavirus infections in children diagnosed by RT-PCR during longitudinal surveillance with weekly sampling: association with symptomatic illness and effect of season. J Med Virol 2006;78:644–650.
79 van der Zalm MM, Wilbrink B, van Ewijk BE, Overduin P, Wolfs TF, van der Ent CK: Highly frequent infections with human rhinovirus in healthy young children: a longitudinal cohort study. J Clin Virol 2011;52:317–320.
80 Annamalay AA, Khoo SK, Jacoby P, Bizzintino J, Zhang G, Chidlow G, Lee WM, Moore HC, Harnett GB, Smith DW, Gern JE, Le-Souef PN, Laing IA, Lehmann D: Prevalence of and risk factors for human rhinovirus infection in healthy aboriginal and non-aboriginal Western Australian children. Pediatr Infect Dis J 2012;31:673–679.
81 van den Bergh MR, Biesbroek G, Rossen JW, de Steenhuijsen Piters WA, Bosch AA, van Gils EJ, Wang X, Boonacker CW, Veenhoven RH, Bruin JP, Bogaert D, Sanders EA: Associations between pathogens in the upper respiratory tract of young children: interplay between viruses and bacteria. PloS One 2012;7:e47711.
82 Bousbia S, Papazian L, Saux P, Forel JM, Auffray JP, Martin C, Raoult D, La Scola B: Repertoire of intensive care unit pneumonia microbiota. PloS One 2012;7:e32486.
83 Sauvage V, Cheval J, Foulongne V, Gouilh MA, Pariente K, Manuguerra JC, Richardson J, Dereure O, Lecuit M, Burguiere A, Caro V, Eloit M: Identification of the first human gyrovirus, a virus related to chicken anemia virus. J Virol 2011;85:7948–7950.
84 Bonvicini F, La Placa M, Manaresi E, Gallinella G, Gentilomi GA, Zerbini M, Musiani M: Parvovirus B19 DNA is commonly harboured in human skin. Dermatology 2010; 220:138–142.
85 Feng H, Shuda M, Chang Y, Moore PS: Clonal integration of a polyomavirus in human merkel cell carcinoma. Science 2008;319: 1096–1100.
86 Shuda M, Feng H, Kwun HJ, Rosen ST, Gjoerup O, Moore PS, Chang Y: T antigen mutations are a human tumor-specific signature for merkel cell polyomavirus. Proc Natl Acad Sci USA 2008;105:16272–16277.
87 Schowalter RM, Pastrana DV, Pumphrey KA, Moyer AL, Buck CB: Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe 2010;7:509–515.
88 Ruer JB, Pepin L, Gheit T, Vidal C, Kantelip B, Tommasino M, Pretet JL, Mougin C, Aubin F: Detection of alpha- and beta-human papillomavirus (HPV) in cutaneous melanoma: a matched and controlled study using specific multiplex PCR combined with DNA microarray primer extension. Exp Dermatol 2009;18: 857–862.
89 Ponti R, Bergallo M, Costa C, Quaglino P, Fierro MT, Comessatti A, Stroppiana E, Sidoti F, Merlino C, Novelli M, Alotto D, Cavallo R, Bernengo MG: Human herpesvirus 7 detection by quantitative real time polymerase chain reaction in primary cutaneous T cell lymphomas and healthy subjects: lack of a pathogenic role. Br J Dermatol 2008;159: 1131–1137.
90 Frenkel N, Schirmer EC, Wyatt LS, Katsafanas G, Roffman E, Danovich RM, June CH: Isolation of a new herpesvirus from human CD4+ T cells. Proc Natl Acad Sci USA 1990; 87:748–752.
91 Nagore E, Ledesma E, Collado C, Oliver V, Perez-Perez A, Aliaga A: Detection of Epstein-Barr virus and human herpesvirus 7 and 8 genomes in primary cutaneous T and B cell lymphomas. Br J Dermatol 2000;143:320–323.
92 Steiner I, Kennedy PG: Herpes simplex virus latent infection in the nervous system. J Neurovirol 1995;1:19–29.
93 Stevens JG, Wagner EK, Devi-Rao GB, Cook ML, Feldman LT: RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. Science 1987; 235:1056–1059.
94 Knaup B, Schunemann S, Wolff MH: Subclinical reactivation of herpes simplex virus type 1 in the oral cavity. Oral Microbiol Immunol 2000;15:281–283.
95 Wald A, Zeh J, Selke S, Warren T, Ryncarz AJ, Ashley R, Krieger JN, Corey L: Reactivation of genital herpes simplex virus type 2 infection in asymptomatic seropositive persons. New Engl J Med 2000;342:844–850.
96 Briese T, Schneemann A, Lewis AJ, Park YS, Kim S, Ludwig H, Lipkin WI: Genomic organization of Borna disease virus. Proc Natl Acad Sci USA 1994;91:4362–4366.
97 Cubitt B, de la Torre JC: Borna disease virus (BDV), a nonsegmented RNA virus, replicates in the nuclei of infected cells where infectious BDV ribonucleoproteins are present. J Virol 1994;68:1371–1381.
98 Pletnikov MV, Gonzalez-Dunia D, Stitz L: Experimental infection: pathogenesis of neurobehavioral disease; in Carbone KM (ed): Borna Disease Virus and Its Role in Neurobehavioral Disease. Washington, ASM, 2002, pp 125–178.
99 Kinnunen PM, Palva A, Vaheri A, Vapalahti O: Epidemiology and host spectrum of Borna disease virus infections. J Gen Virol 2013; 94:247–262.
100 Rott R, Herzog S, Fleischer B, Winokur A, Amsterdam J, Dyson W, Koprowski H: Detection of serum antibodies to Borna disease virus in patients with psychiatric disorders. Science 1985;228:755–756.
101 Bode L, Riegel S, Ludwig H, Amsterdam JD, Lange W, Koprowski H: Borna disease virus-specific antibodies in patients with HIV infection and with mental disorders. Lancet 1988;2:689.
102 Patti AM, Vulcano A, Candelori E, Donfrancesco R, Ludwig H, Bode L: Borna disease virus infection in Italian children: a potential risk for the developing brain? APMIS Suppl 2008:70–73.
103 Kinnunen PM, Billich C, Ek-Kommonen C, Henttonen H, Kallio RK, Niemimaa J, Palva A, Staeheli P, Vaheri A, Vapalahti O: Serological evidence for Borna disease virus infection in humans, wild rodents and other vertebrates in Finland. J Clin Virol 2007;38:64–69.
104 Belyi VA, Levine AJ, Skalka AM: Unexpected inheritance: multiple integrations of ancient bornavirus and ebolavirus/marburgvirus sequences in vertebrate genomes. PLoS Pathog 2010;6:e1001030.
105 Horie M, Honda T, Suzuki Y, Kobayashi Y, Daito T, Oshida T, Ikuta K, Jern P, Gojobori T, Coffin JM, Tomonaga K: Endogenous non-retroviral RNA virus elements in mammalian genomes. Nature 2010;463:84–87.
106 Heim A, Ebnet C, Harste G, Pring-Akerblom P: Rapid and quantitative detection of human adenovirus DNA by real-time PCR. J Med Virol 2003;70:228–239.
107 Zhao Y, Cao X, Zheng Y, Tang J, Cai W, Wang H, Gao Y, Wang Y: Relationship between cervical disease and infection with human papillomavirus types 16 and 18, and herpes simplex virus 1 and 2. J Med Virol 2012;84:1920–1927.
108 Naumenko VA, Tyulenev YA, Yakovenko SA, Kurilo LF, Shileyko LV, Segal AS, Zavalishina LE, Klimova RR, Tsibizov AS, Alkhovskii SV, Kushch AA: Detection of human cytomegalovirus in motile spermatozoa and spermatogenic cells in testis organotypic culture. Herpesviridae 2011;2:7.
109 Bialasiewicz S, Whiley DM, Lambert SB, Nissen MD, Sloots TP: Detection of BK, JC, WU, or KI polyomaviruses in faecal, urine, blood, cerebrospinal fluid and respiratory samples. J Clin Virol 2009;45:249–254.

