THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Ecologie, Evolution, Ressources Génétiques, Paléobiologie

École doctorale GAIA - Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Unité Mixte de recherche UMR 117 ASTRE - Animal, Santé, Territoires, Risques et Ecosystèmes

Mécanismes adaptatifs facilitant l'alternance d'hôte chez le virus de la fièvre catarrhale ovine

Présentée par Yannis MOREAU

Le 3 février 2021

Sous la direction de Jérémy Bouyer et l'encadrement de Serafin GUTIERREZ et Etienne LOIRE

Devant le jury composé de

Frédérick ARNAUD, DR, INRAE, EPHE, Lyon	Rapporteur
Nadia NAFFAKH, DR, CNRS, Institut Pasteur, Paris	Rapporteur
Benoît MOURY, DR, INRAE Avignon	Examinateur
Yannis MICHALAKIS, DR, CNRS, IRD, Montpellier	Examinateur
Emmanuel BREARD, CR, ANSES, Maison d'Alfort	Examinateur
Serafin GUTIERREZ, DR, CIRAD, Montpellier	Co-encadrant

UNIVERSITÉ DE MONTPELLIER

Montpellier SupAgro

UNIVERSITÉ DE MONTPELLIER

École doctorale : GAIA

THÈSE

Pour obtenir le grade de

Docteur de l'Université Montpellier II

Discipline : Écologie, Évolution, Ressources Génétique, Paléobiologie

Présentée et soutenue par

Yannis Moreau

Mécanismes adaptatifs facilitant l'alternance d'hôte chez le virus de la fièvre catarrhale ovine

Thèse dirigée par Jérémy Bouyer, CIRAD et co-encadrée par Serafin Gutierrez, CIRAD et Etienne Loire, CIRAD

Soutenue le 03/02/2021

Devant le jury composé de

Frédérick ARNAUD, DR, INRAE, EPHE, Lyon	**Rapporteur**
Nadia NAFFAKH, DR, CNRS, Institut Pasteur, Paris	**Rapporteur**
Benoît MOURY, DR, INRAE Avignon	**Examinateur**
Yannis MICHALAKIS, DR, CNRS, IRD, Montpellier	**Examinateur**
Emmanuel BREARD, CR, ANSES, Maison d'Alfort	**Examinateur**
Serafin GUTIERREZ, CIRAD, Montpellier	**Co-encadrant**

Résumé

Les arbovirus forment un groupe polyphylétique d'espèces virales. Elles partagent la capacité d'infecter des arthropodes pour être transmise à différentes espèces d'hôte englobant les règnes animaux et végétaux. La convergence de virus aux architectures structurelles et génomiques si différentes vers ce mode de vie implique probablement des adaptations qui favorisent le maintien de ce cycle. Parmi elles, La variation du nombre de copie de gènes (VNCG) est un mécanisme candidat qui a récemment été mise en évidence chez un *nanovirus* de plante. Chez ce virus multipartite - dont les segments génomiques sont encapsidés individuellement - les fréquences de chacun de ces segments convergent vers une valeur spécifique dépendante de l'espèce hôte infectée.

Nous questionnons ici le potentiel de la VNCG à jouer un rôle dans l'adaptation d'un arbovirus à génome segmenté (i.e. les segments sont séparés mais encapsidé au sein d'un même particule virale) : le virus de la fièvre catarrhale ovine (FCO). Le virus de la FCO est un virus de la famille des *Reoviridae* dont le génome se divise en dix segments. Nous avons développé une approche de RT-qPCR pour quantifier chaque segment génomique. Cette approche a été utilisée pour analyser des populations du virus de la FCO de sérotype 4 (FCO-4) issues d'un évènement épizootique naturel. Nous avons étudié les variations de fréquence de chacun de ses dix segments chez deux espèces de ruminant et chez le diptère *Culicoides imicola*, un des vecteurs principaux du virus de la FCO. Ces résultats présentent pour la première fois l'existence du phénomène de VNCG dans des populations naturelles d'un virus animal. De plus, des profils de fréquences de segments (formule génomique) dépendent de l'hôte infecté : ils sont similaires entre les hôtes ruminants et diffèrent entre ruminants et diptère. Ces résultats soutiennent un rôle potentiel de la VNCG dans l'adaptation à l'alternance d'hôte chez le virus de la FCO. Une approche d'infection expérimentale sur *C. imicola* nous a ensuite permit de démontrer que la formule génomique change lors de l'infection d'un nouvel hôte. Nous avons mis en évidence que les populations virales ayant une dissémination complète dans l'insecte ont une formule génomique différente de celles à dissémination incomplète. Ce résultat suggère un possible lien entre la VNCG et la valeur sélective de la population virale. Nos expérimentations *in vitro* montrent également que des formules génomiques spécifiques à un environnement donné peuvent être générées du premier cycle de réplication. Ce résultat suggère que la génération de la formule génomique implique des mécanismes déterministes encore inconnus.

La VNCG constitue donc un phénomène potentiellement associé à l'adaptation au cycle en alternance du virus de la FCO-4. Son rôle plus large dans l'alternance d'hôte chez le virus de la FCO, voire chez d'autres arbovirus segmentés, mériterait d'être exploré.

Mots-clés : virus de la fièvre catarrhale ovine, variation du nombre de copies de gènes, arbovirus, alternance d'hôtes.

Discipline : Écologie, Évolution, Ressources Génétique, Paléobiologie
Cirad, UMR ASTRE - Animal, Santé, Territoires, Risques et Ecosystèmes
Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

Summary

Adaptive mechanisms facilitating host alternation in Bluetongue virus

Arboviruses form a polyphyletic group of viral species. They share the ability to infect arthropods for transmission to different host species encompassing the animal and plant kingdoms. The convergence of viruses with such different structural and genomic architectures towards this way of life probably implies adaptations that favor its maintenance. Among them, variation in gene copy number (GCNV) is a candidate mechanism recently identified in a plant *nanovirus*. In this multipartite virus - whose genomic segments are individually packaged - the frequencies of each of these segments converge towards a specific value depending on the infected host species.

We question here the potential of GCNV to play a role in the adaptation of an arbovirus with a segmented genome (i.e. the segments are separated but packaged within the same viral particle): the bluetongue virus (BTV). BTV is a virus of the Reoviridae family, whose genome is divided into ten segments. We have developed an RT-qPCR approach quantifying each genomic segment. This approach has been used to analyze populations of the serotype 4 BTV (BTV-4) resulting from a natural epizootic event. We studied the frequency variations of each of its ten segments in two ruminant species and in the dipteran midge Culicoides imicola, one of the main vectors of the BTV. These results present for the first time the existence of the phenomenon of GCNV in natural populations of an animal virus. In addition, segment frequency profiles (genomic formula) depend on the infected host: they are similar between ruminant hosts and differ between ruminant and diptera. These results support a potential role of GCNV in adaptation to host alternation in BTV. An experimental infection approach with *C. imicola* then allowed us to demonstrate that the genomic formula changes when a new host is infected. We showed that virus populations with complete dissemination in the insect have a different genomic formula than those with incomplete dissemination. This result suggests a possible link between GCNV and the selective value of the virus population. Our in vitro experiments also show that genomic formulas specific to a given environment can be generated from the first replication cycle. This result suggests that the generation of the genomic formula involves yet unknown deterministic mechanisms.

GCNV is therefore a phenomenon potentially associated with adaptation to the alternating cycle of the BTV-4 virus. Its broader role in host alternation in BTV virus, and even in other segmented arboviruses, is worth exploring.

Key-words: Bluetongue virus, gene copy number variation, arbovirus, host alternation

Remerciements

Je remercie chaleureusement les membres de mon jury d'avoir accepté d'évaluer mon travail de thèse. Merci à Frédérick Arnaud et Nadia Naffakh d'avoir accepté d'être rapporteur pour ce manuscrit. Je remercie également Benoît Moury, Emmanuel Bréard et Yannis Michalakis (joli prénom…) d'être examinateur pour ma soutenance.

Je remercie les financeurs, qui ont permis la réalisation de ce projet et m'ont soutenu dans mes dépenses excessives en réactifs de qPCR, les projets CULIOME et PALE-Blu. Merci aussi au CIRAD de m'avoir accordé une prolongation de contrat pour me soutenir dans la finalisation de ces travaux mais aussi pour son soutien via l'action incitative aux doctorants qui m'a permis à deux reprises de traverser l'Océan pour présenter mes travaux à l'international.

Des remerciements particuliers à Serafin et Etienne mes encadrants qui m'ont soutenu pendant ces trois années malgré mes sautes d'humeur ponctuées d'alternance entre des épisodes dépressifs et des moments de plaisir intense. Serafin, je peux dire que j'ai appris à gérer les ascenseurs émotionnels, enfin j'écris ça un peu vite… merci pour ton soutien, ta disponibilité, ta bienveillance et aussi pour ta patience. J'ai beaucoup appris pendant ces trois années. Etienne, merci pour ton écoute, ton calme à toutes épreuves, une source d'inspiration personnelle. Je suis un peu déçu d'avoir basculé sur un manuscrit 100% qPCR, tu m'as fourni toutes les bases pour travailler à l'étude des génomes. J'espère pouvoir faire honneur à cette formation. Merci pour votre aide dans la réalisation de ce travail.

Je remercie aussi chaleureusement mon comité de suivi, avec lequel j'ai eu la chance d'interagir à plusieurs reprises au cours de ce projet, plus que de coutume grâce à ma prolongation, mais ça a été pour le meilleur. Merci à Samuel Alizon pour sa bienveillance et ses conseils avisés. Merci à Stéphane Blanc pour ses critiques pertinentes qui ont contribué à améliorer ce projet (l'amplification indépendante de l'action des primers). Enfin merci à Thierry Baldet pour sa bonne humeur légendaire et son soutien au cours de ma thèse. Merci aussi à Jérémy Bouyer, d'avoir accepté de

prendre le titre de directeur pour mes premières années de thèse et pour sa participation.

Je tiens à dire un grand merci à l'équipe Virologie de l'ANSES de Maison d'Alfort et à Emmanuel en particulier. Merci pour votre accueil et pour la formation à l'art subtil de la production de virus de la FCO à fort titre, merci d'avoir partagé vos connaissances. Dans le cadre de notre collaboration, vous m'avez offert la possibilité de travailler à partir de vos échantillons, qui font de ce projet ce qu'il est.

Merci aussi à nos collaborateurs de l'IZS de Teramo avec qui j'ai pu interagir au début de mon projet et qui ont partagé avec nous les échantillons qui ont permis la réalisation des travaux présenté au cours de mon deuxième chapitre. Merci aussi pour vos retours constructifs sur mes travaux.

Un remerciement spécial à mon unité d'accueil, ASTRE, et tous les membres qui la composent. Merci à Nathalie Vachiery et Thierry Lefrançois de m'avoir accueilli au sein de l'unité. Merci aussi à Renata qui assure dans la gestion du collectif Microbio. J'ai apprécié l'ensemble des échanges au cours de ces trois années, je ne peux pas tous vous citer mais ça a été un grand plaisir de partager ces moments avec vous.

J'ai eu la chance de travailler dans un groupe dynamique à réunion hebdomadaire. A ce titre, je remercie Antoni, Patricia, Virginie et Serafin pour ces interactions qui, quoi que l'on en pense, font de nous un groupe à part. Patricia, tu as été d'une grande aide, je n'aurais pas pu finir tout ça sans ton aide au labo. J'ai appris à craindre ta fureur et à apprécier ta douceur, merci. Virginie merci pour ta gentillesse et tes conseils.

Antoni, malgré ton goût prononcé pour les ailerons arrière sur les voitures, je suis vraiment heureux d'avoir fait ta connaissance et tu as été un soutien énorme pendant ces trois années. Merci aux doctorants avec qui j'ai eu la chance de partager le bureau, Jeanine et Bernardin, malgré vos prénoms difficiles à porter vous avez su braver les épreuves de la vie, vous avez rendu cette expérience mémorable. Presque dans le même bureau, Cécile, (tu as passé beaucoup de temps sur le canapé donc c'est tout comme) merci pour ta bonne humeur, désolé pour tes lunettes… Aussi à Roger le bien né, notre ami Suisse, qui apparait les soirs de pleine lune au pied des arcs en ciel, Max la menace mais pas trop quand même, à Rémi mon compatriote tourangeau… merci à tous pour les rires et tous les bons moments.

Merci à toute ma famille pour leur soutien permanant, et spécialement à mes parents et mes deux sœurettes. Les amis de Tours et d'ailleurs, merci pour le soutien ou son absence, j'avais besoin de ça aussi. Spécial merci à Alexandre qui a eu le plaisir d'écouter mes complaintes régulières.

Je garde le meilleur pour la fin. Un immense merci à Marine qui m'a suivi à Montpellier, tu m'as soutenu et fait confiance pendant ces trois années. En espérant que ça continue indéfiniment.

C'est presque fini mais je vais remercier une seconde fois Marine et Antoine qui, à l'heure tardive où j'écris ces mots sont en train de gérer tous les bugs de mise en page qui vont de pair avec la remise en temps et en heure du présent manuscrit. Vous êtes le ciment qui fait tenir les briques de ce manuscrit, il commence à être tard.

Valorisations scientifiques

Publications de rang A

Moreau Y, Gil P, Exbrayat A, Rakotoarivony I, Breard E, Sailleau C, Viarouge C, Zientara S, Savini G, Goffredo M, Mancini G, Loire E, Gutierrez S. (2020) **The genome segments of Bluetongue virus differ in copy number in a host-specific manner.** Journal of Virology, Volume 95, Issue 1. https://doi.org/10.1128/JVI .01834-20

Articles publiés hors du contexte de la thèse

Loiseau V, Herniou E A, **Moreau Y**, Lévêque N, Meignin C, Daeffler L, Federici B, Cordaux R, Gilbert C. (2020) **Wide spectrum and high frequency of genomic structural variation, including transposable elements, in large double-stranded DNA viruses**. *Virus Evolution*, Volume 6, Issue 1. https://doi.org/10.1093/ve/vez060

Communication orale

Moreau Y, Gil P, Exbrayat A, Rakotoarivony I, Breard E, Sailleau C, Viarouge C, Zientara S, Savini G, Goffredo M, Mancini G, Loire E, Gutierrez S. **Variations of genes copy number in populations of Bluetongue-virus**. American Society of Virology (ASV), Minesota, Juillet 2019.

Poster

Moreau Y, Gil P, Exbrayat A, Rakotoarivony I, Breard E, Sailleau C, Viarouge C, Zientara S, Savini G, Goffredo M, Mancini G, Loire E, Gutierrez S. **Variations of genes copy number in populations of Bluetongue-virus. Congrès Evolution, Montpellier Août 2018.**

Moreau Y, Gil P, Loire E, Gutierrez S. **Spatial, temporal and host-induced genetic variation in Alphamesonivirus 1 species**. American Society of Virology (ASV), Maryland, Juillet 2018.

Table des matières

Résumé i
Summary ii
Remerciements iii
Valorisations scientifiques vi
Table des illustrations : Figures ix
Table des illustrations : Tableaux xiii
Liste des abréviations xv
INTRODUCTION GENERALE 1
Partie 1 : « Les virus d'origine arthropode » 1
1. Concepts et contexte 1
2. Le cycle complexe en alternance d'hôte 3
3. L'adaptation au cycle en alternance d'hôte 5
Partie 2 : Le virus de la fièvre catarrhale ovine 10
1. Présentation générale 10
2. Biologie du virus de la FCO 11
3. Structure du virus et cycle de réplication 15
4. Génétique évolutive du virus de la FCO 20
Partie 3 : La variation du nombre de copies de gènes (VNCG) chez les virus 27
1. Le concept de VNCG et le monde des virus multipartites 27
2. Régulation de l'expression et adaptations dépendantes de la VNCG 29
3. Une analogie avec la variation du nombre de copies de gènes chez un virus à génome segmenté 30
4. La VNCG chez le virus de la FCO ? 31
Partie 4 : Stratégies, méthodologie et objectifs du projet 33
1. Approches méthodologiques et conceptuelles pour l'étude de la VNCG et de l'adaptation du virus de la FCO-4 33
2. La variation du nombre de copies de gènes chez le virus de la FCO 37
REFERENCES 40
CHAPITRE 1 : Les segments génomiques du virus de la fièvre catarrhale ovine varient en nombre de copies d'une manière hôte spécifique 48
ANNEXES 60
CHAPITRE 2 : La variation du nombre de copies de gènes chez le virus de la FCO-4 est associée au succès de l'infection de l'insecte vecteur *Culicoïdes imicola* 62
RESUME 66
INTRODUCTION 67

MATERIELS ET METHODES 72
RESULTATS 76
DISCUSSION 85
CONCLUSION 90
REFERENCES 92
ANNEXES 96
CHAPITRE 3 : Génération de la formule génomique du virus de la FCO-4 98
RESUME 102
INTRODUCTION 103
MATERIELS ET METHODES 106
RESULTATS 111
DISCUSSION 121
CONCLUSION 127
REFERENCES 129
ANNEXES 131
DISCUSSION GENERALE 138
1. La variation du nombre de copies des segments chez le virus de la FCO-4, un rôle potentiel dans l'adaptation au cycle en alternance ? 139
2. Génération de la VNCG 143
3. Potentielle influence sur l'évolution du virus de la FCO 147
4. Limitations potentielles liées aux approches et aux méthodes 148
5. Perspectives 150
CONCLUSION GENERALE 151
REFERENCES 153
RESUME 158

Table des illustrations : Figures

INTRODUCTION

Figure 1. Distribution des foyers du virus de la fièvre catarrhale ovine de sérotype 4 en Europe, de 2014 à 2017. Cette illustration est tirée du site internet de la plateforme ESA qui a compilé les informations collectées par l'OIE, la FAO et l'ADNS pour sa réalisation. 11
Figure 2. **Représentation schématique en coupe transversale de la particule virale typique du virus de la FCO.** Les 7 protéines structurelles et les segments génomiques ARNdb sont représentés en son centre. Les protéines VP2 et VP5 constituent la couche externe de la capside virale, tandis que les protéines VP7 et VP3 constituent la couche interne. La particule lestée de la capside externe est dénommée particule core. La figure est tirée, sans modification, de l'article de revue de Schwartz-Cornil et al., (71). 16
Figure 3. Représentation schématique des différentes voies de sortie du virus de la FCO infectant des cellules mammifères. Après l'assemblage des particules virales dans le corps d'inclusion la protéine NS3 (en bleu) assure leur transfert vers un corps multi-vésiculaire. Les virions migrent ensuite dans un lysosome dont le pH permet le maintien de l'intégrité membranaire, il assure la sortie du virus sous la forme de vésicules extracellulaires. Les virions peuvent également sortir sous forme enveloppée suite à l'interaction entre NS3 et le complexe ESCRT. Enfin la lyse cellulaire entraîne la sortie de particule virales complète et libres. L'illustration est tirée, sans modifications, de l'article de revue de T. Labadie (78). 19
Figure 4. Effet de la sélection sur la nature des mutations des 10 segments de la FCO-17 en condition de culture cellulaire. Cette figure présente les ratios de dN/dS pour chaque réplicat, pour les 10 passages successifs sur des cellules de *Culicoïdes sonorensis* (CuVa) en vert, des cellules Bovines (BPAEC) en bleu, l'alternance entre ces deux types cellulaires en jaune et la valeur pour l'inocula en noir. Cette figure est tirée de la publication de Kopanke et al. (92), sans modification. 23
Figure 5. Distance génétique par segment pour la FCO-17 entre les conditions de culture cellulaire. Cette figure présente la somme des fréquences de chaque variant nucléotidique par segment normalisé par la longueur de la séquence codante (SNV/CDS) pour chaque réplicat, pour les passages 3, 6 et 9 sur des cellules de Culicoïdes sonorensis (CuVa) en vert, des cellules Bovines (BPAEC) en bleu, l'alternance entre ces deux types cellulaires en jaune et la valeur pour l'inocula en noir. Cette figure est tirée de la publication de Kopanke et al. (92), sans modification. 23
Figure 6. MDS présentant les corrélations de temps pour l'ancêtre commun le plus récent entre paires de virus de la FCO en Inde pour 20 sérotypes différents. Les estimations sont séparées pour chaque segment. Les points représentent l'incertitude sur 500 échantillonnages avec le programme BEAST pour les virus échantillonnés la même année. Cette figure est tirée de la publication de R. Biek et al., (92). 26
Figure 7. MDS présentant les corrélations de temps pour l'ancêtre commun le plus récent entre paires de virus de la FCO en Europe (sérotypes 1 et 8). Les estimations sont séparées pour chaque segment. Les points représentent l'incertitude sur 500 échantillonnages avec le programme BEAST pour les virus échantillonnés la même année. Cette figure est tirée de la publication de K. Nomikou et al., (91). 26
Figure 8. Fréquence relative des segments du FBNSV en fonction des conditions d'inoculation et de la plante hôte. Les infections sur l'espèce *V. faba* sont présentées en bleu clair pour les infections manuelles et en bleu foncé pour les infections par vectorisation par un puceron (3 réplicas

indépendants). Les infections sur *M. truncata* par des pucerons sont présentés en rouge (2 réplicas). Le mode d'inoculation n'influence pas les fréquences, en revanche, des différences entre plantes hôtes sont observées pour tous les segments à l'exception de U1. La figure est tirée de la publication de Sicard et al., (95)..... 28
Figure 9. La formule génomique et son implication sur l'accumulation du FBNSV dans les compartiments foliaires. Le graphique a) montre que l'accumulation du virus est dépendante du développement de la plante mais surtout que l'accumulation du virus augmente en suivant le niveau foliaire infecté. b) Un indice calculé permet de définir une distance à la formule génomique (valeur des fréquences vers lesquelles convergent les segments), la valeur de distance à formule génomique à l'équilibre diminue dans les niveaux foliaires les plus hauts qui correspondent aux feuilles où l'accumulation du virus est la plus importante. La figure est tirée de la publication de Sicard et al., (95) 29

CHAPITRE 2

Figure 1. (A) Les barrières à l'infection, la dissémination et *in fine à* la transmission dans les culicoïdes et les moustiques. Les flèches en gras indiquent le parcours suivit par un arbovirus de l'ingestion à la transmission par voie orale. La figure est tirée du livre Bluetongue par Mellor *et al*. (8), elle a été modifiée à partir de la publication originelle de Hardy et al. (11). **(B) Parties du corps d'un culicoïde avec les glandes salivaires** en 1), le tube digestif divisé en trois sous-parties : l'intestin antérieur en 2), l'intestin moyen en 3) et l'intestin postérieur en 4). L'infection par les arbovirus se produit spécifiquement dans l'intestin moyen. L'illustration est tirée de la publication de Feenstra et al. (13), modifiée à partir de l'originale de A. R. Walker. 69
Figure 2. Nombre cumulé d'ARN viraux pour les 10 segments génomiques représenté par échantillon en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée). 77
Figure 3. Fréquences relatives des segments génomiques du virus de la FCO-4 en fonction de l'état de dissémination du virus. La ligne en pointillés bleue indique la fréquence théorique attendue (0.1) pour les segments s'ils étaient tous présents en quantité équimolaire. Le contrôle correspond aux fréquences calculées à partir du plasmide pGEM-T qui porte une copie de chacune de dix cibles. Les points rouges représentent les valeurs de fréquences de l'inocula tandis que les droites de la même couleur correspondent à leurs valeurs moyennes. Les différences significatives ont été déterminées séparément pour chaque segment avec une ANOVA (fréquence ~ dissémination virale). Les différences affichées correspondent aux résultats du test de Tukey-hsd (p-value < 0.05). 79
Figure 4. Indice de la distance à l'équimolarité en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée) et pour le contrôle pGEM-T portant chaque cible à un ratio équimolaire. Une différence significative est trouvée entre les états de dissémination (Student, p-value = 3.97E-04). 80
Figure 5. Régression linéaire de la distance à l'équimolarité par la charge virale (nombre cumulé de copies ARN) pour chaque condition de dissémination (DC : dissémination complète, DL : dissémination limitée). 80
Figure 6. Indice de la distance à l'inocula en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée). Une différence significative est trouvée entre les deux conditions de dissémination (Welch, p-value = 1.56E-03). 82
Figure 7. Régression linéaire de la distance à l'inocula par la charge virale (nombre cumulé de copies ARN) pour chaque condition de dissémination (DC : dissémination complète, DL : dissémination limitée). 82

Figure 9. Distance à la formule génomique des populations virales de terrain en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée). Une différence significative est trouvée entre les deux conditions de dissémination (Welch, p-value = 1.16E-03). La distance moyenne à la formule génomique de l'hôte culicoïdes est indiqué par la barre rouge pour l'inocula. 84
Figure 9. Comparaison des formules génomiques du virus de la FCO-4 chez son hôte *C. imicola* pour les deux états de dissémination de l'infection expérimentale (DC : dissémination complète, DL : dissémination limitée) et pour les populations virales naturelles. 84

CHAPITRE 3

Figure 1. Comparaison du nombre absolu de copies entre les jeux de données brutes et corrigées. Pour chacune des figures le jeu de données brutes est présenté dans la couleur foncée et le jeu de données corrigées dans la couleur plus claire. A) Comparaison des traitements pour le segment 2, aucune différence significative de nombre de copies absolues n'est observée entre les jeux de données (Wilcoxon, p-value = 0.20). B) Comparaison des jeux de données pour le segment 10. Il n'y a pas de différence significative entre les deux traitements (Wilcoxon, p-value = 0.058). 111
Figure 2. Influence de la MOI sur les fréquences relatives des segments génomiques du virus de la FCO-4 en fonction de la MOI à l'inoculation. La ligne bleue en pointillé indique la fréquence théorique attendue (0,1) pour les segments s'ils étaient tous présents en quantité équimolaire. Les points rouges représentent les valeurs de fréquences de l'inocula tandis que les droites de la même couleur correspondent à leurs valeurs moyennes. Les fréquences du contrôle plasmide ont été inclues pour chaque segment. Les différences significatives ont été déterminées séparément pour chaque segment avec un test de Wilcoxon (p-values < 0.05). Les fréquences pour lesquels une différence entre les conditions de MOI sont observées pour les deux types cellulaires sont marquées par un astérisque noir. Si une différence significative n'est observée que pour un seul type cellulaire, les segments ont été annotés par un astérisque rouge (MDBK) ou bleu (KC). Finalement, les segments pour lesquels aucune différence significative n'est trouvée entre les conditions sont marqués d'un triangle noir (segments 7 et 9). 114
Figure 3. Coefficients de variation calculés pour chaque segment dans nos différentes conditions d'infection. Les coefficients de variation ne présentent pas de variation significative entre les conditions de MOI et de type cellulaire (Welch, p-value > 0,05). 115
Figure 4. (A) Indice de distance à l'équimolarité (dEQ) calculé par échantillon pour chacune de nos conditions par rapport à l'attendu théorique de fréquence en situation d'équimolarité entre segments. Les astérisques indiquent les différences significatives observées entre les conditions de MOI (test de Welch, p-value < 0.05). La MOI a une influence significative sur l'indice de la distance à l'équimolarité. Les points rouges correspondent à l'indice de distance calculé pour l'inocula, la barre de la même couleur correspond à leur valeur moyenne. **(B)** Indice de distance à l'inocula (dIN) défini par rapport aux fréquences moyennes des segments du virus utilisé pour l'infection et représenté pour chaque condition. La MOI a un effet significatif sur la distance à l'inocula (Welch, p-value < 0.05), il est variable en fonction des types cellulaires. Les astérisques marquent la significativité des différences observées par type cellulaire. 116
Figure 5. Indices de distance à la formule génomique calculé par rapport aux valeurs médianes de fréquence des formules génomiques des virus décrits dans des populations naturelles, soit des **(A)** hôtes arthropodes (dGFa), soit **(B)** des bovins (dGFb). Les espèces hôtes correspondent aux types cellulaires testés dans notre approche. Les indices des populations virales de terrain ont été ajoutés afin de rendre compte de la distance à la formule génomique existant en conditions naturelles. Les

astérisques indiquent les différences significatives entre chaque condition en fonction du groupe étudié (Welch, p-value < 0.05). **(C)** Nombre cumulé de copies des segments génomiques pour les différentes conditions d'infection. Pour chaque échantillon les nombres de copies des 10 segment sont additionnés. Entre les conditions de MOI, les différences significatives de charge virale sont indiquées par des astérisques (test de Welch, p-value < 0.05). 118
Figure 6. Formule génomique en fonction des conditions d'infection. Les MOI faibles sont présentées dans une couleur plus claire. (A) Formule génomique des virus pour les infections sur cellules KC. (B) Formules génomiques des populations virales à l'issue de l'infection sur MDBK. 127