110 Burian Z, Szabo H, Szekely G, Gyurkovits K, Pankovics P, Farkas T, Reuter G: Detection and follow-up of torque teno midi virus ('small anelloviruses') in nasopharyngeal aspirates and three other human body fluids in children. Arch Virol 2011;156:1537–1541.
111 Chan PK, Tam WH, Yeo W, Cheung JL, Zhong S, Cheng AF: High carriage rate of TT virus in the cervices of pregnant women. Clin Infect Dis 2001;32:1376–1377.
112 Csoma E, Meszaros B, Asztalos L, Konya J, Gergely L: Prevalence of WU and KI polyomaviruses in plasma, urine, and respiratory samples from renal transplant patients. J Med Virol 2011;83:1275–1278.
113 Kaspersen MD, Larsen PB, Kofod-Olsen E, Fedder J, Bonde J, Hollsberg P: Human herpesvirus-6A/B binds to spermatozoa acrosome and is the most prevalent herpesvirus in semen from sperm donors. PloS One 2012;7:e48810.
114 Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM: Classification of papillomaviruses (PVS) based on 189 PV types and proposal of taxonomic amendments. Virology 2010;401:70–79.
115 Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV: The causal relation between human papillomavirus and cervical cancer. J Clin Pathol 2002;55:244–265.
116 Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N: Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999;189:12–19.
117 Klein F, Amin Kotb WF, Petersen I: Incidence of human papilloma virus in lung cancer. Lung Cancer 2009;65:13–18.
118 Ryerson AB, Peters ES, Coughlin SS, Chen VW, Gillison ML, Reichman ME, Wu X, Chaturvedi AK, Kawaoka K: Burden of potentially human papillomavirus-associated cancers of the oropharynx and oral cavity in the US, 1998–2003. Cancer 2008;113:2901–2909.
119 Hansson BG, Rosenquist K, Antonsson A, Wennerberg J, Schildt EB, Bladstrom A, Andersson G: Strong association between infection with human papillomavirus and oral and oropharyngeal squamous cell carcinoma: a population-based case-control study in southern Sweden. Acta otolaryngol 2005; 125:1337–1344.
120 Hariri S, Unger ER, Sternberg M, Dunne EF, Swan D, Patel S, Markowitz LE: Prevalence of genital human papillomavirus among females in the United States, the national health and nutrition examination survey, 2003–2006. J Infect Dis 2011;204: 566–573.
121 Kreimer AR, Clifford GM, Boyle P, Franceschi S: Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. Cancer Epidemiol Biomarkers Prev 2005;14:467–475.
122 St Guily JL, Jacquard AC, Pretet JL, Haesebaert J, Beby-Defaux A, Clavel C, Agius G, Birembaut P, Okais C, Leocmach Y, Soubeyrand B, Pradat P, Riethmuller D, Mougin C, Denis F: Human papillomavirus genotype distribution in oropharynx and oral cavity cancer in France – the EDiTH VI Study. J Clin Virol 2011;51:100–104.
123 Huibregtse JM, Scheffner M, Howley PM: A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. EMBO J 1991;10: 4129–4135.
124 Crook T, Tidy JA, Vousden KH: Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation. Cell 1991;67: 547–556.
125 Munger K, Werness BA, Dyson N, Phelps WC, Harlow E, Howley PM: Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. EMBO J 1989;8:4099–4105.
126 Giuliano AR, Lu B, Nielson CM, Flores R, Papenfuss MR, Lee JH, Abrahamsen M, Harris RB: Age-specific prevalence, incidence, and duration of human papillomavirus infections in a cohort of 290 US men. J Infect Dis 2008;198:827–835.
127 Steben M, Duarte-Franco E: Human papillomavirus infection: epidemiology and pathophysiology. Gynecol oncol 2007; 107:S2–S5.
128 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.
129 Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, Suzan-Monti M, La Scola B, Koonin EV, Raoult D: Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci USA 2009; 106:21848–21853.
130 Koonin EV, Yutin N: Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. Intervirology 2010;53:284–292.
131 Reed KD, Melski JW, Graham MB, Regnery RL, Sotir MJ, Wegner MV, Kazmierczak JJ, Stratman EJ, Li Y, Fairley JA, Swain GR, Olson VA, Sargent EK, Kehl SC, Frace MA, Kline R, Foldy SL, Davis JP, Damon IK: The detection of monkeypox in humans in the western hemisphere. New Engl J Med 2004; 350:342–350.
132 La Scola B, Marrie TJ, Auffray JP, Raoult D: Mimivirus in pneumonia patients. Emerg Infect Dis 2005;11:449–452.
133 Vincent A, La Scola B, Forel JM, Pauly V, Raoult D, Papazian L: Clinical significance of a positive serology for Mimivirus in patients presenting a suspicion of ventilator-associated pneumonia. Crit Care Med 2009; 37:111–118.
134 Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, Trape JF, Koonin EV, La Scola B, Raoult D: Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–1193.
135 Colson P, Fancello L, Gimenez G, Armougom F, Desnues C, Fournous G, Yoosuf N, Million M, La Scola B, Raoult D: Evidence of the megavirome in humans. J Clin Virol 2013;57:191–200.
136 Popgeorgiev N, Boyer M, Fancello L, Monteil S, Robert C, Rivet R, Nappez C, Azza S, Chiaroni J, Raoult D, Desnues C: Marseillevirus-like virus recovered from blood donated by asymptomatic humans. J Infect Dis 2013, E-pub ahead of print.
137 Colson P, Pagnier I, Yoosuf N, Fournous G, La Scola B, Raoult D: 'Marseilleviridae', a new family of giant viruses infecting amoebae. Arch Virol 2013;158:915–920.
138 Thomas V, Bertelli C, Collyn F, Casson N, Telenti A, Goesmann A, Croxatto A, Greub G: Lausannevirus, a giant amoebal virus encoding histone doublets. Environ Microbiol 2011;13:1454–1466.
139 Allard A, Albinsson B, Wadell G: Detection of adenoviruses in stools from healthy persons and patients with diarrhea by two-step polymerase chain reaction. J Med Virol 1992;37:149–157.
140 Kim MS, Park EJ, Roh SW, Bae JW: Diversity and abundance of single-stranded DNA viruses in human feces. Appl Environ Microbiol 2011;77:8062–8070.
141 Seifi S, Asvadi Kermani I, Dolatkhah R, Asvadi Kermani A, Sakhinia E, Asgarzadeh M, Dastgiri S, Ebrahimi A, Asghari Haggi A, Nadri M, Asvadi Kermani T: Prevalence of oral human papilloma virus in healthy individuals in east Azerbaijan province of Iran. Iran J Public Health 2013;42:79–85.
142 Korup S, Rietscher J, Calvignac-Spencer S, Trusch F, Hofmann J, Moens U, Sauer I, Voigt S, Schmuck R, Ehlers B: Identification of a novel human polyomavirus in organs of the gastrointestinal tract. PloS One 2013; 8:e58021.
143 Csoma E, Sapy T, Meszaros B, Gergely L: Novel human polyomaviruses in pregnancy: higher prevalence of BKPyV, but no WUPyV, KIPyV and HPyV9. J Clin Virol 2012;55:262–265.
144 Lin CL, Kyono W, Tongson J, Chua PK, Easa D, Yanagihara R, Nerurkar VR: Fecal excretion of a novel human circovirus, TT virus, in healthy children. Clin Diagn Lab Immunol 2000;7:960–963.
145 Pride DT, Salzman J, Haynes M, Rohwer F, Davis-Long C, White RA 3rd, Loomer P, Armitage GC, Relman DA: Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. ISME J 2012;6:915–926.

146 Bhattacharya R, Sahoo GC, Nayak MK, Saha DR, Sur D, Naik TN, Bhattacharya SK, Krishnan T: Molecular epidemiology of human picobirnaviruses among children of a slum community in Kolkata, India. Infect Genet Evol 2006;6:453–458.
147 Payne DC, Vinje J, Szilagyi PG, Edwards KM, Staat MA, Weinberg GA, Hall CB, Chappell J, Bernstein DI, Curns AT, Wikswo M, Shirley SH, Hall AJ, Lopman B, Parashar UD: Norovirus and medically attended gastroenteritis in US children. New Engl J Med 2013;368:1121–1130.
148 Edwards KM, Zhu Y, Griffin MR, Weinberg GA, Hall CB, Szilagyi PG, Staat MA, Iwane M, Prill MM, Williams JV: Burden of human metapneumovirus infection in young children. New Engl J Med 2013;368:633–643.
149 Almgren M, Atkinson R, He J, Hilding A, Hagman E, Wolk A, Thorell A, Marcus C, Naslund E, Ostenson CG, Schalling M, Lavebratt C: Adenovirus-36 is associated with obesity in children and adults in Sweden as determined by rapid ELISA. PloS One 2012; 7:e41652.
150 Goel P, Tailor P, Chande AG, Basu A, Mukhopadhyaya R: An infectious HHV-6B isolate from a healthy adult with chromosomally integrated virus and a reporter based relative viral titer assay. Virus Res 2013;173: 280–285.
151 Moens U: Silencing viral microRNA as a novel antiviral therapy? J Biomed Biotechnol 2009;2009:419539.
152 Chen AC, Keleher A, Kedda MA, Spurdle AB, McMillan NA, Antonsson A: Human papillomavirus DNA detected in peripheral blood samples from healthy Australian male blood donors. J Med Virol 2009;81:1792–1796.
153 Karimi-Rastehkenari A, Bouzari M: High frequency of SEN virus infection in thalassemic patients and healthy blood donors in Iran. Virol J 2010;7:1.
154 Bernardin F, Operskalski E, Busch M, Delwart E: Transfusion transmission of highly prevalent commensal human viruses. Transfusion 2010;50:2474–2483.
155 Afkari R, Pirouzi A, Mohsenzadeh M, Azadi M, Jafari M: Molecular detection of TT virus and SEN virus infections in hemodialysed patients and blood donors in south of Iran. Indian J Pathol Microbiol 2012;55:478–480.
156 Grossman Z, Mendelson E, Brok-Simoni F, Mileguir F, Leitner Y, Rechavi G, Ramot B: Detection of adeno-associated virus type 2 in human peripheral blood cells. J Gen Virol 1992;73:961–966.
157 Juhl D, Baylis SA, Blumel J, Gorg S, Hennig H: Seroprevalence and incidence of hepatitis E virus infection in German blood donors. Transfusion 2013, E-pub ahead of print.
158 Vollmer T, Diekmann J, Johne R, Eberhardt M, Knabbe C, Dreier J: Novel approach for detection of hepatitis E virus infection in German blood donors. J Clin Microbiol 2012;50:2708–2713.
159 Utba NM: The prevalence of hepatitis E virus in Al-Sadr city – Baghdad. Clin Lab 2013; 59:115–120.
160 Allering L, Jost H, Emmerich P, Gunther S, Lattwein E, Schmidt M, Seifried E, Sambri V, Hourfar K, Schmidt-Chanasit J: Detection of Usutu virus infection in a healthy blood donor from south-west Germany, 2012. Euro Surveill 2012;17:20341.
161 El-Zayadi AR, Abe K, Selim O, Naito H, Hess G, Ahdy A: Prevalence of GBV-C/hepatitis G virus viraemia among blood donors, health care personnel, chronic non-B non-C hepatitis, chronic hepatitis C and hemodialysis patients in Egypt. J Virol Methods 1999;80:53–58.
162 Gaibani P, Pierro A, Lunghi G, Farina C, Toschi V, Matinato C, Orlandi A, Zoccoli A, Almini D, Landini MP, Torresani E, Sambri V: Seroprevalence of West Nile virus antibodies in blood donors living in the metropolitan area of Milan, Italy, 2009–2011. New Microbiol 2013;36:81–83.
163 Gozalan A, Kalaycioglu H, Uyar Y, Sevindi DF, Turkyilmaz B, Cakir V, Cindemir C, Unal B, Yagci-Caglayik D, Korukluoglu G, Ertek M, Heyman P, Lundkvist A: Human puumala and dobrava hantavirus infections in the Black Sea region of Turkey: a cross-sectional study. Vector Borne Zoonotic Dis 2013;13:111–118.
164 Ergunay K, Aydogan S, Ilhami Ozcebe O, Cilek EE, Hacioglu S, Karakaya J, Ozkul A, Us D: Toscana virus (TOSV) exposure is confirmed in blood donors from central, north and south/southeast Anatolia, Turkey. Zoonoses and public health 2012;59: 148–154.
165 Moens U, Ludvigsen M, Van Ghelue M: Human polyomaviruses in skin diseases. Pathol Res Int 2011;2011:123491.
166 Silva-Fernandes AT, Travassos CE, Ferreira JM, Abrahao JS, Rocha ES, Viana-Ferreira F, dos Santos JR, Bonjardim CA, Ferreira PC, Kroon EG: Natural human infections with vaccinia virus during bovine vaccinia outbreaks. J Clin Virol 2009;44:308–313.
167 Osiowy C, Sauder C: Detection of TT virus in human hair and skin. Hepatol Res 2000; 16:155–162.
168 Corcioli F, Zakrzewska K, Fanci R, De Giorgi V, Innocenti M, Rotellini M, Di Lollo S, Azzi A: Human parvovirus PARV4 DNA in tissues from adult individuals: a comparison with human parvovirus B19 (B19V). Virol J 2010;7:272.
169 Cone RW, Hobson AC, Brown Z, Ashley R, Berry S, Winter C, Corey L: Frequent detection of genital herpes simplex virus DNA by polymerase chain reaction among pregnant women. JAMA 1994;272:792–796.
170 Rice PS, Mant C, Cason J, Bible JM, Muir P, Kell B, Best JM: High prevalence of human papillomavirus type 16 infection among children. J Med Virol 2000;61:70–75.
171 Kramer T, Enquist LW: Directional spread of alphaherpesviruses in the nervous system. Viruses 2013;5:678–707.
172 Chan PK, Ng HK, Hui M, Cheng AF: Prevalence and distribution of human herpesvirus 6 variants A and B in adult human brain. J Med Virol 2001;64:42–46.
173 Chastel C: Eventual role of asymptomatic cases of dengue for the introduction and spread of dengue viruses in non-endemic regions. Front Physiol 2012;3:70.
174 De La Torre JC, Gonzalez-Dunia D, Cubitt B, Mallory M, Mueller-Lantzsch N, Grasser FA, Hansen LA, Masliah E: Detection of Borna disease virus antigen and RNA in human autopsy brain samples from neuropsychiatric patients. Virology 1996;223:272–282.