Figure S1. Régression linéaire de la distance à l'équimolarité par la charge virale. Les équations des droites, les coefficients de corrélation et la significativité de l'interaction sont indiqués pour chaque condition. Les distances à l'équimolarité sont représentées en bleu pour les virus produits sur cellules de culicoïdes KC, en clair pour l'infection à faible MOI (A) et en foncé pour celle à forte MOI (B). Pour les cellules de vache MDBK, l'infection à faible MOI est représentée en jaune clair (C) et en jaune foncé pour la forte MOI (D). 133
Figure S2. Droites de régression linéaire de la distance à l'inocula par la charge virale. Les équations des droites, les coefficients de corrélation et la significativité de l'interaction sont indiqués pour chaque figure. Les distances à l'équimolarité sont représentées en bleu pour les cellules de culicoïdes KC, (A) en clair pour l'infection à faible MOI et (B) en foncé pour celle à forte MOI. Les distances à l'équimolarité sont représentées en jaune pour les cellules de vache MDBK, (C) en clair pour l'infection à faible MOI et (D) en foncé pour celle à forte MOI. 134
Figure S3. Droites de régression linéaire de l'indice pour la distance à la formule génomique de l'hôte culicoïdes par le nombre cumulé de copies d'ARN viral pour les cellules d'insecte aux MOI de 0.3 (A) et de 3 (B). 134
Figure S4. Droites de régression linéaire de l'indice pour la distance à la formule génomique de l'hôte vache par le nombre cumulé de copies d'ARN viral pour les cellules de vache aux MOI de 0.3 (A) et de 3 (B). 135
Figure S5. Nombre de copies quantifiées pour chacun des dix segments génomiques entre chaque condition d'infection. Les segments sont présentés en haut de la figure. L'échelle des ordonnées a été transformée en échelle logarithmique (10). A l'exception des segments 2 et 10, l'ordre de variation est conservé entre les conditions par rapport à l'ordre cumulé des copies pour les dix segments présentés dans la figure 5. 135
Figure S6. Résultats d'infection expérimentale après 48 heures d'infection sur des cellules KC et VERO à deux MOI inférieures à 1. Chaque condition comprenait 10 réplicas biologiques. **(A)** Nombre cumulé de copies des segments génomiques pour les différentes conditions d'infection. Pour chaque échantillon, les nombres de copies des 10 segments sont additionnés. **(B)** Titre estimé à partir du type cellulaire VERO pour chacune des conditions d'infection. **(C)** Estimation de la proportion de particules virales infectieuses par rapport au nombre de particules virales totales, en considérant que chaque virion contient 10 segments génomiques en moyenne. 136
Figure S7. Influence de la MOI sur les fréquences relatives des segments génomiques du virus de la FCO-4 en fonction du type cellulaire. La ligne en pointillé bleue indique la fréquence théorique attendue (0,1) pour les segments s'ils étaient tous présents en quantité équimolaire. Cette figure présente les fréquences des segments quantifiés après 48 heures d'infection sur des cellules KC et VERO infectées à deux MOI inférieures à 1. Chaque condition comprenait 10 réplicas biologiques. La souche virale utilisée a été isolée sur KC puis amplifiée une fois sur cellule VERO. Les couples de primers utilisés et les conditions de réaction de RT-qPCR ne sont pas identiques à ceux utilisés pour l'approche présentée dans ce chapitre. 137

Table des illustrations : Tableaux

INTRODUCTION

Tableau 1. Familles de virus pour lesquels au moins une espèce a été décrite avec un cycle en alternance d'hôte typique des arbovirus. Les caractéristiques génomiques de ces familles sont présentées. Lorsqu'une famille possède un génome non segmenté cela équivaut à dire qu'il possède une structure génomique linéaire. 2
Tableau 2. Segments génomiques du virus de la FCO, taille (FCO-4), noms et fonction principale des protéines virales. 17

CHAPITRE 1

Tableau S1. Analysis of host effect on segment frequency in three datasets. Analysis of variance per segment (model: frequency ~ host) in three datasets (raw, corrected and Wo-10). Raw: raw copy numbers per segment and sample. Corrected: copy numbers after subtraction of copy numbers found in primer-independent amplifications. Wo-10: raw copy numbers for all segments but segment 10. 60
Tableau S2. Analysis of host effect on segment frequency in three datasets. p-values of Tukey-HSD tests obtained from the analyses of variance in Table A4-A. Raw: raw copy numbers per segment and sample. Corrected: copy numbers after subtraction of copy numbers found in primer-independent amplifications. Wo-10: raw copy numbers for all segments but segment 10. Differences in significance levels were found between the Wo-10 dataset and the others for segment 9 (significant differences between ruminant and midge samples only found with Wo-10; p-value = 0.044). 61

CHAPITRE 2

Tableau 1. Effet de l'amplification aléatoire sur le nombre de copies pour les 5 échantillons du contrôle. Le facteur de différence est calculé en divisant les valeurs du jeu de données brutes par celles du corrigé. 76

Tableau S1. A) Comparaison des données corrigées ou brutes pour l'effet d'amplification aléatoire pour les segments 2 et 10 sur le jeu de données incluant tous les échantillons. Les tests de normalité (Shapiro) et d'homocédasticité (F-test) sont validé. Les résultats des tests de Wilkoxon et de Student ne montrent pas de différence entre les jeux de données corrigés et brute pour les segments 2 et 10 (DC : dissémination complète, DL : dissémination limitée). **B)** Effet de l'amplification aléatoire sur le nombre de copies et les fréquences normalisées pour les données filtrées brutes (Ct < 31) et non corrigées. 97

CHAPITRE 3

Tableau 1. **Conditions d'infection expérimentale pour les deux MOI et types cellulaires.** 107

Tableau S1. Différence des fréquences observées à l'attendu théorique de 0.1 pour les différents segments dans les conditions expérimentales. Le tableau présente les p-values des résultats du test de Welch one-sided appliqué séparément pour chaque segment et corrigé pour le taux de découverte de faux positifs par la méthode Benjamin-Hopkins. 131
Tableau S2. Différence des fréquences observées à l'inocula pour les différents segments dans les conditions expérimentales. Le tableau présente les p-values des résultats du test de Welch one-sided appliqué séparément pour chaque segment et corrigé pour le taux de découverte de faux positifs par la méthode Benjamin-Hopkins. 131
Tableau S4. Comparaison des fréquences des 10 segments génomiques. Les deux colonnes de type cellulaire présentent les résultats (p-values) du test de Wilcoxon entre la MOI de 0.3 et de 3 dans chaque lignée de cellule. Les colonnes MOI présentent les résultats du test de Wilcoxon (p-values) de comparaison entre les cellules pour chacune des deux MOI. 132
Tableau S4. Valeurs médianes des indices de distance à la formule génomique du virus de la FCO-4 pour l'hôte arthropode (dGFa) et bovin (dGFb). 132

Liste des abréviations

ADN(c) : acide désoxyribonucléique (complémentaire)

ARN(sb, db) : acide ribonucléique (simple brin, double brin)

AHSV : African horse sickness virus

BSR : souche cellulaire dérivée de BHK, rein de hamster

CT : complexe de transcription

Ct : nombre de cycles d'amplification

DENV : virus de la Dengue

dEQ : distance à l'équimolarité

dGFa : distance à la formule génomique des échantillons culicoïdes de terrain

dGFb : distance à la formule génomique des échantillons bovins de terrain

dIN : distance à l'inocula

DMEM : milieu dubelcco's modified eagles

DMSO : dymethylsulfoxyde

FBNSV : Faba bean necrotic stun virus

FCO(-) : fièvre catarrhale ovine (sérotype)

FG : formule génomique

FGE : formule génomique à l'équilibre

IAV : virus de la grippe A

KC : cellules dérivées de *Culicoides sonrensis*

MDBK : cellules de rein de vache

MOI : multiplicité d'infection

NC : nombre de copies

qPCR : réactions en chaine de polymérase quantitative

RdRp : RNA dependent RNA polymerase

RT : transcription inverse

RVFV : virus de la fièvre de la vallée du Rift

SDM : milieu schneider drosophila medium

VSV : virus de la stomatite vésiculaire

VNCG : variation du nombre de copies de gènes

WNV : virus du Nil Occidental

INTRODUCTION GENERALE

Partie 1 : « Les virus d'origine arthropode »

1. Concepts et contexte

1.1. Définition(s) générale(s)

Le terme arbovirus est une contraction de l'anglais « arthropod-borne viruses », il désigne un groupe d'espèces virales polyphylétiques qui partagent un trait biologique de leur cycle de transmission. Toutes les espèces d'arbovirus ont en commun un cycle en alternance d'hôte obligatoire impliquant un hôte arthropode, aussi qualifié d'espèce vectrice ou vecteur, dans lequel le virus se multiplie. Dans sa définition, le terme arbovirus englobe environ 500 espèces appartenant à 9 familles virales différentes (1) capables d'infecter, par l'intermédiaire de leurs vecteurs, des espèces hôtes appartenant aux règnes végétal ou animal. L'impact des arbovirus sur la santé humaine à l'échelle mondiale et le nombre croissant d'études sur ces pathogènes mènent parfois à un emploi restreint du terme arbovirus, ne considérant que les virus vectorisés par des espèces hématophages (2). L'estimation la plus récente du nombre d'arbovirus décrit comme étant capable d'infecter le genre humain est de 140 espèces (3). A titre de comparaison, nous ne connaissons qu'une cinquantaine d'espèces d'arbovirus animaux (humains exclus) (1). Au cours de ce projet, nous avons étudié un arbovirus animal en particulier. Cependant les mécanismes adaptatifs questionnés nous conduisent à inclure dans notre réflexion des éléments du cycle biologique des virus au sens large.

1.2. Les arbovirus de vecteurs hématophages

Les arbovirus qui incluent exclusivement des hôtes du règne animal dans leur cycle biologique appartiennent à 7 familles virales différentes. Elles sont présentées dans le Tableau 1 avec une description simple de leurs principales caractéristiques génomiques. Les génomes de ces espèces peuvent avoir différentes formes. Les virus

à génome linéaire possèdent un brin nucléaire qui porte l'ensemble des gènes. Pour les virus segmentés les gènes sont séparés entre différents fragments, encapsidés au sein de la même particule virale. La majorité des arbovirus possèdent un génome linéaire ARN. Les représentants de la famille des *Flaviviridae* font partis des arbovirus les plus emblématiques avec le virus ZIKA, de la fièvre jaune ou du Nil occidental. Cette famille des *Flaviviridae* est particulière car la majorité de ses membres ont un cycle en alternance d'hôte. La situation est à l'opposé pour la famille des *Asfaviridae* dans laquelle on ne compte qu'une seule et unique espèce avec un génome ADN et un cycle en alternance arbovirus : le virus de la peste porcine africaine. Il existe en revanche un nombre beaucoup plus important d'arbovirus à génome ADN chez les plantes (1). Enfin la famille des *Reoviridae* fait aussi figure d'exception, elle est la seule pour laquelle les espèces ont un génome ARN double brin et possède le niveau de segmentation le plus élevé. Ces espèces sont presque exclusivement des virus ARN, la nature de leur génome est associée à l'absence de système de correction des erreurs de réplication par l'ARN Polymérase ARN-dépendante (RdRp) qui ne possède pas d'activité exonucléase. La RdRp est associée à la génération plus rapide de mutations qui peuvent constituer un avantage dans le contexte d'un changement d'environnement dynamique.

Tableau 1. Familles de virus pour lesquels au moins une espèce a été décrite avec un cycle en alternance d'hôte typique des arbovirus. Les caractéristiques génomiques de ces familles sont présentées. Lorsqu'une famille possède un génome non segmenté cela équivaut à dire qu'il possède une structure génomique linéaire.

Famille	Génome	Forme	Segmentation
Flaviridae	ARN	simple brin +	non
Togaviridae	ARN	simple brin +	non
Bunyaviridae	ARN	simple brin -	3 segments
Reoviridae	ARN	double brin	10-12 segments
Orthomyxoviridae	ARN	simple brin -	6-8 segments
Rhabdoviridae	ARN	simple brin -	non
Asfarviridae	ADN	double brin	non

Les arbovirus animaux sont vectorisés par 5 groupes distincts d'arthropodess hématophages. Ils appartiennent au sous-embranchement *Chelicerata* pour les tiques (familles *Ixodidae* et *Argasidae*) ou à la classe *Insecta* pour les moustiques (famille *Culicidae*), les culicoïdes (famille *Ceratopogonidae*), les phlebotomes (sous-famille *Plebotominae*) et les punaises (famille *Cimicidae*). Les arbovirus peuvent causer différents types de symptômes chez les hôtes vertébrés. Parmi eux les signes les plus communs sont les fièvres, généralement associées à l'activation du système immunitaire de l'hôte, ainsi qu'une variété de symptômes cliniques parmi lesquels des troubles neurologiques, des hémorragies, des fausses couches ou des malformations congénitales ainsi que l'apparition d'éruptions cutanées (1). L'impact économique de ces maladies est important (4), et concerne à la fois les populations humaines et de nombreux systèmes d'élevages sur toute la planète (5,6).

2. Le cycle complexe en alternance d'hôte

2.1. Des coûts augmentés

L'existence du cycle en alternance d'hôte des arbovirus dans une telle diversité de familles virales et de systèmes d'interaction questionne sur les mécanismes évolutifs qui mènent à son établissement et à son maintien. Pour les arbovirus la plus forte contrainte pourrait être associée au maintien de la capacité à infecter en alternance deux hôtes très différents. Les populations virales subissent un effet fort de sélection purifiante lié à leur cycle en alternance (6-9). Cet effet implique une pléiotropie antagoniste qui limite la fixation de nouvelles mutations et pourrait expliquer le taux d'évolution plus faible observé chez les arbovirus ARN en comparaison aux virus ARN qui ne possèdent qu'un seul hôte (3).

La transmission entre hôtes implique le coût le plus évident du cycle parasitaire des virus. Les coûts présentés dans ce paragraphe peuvent s'appliquer à l'ensemble des virus, cependant leurs effets sont amplifiés dans le contexte du cycle en alternance d'hôte des arbovirus. La transmission du virus entre deux individus est généralement synonyme de goulot d'étranglement et de dérive génétique (9-11). Le changement d'environnement et les barrières physiques auxquels sont soumises les populations virales entraînent l'échantillonnage stochastique d'une fraction de la population virale

au moment de la transmission et de l'infection (10,12). C'est par cet intermédiaire que certains allèles peuvent se propager par l'effet de la dérive génétique (12-13). Des facteurs liés à l'hôte influence la pression de sélection au sein d'une même espèce. La composante génétique de l'hôte va agir sur les réponses immunitaires innées ou acquise et, plus généralement, sur la susceptibilité de l'hôte au pathogène (14-15). La physiologie de l'hôte peut varier en fonction de son âge, du sexe de l'individu et de facteurs abiotiques liés à la saisonnalité. Enfin les phénomènes de goulot d'étranglement et de dérive génétique ne se limitent pas au seul évènement de transmission et un ensemble de barrières physiques au sein de chaque hôte peut influencer le cours de l'infection (16). Pour pouvoir être transmis par leur hôte insecte, les arbovirus doivent passer un ensemble de barrières les menant de l'intestin moyen aux glandes salivaires. Les arbovirus qui possèdent un génome ARN ont conservé leur capacité d'adaptation spécifique à un environnement (17,18). L'absence de correction des erreurs de réplication pourrait mener à la génération de mutations au cours de l'infection des différents compartiments hôte. En cas de fixation, celles-ci pourraient s'avérer délétère dans l'hôte alternatif (13).

2.2. Les bénéfices d'un cycle en alternance d'hôtes ?

Les arbovirus ne forment pas un complexe monophylétique d'espèces, l'évolution a donc conduit des espèces virales appartenant à des familles très éloignée à converger vers ce cycle en alternance d'hôte (1). Comme nous l'avons évoqué, il existe une diversité d'architectures génomiques chez les arbovirus animaux qui trouvent leurs origines dans des familles virales distantes les unes des autres. Nous pouvons associer le cycle des arbovirus à un trait d'histoire de vie qui est apparu indépendamment dans différents taxons et par extension considérer ce cycle comme une convergence évolutive.

Le cycle des arbovirus peut être associé à différents bénéfices. D'un point de vue mécanistique, le cycle de vie biologique des insectes hématophage est directement dépendant de la consommation de sang par les hôtes arthropodes. Cette interaction obligatoire et répétée entre les vecteurs et les hôtes animaux bénéficie directement au virus en augmentant sa probabilité de transmission. Les niches écologiques occupées par les arthropodes hématophages incluent des environnement urbains ou ruraux

favorisant les contacts entre espèces (5). Pour certaines espèces de culicoïdes, la pontes des œufs se fait à même le sol dans les élevages. La proximité entre espèces et les interactions obligatoires entre le vecteur et l'hôte alternatif permettent de limiter l'un des coûts majeurs du cycle viral, la transmission. Si peu de particules sont transmises, les interactions répétées augmentent la probabilité de transmission. Cette interaction prévient la libération des particules virales dans l'environnement extérieur (à l'exception du tube digestif de l'arthropode). La transmission vectorielle, en plus de limiter ce risque d'extinction, soustrait les arbovirus aux pressions exercées par plusieurs facteurs abiotiques comme l'exposition aux UV, la dégradation de la vésicule virale, les variations fortes de température.

Plus généralement l'activité humaine favorise le transport d'animaux et de végétaux à une échelle mondiale. Les systèmes de production agro-industriels favorisent la concentration d'espèces uniques dans des espaces réduits qui constituent un foyer privilégié à l'apparition et à la transmission de ces maladies (5, 6). Par l'intermédiaire des activités humaines, la dispersion des espèces vectrices peut aussi se faire à grande échelle, l'importation du moustique tigre sur le continent Européen en constitue le meilleur exemple (19). Si la propagation des arboviroses est facilitée par les activités humaines, la biologie de certaines espèces vectrices comme la dispersion des culicoïdes par le vent à grande échelle constitue également des facteurs bénéfiques à la dissémination des arbovirus (20).

3. L'adaptation au cycle en alternance d'hôte

3.1. Le(s) concept(s) de quasi-espèce

Au sein d'un hôte, une population virale peut être définie comme un ensemble de génomes ou un nuage de variants génétiques. La nature même des génomes ARN favorise la génération de mutations : la RdRp ne possède pas de système de correction des erreurs de réplication et la taille réduite des génomes favorise leur réplication rapide. De nombreuses études associent cette diversité génétique propre à la nature des virus ARN au concept de la quasi-espèce. Actuellement sur PubMed, 2733 résultats sont associés à ce terme. Le concept de la quasi-espèce virale émerge dans les années 70 (21), il est originellement proposé par Manfred Eigen. Il implique

pour les virus, au sein d'une population, la sélection de différents génotypes autour desquels l'espace des séquences peut être exploré (22). Le concept a aussi été appliqué à l'évolution des virus ARN par Esteban Domingo à la fin des années 70, puis au fil de sa carrière scientifique (23). E. Domingo a orienté sa définition du concept à l'existence de variants génétiques autour d'une séquence, avec la valeur sélective la plus élevée, dans une population soumise à la sélection naturelle. Un point de clivage entre virologiste sur la définition du concept tient à cette différence. Initialement la théorie prévoit que des génotypes coexistent, ils sont soumis à sélection et participent à l'exploration des séquences autour de leurs génomes. Le spectre de mutants autour d'une séquence consensus est aussi soumis à une sélection de groupe mais ne correspond pas à cette définition initiale, nous qualifierons cette dernière de spectre de variant et nous décrirons son implication potentielle dans l'adaptation à l'alternance d'hôte dans le paragraphe suivant. Néanmoins ces deux visions s'accordent sur l'effet de la sélection naturelle qui affecte la population virale dans son ensemble, au sein d'un hôte à un instant donné. La démonstration du concept de quasi-espèce défini par Eigen est difficilement envisageable dans des conditions naturelles d'infection avec une alternance d'hôte, comme pour les arbovirus. Les sélections purifiantes au sein de chaque hôte, les goulots d'étranglement répétés constituent des limites qui semble incompatible avec la théorie de la quasi-espèce. La présence de génotypes complémentaires indispensables au succès de l'infection au sein d'une même espèce virale n'a été démontrée qu'une seule fois, à ma connaissance, chez le poliovirus par Vignuzzi *et al.* (24). Au cours de leur approche expérimentale, ce groupe a été en mesure de déterminer que le succès d'une infection virale « complète », menant à l'infection neuronale, était dépendante de la présence de différents génotypes au sein de la population virale. Cette piste reste à explorer chez les arbovirus mais actuellement les principaux travaux qui ont trait à l'évolution et à l'adaptation des arbovirus se concentrent sur la diversité génétique générée autour d'un génotype central.

3.2. Le spectre de variants et l'adaptation par mutation

L'existence et la génération d'un nuage de variants impliquent des mécanismes identiques à ceux ayant mené à la théorisation de la quasi-espèce chez les virus ARN. La polymérase virale et l'absence de correction des erreurs de réplication, combinée

à une réplication rapide liée à la petite taille de ces génomes viraux, entraînent l'apparition récurrente de nouvelles mutations (25,26). Ces caractéristiques permettent aux virus de s'adapter rapidement aux changements d'environnement entre des hôtes différents mais aussi aux différents compartiments au cours de l'infection d'un hôte (8,27).

Pour les arbovirus, une contrainte supplémentaire s'applique : ils doivent conserver leur capacité à infecter des hôtes distants pour maintenir leur cycle malgré un spectre de variant continuellement généré autour d'une séquence constituant un optimum au sein de l'hôte, voire du compartiment. Au cours de l'infection les virus sont soumis à des variations de pression de sélection, l'infection d'un hôte peut induire un effet de sélection purifiante. Cet effet a été montré chez différentes espèces d'arbovirus aussi bien au cours de l'infection des hôtes vertébrés (28-31), que des vecteurs arthropodes (27,32).

Les arbovirus ont conservé la capacité d'adaptation à des environnement fixes lorsqu'ils sont expérimentalement contraints. Cette adaptation se traduit par la génération de mutations adaptatives et par une augmentation de la valeur sélective aussi bien en condition d'infection expérimentale *in vivo (18, 29)* que *in vitro* (33,34), parfois différemment en fonction des hôtes (27, 29, 31). Les espèces vectrices ou hôtes peuvent également influencer la valeur sélective d'une même espèce arbovirus (11, 30).

Il n'existe pas de consensus clairement défini qui explique l'adaptation des arbovirus à l'alternance d'hôtes. La majorité des arbovirus sont capables de s'adapter à des environnements spécifiques, mais sont soumis à une sélection purifiante au cours de l'alternance d'hôte limitant la génération de mutations adaptatives fixées par pléiotropie antagoniste. Malgré l'existence des goulots d'étranglement, des phénomènes de restauration de la diversité ont été observés pour différents arbovirus au cours de l'infection de l'hôte arthropode (9, 15).

La diversité génétique joue un rôle central dans le succès de la dissémination d'un virus dans son hôte mais la génération de mutation spécifiques à l'hôte au cours du cycle de transmission des arbovirus n'a jamais été montrée.

3.3. La plasticité des génomes viraux

Une alternative (non exclusive) aux adaptations par mutation pourrait résider dans la capacité des populations virales à produire différents phénotypes en conservant leurs diversités d'origine. De nombreuses espèces d'arbovirus présentent des génomes porteurs de larges délétions (35-37). Les particules qui portent ces génomes sont appelées particules défectives. L'existence de ces génomes tronqués, dans les populations naturelles d'arbovirus, a été montré chez différentes espèces (35, 36). La génération de délétion est généralement liée à des conditions de forte concentration de génomes viraux au sein d'une même cellule (38), la position de ces délétions est plus ou moins fixe en fonction des espèces. Si les délétions génomiques peuvent être associées à un effet délétère au niveau du virion, il est aujourd'hui établi que la sélection naturelle affecte la population virale dans son ensemble (24). Dès lors ces délétions pourraient être sélectionnées pour favoriser la population virale.

Historiquement les particules défectives ont été décrites chez un virus ADNdb circulaire (39) : *Autographa californica polyhedrovirus* (AcMNPV), il appartient à la famille des *baculoviridae*. Au cours d'une infection expérimentale pour tester la létalité de différentes combinaisons de variant génétique, il a été établi qu'une certaine proportion de génomes défectifs au sein de la population virale augmentait la valeur sélective du virus (40).

Cette piste n'a jamais été exploré, sous cet angle, chez les arbovirus, elle converge en terme fonctionnel avec un autre mécanisme inféodé aux virus à génome segmenté : la variation du nombre de copies de gène, qui peut aussi bien être causée par l'encapsidation surnuméraire de segments, que par la sélection de génomes défectifs. Ce mécanisme est introduit plus en détail dans la partie 4.2 de l'introduction.

3.4. Exploration de ces stratégies adaptatives chez le virus de la FCO

Le virus de la fièvre catarrhale ovine possède plusieurs caractéristiques génétiques et biologiques assez uniques parmi les arbovirus. Il existe environ 100 espèces de culicoïdes compétentes pour le transmettre à une gamme d'hôtes mammifères très diverse (41-43). La segmentation de son génome a conduit à l'étude de chacun de ces segments et a participé à montrer que le réassortiment conduit à un

mélange génomique à l'origine de l'apparition de nouveaux variants avec une valeur sélective améliorée (44, 45). Si l'encapsidation des segments est assez spécifique (46, 47), chaque segment peut être considéré comme un gène physiquement séparé, et soumis à des pressions de sélections différentes qui induisent des taux d'évolution variables pour chacun d'entre eux (48, 49), ces aspects sont présentés plus en détail dans la partie 4.1 de l'introduction. Jusqu'alors, aucune approche n'a montré l'existence de mécanismes adaptatifs génétiques qui pourraientt favoriser l'adaptation du virus de la FCO à l'alternance d'hôte.

Partie 2 : Le virus de la fièvre catarrhale ovine

1. Présentation générale

Les premiers témoignages de l'existence virus de la FCO ont été rapportés par des éleveurs en Afrique du Sud dans la première moitié du 19ème siècle. Ces témoignages coïncident avec l'introduction de moutons de la race Mérinos, que nous connaissons aujourd'hui pour leur sensibilité au virus de la FCO (41). Les premières descriptions cliniques de la maladie ont été faites par Hutcheon et Spreull en 1902 (41). Cette maladie est historiquement inféodée au continent Africain où sa présence est probablement enzootique, sa présence dans certains pays est liée à l'import de différentes races européennes de moutons. Les premières incursions du virus, en dehors du continent Africain, ont été décrites sur l'île de Chypre en 1943. En 1952, la maladie apparaît dans l'état Californien sur le continent Américain. Aujourd'hui le virus est présent sur tous les continents à l'exception de l'Antarctique. De nouvelles incursions se produisent régulièrement en France (42).

Le virus de la fièvre catarrhale ovine est enzootique dans les zones tempérées à tropicales du globe. Les changements de température ont favorisé des changements dans la distribution de nombreuses espèces ectothermes (43) dont certaines espèces d'insectes vectrices (44). Plusieurs travaux témoignent d'une augmentation de l'aire de distribution de *Culicoïdes imicola*, principale espèce vectrice de la FCO en Afrique, vers le Nord. En Europe, cela s'est traduit par la fixation de foyers du virus de la FCO depuis 1998. A ce jour, 27 sérotypes confirmés (45) du virus de la FCO ont été décrits à travers le monde (29 ont été décrits au total). Parmi eux, 10 ont circulé ou circulent sur le continent Européen : les sérotypes 1, 2, 4, 6, 8, 9, 11, 16, 25 et 27.

Au cours de ce projet, nous avons étudié le sérotype 4 du virus de la FCO en particulier. Le sérotype 4 est à l'origine de plusieurs foyers récents de la maladie dans les Balkans en 2014. Il s'est ensuite étendu par l'Italie, l'île de Sardaigne puis la Corse entre 2015 et 2016. Aujourd'hui le sérotype 4 circule toujours en Europe et touche de nombreux pays (Figure 1).

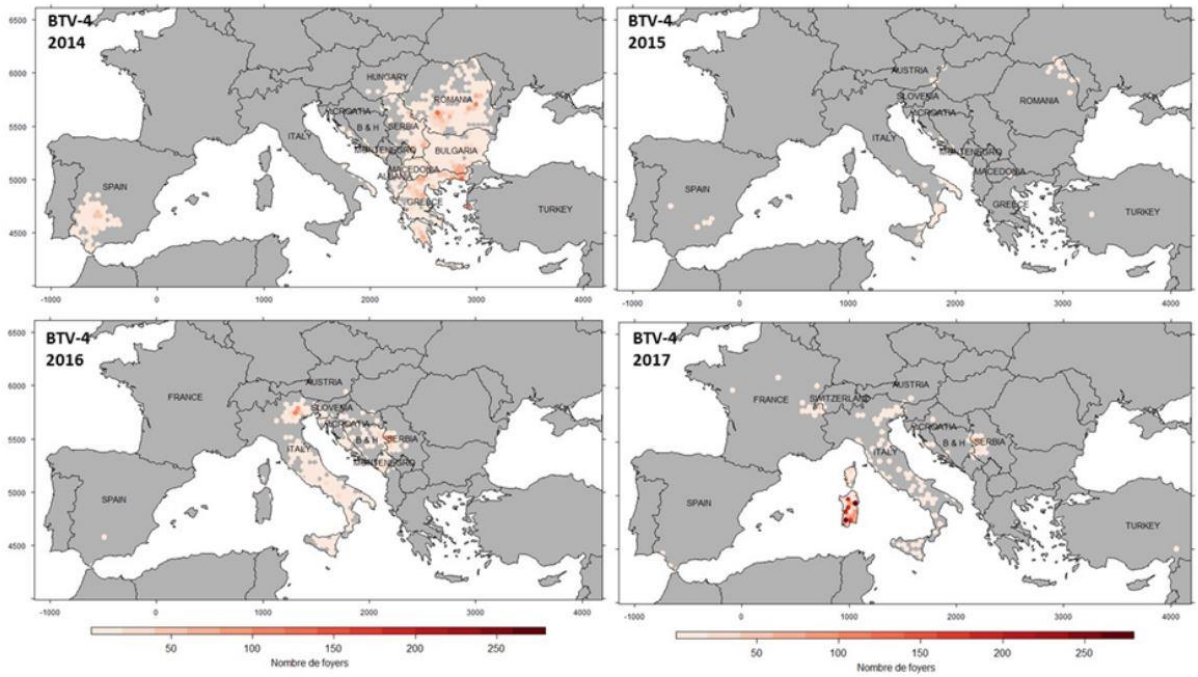


Figure 1. Distribution des foyers du virus de la fièvre catarrhale ovine de sérotype 4 en Europe, de 2014 à 2017. Cette illustration est tirée du site internet de la plateforme ESA qui a compilé les informations collectées par l'OIE, la FAO et l'ADNS pour sa réalisation.