Article n°7: "First molecular detection of *Rickettsia africae* in ticks from the Union of the Comoros."

First molecular detection of *Rickettsia africae* in ticks from the Union of the Comoros

Amina Yssouf[1], Cristina Socolovschi[1], Tahar Kernif[1], **Sarah Temmam**[1], Erwan Lagadec[2], Pablo Tortosa[2,3], Philippe Parola[1,a]

Published in: Intervirology. 2013;56(6):395-412. doi: 10.1159/000354561.

[1] Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1095, Aix-Marseille Université, Marseille, France.

[2] Centre de Recherche et de Veille sur les Maladies Emergentes dans l'Océan Indien (CRVOI)-Plateforme de Recherche CYROI- 2, rue Maxime Rivière, 97490 Ste Clotilde, La Reunion, France

[3] Université de La Réunion, Ste Clotilde, La Réunion, France

[a] Corresponding author: Prof. Philippe PAROLA, tel: 33 (0) 4 91 38 55 17, fax : 33 (0) 4 91 38 77 72, e-mail: philippe.parola@univ-amu.fr

Dans cet article j'ai pu collaborer avec le Dr. Yssouf de l'équipe du Pr. Parola sur la détection de bactéries *Rickettsia Africae* dans des tiques de bovins des Comores en réalisant les analyses phylogénétiques de l'étude.

Yssouf et al. Parasites & Vectors 2014, 7:444
http://www.parasitesandvectors.com/content/7/1/444

Parasites & Vectors

RESEARCH Open Access

First molecular detection of *Rickettsia africae* in ticks from the Union of the Comoros

Amina Yssouf[1], Cristina Socolovschi[1], Tahar Kernif[1], Sarah Temmam[1], Erwan Lagadec[2], Pablo Tortosa[2,3] and Philippe Parola[1*]

Abstract

Background: *Rickettsia africae* is the agent of African tick bite fever, a disease transmitted by ticks in sub-Saharan Africa. In Union of the Comoros, a recent study reported the presence of a *Rickettsia africae* vector but no information has been provided on the circulation of the pathogenic agent in this country.

Methods: To evaluate the possible circulation of *Rickettsia* spp. in Comorian cattle, genomic DNA was extracted from 512 ticks collected either in the Union of the Comoros or from animals imported from Tanzania and subsequently tested for *Rickettsia* infection by quantitative PCR.

Results: *Rickettsia africae* was detected in 90% (60/67) of *Amblyomma variegatum*, 1% (1/92) of *Rhipicephalus appendiculatus* and 2.7% (8/296) of *Rhipicephalus* (*Boophilus*) *microplus* ticks collected in the Union of the Comoros, as well as in 77.14% (27/35) of *Amblyomma variegatum* ticks collected from imported cattle. Partial sequences of both bacterial *gltA* and *ompA* genes were used in a phylogenetic analysis revealing the presence of several haplotypes, all included within the *Rickettsia africae* clade.

Conclusions: Our study reports the first evidence of *Rickettsia africae* in ticks collected from the Union of the Comoros. The data show a significant difference of infection rate of *Rickettsia africae* infected ticks between the Islands, with maximum rates measured in Grande Comore Island, sheltering the main entry port for live animal importation from Tanzania. The high infection levels reported herein indicate the need for an in-depth assessment of the burden of rickettsioses in the Union of the Comoros, especially among those at risk of infection, such as cattle herders.

Keywords: Cattle ticks, *Rickettsia africae*, *Amblyomma variegatum*, *Rhipicephalus*, Comoros

Background

Tick-borne rickettsioses are considered among the oldest known vector-borne zoonotic diseases; they are caused by obligate intracellular Gram-negative bacteria belonging to the spotted fever group (FSG) of the genus *Rickettsia* [1]. Many species of this genus are considered to be vertically transmitted symbionts of invertebrates, suggesting that the arthropod vectors act as reservoirs or amplifiers of rickettsiae in the wild [2-4]. In sub-Saharan Africa, several rickettsial strains have been isolated and detected from ticks and vertebrate animals [5], among which *Rickettsia africae*, the etiological agent of African tick-bite fever (ATBF), is the most common [3,6,7]. The main tick-vectors of *Rickettsia africae* are *Amblyomma hebraeum* in southern Africa and *Amblyomma variegatum* in West, Central and Eastern Africa, as well as in the eastern Caribbean [7-11]. In the Indian Ocean, *Rickettsia africae* has been previously detected in *Amblyomma variegatum* in La Reunion and Madagascar [4,12] but never in arthropod and human samples from other islands of the region, including the Comorian Islands.

The Union of the Comoros is composed of three volcanic islands: Grande Comore (the youngest and most elevated island), Anjouan and Moheli. The archipelago is located in the western Indian Ocean, at the northern entrance of the Mozambique Channel between Madagascar

* Correspondence: philippe.parola@univ-amu.fr
[1]Aix Marseille Université, Unité de Recherche en Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63, CNRS 7278, IRD 198, Inserm 1095, WHO Collaborative Center for Rickettsioses and Other Arthropod-borne Bacterial Diseases, Faculté de Médecine, 27 bd Jean Moulin, 13385 Marseille cedex 5, France
Full list of author information is available at the end of the article

and the East African coast, and is characterized by a warm and humid tropical climate. Cattle are imported as a food source, mainly from Madagascar. In 2000, a free trade agreement was signed between Comoros and Tanzania, facilitating the reciprocal travel and flow of cattle. Consequently, the export of cattle from Tanzania to Comoros, and particularly to Grande Comore Island, has steadily increased [13]. The likelihood of an increase in the number of pathogens introduced through cattle-associated ticks has also risen.