2. Biologie du virus de la FCO

2.1. Les espèces hôtes et le cycle du virus de la FCO-4

2.1.1. Les culicoïdes, moucherons piqueurs

Les culicoïdes sont parmi les plus petits insectes hématophages avec une taille adulte qui varie entre 1 à 3mm de longueur. Les larves de ces espèces ont un développement qui se divise en quatre phases larvaires avant le stade pupe puis l'émergence des adultes (46). En Europe le cycle de vie des culicoïdes est saisonnier. La durée du stade larvaire est dépendante des conditions environnementales et peut varier de deux semaines à plusieurs mois. Les larves pourraient survivre aux températures hivernales au sein des élevages ou par diapause. Les émergences des adultes se produisent principalement au printemps, ici aussi avec une certaine disparité entre les espèces et un second cycle d'émergence peut mener à la sortie d'adulte à la fin de l'été (47). Pour les espèces vectrices d'arboviroses, les œufs ont besoin de moisissures pour se développer, le site de ponte est généralement associé

aux élevages animaux. Les œufs peuvent être déposés à même le substrat souillé, au sol (47). Les sites de ponte et le comportement alimentaire des femelles varient entre les espèces. Chez les culicoïdes, seules les femelles se nourrissent de sang, ce repas est obligatoire pour la maturation des œufs. L'alimentation des individus adultes est principalement constituée de nectar d'origine végétale. En Europe, différentes espèces sont responsables de la propagation du virus de la FCO (46). Sur le pourtour Méditerranéen *Culicoïdes imicola* est le principal vecteur. Pour la partie centrale et Nord de l'Europe, les espèces du complexe *Culicoïdes obsoletus* et *scoticus* et *Culicoïdes pulicaris* sont impliquées (46).

2.1.2. Cycle d'infection dans l'hôte insecte

Les principaux travaux visant à décrire l'infection du virus de la FCO dans l'hôte culicoïde ont été menés en Amérique du Nord où le vecteur principal de la FCO est l'espèce *Culicoïdes sonorensis* (16,46,48). Au cours du gorgement, les femelles peuvent ingérer jusqu'à 10^{-4} ml de sang. La chronologie du cycle d'infection qui suit correspond au cycle du virus à 25°C pour des femelles *C. sonorensis* gorgées sur le virus BTV-1 produit en culture cellulaire, à un titre de 10^6 TCID50/ml. Chaque culicoïde est donc initialement infecté par environ 70 particules virales. Lors des deux premiers jours, la quantité de virus diminue jusqu'à passer sous le seuil de détectabilité, il s'agit de la phase d'éclipse. Cette « disparition » du virus pourrait être associée à la perte des particules virales dans l'intestin moyen des vecteurs par dégradation ou excrétion. Suite à la phase d'éclipse, la concentration en virus augmente jusqu'à atteindre une phase de plateau entre les jours 7 et 9 (16). Les cellules de l'intestin moyen sont les premières infectées, rapidement suivies par les cellules composant les corps gras, au cours du premier jour de l'infection. Les tissus neuraux sont infectés à 3 jours post-infection (j.p.i.) puis les glandes salivaires à 5 j.p.i.. Le virus s'accumule dans les glandes acineuses avant d'être relâchées dans les canaux adjacents où il s'accumule alors en formant des structures para-cristallines (46). La transmission est possible après une période extrinsèque d'incubation, qui correspond à la durée que met le virus entre l'ingestion jusqu'à sa libération dans la salive. Le titre de virus relâché dans la salive a été estimé entre 0.32 et 7.79 TCID50 (16). Les estimations de taux de transmission supportent un fort succès de l'infection de l'hôte vertébré lors de la morsure d'un vecteur infecté (49). La phase de plateau est potentiellement liée aux

nombres de cellules qui peuvent être infectée dans l'hôte, il est dépendant de la taille des individus et varient donc entre les espèces (50). Une fois atteinte, elle restera constante au cours de la vie du culicoïde.

2.1.3. Les hôtes ruminants et la susceptibilité des hôtes mammifères

Historiquement les moutons et plus particulièrement les races européennes constituent les hôtes mammifères les plus touchés par le virus de la (51). Des races bovines peuvent également être infectées avec des effets extrêmement variables dépendants de la virulence de la souche virale impliquée. Les bovins ont une phase de virémie plus longue que les ovins et sont suspecté de jouer un rôle de réservoir pour le virus (51). Les caprins peuvent également être touchés par la maladie, et les infections sont de moindre intensité avec une guérison généralement spontanée (51). Les mortalités observées pour ces hôtes sont attribuées aux conséquences d'œdèmes pulmonaires et de surinfection bactérienne causée par l'affaiblissement général de l'organisme et des hémorragies (51). Certains carnivores peuvent produire des réponses immunitaires contre le virus de la FCO suite à l'ingestion de viandes contaminées, comme les chiens qui ne présentent cependant pas les symptômes de la maladie (52). Un cas exceptionnel de mortalité a été observé chez des lynx maintenus en captivité (53). De nombreuses espèces sauvages pourraient être impactées par le virus ou agir comme des réservoirs potentiels, notamment dans les zones où le virus est endémique.

2.1.4. Le cycle d'infection chez les hôtes ruminants

La pathogénèse est probablement proche entre les hôtes ruminants, malgré des différences entre les phénotypes d'infection (54). Le cycle du virus de la FCO chez les hôtes mammifères est initié par la libération de particules présentes dans la salive des femelles culicoïdes qui se nourrissent au niveau de l'épithélium, sur les capillaires sanguins. Les cellules endothéliales des vaisseaux sanguins et des capillaires sanguins constituent des sites privilégiés de réplication pour le virus (55). Certaines cellules circulantes dans le système vasculaire sont sensibles aux infections du virus et pourraient agir comme des transporteurs favorisant la dispersion du virus à tout

l'organisme (51). A partir du point de morsure, le virus circule jusqu'aux ganglions lymphatiques, qui constituent les sites principaux de la réplication initiale. Il s'agit d'un site central de régulation des défenses immunitaires (56). A partir de ces ganglions, le virus est disséminé à l'ensemble de l'organisme par l'intermédiaire des cellules phagocytes mononuclées et des cellules endothéliales. Parmi les cellules circulantes, l'association du virus avec les érythrocytes est supposée être à l'origine des virémies prolongées. En effet, les virus attachés ne sont pas détectés par le système immunitaire (57). Cette association FCO-érythrocyte pourrait également jouer un rôle dans l'infection des espèces vectrices. Les cellules endothéliales sont à l'origine des principaux symptômes visibles, les cellules lysées induisent l'apparition d'œdèmes multiples, particulièrement visibles au niveau des muqueuses.

2.2. Modes de transmission alternatifs du virus de la FCO

Parmi les symptômes provoqués par le virus de la FCO, les malformations congénitales et avortements sont fréquemment observés pour les ruminants. La propagation du virus au placenta est associée à certaines souches et semble être facilitée pour les virus réassortis avec les vaccins atténués produits en culture cellulaire (58). La consommation du placenta contaminé peut aussi mener à l'infection d'un hôte ; des cas d'infection d'une génisse (séronégative 2 semaines précédant l'infection) et de veau au BTV de sérotype 8 ont été rapporté en période hivernale et en l'absence de circulation du virus (59). Des transmissions verticales en condition d'isolement sont aussi décrits et associés à la consommation de colostrum par les veaux (60). La transmission placentaire du virus a été validé expérimentalement sur mouton et vache (61,62). La transmission sexuelle de la maladie entre individus adultes est possible, la réémergence du sérotype 8 en France est associé à l'insémination artificielle de femelle à partir de semences contaminées (42). Enfin le sérotype 27, décrit récemment, présente des aptitudes nouvelles et semble être transmissible par contagion (45).

La transmission verticale du virus de la FCO est soumise à débat. En l'état, il existe des travaux qui rapportent son existence, avec des pools de larves positives au virus (63) tandis que des expérimentations plus récentes échouent à le démontrer (64). Des preuves de la présence du virus dans des larves soutiennent la capacité du virus de

la FCO à être transmis verticalement chez le culicoïdes. Cependant, ces expériences démontrent que ce mécanisme pourrait être rare, dépendant du virus, des sérotypes ou d'autres facteurs (saison, température).

3. Structure du virus et cycle de réplication

3.1. La particule virale du virus de la FCO

Le virus de la FCO possède une capside de symétrie icosaédrique d'environs 90 nm, elle est constituée de 3 couches de protéines virales (Figure 2), typique de la famille des *Reoviridae* (65). Son génome est constitué de 10 segments ARN double brin. Les particules virales sont constituées de 7 protéines structurelles organisées en deux couches concentriques : la capside externe et la capside interne. La couche externe est constituée des protéines VP2, qui forment des structures à triple ramification, et VP5 qui possèdent une forme globulaire. VP2 est la protéine de capside la plus variable du virus de la FCO (66), cible principale des défenses immunitaires, elle permet de définir le sérotype et de distinguer les différentes souches du virus. Des réponses antigéniques aux protéines de capside VP5 et VP7 existent également (67, 68). La bicouche capsidique externe pourrait être dégradée partiellement ou intégralement, lors de la circulation du virus dans l'environnement extracellulaire. La partie la plus interne de la capside est constituée par la protéine VP3 ; elle est en contact direct avec les complexes de transcriptions porteurs des segments génomiques. Les complexes de transcription sont constitués de 3 protéines : VP4, rattaché à la partie interne de la capside, VP1 et VP6. Ce complexe permet la condensation du génome et sa transcription au sein de la particule virale. Dans une particule parfaite se trouvent 10 complexes de transcription auxquels est attaché un exemplaire de chacun des segments génomiques (69). Cependant l'espace et la conformation interne au sein de la particule virale de la FCO pourrait permettre l'encapsidation de segments surnuméraires (70). Les caractéristiques des 10 segments génomiques du virus de la FCO-4 sont présentées dans le Tableau 2, les fonctions des différentes protéines structurelles ou non sont présentées dans le paragraphe suivant.

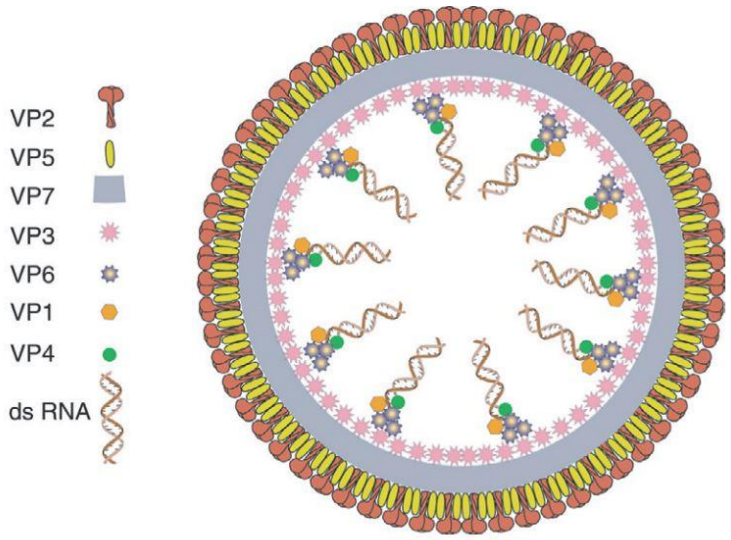


Figure 2. **Représentation schématique en coupe transversale de la particule virale typique du virus de la FCO.** Les 7 protéines structurelles et les segments génomiques ARNdb sont représentés en son centre. Les protéines VP2 et VP5 constituent la couche externe de la capside virale, tandis que les protéines VP7 et VP3 constituent la couche interne. La particule lestée de la capside externe est dénommée particule core. La figure est tirée, sans modification, de l'article de revue de Schwartz-Cornil et al., (71).

Tableau 2. Segments génomiques du virus de la FCO, taille (FCO-4), noms et fonction principale des protéines virales.

Segment n°	Protéine	Taille (pb)	Fonction
1	VP1	3944	RNA dependant RNA Polymerase
2	VP2	2926	Capside I, attachement à la cellule
3	VP3	2772	Capside III, interne
4	VP4	1981	Enzyme protection des extrémités ARN
5	NS1	1774	Formation de tubules cytoplasmique
6	VP5	1637	Capside I, fusion avec la membrane cellulaire
7	VP7	1156	Capside II
8	NS2	1125	Matrice corps d'inclusion viral
9	VP6, NS4	1039	Helicase, réponse ARNi
10	NS3, NS3a	822	Glycoprotéines, sortie de la cellule

3.2. Cycle d'infection cellulaire et réplication

Le mécanisme le mieux décrit et vraisemblablement le plus commun d'attachement de la particule virale intacte se fait par l'intermédiaire de la protéine de capside externe VP2 (72, 73). Cette dernière se lie à des protéines membranaires de la cellule, de nature inconnue, ces protéines cellulaires sont caractérisées par la présence de clathrine adjacente (74). La voie principale d'entrée de la particule virale dans la cellule est celle de l'endocytose par l'intermédiaire des clathrines (73). La membrane cellulaire s'invagine jusqu'au détachement d'une vésicule contenant la particule virale dans le cytoplasme. Cet endosome perd rapidement la couche externe de clathrine, puis, sous l'effet du pH, la protéine de la capside VP5 interagit avec la membrane vésiculaire et la rend perméable (75). Ces réactions mènent à la libération de la particule core, dont la couche externe est constituée de VP7, dans le cytoplasme. La majorité des virions présents dans le cytoplasme sont des particules core, leur présence est détectée environ 1 heure après l'initiation de l'infection (76). La particule core initie la transcription une fois relâchée dans le cytoplasme, sous l'action d'ions Magnésium et de NTP. Les couches de la capside VP7 et VP3 changent de conformation et permettent l'ouverture de pores. Les complexes de transcription

composés de l'hélicase, de l'enzyme de protection des extrémités 5' et de la polymerase virale initient la transcription à partir du brin négatif. L'intégrité de la capside est déterminante dans le succès de la transcription, les ARNdb sont maintenues au sein de la particule core et les ARNm synthétisés correspondent à la séquence complète du brin ARN positif. Les ARN possèdent un chapeau en 5' qui prévient la dégradation des extrémités. Sous l'influence de la protéine NS2, des corps d'inclusion viraux se forment à proximité du noyau. Les protéines du complexe de transcription et les protéines de la particule core sont trouvées au sein du corps d'inclusion. La protéine NS2 possède également une forte activité de liaison aux ARNsb qui favorise l'interaction avec les ARNm viraux au détriment des ARN non spécifiques. Les étapes qui mènent à l'assemblage des particules core, au sein du corps d'inclusion, ne sont pas clairement définies. Les segments possèdent des séquences conservées à leurs extrémités 5' et 3' ; un réseau d'interactions est aussi possible grâce à des motifs présents sur l'ensemble des segments. Il est possible que ces interactions mènent à la formation de macros complexes ARN qui pourraient initier l'assemblage des particules virales. La réplication du brin négatif pourrait se faire au sein des particules core comme cela a été démontré pour le genre *Rotavirus*, appartenant à la même famille que le virus de la FCO. Récemment, il a été démontré que les protéines VP5 et NS3 sont aussi présentes au niveau du corps d'inclusion ; elles pourraient jouer un rôle dans la maturation et la sortie non-lytique des virions (77). Les auteurs suggèrent que la protéine NS3 pourrait être à l'origine d'une modification structurelle du réticulum endoplasmique menant à son association avec le corps d'inclusion viral (77). Dans le cytoplasme, les particules core acquièrent les protéines de la capside virale VP2 potentiellement par l'intermédiaire de filaments intermédiaires de vimentine. Les particules complètes peuvent sortir de la cellule par différents mécanismes (Figure 3). La sortie non lytique par l'intermédiaire des protéines NS3, NS3a qui perméabilisent les membranes cellulaires et mènent à la formation du complexe vésiculaire C'est la voie de sortie principale du virus dans les cellules d'insecte. La sortie non-lytique est également possible au début de l'infection pour les cellules mammifères. Pour ces dernières, la libération des particules virales est principalement associée à la lyse de la cellule par nécrose.

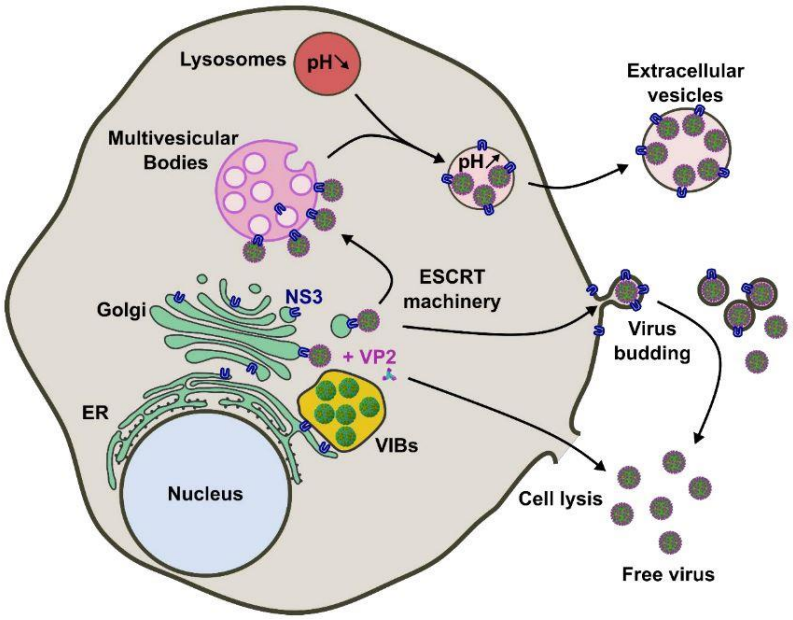


Figure 3. Représentation schématique des différentes voies de sortie du virus de la FCO infectant des cellules mammifères. Après l'assemblage des particules virales dans le corps d'inclusion la protéine NS3 (en bleu) assure leur transfert vers un corps multi-vésiculaire. Les virions migrent ensuite dans un lysosome dont le pH permet le maintien de l'intégrité membranaire, il assure la sortie du virus sous la forme de vésicules extracellulaires. Les virions peuvent également sortir sous forme enveloppée suite à l'interaction entre NS3 et le complexe ESCRT. Enfin la lyse cellulaire entraîne la sortie de particule virales complète et libres. L'illustration est tirée, sans modifications, de l'article de revue de T. Labadie (78).

3.3. Variation du cycle d'infection dépendante de l'hôte

Des différences dans la liaison et dans l'entrée cellulaire existent en fonction du type de cellules infectées (73). Le clivage de la protéine VP2 induit une augmentation importante de l'infectivité dans les cellules d'insecte et chez les culicoïdes adultes (79). La particule reste infectieuse pour les cellules mammifères, sans changer sa virulence (79), mais elle ne peut plus se lier aux érythrocytes. L'augmentation de l'infectivité de ces formes de virions pourrait jouer un rôle important pour l'infection des deux hôtes. Les virions pourraient notamment être soumis à l'effet de protéases dans le tube digestif des culicoïdes (79). La particule core est aussi infectieuse pour les culicoïdes adultes, sans changer sa capacité d'infection (79). Certaines cellules mammifères sont aussi sensibles à l'infection par la particule core mais cette dernière est associée à une baisse d'infectivité (79). Ces observations suggèrent l'existence de différents mécanismes permettant l'entrée des particules virales dans les cellules. Dans la

famille des Reoviridae, le genre des *cypoviridae* présente des virions qui ne possèdent pas de capside externe et sont spécifiques des hôtes insecte (80). Il est possible que l'acquisition de la couche externe de la capside soit le résultat d'une adaptation aux cellules mammifères.

Une différence majeure du cycle d'infection cellulaire entre insectes et mammifères réside dans l'absence de lyse cellulaire chez les insectes. La sortie non lytique de la cellule est dépendante des protéines NS3 et NS3a codées par le segment 10, ce dernier joue un rôle central dans la pathogénicité des souches virales de la FCO. Son rôle dans l'inhibition de la réponse interféron de type I a été démontré dans des cellules mammifères (81). Lors d'une expérience visant à évaluer l'influence de la virulence des différents segment génomiques, le segment 10 a été montré comme central en modulant la virulence pour le sérotype 8 du virus de la FCO (82). Dans le type cellulaire BHK-21, l'absence des protéines NS3 n'empêche pas la réplication du virus (83), elle se traduit en revanche par une diminution importante de la réplication virale dans les cellules d'insectes. Enfin les mutations présentes sur le segment 10 peuvent influencer la production de la protéine NS3 en fonction du type cellulaire, insectes ou mammifères, *in vitro* (84).

4. Génétique évolutive du virus de la FCO

4.1. L'évolution du virus de la FCO soumise à un régime de sélection purifiante dépendant des segments (ou des conditions expérimentales)

Comme pour la majorité des arbovirus, l'évolution du virus de la FCO est soumise à une pression de sélection purifiante. Elle est une conséquence supposée de son cycle obligatoire qui implique l'alternance entre deux hôtes appartenant à des espèces distantes. Cela se traduit par une diminution du nombre de substitutions, et plus particulièrement des substitutions non-synonymes (85). Certains segments semblent particulièrement affectés par cet effet, le segment 3 possède la séquence génomique la plus conservée parmi les 10 segments génomiques du virus de la FCO. Le gène du segment 3 est utilisé pour distinguer les topotypes du virus qui permettent d'assigner une origine géographique ancienne (Est ou Ouest) aux différents sérotypes européens (86). Plus généralement, le taux de substitution moyen par nucléotide (pour

son estimation la plus récente) chez le BTV est compris dans un intervalle allant de 1.46E-02 (segment 2) à 7.81E10-4 (segment 9) substitutions de nucléotide par base, par année (subst/site/année) (87). Cette valeur change en fonction des segments, de la géographie et des sérotypes étudiés (48, 92, 93). Les travaux de Carpi *et al.* (88), qui se concentrent sur l'étude des segments 2, 3, 6 et 10, présentent des valeurs de taux de substitution moins élevés, allant de 6.90E10-4 (segment 10) à 5.22E-05 (segment 6) subst/site/année. Le taux de substitution adaptative (dN/dS) indique ici un effet marqué de sélection purifiante sur le segment 6 (dN/dS = 0.032) et fait écho à la valeur faible du taux d'évolution (5.23E-05 substitutions/site/année) pour ce segment. En revanche le segment 2 présente un taux d'évolution plus élevé (4.37E-04 substitutions/site/année) et un effet moins marqué de sélection purifiante (dN/dS = 0.122). Le résultat le plus surprenant concerne le segment 10 qui, malgré un effet marqué de sélection purifiante (dN/dS = 0.060), possède un taux d'évolution élevé (6.94E04 subst/site/année).

Sur le continent Indien, parmi les 20 sérotypes circulants, les gènes portés par les segments 2 et 6 qui codent respectivement pour les protéines externes de la capside virale VP2 et VP5 possèdent un taux d'évolution supérieur aux autres gènes (respectivement de 1.46E-02 et 9.36E-03 subst/site/année). Les protéines virales de capside ou d'enveloppe régissent les interactions avec le système immunitaire de l'hôte. Chez un autre virus à génome segmenté, la grippe A, la variabilité génétique a également été montré pour le gène de l'Hémagglutinine et de la Neuraminidase (89). Pour les sérotypes indiens, les valeurs des D de Tajima diffèrent de la neutralité pour les deux gènes codants pour les protéines externes de la capside mais aussi pour le segment 7 qui code pour la protéine de capside de la particule core. Les trois gènes codant pour les protéines de capside n'évoluent pas selon un processus aléatoire. Les autres gènes semblent être soumis à une combinaison d'effets neutres et purifiants de la sélection. Ainsi les travaux plus récents sur le virus de la FCO évaluent le taux d'évolution moyen des segments génomiques à des valeurs proches de celles observées chez les arbovirus possédant un génome ARNsb (90). A l'échelle Européenne, pour les sérotypes 1 et 8, 10 mutations ont été trouvées sur le segment 9, elles marquent potentiellement un phénomène d'évolution adaptative suite à un épisode de réassortiment et sont les seules mutations détectées sous sélection positive (91). Cette approche montre que les dN/dS les plus forts sont observés pour

les segments 9 (pour l'orf VP6, dN/dS = 0.333), en comparaison au gène codant pour VP2 (dN/dS = 0.160) et NS2 (dN/dS = 0.109) (91).

Enfin des travaux récemment publiés par Kopanke *et al.* (92) présentent l'évolution du virus de la FCO en condition *in vitro*. Dans cet environnement, les pressions de sélection sont faibles. Leurs résultats soulignent une absence d'effet du type cellulaire ou du passage en alternance sur la génération de SNV. Des mutations peuvent apparaitre sur chacun des 10 segments génomiques. En revanche, les segments 5 et 10 présentent un taux de SNV/CDS plus important qu'aucun autre segment (Figure 4) qui se traduit par une divergence plus importante par rapport à l'inocula (Fst = 0.67 pour le segment 5 et Fst = 0.73 pour le segment 10). Une augmentation des ratios de dN/dS est aussi observée par rapport aux estimations faites sur les populations naturelles marquant des effets de sélection dépendant des segments (Figure 5). Ce résultat souligne la différence majeure entre les conditions naturelles d'évolution du virus et les conditions des systèmes expérimentaux *in vitro* pour le virus de la FCO. Enfin les auteurs soulignent l'existence d'insertions/délétions (indels) qui sont apparus sur chacun des segments à une fréquence environ dix fois inférieure à celle de l'apparition de SNP. Certaines de ces indels avaient la caractéristique particulière de réapparaitre dans une forme identique indépendamment des échantillons ou des passages (92).

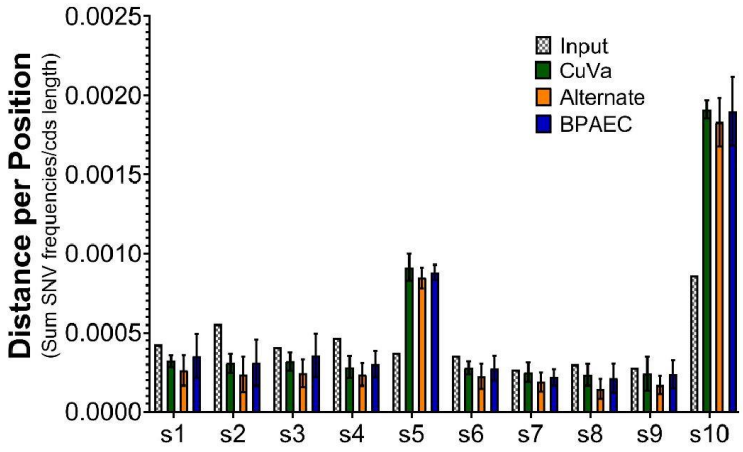


Figure 5. Distance génétique par segment pour la FCO-17 entre les conditions de culture cellulaire. Cette figure présente la somme des fréquences de chaque variant nucléotidique par segment normalisé par la longueur de la séquence codante (SNV/CDS) pour chaque réplicat, pour les passages 3, 6 et 9 sur des cellules de Culicoïdes sonorensis (CuVa) en vert, des cellules Bovines (BPAEC) en bleu, l'alternance entre ces deux types cellulaires en jaune et la valeur pour l'inocula en noir. Cette figure est tirée de la publication de Kopanke et al. (92), sans modification.

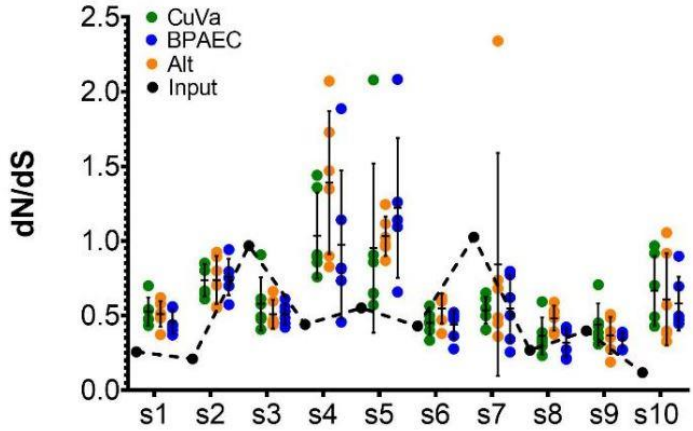


Figure 4. Effet de la sélection sur la nature des mutations des 10 segments de la FCO-17 en condition de culture cellulaire. Cette figure présente les ratios de dN/dS pour chaque réplicat, pour les 10 passages successifs sur des cellules de *Culicoïdes sonorensis* (CuVa) en vert, des cellules Bovines (BPAEC) en bleu, l'alternance entre ces deux types cellulaires en jaune et la valeur pour l'inocula en noir. Cette figure est tirée de la publication de Kopanke et al. (92), sans modification.

Une certaine hétérogénéité entre ces résultats complique leur extrapolation aux modèles du virus de la FCO en général. Une distinction doit être faite entre les observations de terrain et les expérimentations *in vitro*. En condition naturelle, nous pouvons souligner la variabilité constamment décrite du segment 2 au travers de ces approches qui justifie son utilisation comme discriminant entre sérotype (66). Dans l'ensemble de ces travaux, les taux de substitution sont strictement inférieurs à 1 et indique l'effet d'une pression de sélection purifiante s'exerçant sur l'ensemble des segments génomiques. Un point critique de ces résultats tient à l'utilisation de séquences consensus pour l'ensemble des analyses. Malgré la rigueur méthodologique des approches, elles pourraient masquer une partie des variations génétiques présentes au sein de populations naturelles. En condition d'infection expérimentale *in vitro*, la pression de sélection est relâchée, les segments 5 et 10 sont les plus divergents et des indels caractéristiques et reproductibles apparaissent.