Recently, an entomological survey carried out on all three islands of the Union of the Comoros showed the presence of *Amblyomma variegatum*, *Rhipicephalus appendiculatus* and *Rhipicephalus microplus* on cattle [13]. Thus, the aim of this study was to detect and determine the prevalence of *Rickettsia* species that infect ticks on autochthonous cattle and on cattle imported from Tanzania and to evaluate the role of cattle importation in the introduction and of the human pathogen, *Rickettsia*, throughout the country.

Methods

Study sites and tick sampling

The present study used adult ticks that were previously collected to describe cattle tick diversity and distribution in the country. Briefly, the ticks were collected from animals on the three islands of the Union of the Comoros, including 16 of the 17 districts of the country. Adult ticks were collected on cows and goats from the three islands during the 2010 rainy season. For each district, between one and five animals were sampled and identified, and the number of collected ticks was recorded for each site of collection. Following the same protocol, ticks were also collected from cattle imported from Tanzania that were held in the quarantine enclosure located in the harbor vicinity, or in any of the three other quarantine enclosures located in the capital, Moroni (Grande Comore). All ticks were immediately stored in 70% ethanol until morphological and molecular analyses. Tick species were determined morphologically using standard identification keys [14].

DNA extraction and PCR detection of *Rickettsia*

Each tick was sliced longitudinally with a disposable scalpel, and each half was crushed in a buffered solution (G2) with proteinase K (Qiagen Hilden, Germany) and incubated at 56°C overnight. Total DNA from half of each tick was extracted in 50 μl of eluate using the EZ1 DNA Tissue kit (Qiagen, Hilden, Germany). Rickettsial DNA detection was performed by quantitative PCR using the Eurogentec MasterMix Probe PCR kit (Qiagen, Hilden, Germany) following the manufacturer's instructions with a final volume of 10 μl in each reaction as previously described [15]. Each DNA sample was tested by quantitative (q) real-time PCR using a CFX 96 Real Time System (BIO-RAD, Singapore). The presence of SFG *Rickettsia* from was determined with a Taqman probe (Eurogentec, Seraing, Germany) and RKND03F and RKND03R specific primers targeting the citrate synthase A (*glt*A) encoding gene [16,17]. Positive samples were subsequently screened with a previously described *Rickettsia africae*-specific qPCR method. Samples with fewer than 35 cycle thresholds (Ct) were considered positive [18].

In order to generate sequence data allowing phylogenetic analyses of infecting *Rickettsia*, *gltA* and *ompA* were amplified and subsequently sequenced from a subset of randomly selected *Rickettsia*-positive tick samples. A fragment of *gltA* gene was amplified using the Rp CS.409d and CS.1258n primers, previously reported to amplify a 750-bp fragment from all known *Rickettsia* species [6], and *ompA* gene was targeted by using the primers 190.70,190.180, and 190.701, amplifying a 629–632-bp fragment of SFG *Rickettsia* [2,6].

For each PCR reaction, one positive control (*Rickettsia montanensis* DNA) and 2 negative controls (sterile water containing DNA extracted from uninfected ticks maintained in laboratory colonies) were included, with the exception of the *Rickettsia africae*-specific qPCR reaction in which *Rickettsia africae* DNA was used as positive control.

The resulting PCR products were purified and directly sequenced using a BigDye Terminator Cycle Sequencing Kit (Perkin Elmer Applied Biosystems, USA) and an ABI PRISM automated sequencer (Applied Biosystems, USA), as previously described [6].

Sequence editing and phylogenetic analyses

Sequences were analyzed using Chromas Pro (version 1.49 beta Technelysium Pty Ltd, Tewantin, Australia) and compared with sequences available in the GenBank database using NCBI BLAST. Multiple sequence alignments were noted at the nucleotide and amino acid levels using ClustalW implemented in BioEdit software. The sequences matched with sequences from other *gltA* and *ompA* rickettsiae retrieved from the GenBank database. *Rickettsiaconorii* and *Rickettsia rickettsii* sequences have been chosen as root sequences. Accession numbers of the GenBank sequences used for the genetic analyses are indicated in the phylogenetic trees. The selection of the DNA substitution model that best fit the data was performed with MEGA 5.2 and was considered for phylogenetic analyses. We selected different models of nucleotide substitution using the corrected Akaike information criterion. Bayesian phylogenetic inference (BI) was carried out using Mr Bayes 3.1.2 [19] with two independent runs of four incrementally heated, Metropolis Coupled Markov chain Monte Carlo (MC) inquiries beginning with a random tree. MC queries were run for

1,000,000 generations with trees and associated model parameters sampled every 200 generations. The initial 1000 trees in each run were discarded as burn-in samples, and the harmonic mean of the likelihood was calculated by combining each of the two independent runs.

Statistical analysis

A statistical analysis was performed to compare the prevalence of *Rickettsia africae* between the islands using the Mantel-Haenszel test implemented in Epi-info version 3.5 followed by the Yates correction option. Differences were considered statistically significant for P values <0.05.

Results

PCR detection of *rickettsia*

Ticks collected on autochthonous cattle

Rickettsial DNA was detected in 14% (67/477) of ticks from the Union of the Comoros. Table 1 presents the distribution of samples positive for *Rickettsia* by tick species and location. The mean Ct ± SD value of *glt*A amplification by qPCR of positive tick samples was 28.34 ± 3.11. All positive samples were also positive for an *Rickettsia africae*-specific qPCR with a mean Ct ± SD value of 28.51 ± 2.03. On Grande Comore Island, 90% (60/67) of *Amblyomma variegatum* ticks tested positive for *Rickettsia africae*, with Ct values averaging 28.67 ± 1.95, while 1.9% (2/105) of *Rhipicephalus* spp. ticks were positive, with a Ct averaging 28.88 ± 1.8 (Figure 1). The prevalence of *Rickettsia africae* on Anjouan Island was 8.12% (13/160), and a significant difference in prevalence was observed between Anjouan and Grande Comore 32.48% (51/157) (Mantel-Haenszel test, P value < 0.001). Among the positive samples from Anjouan, 61.5% (8/13) were obtained from *Amblyomma variegatum* ticks, while the remainders were from *Rhipicephalus microplus* (Table 1). In Mohéli, one of the five regions we visited had cattle carrying infected ticks (Figure 1). The prevalence of *Rickettsia africae* on this island was 1.9% (3/160), which was also significantly different from that found on Grande Comore (Mantel-Haenszel test, P-value <0.001) and Anjouan (Mantel-Haenszel test, P-value = 0.016). All positive ticks from Mohéli were *Amblyomma variegatum* (Table 1), with positive PCRs displaying an average Ct of 25.24 ± 1.97.