4.2. Le réassortiment, un mécanisme évolutif complexe

La segmentation des génomes présente plusieurs avantages pour les populations virales. Par exemple, la séparation physique des gènes a permis une augmentation de la taille des génomes malgré l'absence de correction des erreurs de réplication de la polymérase, grâce au mécanisme de réassortiment (85). Mais le bénéfice principal associé à la segmentation des génomes est lié à la possibilité de réassortiment entre les souches lors de la co-infection d'une même cellule. Le réassortiment permet le mélange génomique entre plusieurs virions. Pour le virus de la grippe A (IAV), il a été associé à des changements de gamme d'hôte et à l'émergence de nouvelle souches virales (93). Ce processus favorise la propagation des segments porteurs d'adaptations et l'élimination de segments qui ont un effet délétère, en ce sens le réassortiment constitue un type d'échange génétique proche de la reproduction sexuée, avec un mélange (cependant non-aléatoire) des gènes viraux parentaux. Chez le virus de la FCO le réassortiment constitue indéniablement un mécanisme évolutif commun qui est à l'origine de l'émergence de plusieurs souches (94).

Le groupe de K. Nomikou s'est récemment intéressé au phénomène du réassortiment chez BTV à l'échelle Européenne. Ils ont montré que l'association entre les segments

au cours du réassortiment était un phénomène commun, non aléatoire, qui pouvait être associé à une augmentation de la probabilité de mutations non-synonymes (91). Le réassortiment peut être à l'origine de différences génétiques au sein d'un même sérotype. Pour son approche, le groupe de K. Nomikou montre que des évènements de réassortiments se sont produits pour l'ensemble des virus séquencés. Certains segments sont plus fréquemment acteurs du réassortiment comme les segments 7 et 10 qui varient indépendamment des autres segments. En revanche, les segments 2 et 6 semblent être associés lors de ces évènements (Figure 6) (44, 93). Une approche similaire a été conduite pour décrire le réassortiment à l'échelle du continent Indien (92). Des patterns différents d'association entre les segments sont observés entre les 20 sérotypes différents circulants sur le continent. Le groupe de segments 1, 2, 6 et 7 et le groupe de segments 3, 9 et 10 ont une probabilité inférieure de réassortiment que les segments 4 et 5. Leurs résultats mettent en avant la sélection d'un génotype particulier du segment 5 qui s'est propagé dans la majorité des souches virales étudiées. Néanmoins en l'absence de preuve d'une sélection positive pour ce segment, les auteurs avancent deux hypothèses principales pour expliquer sa dissémination : un changement dans l'environnement qui a favorisé la sélection de ce variant ou un phénomène d'épistasie qui impliquerait son interaction avec d'autres segments.

Le réassortiment est un mécanisme au cœur de l'évolution du virus de la FCO. Des interactions existent entre les segments qui favorisent leur association au cours de ces épisodes. Il existe des différences importantes dans les groupes de segments associés en fonction des sérotypes ou de la zone géographique étudiée.

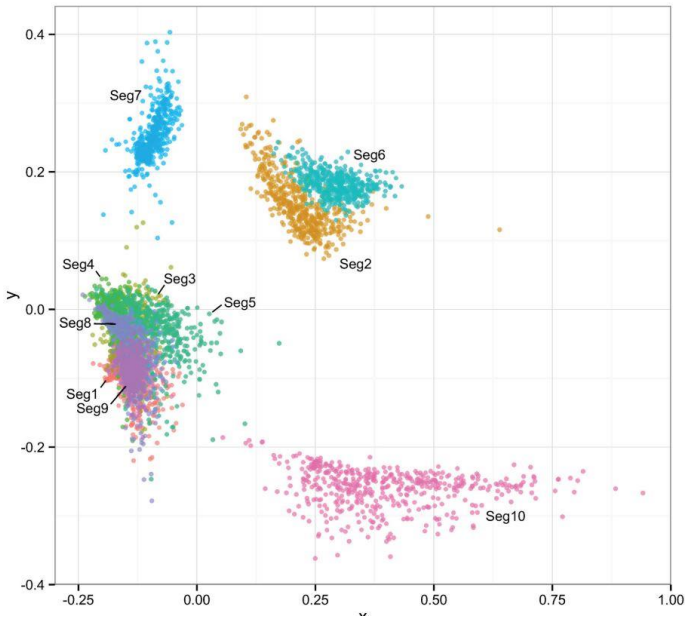


Figure 7. MDS présentant les corrélations de temps pour l'ancêtre commun le plus récent entre paires de virus de la FCO en Europe (sérotypes 1 et 8). Les estimations sont séparées pour chaque segment. Les points représentent l'incertitude sur 500 échantillonnages avec le programme BEAST pour les virus échantillonnés la même année. Cette figure est tirée de la publication de K. Nomikou et al., (91).

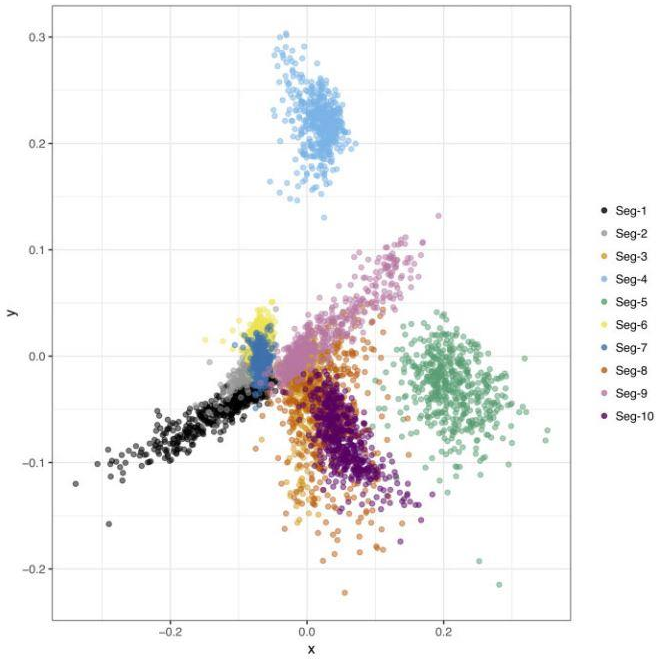


Figure 6. MDS présentant les corrélations de temps pour l'ancêtre commun le plus récent entre paires de virus de la FCO en Inde pour 20 sérotypes différents. Les estimations sont séparées pour chaque segment. Les points représentent l'incertitude sur 500 échantillonnages avec le programme BEAST pour les virus échantillonnés la même année. Cette figure est tirée de la publication de R. Biek et al., (92).

Partie 3 : La variation du nombre de copies de gènes (VNCG) chez les virus

1. Le concept de VNCG et le monde des virus multipartites

La variation du nombre de copies de gènes a été décrite pour la première fois dans le cadre de l'adaptation à différents hôtes chez un virus de plante multipartite du genre *Nanovirus*, le *Faba bean necrotic stun virus* (FBNSV) (95). Les virus multipartites possèdent un génome segmenté divisé en deux à huit segments ADN ou ARN encapsidés individuellement. Ils constituent une part importante des familles virales infectant les plantes et les champignons mais ont aussi été décrit à deux reprises dans des hôtes insectes (96, 97). La compartimentation des segments génomiques a soulevé de nombreuses questions quant aux mécanismes permettant à ces formes virales de produire une descendance en maintenant leur intégrité génétique. Le groupe de Sicard *et al.* a suggéré que ce mode de transmission puisse être bénéfique à la population virale en lui permettant de réguler le nombre de copies de gènes (95). Pour tester cette hypothèse une approche expérimentale a été mise en place pour permettre un suivit des fréquences relatives de chacun des 8 segments du FBNSV au cours de l'infection de différentes plantes hôtes. Leurs résultats montrent que non seulement les gènes du FBNSV s'accumulent à des fréquences différentes mais aussi qu'elles tendent vers une situation d'équilibre, dépendante de la plante hôte et reproductible indépendamment de la fréquence initiale des segments (Figure 8). Leurs résultats suggèrent également que l'état d'équilibre des fréquences observées, baptisé formule génomique, est positivement corrélé à une accumulation du virus plus importante et à des symptômes accrus (Figure 9).

La VNCG avait déjà été décrite chez des virus multipartites, sans pour autant considérer le caractère adaptatif que pouvait revêtir ce mécanisme (98, 99). Suite aux travaux de Sicard et al. (95) au moins deux groupes ont décrit un phénomène de VNCG dépendant de l'environnement chez deux espèces différentes de virus multipartites. Pour le *Bombyx mori bidensovirus* (BmBDV) *qui* possède un génome ADNdb divisé en deux segments (famille *Nanoviridae*), la fréquence des deux segments varie entre l'inoculum, les cellules de l'intestin moyen infectées dans l'hôte

lépidoptère et les virus présents dans les excréments (96). Pour l'*Alfalfa mosaic virus* (famille *Bromoviridae*) qui possède un génome ARN divisé en trois segments, le mécanisme de VNCG a été associé pour ce virus aussi à une formule génomique spécifique vers laquelle convergent les fréquences des trois segments, avec un profil de fréquences sélectionnées dépendant de la plante hôte infectée (100).

Le mécanisme de variation du nombre de copies de gènes existe dans plusieurs systèmes viraux multipartites. Plusieurs hypothèses ont été proposées pour expliquer sa fonction et son rôle dans l'adaptation à différents environnements, dans la littérature ce mécanisme est associé aux mécanismes de régulation de l'expression des gènes chez les eucaryotes ou à divers mécanismes adaptatifs chez les virus.

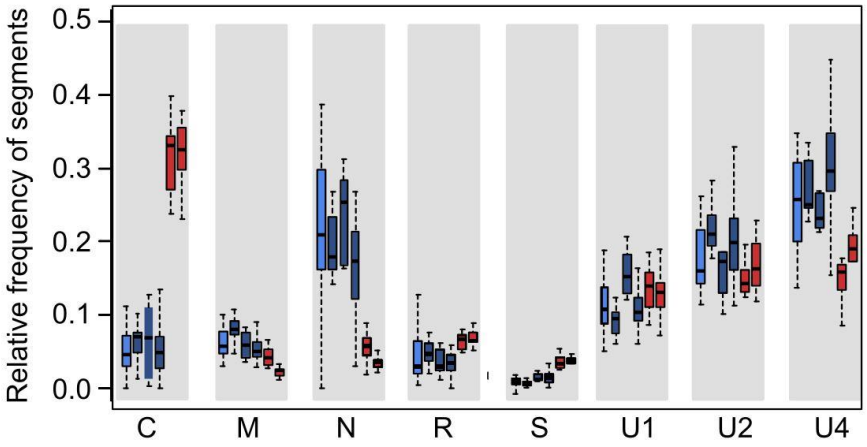


Figure 8. Fréquence relative des segments du FBNSV en fonction des conditions d'inoculation et de la plante hôte. Les infections sur l'espèce *V. faba* sont présentées en bleu clair pour les infections manuelles et en bleu foncé pour les infections par vectorisation par un puceron (3 réplicas indépendants). Les infections sur *M. truncata* par des pucerons sont présentés en rouge (2 réplicas). Le mode d'inoculation n'influence pas les fréquences, en revanche, des différences entre plantes hôtes sont observées pour tous les segments à l'exception de U1. La figure est tirée de la publication de Sicard et al., (95)

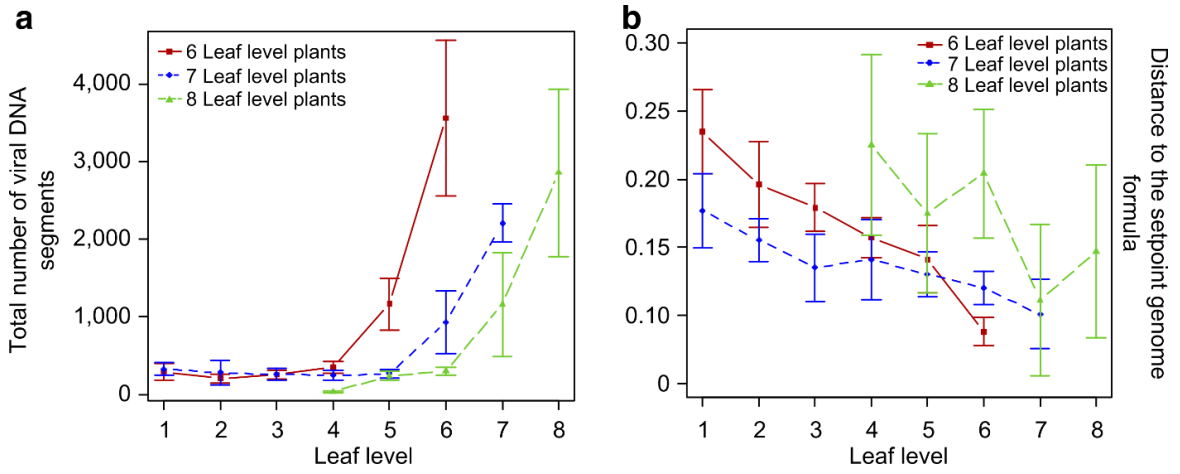


Figure 9. La formule génomique et son implication sur l'accumulation du FBNSV dans les compartiments foliaires. Le graphique a) montre que l'accumulation du virus est dépendante du développement de la plante mais surtout que l'accumulation du virus augmente en suivant le niveau foliaire infecté. b) Un indice calculé permet de définir une distance à la formule génomique (valeur des fréquences vers lesquelles convergent les segments), la valeur de distance à formule génomique à l'équilibre diminue dans les niveaux foliaires les plus hauts qui correspondent aux feuilles où l'accumulation du virus est la plus importante. La figure est tirée de la publication de Sicard et al., (95)

2. Régulation de l'expression et adaptations dépendantes de la VNCG

Dans la littérature, la VNCG est couramment associée aux mécanismes de régulation de l'expression des gènes (101, 102). Elle constitue une des bases de la régulation de l'expression des gènes et peux moduler les voies d'interactions entre gènes (102). En terme fonctionnel, ce mécanisme agit en amont des systèmes de régulation de transcription et de post-transcription. Ce mécanisme peut aussi jouer un rôle dans le développement de pathologies, notamment chez l'Homme (103). Chez le protozoaire *Plasmodium falciparum,* la VNCG permet de compenser des stress environnementaux. Elle influence directement la régulation de l'expression des gènes du parasite (104). Des mécanismes similaires sont décrits chez les bactéries (105).

Chez les virus, au moins deux fonctions adaptatives différentes ont été associées à la variation du nombre de copies de gènes. Une souche vaccinale de poxvirus, qui possède un génome monopartite, est capable d'augmenter le nombre de copies d'un gène particulier impliqué dans la réponse aux facteurs cellulaires sous une condition

d'infection impliquant une pression de sélection spécifique pour cette fonction (106). Le gène qui se duplique peut induire une augmentation de la taille du génome allant jusqu'à 10%. Ce phénomène a été associé à la génération d'une mutation permettant l'adaptation à la nouvelle pression de sélection, l'apparition de la mutation et sa sélection est associée à une perte rapide des copies du gène dupliqué. Ce phénomène séquentiel de déploiement du génome, la génération de mutation suivie d'une contraction du génome, a été baptisé : génome accordéon (106) et apporte un élément nouveau aux mécanismes pouvant favoriser l'adaptation d'un virus à son environnement. Un second mécanisme adaptatif régulé par la variation du nombre de copies intervient dans la définition de l'issue des infections par les Bactériophages λ (107). Ces derniers possèdent un génome ADNdb circulaire. Le nombre de copies de gènes viraux au sein de la bactérie infectée module l'issue de l'infection vers la phase lytique ou lysogénie.

Le mécanisme de VNCG a des implications multiples pour un nombre important d'entités. Son rôle dans la régulation de l'expression n'a jamais été démontré pour un virus, mais une littérature dense présente son lien avec l'expression chez différents organismes eucaryotes ou procaryotes.

3. Une analogie avec la variation du nombre de copies de gènes chez un virus à génome segmenté

Le virus de la grippe A (IAV) possède un génome ARNsb divisé en 8 segments génomiques, qui similairement au virus de la FCO (69), sont considérés comme encapsidés avec une précision qui permet à chaque virion de porter un exemplaire de chaque segment (108). Cette vision est questionnée depuis plusieurs années par différents travaux qui font état de l'absence de certaines protéines virales au sein de cellules infectées par IAV (109) ou qui apportent des preuves d'éléments chimériques portant de larges délétions dans les gènes de IAV (110). Parmi les recherches à l'origine de ces questionnements les travaux de Brooke et al. (111) qui fournissent un exemple robuste de diminution de la présence du gène intègre de la Neuraminidase. Son approche montre que la séquence génomique complète de ce segment est moins présente pour une condition particulière d'infection, avec l'impossibilité de dissocier la

non-incorporation du segment dans son ensemble de la présence d'un segment défectif de la Neuraminidase. Les observations de la diminution de l'encapsidation de ce segment sont reproductibles et se produisent dans des conditions précises (infection des cellules hautes de la trachée de cochons d'Inde). Ces résultats apportent également la preuve que la diminution de la présence du gène est liée à une diminution de son expression et à une diminution de la production des protéines associées. Ce changement est également associé à une augmentation de la valeur sélective de la population virale. Par une approche fonctionnelle, Brooke et al. (111) ont démontré le lien entre la diminution de la présence de ce segment et une mutation portée par le gène de l'Hémagglutinine (la seconde protéine virale de l'enveloppe), fournissant ainsi une explication fonctionnelle au phénomène décrit.

4. La VNCG chez le virus de la FCO ?

Les résultats de Brooke et al. (111) mettent en lumière de nombreux éléments nouveaux qui nourrissent une réflexion plus large sur l'importance que pourrait avoir la VNCG chez les virus à génomes segmentés. Le niveau de segmentation joue certainement un rôle important dans la régulation de l'encapsidation des segments (69), il est moins spécifique pour les génomes avec un nombre de segments plus faible (112). Il a été montré par exemple que le virus de la Fièvre de la vallée du Rift (RVF), un arbovirus avec génome ARNsb divisé en trois segments, possède un système lâche de régulation de l'encapsidation de ses trois segments génomiques (112). Toutes les cellules infectées ne portent pas un set génomique complet de la RVF. Si pour les virus multipartites le niveau de segmentation ne semble pas influencer la capacité de l'espèce virale à générer une formule génomique dépendante des conditions d'infection (95, 100), les connaissances fonctionnelles sur les mécanismes d'encapsidation des génomes des virus segmentés au sein d'une même particule montrent qu'il existe des mécanismes qui contraignent l'encapsidation de sets génomiques complets (69, 108). Ces différences sont associées aux probabilités d'infection d'un génome viral complet qui est nécessaire à la production de particules virales fonctionnelles et au déroulement du cycle d'infection cellulaire. Les segments des virus qui possèdent un génome fortement segmenté peuvent interagir entre eux notamment grâces aux extrémités conservées en 5' et en 3'. Pour la FCO, d'autres

sites d'interaction existe dans les parties codantes des gènes (113). Les travaux de Brooke et al. (111) soutiennent que pour IAV, un virus avec un fort niveau de segmentation (et des mécanismes moléculaires régissant l'intégration des segments viraux au sein du virion, proches du virus de la FCO), les interactions inter-segments pourraient être modulées par des mutations présentent dans les séquences codantes. Des éléments nouveaux apportés par Labadie et al. (114) pourraient également faciliter la co-infection et la complémentation entre génome pour le virus de la FCO. Le virus de la FCO peut être transmis sous la forme de vésicules extracellulaires portant plusieurs particules virales et formant ainsi la première forme d'unité infectieuse multiple décrite chez un arbovirus à génome segmenté.

Partie 4 : Stratégies, méthodologie et objectifs du projet

1. Approches méthodologiques et conceptuelles pour l'étude de la VNCG et de l'adaptation du virus de la FCO-4

Les travaux présentés au cours de ce manuscrit ont impliqué un développement méthodologique important. L'essentiel du travail mené au cours de la thèse est associé au développement des approches méthodologiques dans l'objectif de tester l'existence de deux mécanismes adaptatifs potentiellement liés à l'adaptation au cycle en alternance d'hôte pour le virus de la FCO de sérotype 4. Nous voulions tester deux hypothèses principales, pour déterminer si **(H1) l'adaptation du virus à l'hôte pouvait se traduire, après chaque évènement de transmission, par la génération de mutations spécifiques à l'hôte et dès lors si le taux d'évolution permettait de rétablir ces génotypes à chaque étape du cycle dans les conditions naturelles d'infection**. De manière non-exclusive nous voulions étudier **si (H2) la variation du nombre de copies de gènes pouvait exister chez le virus de la FCO-4, agir comme un système de régulation de l'expression des gènes au niveau populationnel et jouer un rôle dans l'adaptation du virus de la FCO-4 à l'alternance d'hôte**. Pour étudier la génétique du virus dans les conditions naturelles d'infection (H1) nous avons développé une méthode de préparation de librairies spécifiques au virus de la FCO-4. Un pipeline d'analyse bio-informatique a été développé pour préparer les données à l'analyse en considérant les spécificités méthodologiques de la préparation des librairies. L'étude du mécanisme de variation du nombre de copies de gènes (H2) à lui nécessité le développement d'une approche de RT-qPCR spécifique au virus de la FCO-4 en s'assurant de l'absence de biais au cours des différentes étapes de préparation et de quantification. Afin de traduire le temps impliqué au cours du projet au développement de ces méthodes, je vais présenter ces méthodes finales et les principaux facteurs qui ont guidé leur application en tant que telles.

1.1. VNCG et RT-qPCR

1.1.1. Analogies entre approches

Des approches de qPCR et de RT-qPCR pour la quantification précise des segments génomiques de virus dans le cadre du test de l'hypothèse de la variation du nombre de copies de gènes ont été développées pour plusieurs modèles viraux multipartites (95, 96, 100). La méthode de quantification utilisée est le SYBR-GREEN développée par ROCHE. Sa simplicité d'application a permis une optimisation avancée des paramètres de la réaction de qPCR et le test d'un nombre important de primers candidats pour les quantifications. En termes de compromis qualité-prix, la méthode SYBR-GREEN nous a permis d'analyser un nombre massif d'échantillon pour chacune de nos cibles en ajoutant des contrôles continus de mesure pour chaque quantification. Plus concrètement nous avons estimé le nombre de réactions nécessaires à l'obtention des résultats finaux du chapitre 1 à plus de 8000 sans prendre en compte les ajustements et le développement. L'analyse des variations se fait avec les fréquences, calculées à partir de nombres de copies absolues. Pour chaque échantillon, le nombre de copies détectées est contrôlé par l'intermédiaire d'un plasmide portant chaque cible en quantité équimolaire, il est utilisé pour la gamme et comme contrôle pour chaque réaction de qPCR. Les méthodes de quantification sont donc similaires aux approches développées pour les virus multipartites.

1.1.2. Application au virus de la FCO

Le virus de la FCO possède un génome ARNdb divisé en 10 segments distincts. Pour la transcription le virus de la FCO utilise le brin négatif de l'ARNdb comme matrice. Les ARNm sont complémentaires du brin négatif pour chacun des 10 segments (72). Notre stratégie consiste à cibler le brin génomique négatif de chaque segment avec un primer unique à l'étape de RT. Ce brin n'est présent qu'au sein des particules virales assemblées ou aux étapes ultimes de l'assemblage des particules virales. La conceptualisation du mécanisme d'encapsidation prévoit une interaction entre chacun des brins positifs au sein du corps d'inclusion virale comme facteur déclenchant l'assemblage de la particule ou un système de transport des brins génomiques positifs dans des particules virales préformées. Dans ces deux conditions,

la séquence du brin négatif correspond aux génomes encapsidés au sein de particules potentiellement transmissibles.

1.1.3. Méthode, contrôle et optimisation

Pour les primers sélectionnés (Chapitre 1, Tableau 2) une série de validations a été effectuée afin de contrôler la robustesse de chaque couple. Les premières validations ont été faites à partir du virus de la FCO-4 amplifié en culture cellulaire. L'efficacité des amplifications a été contrôlée pour chaque primers à partir d'un extrait ARN dilué en gamme avec une concentration constante en ARN pour chacun des 6 points testés (5 réplicats par point). A cette étape les valeurs de r^2, de pente et des efficacités ont été validé sur matrice ARN et les conditions de réactions optimisées. La spécificité des amplifications a ensuite été contrôlée pour chaque couple de primers par RT-qPCR sur chaque matrice exploitée au cours du projet (in vivo et vitro, Chapitre 1, Tableau 2). Une séquence portant chacune des 10 cibles de qPCR a été insérée dans un plasmide Pgemt par la société EUROGENETEC. L'efficacité a été validée pour chaque couple sur le plasmide à l'étape de qPCR. Les conditions de concentration des primers à l'étape de qPCR ont été optimisées pour améliorer l'efficacité des réactions. Les conditions de dénaturation des ARNdb ont été ajustées afin de favoriser la quantification du virus notamment à partir de matrice complexe. Au passage des premiers échantillons de terrain, des pics doubles de courbe de dissociation sont apparus pour 2 des 10 couples, nous obligeant à valider de nouveau cette approche pour de nouveaux couples. Les effets d'amplifications indépendant de l'amplification des primers ont initialement été contrôlés pour chaque couple de segment et échantillon. La limite de détection n'a pas été calculée pour nos primers, en revanche, une perte de précision entre réplica pour les Ct supérieurs à 36 ainsi qu'une diminution de la spécificité au-delà de 35 cycles d'amplification nous ont conduit à n'inclure que les échantillons possédant un Ct inférieur à 31 dans nos analyses.

1.2. Préparation de librairies pour le séquençage spécifique du virus

Ces aspects sont présentés sous une forme simplifiée ici. Le séquençage des échantillons a été retardé par le contexte particulier de la fin de ce projet de thèse. Nous ne sommes malheureusement pas en mesure de présenter ces résultats au cours de ce manuscrit

1.2.1. Origine de la méthode

Pour cette méthode notre objectif principal visait à séquencer des populations sauvages du virus de la FCO sans passer par une étape d'isolation en culture cellulaire. Nous avons analysé les résultats d'un premier séquençage afin de valider notre approche. Notre approche méthodologique se base sur une méthode de préparation des librairies spécifiques au virus qui avait déjà été appliquée par notre groupe (présentation poster au congrès ASV, « Spatial, temporal and host induced genomic variations in *Alphamesonivirus*1 species », 2018) sur l'espèce Alphamesonivirus-1 (Ordre des *Nidovirales*, génome linéaire ARN positif). Elle consiste à amplifier le génome ARN au cours d'une étape de RT avec des primers spécifiques qui criblent l'ensemble du génome. Elle permet de limiter le nombre de cycle de PCR à 4 au cours de la préparation des librairies et permet ainsi le séquençage de variants minoritaires.

1.2.2. Application au modèle et validation de l'approche

Pour le développement de cette méthode et son application spécifique au virus de la FCO-4, nous avons utilisé les séquences consensus obtenues pour ce sérotype par l'unité de virologie de l'ANSES (115). Des primers spécifiques ont été conçus environ toutes les 200 paires de base afin de couvrir l'intégralité du génome du brin négatif et des extrémités. Pour s'assurer de la robustesse de notre méthode nous avons comparé plusieurs résultats de séquençage à partir d'un même isolat viral. Trois échantillons d'une même culture cellulaire infectée (A, B et C) ont été extraits en parallèle à trois dilutions différentes (10-1, 10-3 et 10-4). Chacun des 9 extraits a servi à préparer une librairie et décrire l'efficacité de notre méthode. Nous avons obtenu une couverture complète du génome avec une spécificité supérieure à 98% de reads pour

toutes les dilutions. Des différences de profondeurs sont observées en fonction des dilutions, moins entre réplicas. Nous avons modifié notre approche à l'obtention de ces résultats en diminuant la concentration des primers utilisés en 3' et en désignant de nouveaux primers pour les zones du génome avec une couverture inférieure à 100X.

2. La variation du nombre de copies de gènes chez le virus de la FCO

Au cours de ce manuscrit je vais m'attacher à présenter différents aspects du phénomène de variation du nombre de copies de gènes qui pourraient avoir un lien avec l'adaptation du virus de la FCO au cycle en alternance d'hôte. Certains des objectifs présentés ci-dessous ont dû être abandonnés, faute de temps. Ces paragraphes précisent les questions et méthodes qui leurs sont associées dans le manuscrit. Les Chapitres correspondant sont renseignés à la fin de chaque paragraphe.

2.1. L'existence d'un phénomène

Chez les virus, la variation du nombre de copies de gènes était jusqu'à peu un mécanisme inféodé aux virus multipartites. Notre premier objectif a donc été de tester si les segments génomiques de la FCO pouvaient varier en fréquence au sein d'une population virale. Nous avons décidé de tester l'existence de ce mécanisme à partir d'échantillons de terrains collectés au cours de l'épizootie la plus récente du virus de la FCO-4 qui a eu lieu en Corse et en Sardaigne. Ces résultats ont fait l'objet d'une publication dans Journal of Virology. Elle constitue le premier Chapitre de ce manuscrit et sera exploitée dans les Chapitres 2 et 3.

2.2. L'effet de l'environnement, implication des hôtes

Les échantillons de terrain collectés lors de l'épizootie correspondent à trois hôtes naturels différents du virus de la FCO-4. Des échantillons sanguins ont été collectés pour les moutons et les vaches, des culicoïdes ont été triés et rassemblés par pools de 50 femelles culicoïdes pares de l'espèce *C. imicola*. Nous avons ainsi pu

comparer les fréquences de chacun des 10 segments génomiques du virus de la FCO-4 en fonction de l'espèce hôte infecté. Ces résultats sont inclus dans la publication citée précédemment et présentés dans le chapitre 1.

2.3. Régulation de l'expression

Nous avions prévu d'étudier l'expression de chacun des gènes du virus en utilisant une approche de qPCR similaire à celle utilisée pour quantifier les segments génomiques mais ciblant le brin positif de l'ARN viral. L'objectif était de tester si les profils de variation des fréquences des segments génomiques pouvaient être liés à une régulation de l'expression dépendante de l'hôte. Nous avons testé cette approche sur les échantillons de terrain mais malheureusement le changement des primers utilisés à l'étape de RT n'a pas permis de valider la précision de la méthode et nous avons manqué de temps pour optimiser la quantification des brins positifs du virus de la FCO-4.

2.4. Influence sur la valeur sélective

Chez les virus multipartites, une corrélation entre la formule génomique des segments à l'équilibre et leur nombre de copies cumulé a été établi. Tester l'influence de la formule génomique sur la valeur sélective s'est avéré plus compliqué pour notre modèle. Nous avons utilisé des échantillons issus d'une infection expérimentale de *Culicoides imicola* avec le virus de la FCO-4 amplifié en culture cellulaire. Les corps des culicoïdes ont été traités individuellement et séparés en deux groupes en fonction de la dissémination du virus jusqu'à la tête ou non. La dissémination du virus jusqu'à la tête constitue un proxy de valeur sélective positive puisqu'il s'agit d'une condition *sine qua non* à la transmission du virus. Les formules génomiques des populations virales ont été comparées en fonction de cet état et leurs variations par rapport aux fréquences des segments de l'inocula ont également pu être étudiées. Ces résultats sont présentés au cours du chapitre 2.

2.5. Mécanismes régulateurs à l'origine de la VNCG

Le caractère aléatoire ou déterministe de la génération de la formule génomique a été évalué au cours d'une infection expérimentale en culture cellulaire. Deux conditions de multiplicité d'infection : faible et forte, ont été comparées sur deux types cellulaires correspondant aux hôtes naturels. Le changement de fréquence des segments génomiques a été évalué après un cycle de réplication virale. Ces résultats sont présentés dans le Chapitre 3.

REFERENCES

1. Hubalek Z, Rudolf I, Nowotny N. 2014. Arboviruses pathogenic for domestic and wild animals. Adv Virus Res 89:201-75.
2. Ciota AT, Kramer LD. 2010. Insights into arbovirus evolution and adaptation from experimental studies. Viruses 2:2594-617.
3. Ciota AT. 2019. The role of co-infection and swarm dynamics in arbovirus transmission. Virus Res 265:88-93.
4. Thompson R, Martin Del Campo J, Constenla D. 2020. A review of the economic evidence of Aedes-borne arboviruses and Aedes-borne arboviral disease prevention and control strategies. Expert Rev Vaccines 19:143-162.
5. Kfe. 2019. Plant arboviruses: major threats to food security. Microbiology
6. Zahouli JBZ, Koudou BG, Muller P, Malone D, Tano Y, Utzinger J. 2017. Effect of land-use changes on the abundance, distribution, and host-seeking behavior of Aedes arbovirus vectors in oil palm-dominated landscapes, southeastern Cote d'Ivoire. PLoS One 12:e0189082.
7. Braun T, Borderia AV, Barbezange C, Vignuzzi M, Louzoun Y. 2019. Long-term context-dependent genetic adaptation of the viral genetic cloud. Bioinformatics 35:1907-1915.
8. Patterson EI, Khanipov K, Rojas MM, Kautz TF, Rockx-Brouwer D, Golovko G, Albayrak L, Fofanov Y, Forrester NL. 2018. Mosquito bottlenecks alter viral mutant swarm in a tissue and time-dependent manner with contraction and expansion of variant positions and diversity. Virus Evol 4:vey001.
9. Grubaugh ND, Ruckert C, Armstrong PM, Bransfield A, Anderson JF, Ebel GD, Brackney DE. 2016. Transmission bottlenecks and RNAi collectively influence tick-borne flavivirus evolution. Virus Evol 2:vew033.
10. Benoît Moury FF, and Rachid Senoussi. 2007. Estimation of the number of virus particles transmitted by an insect vector.
11. Forrester NL, Coffey LL, Weaver SC. 2014. Arboviral bottlenecks and challenges to maintaining diversity and fitness during mosquito transmission. Viruses 6:3991-4004.
12. Grubaugh ND, Fauver JR, Ruckert C, Weger-Lucarelli J, Garcia-Luna S, Murrieta RA, Gendernalik A, Smith DR, Brackney DE, Ebel GD. 2017. Mosquitoes Transmit Unique West Nile Virus Populations during Each Feeding Episode. Cell Rep 19:709-718.
13. Grubaugh ND, Weger-Lucarelli J, Murrieta RA, Fauver JR, Garcia-Luna SM, Prasad AN, Black WCt, Ebel GD. 2016. Genetic Drift during Systemic Arbovirus Infection of Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching. Cell Host Microbe 19:481-92.
14. Hoye BJ, Fouchier RA, Klaassen M. 2012. Host behaviour and physiology underpin individual variation in avian influenza virus infection in migratory Bewick's swans. Proc Biol Sci 279:529-34.
15. Lequime S. Interactions flavivirus-moustiques : diversité et transmission. Virologie. Université Pierre et Marie Curie - Paris VI, 2016.
16. Fu H, Leake CJ, Mertens PPC, Mellor PS. 1999. The barriers to bluetongue virus infection, dissemination and transmission in the vector, Culicoïdes variipennis (Diptera: Ceratopogonidae). Archives of Virology 144:747-761.

17. Ciota AT, Jia Y, Payne AF, Jerzak G, Davis LJ, Young DS, Ehrbar D, Kramer LD. 2009. Experimental passage of St. Louis encephalitis virus in vivo in mosquitoes and chickens reveals evolutionarily significant virus characteristics. PLoS One 4:e7876.
18. Reyes-Ruiz JM, Osuna-Ramos JF, Bautista-Carbajal P, Jaworski E, Soto-Acosta R, Cervantes-Salazar M, Angel-Ambrocio AH, Castillo-Munguia JP, Chavez-Munguia B, De Nova-Ocampo M, Routh A, Del Angel RM, Salas-Benito JS. 2019. Mosquito cells persistently infected with dengue virus produce viral particles with host-dependent replication. Virology 531:1-18.
19. Brady OJ, Hay SI. 2020. The Global Expansion of Dengue: How Aedes aegypti Mosquitoes Enabled the First Pandemic Arbovirus. Annual Review of Entomology 65:191-208.
20. A. Mignotte CGea. 2020. Important dispersal capacity of one main species vector of Bluetongue and Schmallenberg viruses Culicoïdes obsoletus (Diptera: Ceratopogonidae) revealed by landscape genetic analys... Research Square doi:10.21203/rs.3.rs-75603/v1.
21. Eigen M. 1971. Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465-523.
22. Eigen M. 1996. On the nature of virus quasispecies. Trends in Microbiology 4:216-218.
23. Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. Microbiol Mol Biol Rev 76:159-216.
24. Marco Vignuzzi JKS, Jamie J. Arnold, Craig E. Cameron, and Raul Andino. 2006. Quasispecies diversity determines pathogenesis through cooperative interactions within a viral population. Nature.
25. Elena SF, Sanjuan R. 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J Virol 79:11555-8.
26. Pfeiffer JK, Kirkegaard K. 2005. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. PLoS Pathog 1:e11.
27. Sim S, Aw PP, Wilm A, Teoh G, Hue KD, Nguyen NM, Nagarajan N, Simmons CP, Hibberd ML. 2015. Tracking Dengue Virus Intra-host Genetic Diversity during Human-to-Mosquito Transmission. PLoS Negl Trop Dis 9:e0004052.
28. Caporale M, Di Gialleonorado L, Janowicz A, Wilkie G, Shaw A, Savini G, Van Rijn PA, Mertens P, Di Ventura M, Palmarini M. 2014. Virus and host factors affecting the clinical outcome of bluetongue virus infection. J Virol 88:10399-411.
29. Deardorff ER, Fitzpatrick KA, Jerzak GV, Shi PY, Kramer LD, Ebel GD. 2011. West Nile virus experimental evolution in vivo and the trade-off hypothesis. PLoS Pathog 7:e1002335.
30. Grubaugh ND, Smith DR, Brackney DE, Bosco-Lauth AM, Fauver JR, Campbell CL, Felix TA, Romo H, Duggal NK, Dietrich EA, Eike T, Beane JE, Bowen RA, Black WC, Brault AC, Ebel GD. 2015. Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness. PLoS Pathog 11:e1004874.
31. Mohamed Ali S, Amroun A, de Lamballerie X, Nougairede A. 2018. Evolution of Chikungunya virus in mosquito cells. Sci Rep 8:16175.
32. Vasilakis N, Deardorff ER, Kenney JL, Rossi SL, Hanley KA, Weaver SC. 2009. Mosquitoes put the brake on arbovirus evolution: experimental evolution reveals slower mutation accumulation in mosquito than vertebrate cells. PLoS Pathog 5:e1000467.

33. Coffey LL, Vignuzzi M. 2011. Host alternation of chikungunya virus increases fitness while restricting population diversity and adaptability to novel selective pressures. J Virol 85:1025-35.
34. Greene IP, Wang E, Deardorff ER, Milleron R, Domingo E, Weaver SC. 2005. Effect of alternating passage on adaptation of sindbis virus to vertebrate and invertebrate cells. J Virol 79:14253-60.
35. A BM. 1982. Analysis of Extracellular West Nile Virus Particles Produced by Cell Cultures from Genetically Resistant and Susceptible Mice Indicates Enhanced Amplification of Defective Interfering Particles by Resistant Cultures. Journal of Virology.
36. Langsjoen RM, Muruato AE, Kunkel SR, Jaworski E, Routh A. 2020. Differential Alphavirus Defective RNA Diversity between Intracellular and Extracellular Compartments Is Driven by Subgenomic Recombination Events. mBio 11.
37. Li D, Lott WB, Lowry K, Jones A, Thu HM, Aaskov J. 2011. Defective interfering viral particles in acute dengue infections. PLoS One 6:e19447.
38. Bull JC, Godfray HC, O'Reilly DR. 2001. Persistence of an occlusion-negative recombinant nucleopolyhedrovirus in Trichoplusia ni indicates high multiplicity of cellular infection. Appl Environ Microbiol 67:5204-9.
39. Pijlman GP, Dortmans JC, Vermeesch AM, Yang K, Martens DE, Goldbach RW, Vlak JM. 2002. Pivotal role of the non-hr origin of DNA replication in the genesis of defective interfering baculoviruses. J Virol 76:5605-11.
40. Lopez-Ferber M, Simon O, Williams T, Caballero P. 2003. Defective or effective? Mutualistic interactions between virus genotypes. Proc Biol Sci 270:2249-55.
41. Erasmus BJ, Christiaan Potgieter A. 2009. Chapter 2 - The history of bluetongue, p 7-21. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50006-X. Academic Press, London.
42. Pascall DJ, Nomikou K, Breard E, Zientara S, Filipe ADS, Hoffmann B, Jacquot M, Singer JB, De Clercq K, Botner A, Sailleau C, Viarouge C, Batten C, Puggioni G, Ligios C, Savini G, van Rijn PA, Mertens PPC, Biek R, Palmarini M. 2020. "Frozen evolution" of an RNA virus suggests accidental release as a potential cause of arbovirus re-emergence. PLoS Biol 18:e3000673.
43. Curtis A. Deutsch JJT, Raymond B. Huey, Kimberly S. Sheldon, Cameron K. Ghalambor, and David C. Haak, and Paul R. Martin. 2008. Impacts of climate warming on terrestrial ectotherms across latitude. PNAS.
44. Jacquet S, Huber K, Pages N, Talavera S, Burgin LE, Carpenter S, Sanders C, Dicko AH, Djerbal M, Goffredo M, Lhor Y, Lucientes J, Miranda-Chueca MA, Pereira Da Fonseca I, Ramilo DW, Setier-Rio ML, Bouyer J, Chevillon C, Balenghien T, Guis H, Garros C. 2016. Range expansion of the Bluetongue vector, Culicoïdes imicola, in continental France likely due to rare wind-transport events. Sci Rep 6:27247.
45. Breard E, Schulz C, Sailleau C, Bernelin-Cottet C, Viarouge C, Vitour D, Guillaume B, Caignard G, Gorlier A, Attoui H, Gallois M, Hoffmann B, Zientara S, Beer M. 2018. Bluetongue virus serotype 27: Experimental infection of goats, sheep and cattle with three BTV-27 variants reveal atypical characteristics and likely direct contact transmission BTV-27 between goats. Transbound Emerg Dis 65:e251-e263.
46. Mellor PS, Carpenter S, White DM. 2009. Chapter 14 - Bluetongue virus in the insect host, p 295-320. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue

doi:https://doi.org/10.1016/B978-012369368-6.50018-6. Academic Press, London.
47. P. RJ. 1982. Contribution à la connaissance de l'écologie et de la biologie des Cératopogonidés (diptères, nématocères). HAL.
48. Foster Nm Fau - Jones RH, Jones Rh Fau - Mccrory BR, McCrory BR. PRELIMINARY INVESTIGATIONS ON INSECT TRANSMISSION OF BLUETONGUE VIRUS IN SHEEP.
49. Baylis M, O'Connell L Fau - Mellor PS, Mellor PS. Rates of bluetongue virus transmission between Culicoïdes sonorensis and sheep.
50. Carpenter S, Lunt Hl Fau - Arav D, Arav D Fau - Venter GJ, Venter Gj Fau - Mellor PS, Mellor PS. Oral susceptibility to bluetongue virus of Culicoïdes (Diptera: Ceratopogonidae) from the United Kingdom.
51. Maclachlan NJ, Gard G. 2009. Chapter 13 - Clinical signs and pathology, p 285-293. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50017-4. Academic Press, London.
52. Oura CA, El Harrak M. 2011. Midge-transmitted bluetongue in domestic dogs. Epidemiol Infect 139:1396-400.
53. Jauniaux TP, De Clercq KE, Cassart DE, Kennedy S, Vandenbussche FE, Vandemeulebroucke EL, Vanbinst TM, Verheyden BI, Goris NE, Coignoul FL. 2008. Bluetongue in Eurasian lynx. Emerging infectious diseases 14:1496-1498.
54. Barratt-Boyes SM, Rossitto PV, Taylor BC, Ellis JA, MacLachlan NJ. 1995. Response of the regional lymph node to bluetongue virus infection in calves. Veterinary Immunology and Immunopathology 45:73-84.
55. Karin E Darpel PM, Jennifer Simpson, Simon J Anthony, Eva Veronesi, Harriet W Brooks, Heather Elliott, Joe Brownlie, Haru-Hisa Takamatsu, Philip S Mellor and Peter PC Mertens. 2012. Involvement of the skin during bluetongue virus infection and replication in the ruminant host. VETERINARY RESEARCH.
56. D. M. HAIG JHHRPM. 1999. Local immune responses in afferent and efferent lymph. Immunology.
57. MacLachlan NJ, Nunamaker RA, Katz JB, Sawyer MM, Akita GY, Osburn BI, Tabachnick WJ. 1994. Detection of bluetongue virus in the blood of inoculated calves: comparison of virus isolation, PCR assay, and in vitro feeding ofCulicoïdes variipennis. Archives of Virology 136:1-8.
58. Hawkes PDKRA. 2004. A comparison of laboratory and ‘wild’ strains of bluetongue virus – is there any difference and does it matter? Veterinaria Italiana.
59. Santman-Berends IM, van Wuijckhuise L, Vellema P, van Rijn PA. 2010. Vertical transmission of bluetongue virus serotype 8 virus in Dutch dairy herds in 2007. Vet Microbiol 141:31-5.
60. Mayo CE, Crossley BM, Hietala SK, Gardner IA, Breitmeyer RE, Maclachlan NJ. 2010. Colostral transmission of bluetongue virus nucleic acid among newborn dairy calves in California. Transbound Emerg Dis 57:277-81.
61. Backx A, Heutink R, van Rooij E, van Rijn P. 2009. Transplacental and oral transmission of wild-type bluetongue virus serotype 8 in cattle after experimental infection. Vet Microbiol 138:235-43.
62. G. Worwa MH, F. Ehrensperger, V. Chaignat, M. A. Hofmann, C. Griot, N. J. Maclachlan, B. Thuer. 2009. Experimental transplacental infection of sheep with bluetongue virus serotype 8. Vetarinary record.

63. White DM, Wilson WC, Blair CD, Beaty BJ. 2005. Studies on overwintering of bluetongue viruses in insects. J Gen Virol 86:453-462.
64. Osborne CJ, Mayo CE, Mullens BA, McDermott EG, Gerry AC, Reisen WK, MacLachlan NJ. 2015. Lack of Evidence for Laboratory and Natural Vertical Transmission of Bluetongue Virus in Culicoïdes sonorensis (Diptera: Ceratopogonidae). J Med Entomol 52:274-7.
65. D. W. VERWOERD HL, AND R. A. OELLERMANN. 1970. Characterization of Bluetongue Virus Ribonucleic Acid. Journal of Virology.
66. Maan S, Maan NS, Samuel AR, Rao S, Attoui H, Mertens PPC. 2007. Analysis and phylogenetic comparisons of full-length VP2 genes of the 24 bluetongue virus serotypes. J Gen Virol 88:621-630.
67. Bouet-Cararo C, Contreras V, Caruso A, Top S, Szelechowski M, Bergeron C, Viarouge C, Desprat A, Relmy A, Guibert JM, Dubois E, Thiery R, Breard E, Bertagnoli S, Richardson J, Foucras G, Meyer G, Schwartz-Cornil I, Zientara S, Klonjkowski B. 2014. Expression of VP7, a Bluetongue virus group specific antigen by viral vectors: analysis of the induced immune responses and evaluation of protective potential in sheep. PLoS One 9:e111605.
68. Lobato ZIP, Coupar BEH, Gray CP, Lunt R, Andrew ME. 1997. Antibody responses and protective immunity to recombinant vaccinia virus-expressed bluetongue virus antigens. Veterinary Immunology and Immunopathology 59:293-309.
69. Roy P. 2017. Bluetongue virus structure and assembly. Curr Opin Virol 24:115-123.
70. Desselberger U. 2020. What are the limits of the packaging capacity for genomic RNA in the cores of rotaviruses and of other members of the Reoviridae? Virus Res 276:197822.
71. Schwartz-Cornil I, Mertens PP, Contreras V, Hemati B, Pascale F, Breard E, Mellor PS, MacLachlan NJ, Zientara S. 2008. Bluetongue virus: virology, pathogenesis and immunity. Vet Res 39:46.
72. Noad R, Roy P. 2009. Chapter 4 - Bluetongue virus replication and assembly, p 53-76. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50008-3. Academic Press, London.
73. Ross-Smith N, Darpel KE, Monaghan P, Mertens PPC. 2009. Chapter 5 - Bluetongue virus: cell biology, p 77-99. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50009-5. Academic Press, London.
74. and SSH, Roy P. 1999. Expression and Functional Characterization of Bluetongue Virus VP2 Protein: Role in Cell Entry. Journal of Virology.
75. Hassan SH, Wirblich C, Forzan M, Roy P. 2001. Expression and functional characterization of bluetongue virus VP5 protein: role in cellular permeabilization. J Virol 75:8356-67.
76. Forzan M, Marsh M, Roy P. 2007. Bluetongue virus entry into cells. J Virol 81:4819-27.
77. Mohl BP, Kerviel A, Labadie T, Matsuo E, Roy P. 2020. Differential Localization of Structural and Non-Structural Proteins during the Bluetongue Virus Replication Cycle. Viruses 12.
78. Labadie T, Sullivan E, Roy P. 2020. Multiple Routes of Bluetongue Virus Egress. Microorganisms 8.

79. P. Mertens JNB, A. Walton, M. P. Wellby, H. Fu, R. S. O'Hara, S. M. Brookes, and P. S. Mellor. 1996. Enhanced Infectivity of Modified Bluetongue Virus Particles for Two Insect Cell Lines and for Two Culicoïdes Vector Species. Virology.
80. Shapiro A, Green T, Rao S, White S, Carner G, Mertens PP, Becnel JJ. 2005. Morphological and molecular characterization of a Cypovirus (Reoviridae) from the mosquito Uranotaenia sapphirina (Diptera: Culicidae). J Virol 79:9430-8.
81. Vitour D, Doceul V, Ruscanu S, Chauveau E, Schwartz-Cornil I, Zientara S. Induction and control of the type I interferon pathway by Bluetongue virus.
82. Janowicz A, Caporale M, Shaw A, Gulletta S, Di Gialleonardo L, Ratinier M, Palmarini M. 2015. Multiple genome segments determine virulence of bluetongue virus serotype 8. J Virol 89:5238-49.
83. van Gennip RG, van de Water SG, van Rijn PA. 2014. Bluetongue virus nonstructural protein NS3/NS3a is not essential for virus replication. PLoS One 9:e85788.
84. Ftaich N, Ciancia C, Viarouge C, Barry G, Ratinier M, van Rijn PA, Breard E, Vitour D, Zientara S, Palmarini M, Terzian C, Arnaud F. 2015. Turnover Rate of NS3 Proteins Modulates Bluetongue Virus Replication Kinetics in a Host-Specific Manner. J Virol 89:10467-81.
85. C HE. 2009. The Evolution and Emergence of RNA Viruses. 16:899-899.
86. Nomikou K, Dovas CI, Maan S, Anthony SJ, Samuel AR, Papanastassopoulou M, Maan NS, Mangana O, Mertens PP. 2009. Evolution and phylogenetic analysis of full-length VP3 genes of Eastern Mediterranean bluetongue virus isolates. PLoS One 4:e6437.
87. Jacquot M, Rao PP, Yadav S, Nomikou K, Maan S, Jyothi YK, Reddy N, Putty K, Hemadri D, Singh KP, Maan NS, Hegde NR, Mertens P, Biek R. 2019. Contrasting selective patterns across the segmented genome of bluetongue virus in a global reassortment hotspot. Virus Evolution 5:1-14.
88. Carpi G, Holmes EC, Kitchen A. 2010. The evolutionary dynamics of bluetongue virus. Journal of Molecular Evolution 70:583-592.
89. Worobey M, Han GZ, Rambaut A. 2014. A synchronized global sweep of the internal genes of modern avian influenza virus. Nature 508:254-7.
90. Hanada K, Suzuki Y, Gojobori T. 2004. A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. Mol Biol Evol 21:1074-80.
91. Nomikou K, Hughes J, Wash R, Kellam P, Breard E, Zientara S, Palmarini M, Biek R, Mertens P. 2015. Widespread Reassortment Shapes the Evolution and Epidemiology of Bluetongue Virus following European Invasion. PLoS Pathogens 11:1-23.
92. Kopanke JH, Lee JS, Stenglein MD, Mayo CE. 2020. The Genetic Diversification of a Single Bluetongue Virus Strain Using an In Vitro Model of Alternating-Host Transmission. Viruses 12.
93. Pohlmann A, Starick E, Harder T, Grund C, Hoper D, Globig A, Staubach C, Dietze K, Strebelow G, Ulrich RG, Schinkothe J, Teifke JP, Conraths FJ, Mettenleiter TC, Beer M. 2017. Outbreaks among Wild Birds and Domestic Poultry Caused by Reassorted Influenza A(H5N8) Clade 2.3.4.4 Viruses, Germany, 2016. Emerg Infect Dis 23:633-636.
94. Shaw AE, Ratinier M, Nunes SF, Nomikou K, Caporale M, Golder M, Allan K, Hamers C, Hudelet P, Zientara S, Breard E, Mertens P, Palmarini M. 2013.

Reassortment between two serologically unrelated bluetongue virus strains is flexible and can involve any genome segment. J Virol 87:543-57.
95. Sicard A, Yvon M, Timchenko T, Gronenborn B, Michalakis Y, Gutierrez S, Blanc S. 2013. Gene copy number is differentially regulated in a multipartite virus. Nature Communications 4:1-8.
96. Hu Z, Zhang X, Liu W, Zhou Q, Zhang Q, Li G, Yao Q. 2016. Genome segments accumulate with different frequencies in Bombyx mori bidensovirus. J Basic Microbiol 56:1338-1343.
97. Ladner JT, Wiley MR, Beitzel B, Auguste AJ, Dupuis AP, 2nd, Lindquist ME, Sibley SD, Kota KP, Fetterer D, Eastwood G, Kimmel D, Prieto K, Guzman H, Aliota MT, Reyes D, Brueggemann EE, St John L, Hyeroba D, Lauck M, Friedrich TC, O'Connor DH, Gestole MC, Cazares LH, Popov VL, Castro-Llanos F, Kochel TJ, Kenny T, White B, Ward MD, Loaiza JR, Goldberg TL, Weaver SC, Kramer LD, Tesh RB, Palacios G. 2016. A Multicomponent Animal Virus Isolated from Mosquitoes. Cell Host Microbe 20:357-367.
98. Feng JL, Chen SN, Tang XS, Ding XF, Du ZY, Chen JS. 2006. Quantitative determination of cucumber mosaic virus genome RNAs in virions by real-time reverse transcription-polymerase chain reaction. Acta Biochim Biophys Sin (Shanghai) 38:669-76.
99. Kwon SJ, Seo Jk Fau - Rao ALN, Rao AL. Detection and quantification of viral and satellite RNAs in plant hosts.
100. Wu B, Zwart MP, Sánchez-Navarro JA, Elena SF. 2017. Within-host Evolution of Segments Ratio for the Tripartite Genome of Alfalfa Mosaic Virus. Scientific Reports 7:1-15.
101. Yuriy Mileykoa RIJ, and Joshua S. Weitza. 2008. Small-scale copy number variation and large-scale changes in gene expression. PNAS.
102. Gamazon ER, Stranger BE. 2015. The impact of human copy number variation on gene expression. Brief Funct Genomics 14:352-7.
103. Almal SH, Padh H. Implications of gene copy-number variation in health and diseases.
104. Nair S, Miller B, Barends M, Jaidee A, Patel J, Mayxay M, Newton P, Nosten F, Ferdig MT, Anderson TJ. 2008. Adaptive copy number evolution in malaria parasites. PLoS Genet 4:e1000243.
105. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. 2009. Mechanisms of change in gene copy number. Nature Reviews Genetics 10:551-564.
106. Elde NC, Child SJ, Eickbush MT, Kitzman JO, Rogers KS, Shendure J, Geballe AP, Malik HS. 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. Cell 150:831-841.
107. Zeng L, Skinner SO, Zong C, Sippy J, Feiss M, Golding I. 2010. Decision making at a subcellular level determines the outcome of bacteriophage infection. Cell 141:682-91.
108. Haralampiev I, Prisner S, Nitzan M, Schade M, Jolmes F, Schreiber M, Loidolt-Kruger M, Jongen K, Chamiolo J, Nilson N, Winter F, Friedman N, Seitz O, Wolff T, Herrmann A. 2020. Selective flexible packaging pathways of the segmented genome of influenza A virus. Nat Commun 11:4355.
109. Brooke CB, Ince WL, Wrammert J, Ahmed R, Wilson PC, Bennink JR, Yewdell JW. 2013. Most influenza a virions fail to express at least one essential viral protein. J Virol 87:3155-62.
110. Martin MA, Kaul D, Tan GS, Woods CW, Koelle K. 2019. doi:10.1101/814673.

111. Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. 2014. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. Proceedings of the National Academy of Sciences 111:16854-16859.
112. Wichgers Schreur PJ, Kortekaas J. 2016. Single-Molecule FISH Reveals Non-selective Packaging of Rift Valley Fever Virus Genome Segments. PLoS Pathog 12:e1005800.
113. Sung PY, Roy P. 2014. Sequential packaging of RNA genomic segments during the assembly of Bluetongue virus. Nucleic Acids Res 42:13824-38.
114. Labadie T, Roy P. 2020. A non-enveloped arbovirus released in lysosome-derived extracellular vesicles induces super-infection exclusion. doi:10.1101/2020.06.11.146357.
115. Sailleau C, Breard E, Viarouge C, Gorlier A, Quenault H, Hirchaud E, Touzain F, Blanchard Y, Vitour D, Zientara S. 2018. Complete genome sequence of bluetongue virus serotype 4 that emerged on the French island of Corsica in December 2016. Transboundary and Emerging Diseases 65:e194-e197.

CHAPITRE 1 : Les segments génomiques du virus de la fièvre catarrhale ovine varient en nombre de copies d'une manière hôte spécifique

Le virus de la fièvre catarrhale ovine (FCO) est un arbovirus qui se distingue par son architecture génomique singulière. Son génome est divisé en 10 segments ARN double brin faisant de lui l'arbovirus au niveau de segmentation le plus élevé. Chez le virus de la FCO au moins 27 sérotypes différents ont été décrit et témoigne de sa capacité d'adaptation à plusieurs gammes d'hôtes. Nous proposons l'hypothèse que la variation du nombre de copie de gènes (VNCG) puisse faciliter son adaptation à ce système en alternance d'hôte. La VNCG au sein d'une population virale est un phénomène qui pourrait avoir un rôle central dans l'adaptation de certaines espèces de virus multipartites de plante à différents hôtes. Par ailleurs, un phénomène similaire a été identifié par Brooke *et al.* chez le virus de la grippe A : la diminution du nombre de copies d'un segment joue un rôle direct sur la régulation de son expression et sur la production de la protéine d'enveloppe qu'il code.