The amplification, sequencing and BLAST analyses of *ompA* and *gltA* genes from positive DNA samples extracted from *Amblyomma variegatum*, *Rhipicephalus microplus* and *Rhipicephalus appendiculatus* ticks confirmed the presence of *Rickettsia africae* in 14% of our samples. The sequence analysis of *ompA* genes obtained from *Rhipicephalus* spp. samples revealed 99.0 to 99.6% nucleotide identity with *Rickettsia africae* detected from *Rhipicephalus evertsi evertsi* in Senegal (GenBank accession numbers JN043509), while *R. africae* sequenced from *Amblyomma variegatum* showed 98 to 99.7% average identity with *R. africae* detected in Ethiopia from the same tick species (GenBank accession numbers CP001612). The sequencing of *gltA* genes obtained from *Amblyomma variegatum* and *Rhipicephalus* spp. showed 99.1 to 99.42% average identity with the published sequences of *Rhipicephalus africae* that were amplified from *Amblyomma variegatum* collected in the West Indies (GenBank accession number HM050288).

Ticks collected from imported cattle

Of the ticks collected from cattle imported from Tanzania, 77.14% (27/35) were positive for *Rickettsia* spp., with a Ct average of 27.75 ± 4.13. All of the samples detected positive for *Rickettsia* spp. were also positive for *Rickettsia africae* by qPCR (Figure 1), with Ct averages of 26.7 ± 4.67.

The sequencing of *ompA* genes obtained from *Amblyomma variegatum* collected from imported cattle showed an average of 98% alignment with published sequences of *Rickettsia africae* from *Amblyomma variegatum* collected in Antigua (GenBank accession number EU622980).

The sequence analysis of the *gltA* gene obtained from *Amblyomma variegatum* ticks revealed 98.0 to 99.0% nucleotide identity with the *gltA* gene from *Rickettsia africae* detected in *Amblyomma variegatum* from Ethiopia (GenBank accession numbers CP001612) and in *Rhipicephalus evertsi evertsi* from Senegal (GenBank accession numbers HM050288), respectively.

Phylogenetic analysis

The best DNA substitution model fitting the data was determined to be HKY for both the *gltA* and *ompA* sequences. The Bayesian Inference tree based on *gltA* sequences

Table 1 Prevalence of *Rickettsia africae* by species and regions

Species Regions	*Amblyomma variegatum*	*Rhipicephalus microplus*	*Rhipicephalus appendiculatus*	Number of ticks tested by region
Grande Comore	94.23% (49/52)	8% (1/13)	1% (1/92)	157
Anjouan	89% (8/9)	3.3% (5/151)	-	160
Moheli	7% (3/28)	0% (0/132)	-	160
Ticks collected from imported cattle	77.14% (27/35)	-	-	35
Total from the Union of the Comoros	65.17% (58/89)	2.7% (8/296)	1% (1/95)	477

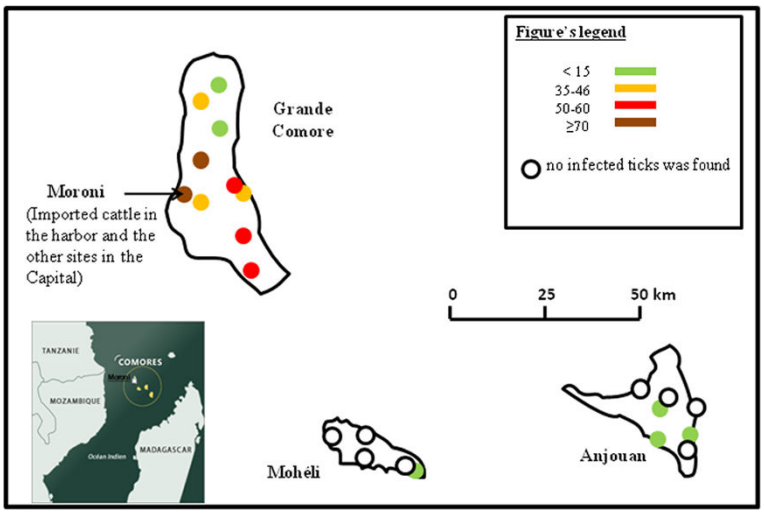


Figure 1 Prevalence of infected ticks by island and regions.

(Figure 2A) showed a distinct phylogenetic Comorian clade (posterior probability > 0.968) in which one well-supported group of sequences (posterior probability > 0.999) included sequences from *Rickettsia africae* infecting ticks sampled in Grande Comore and Anjouan but not Mohéli, and a *Rickettsia africae* sequence obtained from a tick imported from Tanzania. The Bayesian Inference tree based on *ompA* sequences (Figure 2B) showed a similar tree topology, which consolidate our analysis.

Consensus sequences of *Rickettsia africae* infecting ticks sampled on Grande Comore, Mohéli, Anjouan and Tanzania were generated and aligned to determine nucleotide mutations specific to a geographic origin. No specific mutation was observed for the *ompA* genes (data not shown); interestingly, the *gltA* sequences originating from Mohéli have 5 specific non-synonymous mutations (Table 2). The same analysis was conducted to construct consensus sequences of *Rickettsia africae* isolated from

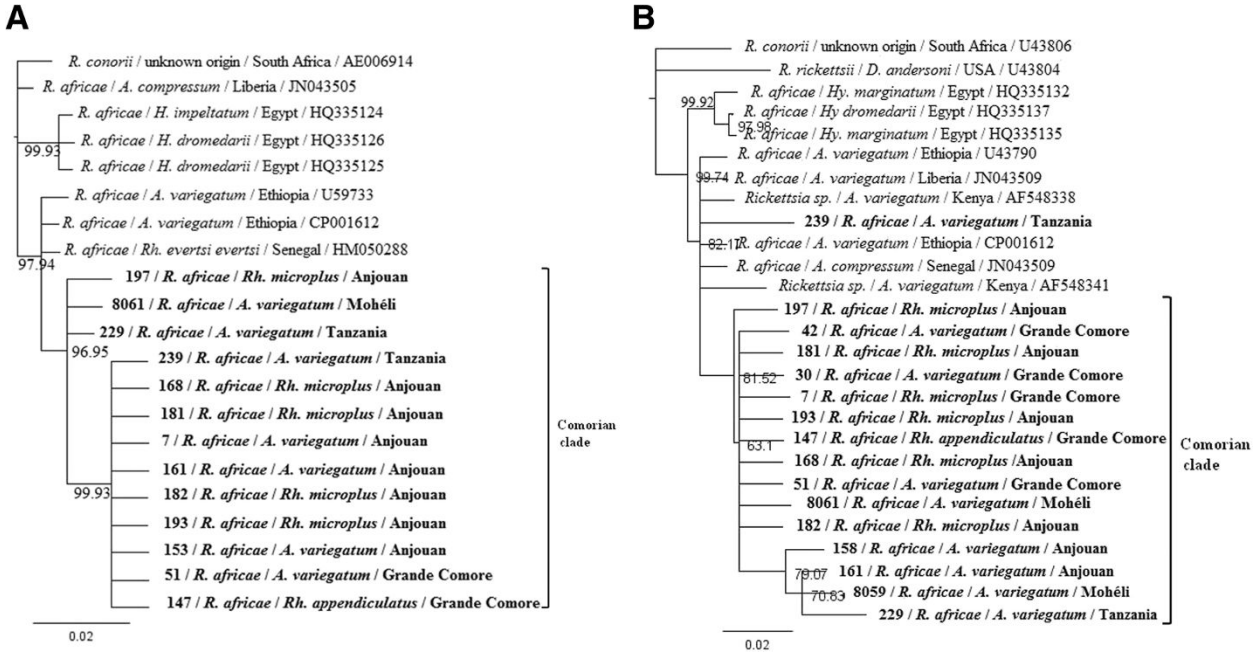


Figure 2 Phylogenetic tree inferred from the comparison of *Rickettsia africae* strains from specimens tested and selected GenBank *Rickettsia africae* sequences. Posterior probabilities are expressed in percentages and indicated at branch nodes. The sequences generated in this study are highlighted in bold. **A**: Phylogenetic tree based on the *gltA* citrate synthase-encoding gene. **B**: Phylogenetic tree based on the *ompA* outer membrane protein A gene. Abbreviation: *Rh = Rhipicephalus*; *Hy = Hyalomma*; *D = Dermacentor.*