Au cours de ce chapitre avons étudié ce phénomène grâce à une approche de RT-qPCR développée spécifiquement pour quantifier les segments génomiques du virus de la FCO de sérotype 4 (FCO-4). Nous avons suivi la variation du nombre de copie de gènes de ce virus dans des populations virales circulant pendant l'épizootie qui a eu lieux en Corse et en Sardaigne en 2016. Nos résultats montrent que les fréquences des segments génomiques du virus de la FCO-4 diffèrent entre elles de façon reproductible au sein des hôtes naturellement infectés. Les profils de variation des fréquences observés correspondent à la définition de la formule génomique initialement proposée pour décrire la convergence des fréquences des segments vers un équilibre différent de l'équimolarité chez un virus de plantes multipartite.

La mesure de la formule génomique du virus de la FCO sur des échantillons de terrain nous as permis de mettre en avant l'existence du phénomène de VNCG chez le virus de la FCO-4 et une variation du profil de fréquences dépendant des hôtes infectés. Pour le virus présent chez les deux espèces de ruminant *Ovis aries* et *Bos taurus* un profil de fréquences similaire est observé. En revanche la FG observée dans les insectes vecteurs, de l'espèce *Culicoides imicola*, diffère significativement des profils de fréquences génomiques des mammifères. Chez les culicoïdes, la variation maximale d'abondance entre deux segments est de trois.

Ces résultats présentent pour la première fois la variation du nombre de copies de gènes (VNCG) chez le virus de la FCO-4. Les profils de fréquence des segments

génomiques sont dépendant de l'hôte suggérant que la VNCG pourrait être un mécanisme impliqué dans l'adaptation à l'alternance d'hôte chez le virus de la FCO.

Ce chapitre a été accepté et publié dans *Journal of Virology* :

DOI: 10.1128/JVI.01834-20

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENETIC DIVERSITY AND EVOLUTION

The Genome Segments of Bluetongue Virus Differ in Copy Number in a Host-Specific Manner

Yannis Moreau,[a,b] Patricia Gil,[a,b] Antoni Exbrayat,[a,b] Ignace Rakotoarivony,[a,b] Emmanuel Bréard,[c] Corinne Sailleau,[c] Cyril Viarouge,[c] Stephan Zientara,[c] Giovanni Savini,[d] Maria Goffredo,[d] Giuseppe Mancini,[d] Etienne Loire,[a,b] Serafin Gutierrez[a,b]

[a]CIRAD, UMR ASTRE, Montpellier, France
[b]ASTRE, CIRAD, INRA, Univ Montpellier, Montpellier, France
[c]UMR1161 Virologie, ANSES, INRA, Ecole Nationale Vétérinaire d'Alfort, Université Paris-Est, Maisons-Alfort, France
[d]Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise G. Caporale, Teramo, Italy

ABSTRACT Genome segmentation is mainly thought to facilitate reassortment. Here, we show that segmentation can also allow differences in segment abundance in populations of bluetongue virus (BTV). BTV has a genome consisting in 10 segments, and its cycle primarily involves periodic alternation between ruminants and *Culicoides* biting midges. We have developed a reverse transcription-quantitative PCR (RT-qPCR) approach to quantify each segment in wild BTV populations sampled in both ruminants and midges during an epizootic. Segment frequencies deviated from equimolarity in all hosts. Interestingly, segment frequencies were reproducible and distinct between ruminants and biting midges. Beyond a putative regulatory role in virus expression, this phenomenon could lead to different evolution rates between segments.

IMPORTANCE The variation in viral gene frequencies remains a largely unexplored aspect of within-host genetics. This phenomenon is often considered to be specific to multipartite viruses. Multipartite viruses have segmented genomes, but in contrast to segmented viruses, their segments are each encapsidated alone in a virion. A main hypothesis explaining the evolution of multipartism is that, compared to segmented viruses, it facilitates the regulation of segment abundancy, and the genes the segments carry, within a host. These differences in gene frequencies could allow for expression regulation. Here, we show that wild populations of a segmented virus, bluetongue virus (BTV), also present unequal segment frequencies. BTV cycles between ruminants and *Culicoides* biting midges. As expected from a role in expression regulation, segment frequencies tended to show specific values that differed between ruminants and midges. Our results expand previous knowledge on gene frequency variation and call for studies on its role and conservation beyond multipartite viruses.

KEYWORDS bluetongue virus, gene copy number, population genetics, virus-host interactions

Variations in segment frequencies have been recently uncovered in within-host populations of multipartite viruses (1). Multipartite viruses are a diverse group of viruses, including DNA and RNA viruses, that infect mainly plants (2). Their genomes are composed of several segments, and in contrast to segmented viruses, each segment is encapsidated individually. Thus, each segment is thought to be physically independent from the others. Remarkably, certain segments are more abundant than others in within-host populations of these viruses (1, 3, 4). The term "genome formula" (GF) was thus coined to define the relative frequencies of all the segments in a within-host population. Moreover, segments tend to converge toward a specific frequency profile,

Citation Moreau Y, Gil P, Exbrayat A, Rakotoarivony I, Bréard E, Sailleau C, Viarouge C, Zientara S, Savini G, Goffredo M, Mancini G, Loire E, Gutierrez S. 2021. The genome segments of bluetongue virus differ in copy number in a host-specific manner. J Virol 95:e01834-20. https://doi.org/10.1128/JVI.01834-20.

Editor Colin R. Parrish, Cornell University

Copyright © 2020 Moreau et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Etienne Loire, etienne.loire@cirad.fr, or Serafin Gutierrez, serafin.gutierrez@cirad.fr.

Received 22 September 2020
Accepted 29 September 2020
Accepted manuscript posted online 7 October 2020
Published 9 December 2020

TABLE 1 Samples per host species and site[a]

Host	Island	Samples (no.)	Curated (no.)
Sheep	Corsica	24	11
Sheep	Sardinia	7	4
Midge	Corsica	39	5
Midge	Sardinia	10	4
Cow	Corsica	26	12

[a]Sheep, *Ovis aries*; cow, *Bos taurus*; midge, *Culicoides imicola*; Curated, samples in which RT-qPCRs of all segments provided a C_T value below 31.

the so-called "set-point GF" (i.e., the median segment frequencies in a population of infected hosts) (1). The fact that GFs deviate from segment equimolarity is puzzling from an evolutionary standpoint. Such population genetics should imply a cost derived from the higher probability of losing scarce segments during population bottlenecks (5). However, populations close to the set-point GF tend to accumulate at higher rates, suggesting a link with virus fitness (1, 4), probably through optimization of gene expression (1, 6). As expected from a role in expression regulation, the set-point GF has been shown to be host specific (1, 4). However, the putative relationship between GF and gene expression has not yet been directly demonstrated in multipartite viruses (7).

Currently, there are two examples of nonmultipartite viruses with a behavior reminiscent of set-point GFs in multipartite viruses. In the first example, populations of a monopartite baculovirus bore specific differences in gene frequencies that are due to certain genotypes with a deletion in the same genomic region (8). In the second example, a lower encapsidation rate of segment N led to unequal gene frequencies in populations of an experimentally evolved clone of influenza A virus, a segmented virus (9). In both cases, increases in fitness were observed despite the essential role of the genes with lower frequencies (9, 10). Moreover, fitness improvements were probably due to regulation of virus expression (9, 10).

These observations raise the question on how widespread gene frequency variation is. We have explored this question in bluetongue virus (BTV; *Orbivirus*, *Reoviridae*), a segmented virus affecting livestock worldwide. BTV possesses a double-stranded RNA (dsRNA) genome divided into 10 segments, a segment number among the highest in viruses. Most segments code for a single protein (11). Moreover, BTV is mainly an arthropod-borne virus infecting ruminants and *Culicoides* biting midges. This complex cycle probably involves yet-unknown mechanisms allowing expression regulation in phylogenetically distant hosts.

RESULTS

We first built a collection of samples from naturally infected hosts during the 2014/2019 epizooty of BTV serotype 4 (BTV-4) in Europe (12). Blood samples were collected from sheep and cows from Corsica (France) and sheep from Sardinia (Italy), two Mediterranean islands (Table 1). The collection also included pools of 50 adult females of *Culicoides imicola* from the same islands and period (Table 1).

We then developed a SYBR green quantitative PCR (qPCR) approach to quantify genome segments similar to those used with multipartite viruses (1, 3, 4). The SYBR green assay was chosen over other qPCR assays due to its accuracy, simplicity, and cost efficiency in schemes requiring the validation of many primer pairs in parallel and the analysis of large sample numbers. Importantly, BTV segments can be quantified apart from viral transcripts because the negative strand of the dsRNA segments is generated only within fully formed capsids (13). Hence, we developed a two-step reverse transcription (RT)-qPCR assay for each segment of BTV-4 (strain BTV4-16-03) (12) in which the RT primer targets only the negative strand. We tested a large panel of primer couples to identify couples with similar efficiencies (minimum/maximum = 91.8/95.2) and specificity when using BTV-free templates (Table 2).

Each RNA sample was screened with the 10 RT-qPCR assays. To ensure exact calibration of the 10 qPCRs, we generated and used a single plasmid as the template

TABLE 2 Primers selected for the quantification of the negative strand of the BTV-4 segments

Segment	Direction	Sequence	5′ position	Amplicon size (bp)	Efficiency (%)	Mismatch[a]	Specificity threshold (C_T)		Final concn (μM)
							Biting midges	Sheep blood	
1	Forward	AACATGGCTATTGGGACC	2213	125	91.9	>5	>39.5	>38.5	0.5
	Reverse	TCTGTAGTGTGTAGCTTTGTGTAA	2337						
2	Forward	TCTAGCTTCTCTATGTTTAGGGC	2541	70	93.0	>5	>39.8	>40	0.4
	Reverse	TTTCTTTGGATGCGCGAC	2610						
3	Forward	ATTCCGAGCGGCTTTAAGA	2287	75	94.5	>5	>40	>40	0.5
	Reverse	CGTACAGTGCGTAATACACC	2361						
4	Forward	TATGCGTAAGGGATTTGGTG	343	88	94.2	>5	>35	>35.3	0.5
	Reverse	GTTCAACGTCTCCGCTTC	430						
5	Forward	TGATCGCGGCAACTGAC	189	100	92.4	5	>40	>40	0.8
	Reverse	CAGACTGTTTCCCGATCATAC	279						
6	Forward	GTGAATCTTATGGAGAGTCGG	212	70	91.8	>5	>37.4	>40	0.5
	Reverse	GGTAATTCTTCACCTGTACCCAA	281						
7	Forward	ATGGCTACGATTGGTGTACTA	278	74	95.2	>5	>38.4	>38.7	0.4
	Reverse	ACGCGAGCAATCTCATTC	351						
8	Forward	AAGAGATGATTCCGGGAACTA	608	78	91.8	>5	>34.2	>34.8	0.4
	Reverse	CCAGCTTCCACCTCCTTA	686						
9	Forward	GGAACCCAAAGAGGAAGACA	133	110	95.1	>5	>37.5	>36.4	0.6
	Reverse	CCACATCTGCATCTTTAGC	242						
10	Forward	AAGGCTGCATTCGCATCGTA	242	115	93.5	>5	>40	>40	0.5
	Reverse	AGCCTCCTAGGTCGCTTTTC	356						

[a]Number of positions differing from sequence of the segment positive strand.

for all the standard curves. This plasmid harbors a concatenate of the 10 regions targeted by our RT-qPCR assays, thus ensuring exactly the same initial number of copies per amplicon. We selected for further analysis only those samples showing a threshold cycle (C_T) value below 31 for all segments (approximately 1,000 copies). The final data set consisted of 15, 12, and 9 samples from sheep, cows, and biting midges, respectively (Table 1).

We tested for primer-independent cDNA synthesis, a phenomenon in which RNA secondary structures or nucleic acid molecules (e.g., tRNA) can act as primers in the RT reaction (14). This phenomenon can affect strand-specific quantification in viruses (15), but to our knowledge, it is not tested for during quantification of BTV or other reoviruses (e.g., reference 16). We compared RT-qPCR output with or without a primer at the RT step for all segments and samples. RT reactions in the absence of primers for segments 2, 7, and 10 were the only ones yielding mean copy numbers at least 1% of that obtained with an RT primer (9%, 1%, and 26%, respectively; Fig. 1A). Primer-independent amplification could not be abolished without losing sensitivity through changes in the elongation temperature or the primer concentration at the RT step as previously proposed (15) (Fig. 1B). We compared segment frequencies in all samples between two data sets, corrected or not for the copy number obtained without a primer during the RT step. We observed a significant difference for segment 10 (Wilcoxon test, false-discovery rate [FDR] corrected, P value = 0.008), although this effect was detected only in sheep samples when the data set was split per host species (Wilcoxon test, FDR corrected, P value = 0.026 for the sheep data set) (Fig. 2A). Therefore, all statistical analyses were carried out in three data sets in parallel: (i) a data set with the raw copy numbers, (ii) a data set obtained after subtraction of copy numbers found in the absence of an RT primer (corrected data set), and (iii) a data set with raw copy numbers but without data from segment 10. No main difference in the

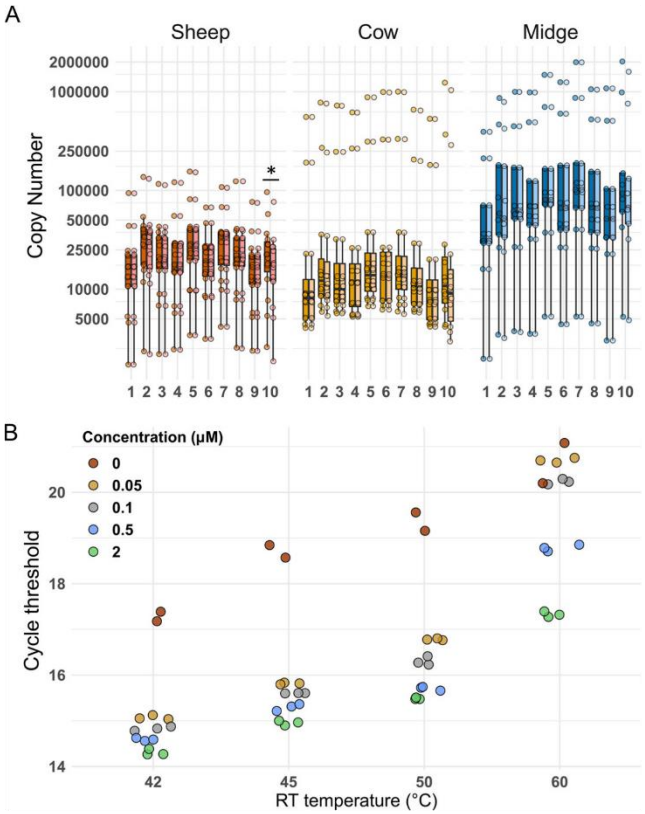


FIG 1 (A) Influence of primer-independent cDNA amplification on segment copy number (CN) in each host. CN values are colored according to host species, with lighter coloring for CN values after correction for primer-independent cDNA amplification. The correction consisted of subtracting the copy number yielded in an RT-qPCR assay without a primer in the RT step, and it was done for each segment and sample. Segment numbers appear at bottom. The asterisk indicates the only segment-host combination (segment 10 in sheep) for which a significant difference was detected between the uncorrected and corrected CNs. (B) Influence of temperature and primer concentration at the RT step on primer-independent cDNA amplification and qPCR sensibility. Data presented were obtained with the RT-qPCR assay for segment 10, the segment showing significant primer-independent amplification. RNA extracted from BTV-infected BHK cells was used as the template. RT-qPCRs were carried out as described in the text. Colors follow final primer concentration (μM) in RT step 1. Red dots indicate PCR output when RT was performed without a primer. The experiment was performed with two or three biological replicates. We selected an RT temperature of 42°C and a primer concentration of 2 μM.

significance of the statistical tests was observed between the three data sets (Table 3). The results presented below were obtained with the corrected data set.

As observed in multipartite viruses, relative segment frequencies were variable and their distributions sometimes deviated from those expected under segment equimolarity (Fig. 2A). Moreover, distribution of segment frequencies in each sample did not suggest multimodality in frequency distributions between samples (for example, a group of samples with equimolar amounts of segments and another group with segment frequencies deviating from equimolarity) (Fig. 2B and Fig. 3). Segments 1 and 9 had lower frequencies and segments 5 and 7 had higher frequencies than expected under equimolarity in all hosts (Student's *t* test one-sided, Benjamini-Hochberg corrected, *P* values <0.01) (Fig. 2A). The frequencies of segments 2, 6, and 10 also significantly deviated from equimolarity but only in sheep, cows, or midges, respectively (Table 4).

An analysis of variance showed a significant effect of the variable "segment" on segment frequencies (Fig. 2A; Table 3). The analysis did not detect a significant effect

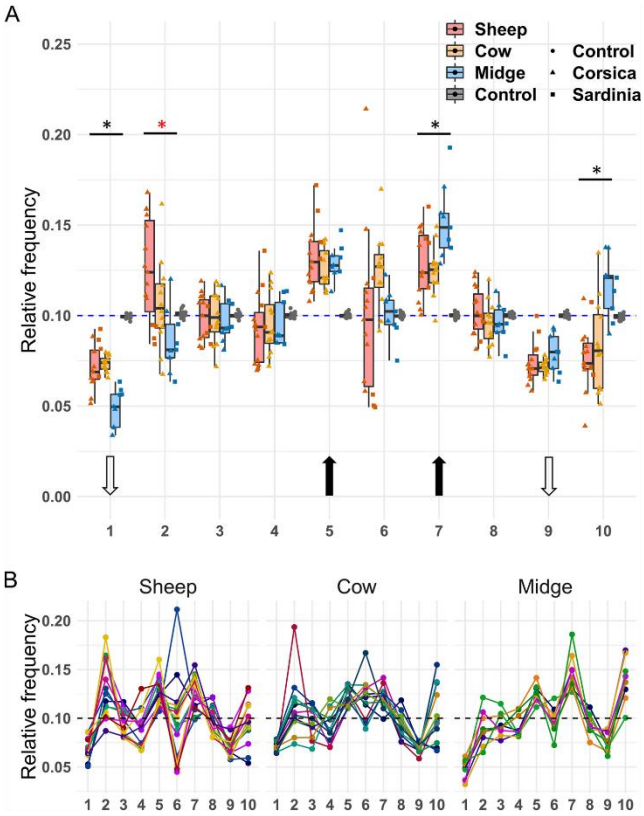


FIG 2 (A) Relative frequencies of BTV segments in naturally infected sheep, cows, and biting midges. The number of each segment is shown below the graph. The blue dashed line indicates the expected frequency under segment equimolarity. The results obtained with the plasmid control ("Control," boxplots in gray) are shown to allow comparison with a template with equimolar amounts of PCR targets. Dot shape indicates sample origin (island) or template type (plasmid control versus field samples). The black and white arrows indicate the segments for which the average relative frequency was significantly higher or lower, respectively, than expected under equimolarity. Significant differences in frequencies among hosts for a given segment are indicated with a star above boxplots. Black asterisks correspond to significant differences between populations in the two ruminant hosts versus those in midges, whereas the red asterisk shows a significant difference between BTV populations from sheep and biting midges. (B) Segment frequencies per sample. Dots represent the relative frequency of a given segment (*x* axis) in a sample. Samples have been split per host in three panels (sheep, cow, and midge). For each host, all dots from a given sample are linked by a line of the same color.

of host or island but found significant interactions between segment and host, and between segment and island (Table 3). The analysis also found a marginally significant interaction among the three independent variables (P value = 0.049) (Table 3). Such crossover interactions are likely to take place when examining frequency data with several independent variables. We thus analyzed the effect of host species on segment frequency as previously done (1). Several factors hampered a robust analysis of a putative island effect, and it was not further explored (i.e., limited sample size, lack of cow samples from Sardinia, and differences in sample processing between islands).

No significant difference in the frequencies of any segment was found between BTV populations from ruminants (Tukey's honestly significant difference [HSD] test for each segment with host as independent variable [Fig. 2A]; the full statistical output can be found at https://github.com/loire/MatSupp_Moreau20JVir). In contrast, segments 1, 7, and 10 showed significant differences in their frequencies between midge and ruminant samples (Fig. 2A). A significant difference was also identified for segment 2 between midge and sheep samples (Fig. 2A).

TABLE 3 Analysis of factors influencing segment frequencies in three data sets[a]

Factor	*F* value			*P* value		
	Raw	Corrected	Wo-10	Raw	Corrected	Wo-10
Segment	52.16	56.77	68.00	<2e−16	<2e−16	<2e−16
Host	0	0	0	1.000	1.000	1.000
Island	0	0	0	1.000	1.000	1.000
Segment:host	6.18	6.21	5.62	6.51e−13	5.47e−13	1.71e−10
Segment:island	5.58	6.57	7.95	3.87e−7	1.38e−8	1.19e−9
Host:island	0	0	0	1.000	1.000	1.000
Segment:host:island	1.92	1.91	2.14	0.048	0.049	0.032

[a]Analyses of variance (model: frequency ~ segment × host × island) with three data sets (raw, corrected, and Wo-10). Raw, raw copy numbers per segment and sample; corrected, copy numbers after subtraction of copy numbers found in primer-independent amplifications; Wo-10, raw copy numbers for all segments but segment 10.

These results supported the existence of host-specific set-point GFs in BTV-4 (Fig. 4). The relative abundances of segments 5, 6, 7, and 10 were between 2- and 3-fold higher than that of segment 1 in midges, whereas differences in ruminants were below 2-fold (Fig. 4). The largest differences were in the same order as those in nonmultipartite viruses with GFs improving fitness (between 3- and 4-fold) (8, 9).

DISCUSSION

Deletions and nonconservative encapsidation have been observed in BTV populations infecting cell culture (17–19). These phenomena, undistinguishable with our PCR approach, could generate the observed differences among segments. Moreover, both phenomena would result in differences in the number of fully functional copies between segments and thus impact gene expression. However, if random, these phenomena alone cannot explain reproducible host-specific GFs. The observed GFs may also be (i) generated by host factors, (iii) encoded in the virus genome, or (iii) the product of selection on randomly generated GF variation (1, 5).

Host-specific GFs are expected under a hypothetical role in viral gene expression. However, we cannot provide a robust functional explanation on the differences between segments or hosts. For example, differences in copy number roughly followed stoichiometry in capsids for segment 7 (Fig. 4). However, this trend was perceptible only in midges whereas capsid structure is supposedly similar in ruminants and midges.

Whatever their role (or absence of), unbalanced GFs should have an impact on BTV evolution. If segments differ in their population sizes (i.e., their copy number), we could expect different evolutionary rates among segments and the genes they carry (5, 20–22). Moreover, a given segment could have different evolutionary rates depending on the host.

Overall, our results strongly suggest that wild BTV populations can have set-point GFs that differ from segment equimolarity and, even more intriguingly, between insect

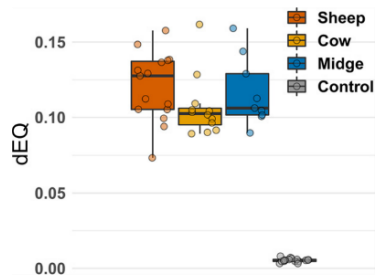


FIG 3 Distributions of segment frequencies per sample and host. Each dot indicates the distance to equimolarity of segment frequencies in a given sample (dEQ; see Materials and Methods for the formula). dEQ estimates obtained with the control plasmid (boxplot in gray) are provided to allow comparison with samples bearing equimolar amounts of PCR targets.

TABLE 4 Deviation from segment equimolarity[a]

	Sheep		Cow		Midges	
Segment	Less	Greater	Less	Greater	Less	Greater
1	5.31e−8	1	1.35e−9	1	1.32e−6	1
2	1	1.87e−3	1	0.295	0.106	1
3	0.293	1	0.431	1	0.117	1
4	7.34e−2	1	0.152	1	0.106	1
5	1	2.61e−5	1	4.88e−5	1	1.63e−4
6	0.571	1	1	9.58e−3	0.449	1
7	1	1.13e−4	1	9.57e−5	1	2.07e−4
8	0.589	1	0.152	1	0.106	1
9	3.97e−8	1	5.14e−10	1	2.09e−4	1
10	0.513	1	1	0.673	1	9.28e−4

[a]Relative segment frequencies were compared to the frequency expected under segment equimolarity (frequency = 0.1) using Student *t* tests (one-sided, correction with Benjamini-Hochberg method). The *P* values shaded in gray correspond to segments for which frequencies significantly deviated from equimolarity only in a single host (segments 2, 6, and 10).

and ruminant hosts. This work expands our limited knowledge on gene frequency variation in viruses and calls for studies on its role and conservation in BTV and beyond.

MATERIALS AND METHODS

Sample collection and RNA extraction. Samples were obtained from two Mediterranean islands, Corsica (France) and Sardinia (Italy), between 2016 and 2017. Blood samples from sheep were collected from symptomatic animals from farms undergoing a BTV outbreak in both Corsica and Sardinia. Blood samples from cows were obtained during routine testing for BTV in slaughterhouses in Corsica. Biting midges were collected using OVI traps as previously described (23), except that the water in the traps that was replaced with a solution limiting RNase activity (24). *Culicoides imicola* females were identified on a cold plate, gathered in pools of 50 individuals, and stored in absolute ethanol at −20°C for further investigation.

RNA extraction from Corsican samples was performed with the RNA viral extraction kit (Macherey-Nagel) following manufacturer's instructions and using linear acrylamide (New England Biolabs) as carrier. Sardinian samples were extracted with the MagMAX core nucleic acid purification kit (Applied Biosystems) following manufacturer's instructions.

Development and validation of the quantitative PCR designs. Primers were designed with the LightCycler probe design software 2.0 (Roche) using as the template the segment sequences of strain BTV4-16-03 (GenBank accession numbers KY654328 to KY654337). Primer pairs located in the central regions of segments were preferred. Primers for the reverse transcription of the complementary strand were tested for potential hybridization on the coding strand using Geneious v10 software, and only primers with at least 5 mismatches were selected (Table 2). Then, primer pairs underwent several

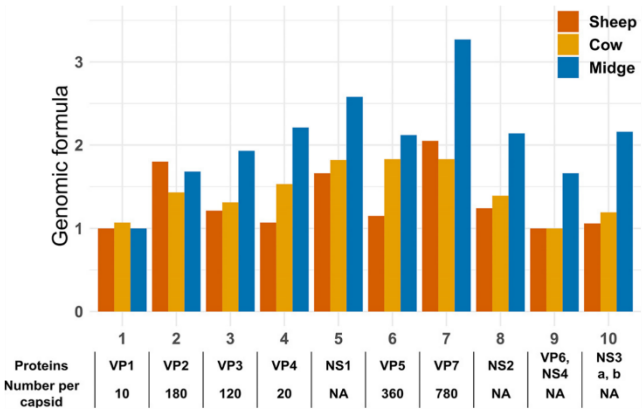


FIG 4 Set-point genome formulas are host specific in BTV-4. Set-point GFs were calculated for each host through dividing the median frequency of a given segment by the lowest median frequency among segments. The proteins encoded by each segment are shown below the graph (VP and NS stand for viral particle and nonstructural protein, respectively). Moreover, the stoichiometry in the capsid for structural proteins is provided when available (11). This information is provided to allow comparison between stoichiometry and set-point GFs.

TABLE 5 Reaction conditions for the RT-qPCR assays: reagents and their volumes[a]

Reaction	Reagents	Vol (μl)
DMSO treatment	DMSO (100%)	2.8
	RNA template	11.2
RT step 1	Denatured RNA template	1
	Forward primer (20 μM)	1
	Nuclease-free water	10
RT step 2	5× reaction buffer	4
	RiboLock RNase inhibitor (20 U/μl)	1
	10 mM dNTP mix	2
	Revert-Aid RT (200 U/μl)	1
qPCR	Nuclease-free water	2
	Forward/reverse primers (10×)	1
	Master mix SYBR (2×)	5
	cDNA template	2

[a]Abbreviations: DMSO, dimethyl sulfoxide; dNTP, deoxynucleoside triphosphate.

experimental validations. First, primer pairs were tested for efficiency using serial dilutions of RNA suspensions from infected Vero cells as a template in the RT step. Dilutions were made in RNA suspensions from noninfected cell culture, to maintain RNA concentration constant. Primer pairs showing PCR efficiencies between 90% and 100% and a single peak during melting curve analysis of amplicons were selected for further validation. Specificity for selected primer pairs was tested against cDNA suspensions obtained from RT-processed RNA from BTV-free sheep blood and *Culicoides* biting midges and with concentrations similar to those of the field samples. Only primer pairs providing C_T values above 34 were selected for further validation. Another round of efficiency validation was then carried out but this time using a plasmid as the template. The plasmid contained a concatenate of the regions targeted by primer pairs (i.e., each amplicon plus 20-bp flanking regions in both extremities of each amplicon), thus ensuring exactly the same initial number of copies per amplicon. Only primer pairs showing similar efficiencies and profiles of the melting temperature analyses between assays using either plasmids or cDNA were selected. Moreover, primer concentration was optimized to improve efficiency so as to obtain primer pairs with efficiencies at most differing in 5 units and, thus, with highly similar PCR dynamics. In total, we tested 61 primer pairs to generate the final set (Table 2; the primers rejected can be found at https://github.com/loire/MatSupp_Moreau20JVir).

Reaction and cycling conditions of RT-qPCRs. Tables 5 and 6 present the reaction and cycling conditions. First, RNA solutions were denatured with a dimethyl sulfoxide (DMSO) treatment. Then, RT and qPCRs were carried out in a two-step process. RT reactions were prepared separately for each segment as presented in Tables 5 and 6. RT reactions were performed with primers at a final concentration of 2 μM for all segments and as described in Table 5 with the Revert-Aid kit (Thermo Fisher Scientific) in a Simpliamp Thermal Cycler (Thermo Fisher Scientific). The RT primer volume was replaced with nuclease-free water in tests for primer-independent cDNA amplification. All qPCRs were carried out in triplicate (primer concentrations are shown in Table 2) with the LightCycler FastStart DNA Master Plus SYBR green I kit (Roche) in a LightCycler 480 thermocycler (Roche).

Statistical analyses. All statistical tests were carried out with the R software (R Development Core Team, 2011, version 1.2.5033). Distances of segment frequencies in a sample to equimolarity (dEQ)

TABLE 6 Cycling conditions for the RT-qPCR assays: temperature and length of each RT-qPCR step

Reaction	Length	Temp (°C)	No. of cycles	Step
Denaturation	3 min	95	1	DMSO treatment
RT step 1	5 min	65	1	Hybridization
RT step 2	60 min	42	1	Reverse transcription
	5 min	70	1	Reverse transcription
qPCR	10 min	95	1	Denaturation
	15 s	95	45	Amplification
	20 s	60	45	Amplification
	20 s	72	45	Amplification
	1 min	95	1	Melting curve
	1 min	58	1	Melting curve
	0.2°C/s	58–97	1	Melting curve

were calculated as $\mathrm{dEQ} = \sum_{i}^{10} |p_i - 0.1|/2$ according to the work of Manly (equation 5.7, p. 68 of reference 25), where *i* is the segment, *p* is the relative frequency of the segment in the sample, and 0.1 is the relative frequency at equimolarity.

ACKNOWLEDGMENTS

This work was supported by research grants CuliOme (ANIHWA, ERA-Net) and PALE-Blu (H2020) from European Union. This work was also funded by the Direction générale de l'alimentation from the French Ministry of Agriculture and Food.

We are grateful to the Directions départementales de la cohésion sociale et de la protection des populations from South Corsica and High Corsica, and to the Groupement Défense Sanitaire Bétail from High Corsica, for their support in collecting biting midges. Finally, we are grateful to Gilles Meyer, Claude Hamers, and Mathieu Chevrier for fruitful discussions during the preliminary steps of this work.

REFERENCES

1. Sicard A, Yvon M, Timchenko T, Gronenborn B, Michalakis Y, Gutierrez S, Blanc S. 2013. Gene copy number is differentially regulated in a multipartite virus. Nat Commun 4:2248. https://doi.org/10.1038/ncomms3248.
2. Sicard A, Michalakis Y, Gutierrez S, Blanc S. 2016. The strange lifestyle of multipartite viruses. PLoS Pathog 12:e1005819. https://doi.org/10.1371/journal.ppat.1005819.
3. Hu Z, Zhang X, Liu W, Zhou Q, Zhang Q, Li G, Yao Q. 2016. Genome segments accumulate with different frequencies in Bombyx mori bidensovirus. J Basic Microbiol 56:1338–1343. https://doi.org/10.1002/jobm.201600120.
4. Wu B, Zwart MP, Sánchez-Navarro JA, Elena SF. 2017. Within-host evolution of segments ratio for the tripartite genome of alfalfa mosaic virus. Sci Rep 7:5004. https://doi.org/10.1038/s41598-017-05335-8.
5. Gutiérrez S, Zwart MP. 2018. Population bottlenecks in multicomponent viruses: first forays into the uncharted territory of genome-formula drift. Curr Opin Virol 33:184–190. https://doi.org/10.1016/j.coviro.2018.09.001.
6. Sicard A, Pirolles E, Gallet R, Vernerey MS, Yvon M, Urbino C, Peterschmitt M, Gutierrez S, Michalakis Y, Blanc S. 2019. A multicellular way of life for a multipartite virus. Elife 8:e43599. https://doi.org/10.7554/eLife.43599.
7. Zwart MP, Elena SF. 2020. Modeling multipartite virus evolution: the genome formula facilitates rapid adaptation to heterogeneous environments. Virus Evol 6:veaa022. https://doi.org/10.1093/ve/veaa022.
8. Lopez-Ferber M, Simon O, Williams T, Caballero P. 2003. Defective or effective? Mutualistic interactions between virus genotypes. Proc Biol Sci 270:2249–2255. https://doi.org/10.1098/rspb.2003.2498.
9. Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. 2014. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. Proc Natl Acad Sci U S A 111:16854–16859. https://doi.org/10.1073/pnas.1415396111.
10. Simón O, Williams T, Cerutti M, Caballero P, López-Ferber M. 2013. Expression of a peroral infection factor determines pathogenicity and population structure in an insect virus. PLoS One 8:e78834. https://doi.org/10.1371/journal.pone.0078834.
11. Roy P. 2017. Bluetongue virus structure and assembly. Curr Opin Virol 24:115–123. https://doi.org/10.1016/j.coviro.2017.05.003.
12. Sailleau C, Breard E, Viarouge C, Gorlier A, Quenault H, Hirchaud E, Touzain F, Blanchard Y, Vitour D, Zientara S. 2018. Complete genome sequence of bluetongue virus serotype 4 that emerged on the French island of Corsica in December 2016. Transbound Emerg Dis 65:e194–e197. https://doi.org/10.1111/tbed.12660.
13. Mertens PPC, Diprose J. 2004. The bluetongue virus core: a nano-scale transcription machine. Virus Res 101:29–43. https://doi.org/10.1016/j.virusres.2003.12.004.
14. Feng L, Lintula S, Ho TH, Anastasina M, Paju A, Haglund C, Stenman UH, Hotakainen K, Orpana A, Kainov D, Stenman J. 2012. Technique for strand-specific gene-expression analysis and monitoring of primer-independent cDNA synthesis in reverse transcription. Biotechniques 52:263–270. https://doi.org/10.2144/0000113842.
15. Plaskon NE, Adelman ZN, Myles KM. 2009. Accurate strand-specific quantification of viral RNA. PLoS One 4:e7468. https://doi.org/10.1371/journal.pone.0007468.
16. Fenaux M, Cuadras MA, Feng N, Jaimes M, Greenberg HB. 2006. Extraintestinal spread and replication of a homologous EC rotavirus strain and a heterologous rhesus rotavirus in BALB/c mice. JVI 80:5219–5232. https://doi.org/10.1128/JVI.02664-05.
17. Anthony SJ, Darpel KE, Belaganahalli MN, Maan N, Nomikou K, Sutton G, Attoui H, Maan S, Mertens PP. 2011. RNA segment 9 exists as a duplex concatemer in an Australian strain of epizootic haemorrhagic disease virus (EHDV): genetic analysis and evidence for the presence of concatemers as a normal feature of orbivirus replication. Virology 420:164–171. https://doi.org/10.1016/j.virol.2011.09.009.
18. Patel A, Roy P. 2014. The molecular biology of Bluetongue virus replication. Virus Res 182:5–20. https://doi.org/10.1016/j.virusres.2013.12.017.
19. Squire KRE, Osburn BI, Chuang RY, Doi RH. 1983. A survey of electropherotype relationships of bluetongue virus isolates from the western United States. J Gen Virol 64:2103–2115. https://doi.org/10.1099/0022-1317-64-10-2103.
20. Jacquot M, Rao PP, Yadav S, Nomikou K, Maan S, Jyothi YK, Reddy N, Putty K, Hemadri D, Singh KP, Maan NS, Hegde NR, Mertens P, Biek R. 2019. Contrasting selective patterns across the segmented genome of bluetongue virus in a global reassortment hotspot. Virus Evol 5:vez027. https://doi.org/10.1093/ve/vez027.
21. Carpi G, Holmes EC, Kitchen A. 2010. The evolutionary dynamics of bluetongue virus. J Mol Evol 70:583–592. https://doi.org/10.1007/s00239-010-9354-y.
22. Nomikou K, Hughes J, Wash R, Kellam P, Breard E, Zientara S, Palmarini M, Biek R, Mertens P. 2015. Widespread reassortment shapes the evolution and epidemiology of bluetongue virus following European invasion. PLoS Pathog 11:e1005056. https://doi.org/10.1371/journal.ppat.1005056.
23. Balenghien T, Pages N, Goffredo M, Carpenter S, Augot D, Jacquier E, Talavera S, Monaco F, Depaquit J, Grillet C, Pujols J, Satta G, Kasbari M, Setier-Rio ML, Izzo F, Alkan C, Delecolle JC, Quaglia M, Charrel R, Polci A, Breard E, Federici V, Cetre-Sossah C, Garros C. 2014. The emergence of Schmallenberg virus across Culicoides communities and ecosystems in Europe. Prev Vet Med 116:360–369. https://doi.org/10.1016/j.prevetmed.2014.03.007.
24. Camacho-Sanchez M, Burraco P, Gomez-Mestre I, Leonard JA. 2013. Preservation of RNA and DNA from mammal samples under field conditions. Mol Ecol Resour 13:663–673. https://doi.org/10.1111/1755-0998.12108.
25. Manly BFJ. 2005. Multivariate statistical methods. A primer, 3rd ed. Chapman & Hall/CRC, Boca Raton, FL.

ANNEXES

Tableau S1. Analysis of host effect on segment frequency in three datasets. Analysis of variance per segment (model: frequency ~ host) in three datasets (raw, corrected and Wo-10). Raw: raw copy numbers per segment and sample. Corrected: copy numbers after subtraction of copy numbers found in primer-independent amplifications. Wo-10: raw copy numbers for all segments but segment 10.

	F-value			p-value		
segment	**Raw**	**Corrected**	**Wo-10**	**Raw**	**Corrected**	**Wo-10**
1	21.04	19.02	15.87	1.29e-6	3.21e-6	1.48e-5
2	6.534	7.113	6.194	4.07e-3	2.7e-3	5.2e-3
3	0.124	0.135	6.194	0.884	0.874	0.933
4	0.076	0.079	0.069	0.927	0.924	0.711
5	0.792	0.570	0.344	0.461	0.571	0.366
6	2.432	2.590	1.037	0.103	0.092	0.080
7	6.990	6.635	2.735	2.94e-3	3.78e-3	6.95e-4
8	0.950	1.110	9.138	0.397	0.341	0.426
9	2.050	1.870	0.876	0.145	0.169	0.041
10	9.548	10.89	/	5.35e-4	2.33e-4	/

Tableau S2. Analysis of host effect on segment frequency in three datasets. p-values of Tukey-HSD tests obtained from the analyses of variance in Table A4-A. Raw: raw copy numbers per segment and sample. Corrected: copy numbers after subtraction of copy numbers found in primer-independent amplifications. Wo-10: raw copy numbers for all segments but segment 10. Differences in significance levels were found between the Wo-10 dataset and the others for segment 9 (significant differences between ruminant and midge samples only found with Wo-10; p-value = 0.044).

			p-value		
Segment	Host1	Host2	Raw	Corrected	Wo-10
1	Cow	Midge	3.68e-6	8.83e-6	3.09e-5
1	Cow	Sheep	0.845	0.866	0.813
1	Midge	Sheep	7.43e-6	1.65e-5	8.05e-5
2	Cow	Midge	0.185	0.167	0.262
2	Cow	Sheep	0.167	0.139	0.138
2	Midge	Sheep	2.97e-3	1.96e-3	4.1e-3
3	Cow	Midge	0.886	0.875	0.990
3	Cow	Sheep	0.996	0.995	0.970
3	Midge	Sheep	0.911	0.905	0.931
4	Cow	Midge	0.999	0.998	0.928
4	Cow	Sheep	0.930	0.927	0.886
4	Midge	Sheep	0.958	0.957	0.695
5	Cow	Midge	0.800	0.820	0.346
5	Cow	Sheep	0.428	0.543	0.626
5	Midge	Sheep	0.877	0.936	0.808
6	Cow	Midge	0.237	0.229	0.325
6	Cow	Sheep	0.110	0.093	0.0686
6	Midge	Sheep	0.983	0.971	0.823
7	Cow	Midge	3.45e-3	4.48e-3	1.16e-3
7	Cow	Sheep	0.789	0.810	0.893
7	Midge	Sheep	0.011	0.013	2.23e-3
8	Cow	Midge	1.000	0.993	0.673
8	Cow	Sheep	0.465	0.377	0.405
8	Midge	Sheep	0.511	0.507	0.951
9	Cow	Midge	0.129	0.152	0.044
9	Cow	Sheep	0.780	0.799	0.885
9	Midge	Sheep	0.319	0.349	0.087
10	Cow	Midge	5.36e-3	2.73e-3	/
10	Cow	Sheep	0.714	0.704	/
10	Midge	Sheep	4.79e-4	2.15e-4	/

CHAPITRE 2 : La variation du nombre de copies de gènes chez le virus de la FCO-4 est associée au succès de l'infection de l'insecte vecteur *Culicoïdes imicola*

Dans le cycle des arbovirus, l'espèce vectrice est responsable de la dynamique de transmission virale. Les interactions entre hôtes sont favorisées par le caractère obligatoire d'interactions trophiques. La capacité du virus à infecter l'hôte vecteur est essentielle mais contrainte par des facteurs liés à l'espèce, sa génétique, sa physiologie et les processus sélectifs auxquels le virus fait face au cours de sa propagation.

Pour les arbovirus d'insectes hématophages le succès du virus à se disséminer jusqu'aux glandes salivaire pour être de nouveau transmis implique un ensemble d'étapes au cours desquels la population virale doit coloniser plusieurs organes en étant soumise à des effets stochastiques de dérive génétique et à la pression de sélection. Pour le virus de la FCO nous avons établi qu'au niveau de la population le nombre de copies de gènes pouvait varier d'une manière dépendante de l'hôte dans les conditions naturelles d'infection.

Dans ce chapitre nous questionnons le caractère adaptatif du phénomène de la variation du nombre de copies de gènes au cours de l'infection de l'espèce vectrice *Culicoides imicola* par le virus de la FCO-4. Nous présentons ici les résultats d'une approche expérimentale d'infection par voie orale mise en place à l'Istituto Zooprofilattico Sperimentale de Teramo qui nous a permis d'étudier la formule génomique de populations virales de la FCO-4 en fonction du succès du virus à se disséminer jusqu'à la tête après dix jours d'infection. Ces résultats montrent des changements de fréquence des segments génomiques pour toutes les populations virales étudiées, confirmant l'existence de la VNCG dans ces conditions d'infection expérimentale et prouvant l'existence d'un changement dynamique au cours du temps. Les formules génomiques des populations virales présentent dans le corps des individus sont différentes entre les deux états de dissémination du virus. Les populations virales pour lesquelles la dissémination est limitée au corps des individus présentent des variations similaires aux virus étudiés sur le terrain au cours du premier chapitre. En revanche les populations virales à dissémination complète présentent des variations pouvant atteindre un facteur dix de différence d'abondance entre segments. Pour ces dernières, le segment 10 est le plus représenté.

Nos travaux établissent pour la première fois une corrélation entre la formule génomique et la valeur sélective de la population virale chez un arbovirus. Ces résultats soutiennent que les formules génomiques peuvent changer de manière

dynamique au cours du temps mais également que la VNCG pourrait jouer un rôle dans le succès de l'infection du virus de la FCO-4.

Ce chapitre présente des résultats qui seront valorisés suite à la thèse.

La variation du nombre de copies de gènes chez le virus de la FCO-4 est associée au succès de l'infection de l'insecte vecteur *Culicoides imicola*

Yannis Moreau[1,2*], Etienne Loire[1,2], Giovanni Savini[3], Maria Goffredo[3], Giuseppe Mancini[3], Serafìn Gutierrez[1,2*]

[1] ASTRE, Univ Montpellier, Cirad, INRA, Montpellier, France.
[2] Cirad, UMR ASTRE, F-34398 Montpellier, France.
[3] Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise G. Caporale, Teramo, Italy

* Corresponding author: Yannis Moreau, UMR 117 ASTRE, Campus international de Baillarguet, 34398 Montpellier cedex 5, yannis.moreau@cirad.fr; or Serafìn Gutierrez, UMR 117 ASTRE, Campus international de Baillarguet, 34398 Montpellier cedex 5, serafin.gutierrez@cirad.fr

RESUME

Dans les populations naturelles du virus de fièvre catarrhale ovine (FCO) des variations dans la fréquence des segments génomiques ont été observées. Ces profils de variations sont associés à l'hôte, par analogie avec les virus multipartites nous considérons ces profils comme des formules génomiques (FG). Notre étude a pour objectifs de déterminer si ces variations sont produites lors de l'infections des culicoïdes et, de plus, si elles ont un lien avec la valeur sélective du virus. Une approche d'infection expérimentale incluant des femelles adultes de l'espèce *Culiloïdes imicola* infectées avec le sérotype 4 du virus de la FCO nous a permis de décrire un lien possible entre la variation du nombre de copies de gènes (VNCG) et la valeur sélective du virus. Nous avons comparé les fréquences des segments présentes initialement dans l'inocula aux fréquences des segments dans les populations virales 10 jours après le repas de sang. La VNCG a été comparée entre les populations de virus en fonction de leur succès à se disséminer jusqu'à la tête qui contient une partie des glandes salivaires. Pour le virus de la FCO-4, les fréquences des segments génomiques peuvent varier au cours du temps. La divergence par rapport à la FG de l'inocula est plus importante dans les populations où le virus s'est disséminé jusqu'à la tête. Le succès de dissémination n'a pas eu d'effet sur la charge virale. En revanche, nous avons trouvé un profil particulier de fréquence des segments corrélé à la capacité du virus à se disséminer jusqu'à la tête. Ces populations du virus de la FCO à dissémination « complète » présentaient des écarts plus importants de fréquences entre segments. Ces résultats établissent pour la première fois chez un arbovirus une corrélation entre la formule génomique et la capacité du virus à se disséminer dans un hôte.

Mots-clefs : FCO - culicoïdes - arbovirus - formule génomique - adaptation

INTRODUCTION

La variation du nombre de copies de gènes (VNCG) pourrait être un mécanisme adaptatif négligé chez les virus, avec un rôle potentiel sur la régulation de l'expression des gène viraux (1-3). Comme cela a été proposé chez le *Faba bean necrotic stun virus*, la VNCG pourrait jouer un rôle dans l'adaptation au changement d'espèce hôte (1). La convergence des segments viraux vers un profil particulier de fréquence, la formule génomique, a été observée chez deux espèces de virus multipartites (1, 4). Chez les virus à génome segmenté le lien entre le nombre de copies des segments (NC), l'expression des gènes et la valeur sélective a été montré pour une souche clonale, sélectionnée en laboratoire, du virus de la grippe A. Pour cette dernière la diminution du NC d'un gène codant pour une protéine d'enveloppe diminue son expression et augmente la valeur sélective de la population virale (2). Nous avons récemment démontré l'existence de la VNCG, pour la première fois chez un arbovirus animal. Les profils de fréquences des segments suggèrent que la variation du nombre de copies de segments pourrait jouer dans le cycle en alternance d'hôtes du virus de la FCO-4 en condition naturelle d'infection (5). Ces résultats soulèvent plusieurs points d'interrogation. Si nous sommes capables de discriminer des variations de fréquences entre hôtes, nous ne pouvons démontrer que le nombre de copies de gènes change au cours de l'infection d'un individu. En effet, la taille de l'échantillonnage étant limitée, une mise en évidence du changement de la formule lors d'un changement d'espèce hôte était nécessaire. De la même manière, nous ne disposions pas d'élément permettant de discuter le caractère adaptatif de ces profils particuliers obtenus pour les différentes espèces hôtes. Notre approche présente vise à répondre à ces questions chez l'hôte arthropode *Culicoïdes imicola*, un des plus importants vecteurs du virus de la FCO. Une approche d'infection expérimentale mise en place par nos collaborateurs de l'IZS de Teramo nous a permis d'étudier la variation des fréquences des segments génomiques lors de l'infection de *C. imicola* mais aussi un possible lien entre VNCG et valeur sélective du virus.

Les arbovirus animaux sont majoritairement vectorisés par des insectes hématophages appartenant à l'ordre des *Diptera* (6). Le mode de transmission horizontal est le plus répandu pour ces arbovirus, il implique la réplication du virus

dans les deux hôtes et une transmission du virus en alternance entre deux hôtes distants (6). La distribution des espèces vectrices est associée à la présence de certaines arboviroses et l'extension de leurs aires de distribution vers le Nord mène à l'apparition en zones tempérées d'arbovirus jusqu'alors exclusivement associés aux régions tropicales (7). Pour le virus de la FCO en particulier, le cycle saisonnier de la transmission du virus est associé aux cycles des culicoïdes vecteurs (8). Malgré l'existence de modes de transmission alternatifs, le rôle des espèces vectrices reste essentiel à la diffusion de la maladie (8).

Le corps des insectes vecteurs peut-être perçu comme une succession de barrières limitant la propagation virale et participant à diminuer la probabilité de transmission du virus (8). D'autres facteurs dépendants de la génétique de l'hôte peuvent aussi moduler les voix de défenses immunes innées (9). Les particules virales sont ingérées par l'insecte au cours de la prise d'un repas sanguin sur un animal infecté. Suite au gorgement du vecteur, différents scénarios sont possibles (8), une arborescence récapitulative est présentée dans la figure 1A. Chez les culicoïdes, le virus de la FCO circule des pièces buccales jusqu'au tube digestif (Figure 1B) et peut initier l'infection au niveau des cellules épithéliales de l'intestin moyen (10). En cas de succès, il est relâché dans le système circulatoire ouvert et peut infecter des cibles secondaires comme les corps gras ou encore les cellules nerveuses (10). Pour compléter son cycle d'infection, le virus de la FCO doit infecter les glandes salivaires et peut alors être libéré pour former des structures agglomérées de particules virales dans les canaux efférents des glandes salivaires (8, 11). Entre l'ingestion et la transmission, une population virale circule dans les trois parties du corps du culicoïdes (Figure 1B). Le temps qui sépare ces deux évènements est la période d'incubation extrinsèque, elle est dépendante du virus, de l'espèce vectrice et des conditions de température, et sa durée varie de 7 à 14 jours en fonction des conditions (8). Une température plus élevée est souvent associée à une vitesse plus rapide de l'infection du virus dans son hôte ectotherme (12).

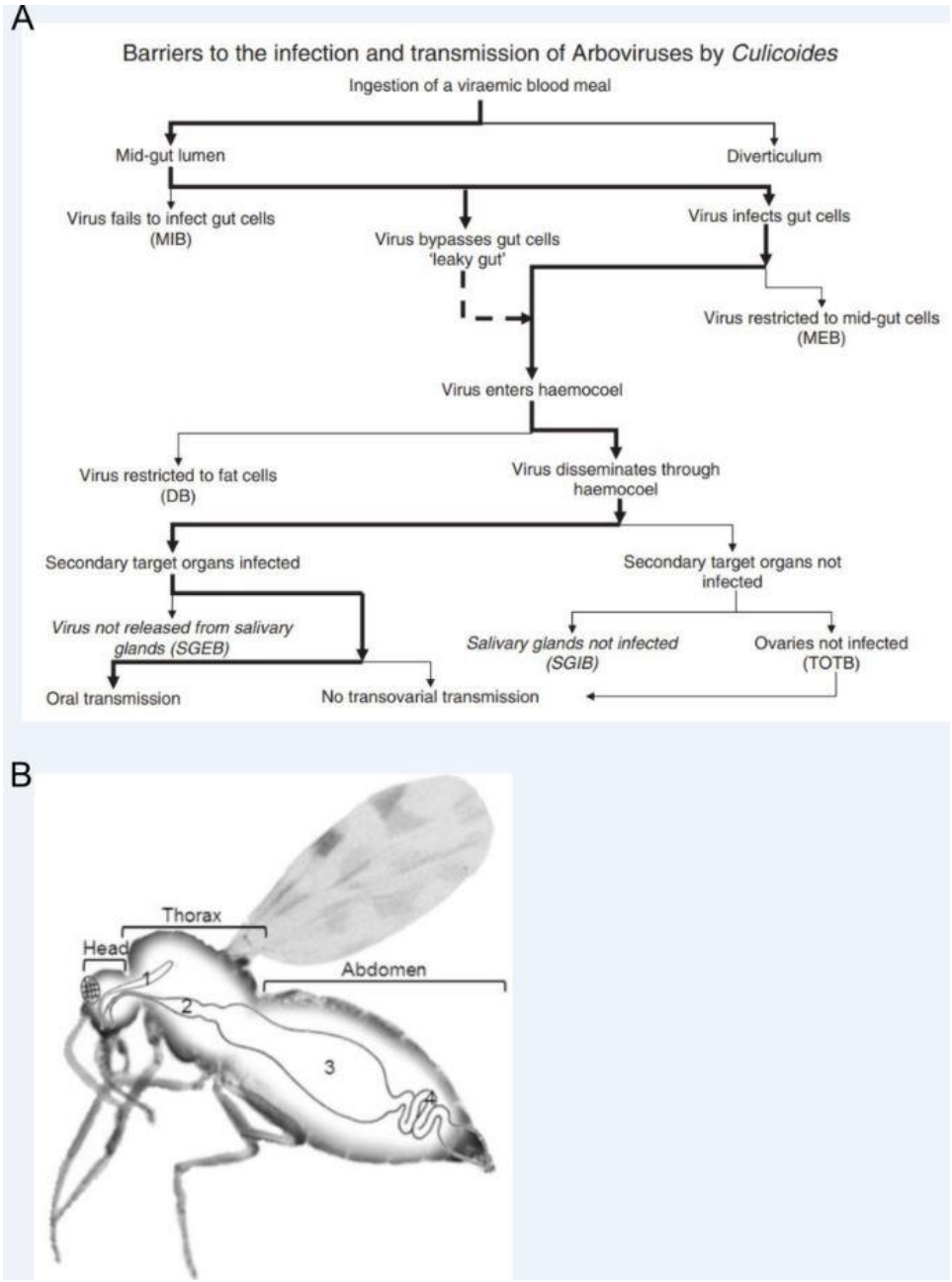


Figure 1. (A) Les barrières à l'infection, la dissémination et *in fine à* la transmission dans les culicoïdes et les moustiques. Les flèches en gras indiquent le parcours suivit par un arbovirus de l'ingestion à la transmission par voie orale. La figure est tirée du livre Bluetongue par Mellor *et al*. (8), elle a été modifiée à partir de la publication originelle de Hardy et al. (11). **(B) Parties du corps d'un culicoïde avec les glandes salivaires** en 1), le tube digestif divisé en trois sous-parties : l'intestin antérieur en 2), l'intestin moyen en 3) et l'intestin postérieur en 4). L'infection par les arbovirus se produit spécifiquement dans l'intestin moyen. L'illustration est tirée de la publication de Feenstra et al. (13), modifiée à partir de l'originale de A. R. Walker.

Certaines protéines du virus de la FCO ont des effets qui varient en fonction de l'hôte (14). Chez ce virus non-enveloppé, le clivage de la couche externe de la capside virale constituée de la protéine VP2 augmente son infectivité dans les cellules d'insectes mais aussi dans les culicoïdes adultes (15). Cette dénaturation pourrait être un phénomène courant après l'ingestion du virus par les culicoïdes, sous l'action de protéases présentes dans le tube digestif (16). La particule core (sans les protéines de capsides VP2 et VP5) possède une capside externe constituée de la protéine VP7, il s'agit de la couche intermédiaire de la capside virale. Cette forme de virion est majoritaire dans les cellules infectées. La particule core est aussi infectieuse pour les cellules d'insectes et les culicoïdes (16). L'infection du vecteur est possible pour différentes formes du virus qui, soit produisent un effet bénéfique, soit ne modifient pas son infectivité pour l'hôte insecte. Plusieurs protéines non structurelles influencent aussi la valeur sélective du virus de la FCO en fonction du type cellulaire infecté, le rôle des protéines non-structurelles NS3 et NS3a semble déterminant (13, 17).

Le segment 10 code pour les protéines non structurelles NS3 et NS3a. Elles sont impliquées dans la sortie non lytique des cellules (18), cruciale lors de l'infection de culicoïdes adultes (13), mais aussi dans l'inhibition de la réponse interféron de type I (19), principalement chez les mammifères. Ainsi le segment 10 a également un impact sur la virulence au cours de l'infection des cellules mammifères (20). La dénaturation du segment 10 n'empêche pas la réplication du virus du virus de la FCO dans les cellules mammifères BSR et insectes KC, mais retarde la sortie du virus avec un effet plus marqué dans le type cellulaire insecte (21). En revanche, au cours d'infection de culicoïdes adultes, son intégrité semble essentielle à la propagation du virus, lors d'infection de l'espèce vectrice *C. sonorensis* par voix intra-thoracique, la dissémination du virus est dépendante de l'expression des protéines NS3 et NS3a en particulier (13). La nature de ces protéines peut influencer la cinétique de réplication du virus de la FCO dans le type cellulaire mammifère (17). Ce segment est le plus petit présent chez les *Orbivirus* avec une taille inférieure à 1Kb et une séquence qui peut varier entre les espèces du genre jusqu'à 36%. E. C. Holmes et ses collègues ont également décrit ce segment comme étant parmi les plus variable existant chez le virus de la FCO (22). Ces éléments soulignent l'importance de la composante génétique d'un segment particulier avec un effet dépendant de l'hôte.