Table 2 Non-synonymous mutations in *Rickettsia africae gltA* Mohéli sequences

Nucleotide position (U59733)	560	566	635	692	1067
Codon sequence	GAT → AAT	CCG → TCG	CCA → GCA	AAT → TAT	AGA → GGA
Amino-acid mutation in Mohéli *gltA* sequence	D → N	P → S	P → A	N → Y	R → G

distinct tick species, but no specific mutation was observed in either the *ompA* or *gltA* gene (data not shown).

Discussion

A. variegatum is the main vector of *Rickettsia africae*, a spotted fever group (SFG) *Rickettsia* bacterium in sub-Saharan Africa [9,20,21], though it is also considered a competent vector for other human and animal pathogens [22], including the highly virulent Crimean-Congo hemorrhagic fever virus [23]. In the present study, we show for the first time the presence of *Rickettsia africae* in cattle ticks collected from the Union of the Comoros. The confirmation of the presence of this *Rickettsia* spp. provides background for further epidemiologic and clinical investigations of tick-borne diseases in the Union of the Comoros. Indeed, other than *Theileria parva*, a parasitic protozoan that is the causative agent of the East Coast fever in cattle and that was previously detected in *Rhipicephalus appendiculatus* [24], no other tick-borne pathogens have been detected in this country.

We provide evidence for *Rickettsia africae* infection in *Amblyomma variegatum* ticks from all three islands of the Union of the Comoros. Although the *Amblyomma variegatum* tick infection rate varied among the islands as observed in Grande Comore and Anjouan, further study needs to confirm this result. The presence of *Rickettsia africae*-infected ticks in Grande Comore and Anjouan is congruent with the geographic distribution of *Amblyomma variegatum* [8]. The high prevalence of *Rickettsia africae* (90%) in *Amblyomma variegatum* collected in all study sites of Grande Comore showed endemicity of this bacterium and that this tick species is the reservoir of *R. africae* in the Archipelago. Ticks of the genus *Amblyomma* are considered to be the main vectors for *Rickettsia africae* although this bacterium has recently been found infecting other genera, including *Rhipicephalus* [21,22,25]. In this study, we also found *Rhipicephalus* spp. ticks carrying *Rickettsia africae* DNA although the presence rate was substantially lower (2%) than the rate measured in *Amblyomma variegatum* (65.17%). However, our data do not provide direct evidence for the vector competence of the *Rhipicephalus* (*Boophilus*) *microplus* genus for *Rickettsia africae* because *Rhipicephalus*-positive ticks were always collected on animals infested with *A. variegatum* ticks that tested positive for *R. africae.* Thus, we can hypothesize that *Rhipicephalus* ticks that were positive for *Rickettsia africae* acquired the pathogen during a blood meal, but this does not prove that these species of ticks are competent vectors for *Rickettsia africae*. Further experimental evaluations of vector competence clearly need to be carried out in order to establish the vector competence of *Rhipicephalus* ticks.

Ticks sampled on cattle imported from Tanzania showed a prevalence of 77.14% for *Rickettsia africae*, which is coherent with serological studies previously carried out in Tanzania that also showed high seroprevalence levels [26]. *Rickettsia africae* polymorphism did not show any island or tick host species structuration. Comorian haplotypes were closely related with African haplotypes, thus strengthening the hypothesis of an African origin for *Rickettsia africae* in the Comoros archipelago. Phylogenetic trees could not determine the origin of *Rickettsia africae* infection in the ticks imported from Tanzanian cattle. However, the polymorphism of *omp*A and *glt*A is notoriously weak and sequencing of additional markers may help in resolving this issue. Citrate synthase is a component of nearly all living cells and is one enzyme of the citric acid cycle, a key metabolic pathway that plays a key role in energy production [27]. The *OmpA* gene plays a role in a protective immune response and is considered as a good candidate for phylogenetic analysis for most of the SFG *Rickettsiae* [2]. These gene sequences showed no clear differences between east African and Comorian bacterial haplotypes, thus supporting an African/Comorian *Rickettsia africae* metapopulation, although the use of additional markers, or the full sequencing of bacterial isolates would obviously provide more robust information.

Conclusion

The PCR assays and sequence analyses provide new information on the epidemiology of ticks infected with *Rickettsia africae* in the Union of the Comoros. The detection of *Rickettsia africae* in ticks collected from live cattle imported from Tanzania represents a risk to local farms. Our results strengthen the need for an evaluation of the burden of ATBF in the human populations in contact with cattle. Together with the previous investigation reporting the introduction of a tick species of veterinary importance new to the country, the present study clearly shows that because of the insularity, ticks and parasites introduction into the country represents a risk of biological invasion. Therefore, increased vigilance is required to limit this risk [28].

Competing interests

The authors declare that they have no competing interests.

Authors' contributions
AY performed the experiments and drafted the manuscript with the input from all co-authors, CS provided constant intellectual support in the course of the experiments, TK participated in the experiments. ST performed the phylogenetic analysis, EL provided the biological material. PT conceived the study. PP conceived the study and gave approval of the final version.
All authors read and approved the final version of the manuscript.

Acknowledgements
We wish to thank Frederic Stachurski for his contribution in the achievement of the prevalence distribution map of *Rickettsia africae* in the Union of the Comoros.

Author details
[1]Aix Marseille Université, Unité de Recherche en Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63, CNRS 7278, IRD 198, Inserm 1095, WHO Collaborative Center for Rickettsioses and Other Arthropod-borne Bacterial Diseases, Faculté de Médecine, 27 bd Jean Moulin, 13385 Marseille cedex 5, France. [2]Centre de Recherche et de Veille sur les Maladies Emergentes dans l'Océan Indien (CRVOI)-Plateforme de Recherche CYROI- 2, rue Maxime Rivière, 97490 Ste Clotilde, La Reunion, France. [3]Université de La Réunion, Ste Clotilde, La Réunion, France.

Received: 4 March 2014 Accepted: 9 September 2014
Published: 22 September 2014

References

1. Parola P, Paddock CD, Socolovschi C, Labruna MB, Mediannikov O, Kernif T, Abdad MY, Stenos J, Bitam I, Fournier PE, Raoult D: **Update on tick-borne rickettsioses around the world: a geographic approach.** *Clin Microbiol Rev* 2013, **26:**657–702.
2. Fournier PE, Roux V, Raoult D: **Phylogenetic analysis of spotted fever group rickettsiae by study of the outer surface protein rOmpA.** *Int J Syst Bacteriol* 1998, **48**(Pt 3):839–849.
3. Parola P, Vestris G, Martinez D, Brochier B, Roux V, Raoult D: **Tick-borne rickettiosis in Guadeloupe, the French West Indies: isolation of *Rickettsia africae* from *Amblyomma variegatum* ticks and serosurvey in humans, cattle, and goats.** *Am J Trop Med Hyg* 1999, **60:**888–893.
4. Parola P, Barre N: ***Rickettsia africae*, the agent of African tick-bite fever: an emerging pathogen in the West Indies and reunion island (Indian ocean).** *Bull Soc Pathol Exot* 2004, **97:**193–198.
5. Parola P: **Rickettsioses in sub-Saharan Africa.** *Ann N Y Acad Sci* 2006, **1078:**42–47.
6. Mura A, Socolovschi C, Ginesta J, Lafrance B, Magnan S, Rolain JM, Davoust B, Raoult D, Parola P: **Molecular detection of spotted fever group rickettsiae in ticks from Ethiopia and Chad.** *Trans R Soc Trop Med Hyg* 2008, **102:**945–949.
7. Ndip LM, Fokam EB, Bouyer DH, Ndip RN, Titanji VP, Walker DH, McBride JW: **Detection of *Rickettsia africae* in patients and ticks along the coastal region of Cameroon.** *Am J Trop Med Hyg* 2004, **71:**363–366.
8. Eldin C, Mediannikov O, Davoust B, Cabre O, Barre N, Raoult D, Parola P: **Emergence of *Rickettsia africae*, Oceania.** *Emerg Infect Dis* 2011, **17:**100–102.
9. Kelly PJ: ***Rickettsia africae* in the West Indies.** *Emerg Infect Dis* 2006, **12:**224–226.
10. Jensenius M, Fournier PE, Kelly P, Myrvang B, Raoult D: **African tick bite fever.** *Lancet Infect Dis* 2003, **3:**557–564.
11. Barre N, Garris G, Camus E: **Propagation of the tick *Amblyomma variegatum* in the Caribbean.** *Rev Sci Tech* 1995, **14:**841–855.
12. Fournier PE, El Karkouri K, Leroy Q, Robert C, Giumelli B, Renesto P, Socolovschi C, Parola P, Audic S, Raoult D: **Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in rickettsia species may be explained by genome reduction.** *BMC Genomics* 2009, **10:**166.
13. Yssouf A, Lagadec E, Bakari A, Foray C, Stachurski F, Cardinale E, Plantard O, Tortosa P: **Colonization of Grande Comore Island by a lineage of *Rhipicephalus appendiculatus* ticks.** *Parasit Vectors* 2011, **4:**38.
14. Walker AR, Bouattour A, Camicas J, Estrada-Pena A, Horak IG, Pegram RG, Preston PM: *Ticks of Domestic Animals in Africa: A Guide to Identification of Species*, Bioscience Reports. Edinburgh, United Kingdom: 2003.
15. Socolovschi C, Pages F, Raoult D: **Rickettsia felis in *Aedes albopictus* mosquitoes, Libreville, Gabon.** *Emerg Infect Dis* 2012, **18:**1687–1689.
16. Bechah Y, Socolovschi C, Raoult D: **Identification of rickettsial infections by using cutaneous swab specimens and PCR.** *Emerg Infect Dis* 2011, **17:**83–86.
17. Socolovschi C, Gomez J, Marie JL, Davoust B, Guigal PM, Raoult D, Parola P: ***Ehrlichia canis* in *Rhipicephalus sanguineus* ticks in the ivory coast.** *Ticks Tick Borne Dis* 2012, **3:**411–413.
18. Kernif T, Djerbouh A, Mediannikov O, Ayach B, Rolain JM, Raoult D, Parola P, Bitam I: ***Rickettsia africae* in *Hyalomma dromedarii* ticks from sub-Saharan Algeria.** *Ticks Tick Borne Dis* 2012, **3:**377–379.
19. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, **19:**1572–1574.
20. Lynen G, Zeman P, Bakuname C, Di Giulio G, Mtui P, Sanka P, Jongejan F: **Cattle ticks of the genera *Rhipicephalus* and *amblyomma* of economic importance in Tanzania: distribution assessed with GIS based on an extensive field survey.** *Exp Appl Acarol* 2007, **43:**303–319.
21. Macaluso KR, Davis J, Alam U, Korman A, Rutherford JS, Rosenberg R, Azad AF: **Spotted fever group rickettsiae in ticks from the Masai Mara region of Kenya.** *Am J Trop Med Hyg* 2003, **68:**551–553.
22. Mediannikov O, Diatta G, Zolia Y, Balde MC, Kohar H, Trape JF, Raoult D: **Tick-borne rickettsiae in Guinea and Liberia.** *Ticks Tick Borne Dis* 2012, **3:**43–48.
23. Camicas JL, Cornet JP, Gonzalez JP, Wilson ML, Adam F, Zeller HG: **Crimean-Congo hemorrhagic fever in Senegal. Latest data on the ecology of the CCHF virus.** *Bull Soc Pathol Exot* 1994, **87:**11–16.
24. De Deken R, Martin V, Saido A, Madder M, Brandt J, Geysen D: **An outbreak of east coast fever on the Comoros: a consequence of the import of immunised cattle from Tanzania?** *Vet Parasitol* 2007, **143:**245–253.
25. Portillo A, Perez-Martinez L, Santibanez S, Blanco JR, Ibarra V, Oteo JA: **Detection of *Rickettsia africae* in *Rhipicephalus* (*Boophilus*) *decoloratus* ticks from the Republic of Botswana, south Africa.** *Am J Trop Med Hyg* 2007, **77:**376–377.
26. Prabhu M, Nicholson WL, Roche AJ, Kersh GJ, Fitzpatrick KA, Oliver LD, Massung RF, Morrissey AB, Bartlett JA, Onyango JJ, Maro VP, Kinabo GD, Saganda W, Crump JA: **Q fever, spotted fever group, and typhus group rickettsioses among hospitalized febrile patients in northern Tanzania.** *Clin Infect Dis* 2011, **53:**e8–e15.
27. Roux V, Rydkina E, Eremeeva M, Raoult D: **Citrate synthase gene comparison, a new tool for phylogenetic analysis, and its application for the rickettsiae.** *Int J Syst Bacteriol* 1997, **47:**252–261.
28. Tortosa P, Pascalis H, Guernier V, Cardinale E, Le Corre M, Goodman SM, Dellagi K: **Deciphering arboviral emergence within insular ecosystems.** *Infect Genet Evol* 2012, **12:**1333–1339.

doi:10.1186/1756-3305-7-444
Cite this article as: Yssouf *et al.*: First molecular detection of *Rickettsia africae* in ticks from the Union of the Comoros. *Parasites & Vectors* 2014 **7**:444.