Dans ce contexte la variation du nombre de copies de gènes pourrait non seulement participer à la régulation de l'expression mais aussi être lié à la génération de diversité génétique aboutissant *in fine* à une augmentation de la probabilité de mutations adaptatives (23). Nous allons étudier son influence sur la capacité du virus de la FCO-4 à se disséminer et à se répliquer au cours de l'infection du vecteur *C. imicola*.

MATERIELS ET METHODES

Préambule

Les travaux présentés dans ce chapitre sont le fruit d'une collaboration entre le groupe de G. Savini, de l'Istituto Zooprofilattico Sperimentale (IZS) de Teramo en Italie et notre équipe du CIRAD de Montpellier. La capture, l'infection expérimentale et le crible des têtes par RT-PCR ont été réalisé par nos collègues de l'IZS. Le crible des corps, la quantification par RT-qPCR et les analyses statistiques ont été faites au CIRAD.

Capture de culicoïdes

Des femelles culicoïdes (Ordre : Diptera, Famille : Ceratopogonidae) appartenant à l'espèce *Culicoides imicola* ont été collectées dans le cadre d'une étude sur la compétence vectorielle de plusieurs espèces de culicoïdes pour les sérotypes du virus de la FCO circulant en Europe (24). Les femelles non-pigmentées utilisées pour l'expérience ont été piégées sur l'île de Sardaigne en 2016. Des pièges de type OVI-TRAP (25) étaient disposés dans les élevages. Les culicoïdes ont été acclimatés en cage à filet dans une enceinte climatique à une température de 25°C, hygrométrie > 40% pendant 48h. Un coton humidifié avec une eau à 10% de sucrose a été ajouté *ad libitum* pour l'alimentation (24).

Infection expérimentale des culicoïdes

Le virus de la FCO-4 utilisé pour l'expérience provient du sang d'un mouton infecté en 2014 dans la région d'Apulie en Italie. Il a été isolé sur des cellules de culicoïdes KC puis amplifié 3 fois sur des cellules mammifères VERO. La souche est nommée en conséquence de ces passages successifs FCO-4_K1V3. Les femelles culicoïdes sélectionnées pour l'expérience ont pu se gorger avec un système Hemotek au travers d'une membrane en parafilm. Le virus FCO-4-K1V3, à un titre de $10^{5.3}$ – $10^{6.54}$ pFU/mL, a été mélangé au 1/3 avec du sang de mouton hépariné et chauffé à 37°C. Après 45 minutes de gorgement, les femelles ont été replacées en cage dans

l'enceinte climatique pendant 10 jours avec un accès *ad libitum*à un coton imbibé d'eau à 10% de sucrose. Au terme, les individus ont été disséqués sur bloc froid afin de collecter séparément les têtes et les corps (Figure 1B). Les échantillons ont ensuite été conservé à -80°C jusqu'à l'extraction des ARN.

Extraction des ARN viraux

Les échantillons ont été suspendus dans un faible volume de PBS 1X et homogénéisé avec un TissueLyser II (Qiagen) à 30 Htz, deux fois 30s. Les broyats obtenus sont centrifugés à 3700 rpm pendant 5 minutes à 4°C et 150 µl du surnageant sont récupérés pour l'extraction. L'extraction a été faite avec le kit 'MagMAX CORE Nucleic Acid Purification Kit' (Applied Biosystem) en suivant les instructions du fournisseur. L'élution des ARN est faite avec de l'eau sans nucléase.

Crible des échantillons par RT-qPCR

Les têtes ont été testées individuellement avec la méthode de qPCR diagnostique de référence pour le BTV à l'IZS (26). Le taux d'individus de l'espèce *C. imicola* avec des têtes positives avait été estimé à 70.4% (27). Parmi ces positifs, nous avons sélectionné 30 individus pour lesquels les têtes ont été trouvées négatives pour les sérotypes 2, 8 et 16 (potentiellement circulant dans la région). Une deuxième analyse par RT-qPCR pour l'ensemble des segments a permis d'identifier dix-sept individus parmi ces 30 individus avec un Ct inférieur à 31 pour les 10 segments. Dans un second temps, les corps des individus ayant des têtes testées négatives au BTV (n = 40) ont été criblés avec l'approche de RT-qPCR segment-spécifique développée au cours du projet de thèse (5) en utilisant le couple de primers ciblant le segment 7. La détection du BTV dans les corps a été faite avec les primers ciblant le segment 7 qui a été trouvé comme étant le segment le plus abondant dans les échantillons de terrain positifs au BTV-4. Vingt-deux individus étaient positifs sur un total de 40. Parmi ces 22 individus, huit individus ont pu être exploités car ils possédaient un Ct inférieur à 31 pour les 10 segments.

Quantification des segments génomiques par RT-qPCR

Notre méthode de quantification des brins négatifs du virus de la FCO-4 (5) a été utilisé pour décrire les variations de fréquences des segments viraux. Les extraits ARN sont dénaturés 3 minutes à 95°C avec 20% de DMSO. L'étape de RT est réalisée avec le kit ReverAid (FicherScientific), elle est réalisée indépendamment pour chacun des 10 segments avec un contrôle pour quantifier l'effet d'amplification indépendamment de l'action des primers. La qPCR utilise la méthode de fluorescence SYBRGREEN, nous avons donc contrôlé la spécificité de nos amplifications par visualisation des courbes de dissociation à l'issue de la réaction de qPCR. Chaque réaction a été réalisée en triplica et les contrôles négatifs sont inclus pour chacune de nos cibles dans chaque plaque. Les cibles sont quantifiées à l'aide d'un plasmide qui porte un insert avec les dix cibles de qPCR encadrées en 5' et 3' par les 20pb correspond à la séquence génomique consensus du virus (28). Ce plasmide sert à contrôler la validité de notre approche en validant la présence équimolaire de chacune de nos cibles pour chaque plaque de qPCR. Il est aussi utilisé pour définir un point de gamme à une concentration fixe. Pour s'assurer de la fiabilité de nos résultats chaque échantillon a été sélectionné si les valeurs de Ct pour chacune des 10 cibles étaient inférieures à 31.

Effet de l'amplification indépendante de l'action des primers à l'étape de transcription inverse

Chaque segment a été testé pour détecter un effet d'amplification indépendante de l'action des primers à l'étape de RT (29). Le contrôle inclut 5 individus différents, sélectionnés aléatoirement. Cet effet peut être dû à de multiples molécules, comme les structures secondaires de l'ARN linéaire simple brin par exemple. Nous avions quantifié et décrit son influence sur nos résultats de RT-qPCR lors de notre étude de la variation du nombre de copies virales dans les échantillons collectés sur le terrain (5). Pour contrôler cet effet, une étape de RT est réalisée en l'absence de primer pour chaque échantillon. Les brins d'ADNc obtenus sont testés avec l'approche commune de qPCR pour chaque segment. Les produits amplifiés de cette manière résultent d'une amplification indépendante de l'action des primers à l'étape de RT. Le jeu de données brute présente les valeurs non modifiées de nombre de copies (NC). Le jeu

de données corrigé est généré par soustraction de la valeur du NC final obtenue sans primer à l'étape de RT à la valeur brute de NC pour chaque segment.

Statistiques

Les résultats ont été analysés avec le logiciel R (R Development Core Team, 2011, 239 version 1.2.5033). Les comparaisons entre conditions ont été faites dépendamment des résultats de normalité de distribution et d'homoscédasticité. Le test de Wilcoxon a été utilisé si la normalité des distributions est rejetée (test de Shapiro), le test de Welch si l'homocédasticité est rejetée (var. Test ou Bartlett) et le test de Student si toutes les conditions précédentes sont validées. Les indices de distance ont été calculés comme suit :

(1) distance à l'équimolarité : dEQ = $\sum_i^{10}|p\mathrm{i} - 0.1|/2$

(2) distance à la formule génomique de l'inocula : dIN = $\sum_i^{10}|pi - pINi|/2$

(3) distance à la formule génomique des culicoïdes de terrain : dGFa = $\sum_i^{10}|pi - pGFi|/2$

suivant l'équation de Manly (30) où *i* est le segment, *p* la fréquence relative du segment dans l'échantillon, *0.1* la fréquence relative à l'équimolarité, *pIN* la fréquence de l'inocula, *pGFa* la fréquence mediane de la formule génomique dans les infections naturelles de culicoïdes (voir Chapitre 1=. La formule génomique dans une population d'hôtes a été calculée en divisant la valeur médiane de fréquence de chaque segment par la valeur médiane de fréquence la plus faible.

RESULTATS

Amplification indépendante de l'action des primers à l'étape de transcription inverse

Parmi les 5 échantillons pour lesquels les 10 segments ont été testé, seuls les segments 2 et 10 ont montré un effet détectable de l'amplification indépendante de l'action des primers à l'étape de transcription inverse (AIAPET) (Tableau 1, Annexes Figure 1A). Cet effet est respectivement de 3% et 8% sur les quantifications de nombre de copies. Nous avons suivi l'AIAPET pour ces deux segments dans l'ensemble de nos mesures. Pour le jeu de données incluant tous les échantillons, aucun effet significatif d'AIAPET n'a été détecté pour les segments 2 et 10 (Annexes Figure 1B, 1C, Annexes Tableau 1). Pour l'ensemble des échantillons, une augmentation de l'effet sur la quantification de nombre de copies est observée par rapport au contrôle, il passe à 8 % pour le segment 2 et à 20 % pour le segment 10. Les résultats obtenus sont similaires aux résultats avec des échantillons de terrain (voir chapitre 1), pour lesquels une variation de 9 et 25% étaient observées pour les segments 2 et 10. Les résultats de RT-qPCR pour les segments 2 et 10 ont été corrigés pour cet effet dans les analyses suivantes.

Tableau 1. Effet de l'amplification aléatoire sur le nombre de copies pour les 5 échantillons du contrôle. Le facteur de différence est calculé en divisant les valeurs du jeu de données brutes par celles du corrigé.

Segment	Rapport (brute/corrigé)
1	1
2	1,03
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1,08

La charge virale, proxy de la valeur sélective d'une population virale

Classiquement la charge virale constitue un proxy de la valeur sélective. La charge virale dans les corps de chaque individu a été comparée selon l'état de dissémination du virus : dissémination complète (dissémination dans la tête) et dissémination limitée (absence de virus dans la tête) (Figure 2). Nous ne sommes pas parvenus à développer une approche quantitative robuste pour un gène de ménage chez l'espèce *C. imicola*. Nous avons donc décidé d'utiliser le nombre cumulé de copies ARN quantifié par échantillon (*i.e.* par corps sans tête). Nous n'avons pas trouvé de différence significative entre les deux états de dissémination (Wilcoxon, p-value = 0.086). La multiplication du virus dans le corps semble indépendante de la capacité des populations du virus de la FCO-4 à se disséminer jusqu'à la tête.

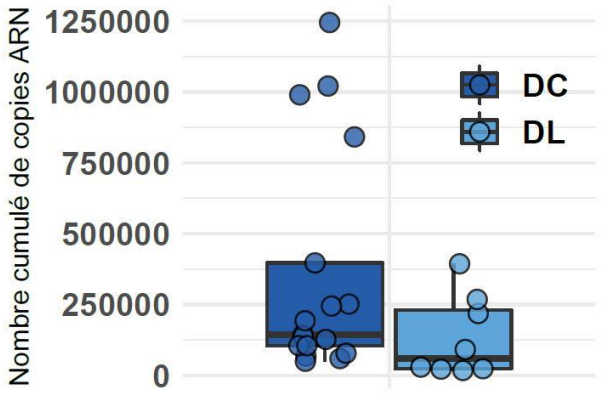


Figure 2. Nombre cumulé d'ARN viraux pour les 10 segments génomiques représenté par échantillon en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée).

La déviation des fréquences à l'équimolarité

La comparaison à la condition d'équimolarité, où chaque segment présenterai un nombre égal de copies, a été contrôlée par deux approches statistiques. Dans un premier temps le nombre de copies des segments ont été comparées à sa valeur théorique attendue en situation d'équimolarité (fréquence = 0,1) pour chaque échantillon en fonction de la dissémination du virus par un test de Chi2. Le nombre de copies déviait significativement de l'attendue théorique pour les deux états de dissémination (p-value < 2.0E-16). Pour préciser ces différences, la déviation des fréquences à cette valeur théorique de 0.1 a ensuite été testée pour chaque segment

avec un test de Student unilatéral. Pour les deux conditions, nous avons observé une diminution significative de la fréquence pour les segments 1, 3, 4 et 9 (p-value < 0.05). A l'inverse une augmentation de la fréquence est observée pour les segments 5, 7 et 10 (p-value < 0.05) (Figure 3). Ces résultats soutiennent une déviation significative de l'équimolarité pour 7 segments et ceci dans les deux états de dissémination. Des différences sont aussi observées en fonction de l'état de dissémination. Le segment 2 diminue par rapport l'attendu théorique lorsque le virus se propage jusqu'à la tête (p-value = 2.76E-14). Le segment 8 augmente significativement pour l'état de dissémination complet (p-value = 1.48E-05). Enfin le segment 6 augmente également lorsque la dissémination est limitée au corps (p-value = 0.047).

Nous avons calculé un indice de distance à l'équimolarité (dEQ) pour chaque échantillon pour déterminer si les variations de fréquences par rapport à l'équimolarité étaient liées à la dissémination du virus. L'indice dEQ a été comparé en fonction des états de dissémination (Figure 4). La distance à l'équimolarité est significativement plus élevée pour les échantillons à dissémination complète.

Nous avons testé si la dEQ pouvait être influencée par la charge virale. Pour chaque état de dissémination, une régression linéaire (dEQ ~ charge virale) n'a pas détecté d'effet significatif de la charge virale sur la dEQ (Figure 5).

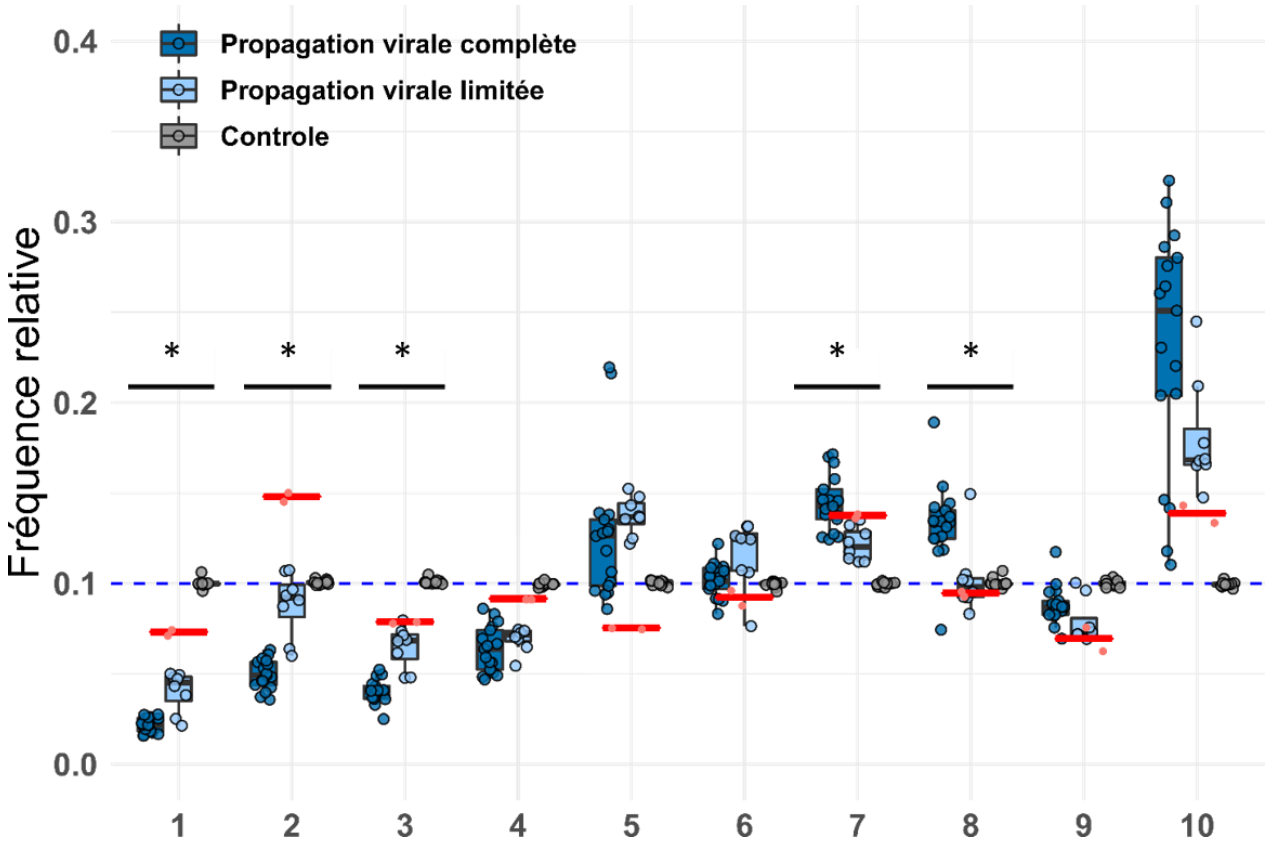


Figure 3. Fréquences relatives des segments génomiques du virus de la FCO-4 en fonction de l'état de dissémination du virus. La ligne en pointillés bleue indique la fréquence théorique attendue (0.1) pour les segments s'ils étaient tous présents en quantité équimolaire. Le contrôle correspond aux fréquences calculées à partir du plasmide pGEM-T qui porte une copie de chacune de dix cibles. Les points rouges représentent les valeurs de fréquences de l'inocula tandis que les droites de la même couleur correspondent à leurs valeurs moyennes. Les différences significatives ont été déterminées séparément pour chaque segment avec une ANOVA (fréquence ~ dissémination virale). Les différences affichées correspondent aux résultats du test de Tukey-hsd (p-value < 0.05).

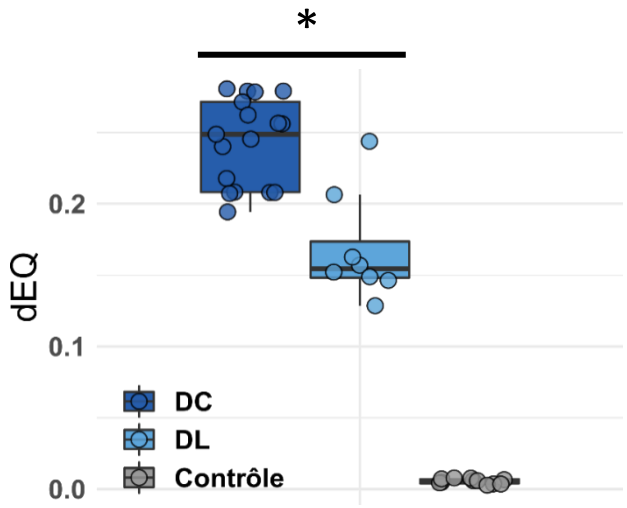


Figure 4. Indice de la distance à l'équimolarité en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée) et pour le contrôle pGEM-T portant chaque cible à un ratio équimolaire. Une différence significative est trouvée entre les états de dissémination (Student, p-value = 3.97E-04).

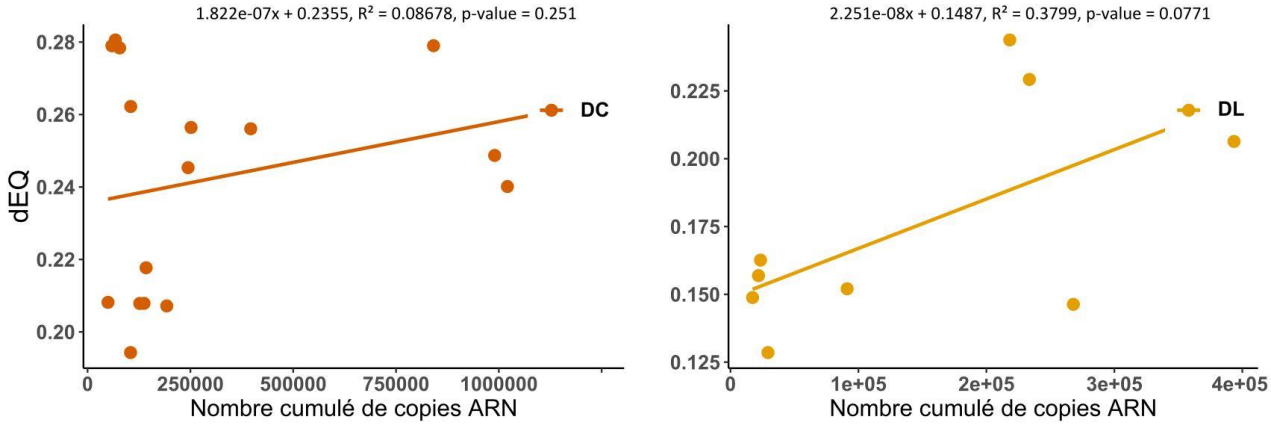


Figure 5. Régression linéaire de la distance à l'équimolarité par la charge virale (nombre cumulé de copies ARN) pour chaque condition de dissémination (DC : dissémination complète, DL : dissémination limitée).

Variation de la fréquence des segments au cours de l'infection

Nous avons quantifié les fréquences des segments de la population virale FCO-4_K1V3 utilisée pour infecter les culicoïdes (Figure 3). La fréquence moyenne de chaque segment dans l'inocula a été utilisé pour calculer une distance à l'inocula (dIN) pour les deux conditions de dissémination (Figure 6). Pour les deux états de

dissémination, la différence par rapport à l'inocula est hautement significative (Chi2, p-value < 2.0E-16). Les fréquences de chaque segment dans chaque condition ont ensuite été comparées à la fréquence de l'inocula par un test de Welch. Les segments 1, 2, 3 et 4 ont vu leur fréquence diminuer significativement par rapport à celle de l'inocula. A l'inverse, les segments 5, 6, 9 et 10 ont vu leurs fréquences augmenter pour les deux conditions. A l'exception du segment 2, ces variations suivent les résultats obtenus pour la déviation à l'équimolarité. Deux différences sont observées entre les états de dissémination. Elles concernent le segment 7 qui diminue significativement par rapport à l'inocula pour la dissémination limitée (p-value = 0.021), et le segment 8 qui augmente par rapport à la valeur de fréquence de l'inocula seulement pour la condition de dissémination complète (p-value = 2.46E-06).

De la même manière que la distance à l'équimolarité, la distance à l'inocula est significativement supérieure pour les échantillons appartenant à la condition dissémination complète (Welch, p-value = 1.56E-03). Les fréquences des segments génomiques ont divergé de celles de l'inocula de manière plus importante lorsque le virus a été capable de se propager jusqu'à la tête. Ce résultat souligne l'importance que pourrait revêtir le mécanisme de variation du nombre de copies de gènes dans un phénomène d'ordre adaptatif.

Comme précédemment pour la distance à l'équimolarité, nous avons voulu contrôler si la charge virale pouvait influencer la distance à l'inocula. Une régression linéaire de formule lm = (dIN ~ charge virale) n'a pas détecté d'effet significatif de la corrélation entre les deux variables (Figure 7).

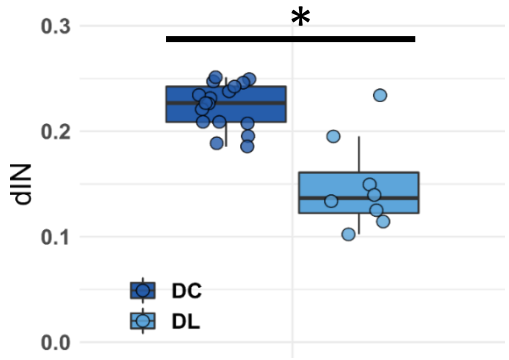


Figure 6. Indice de la distance à l'inocula en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée). Une différence significative est trouvée entre les deux conditions de dissémination (Welch, p-value = 1.56E-03).

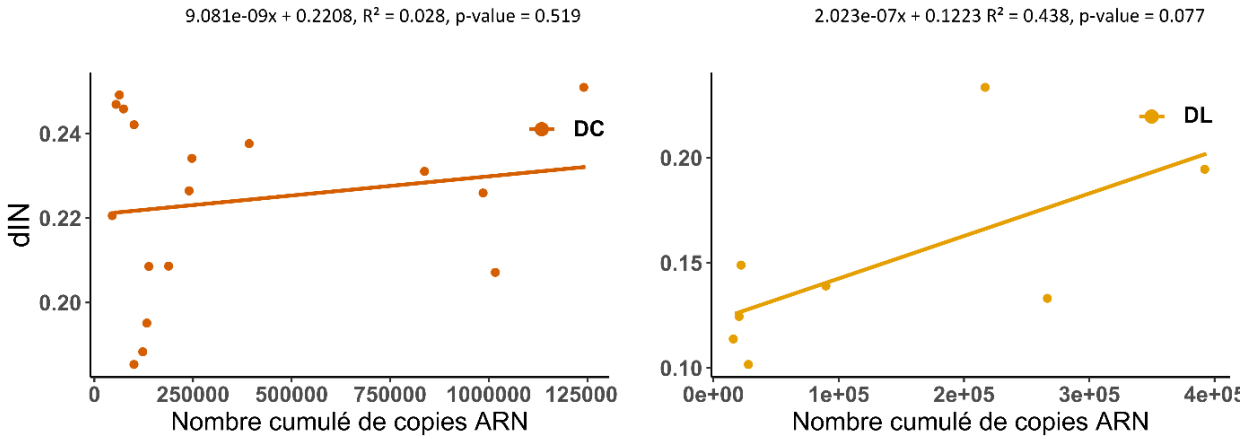


Figure 7. Régression linéaire de la distance à l'inocula par la charge virale (nombre cumulé de copies ARN) pour chaque condition de dissémination (DC : dissémination complète, DL : dissémination limitée).

La variation des fréquences entre les états de dissémination

Les fréquences relatives des conditions de dissémination ont été comparées (ANOVA, fréquence ~ segment x dissémination). Les facteurs « segment » (p-value < 2.0E-16) et l'interaction « segment : dissémination » (p-value = 6.43E-09) avaient tous les deux un effet significatif sur les variations de fréquence. Plusieurs différences de fréquence ont été trouvées entre les deux conditions de dissémination (test de Tukey-hsd avec l'état de dissémination en tant que variable explicative, pour chaque segment). Les segments 1, 2, 3, 6, 7 et 8 ont des fréquences significativement différentes entre nos conditions de propagations (Figure 3).

Comparaison de la formule génomique avec celle de populations naturelles

Nous avons comparé les résultats de cette expérience d'infection expérimentale aux résultats obtenus sur les populations naturelles du virus de la FCO-4 (Chapitre 1). Nous avons calculé un 3ème indice de distance, correspondant ici à la distance à la formule génomique chez *C. imicola* obtenue à partir des échantillons du terrain (dGFa). Notre approche considère toujours séparément les deux conditions de dissémination. La dGFa est significativement différente entre les deux états de dissémination (Figure 8, p-value = 1.16E-03).

Pour visualiser l'ordre des différences de fréquences existant aussi bien entre les conditions expérimentales qu'entre les états de dissémination, nous avons calculé la formule génomique (FG) pour chaque condition (*i.e*. deux niveaux de dissémination et les populations naturelles) (Figure 9). Les FGs des populations de terrain et à dissémination limitée sont extrêmement proches, l'ordre de variation maximale est d'environ 3 entre le segment 1, le plus faiblement représenté, et le segment 10, le plus fortement représenté. La différence par rapport aux populations à dissémination complète est frappante. Pour ces dernières nous observons un ordre de variation de dix entre le segment 1 et le segment 10.

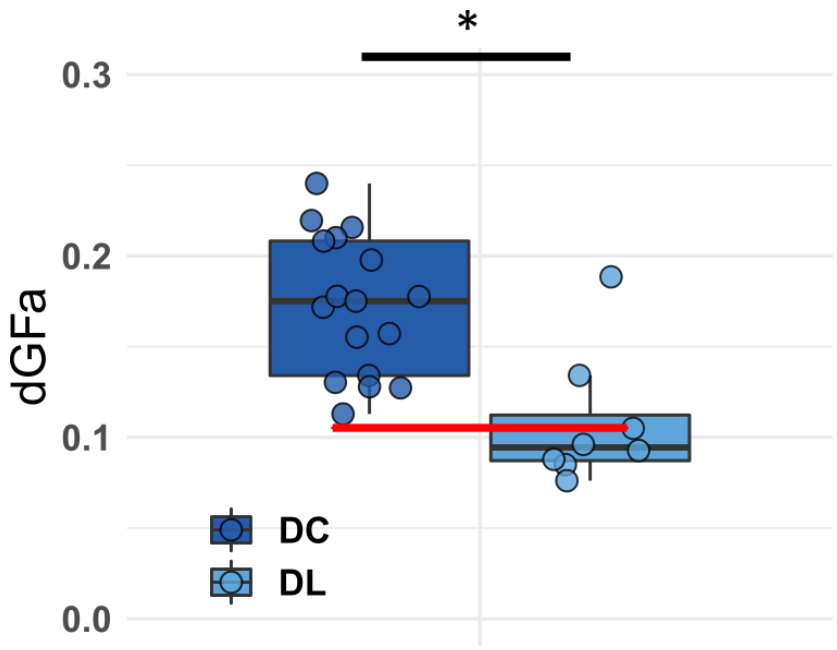


Figure 9. Distance à la formule génomique des populations virales de terrain en fonction de la dissémination du virus (DC : dissémination complète, DL : dissémination limitée). Une différence significative est trouvée entre les deux conditions de dissémination (Welch, p-value = 1.16E-03). La distance moyenne à la formule génomique de l'hôte culicoïdes est indiqué par la barre rouge pour l'inocula.

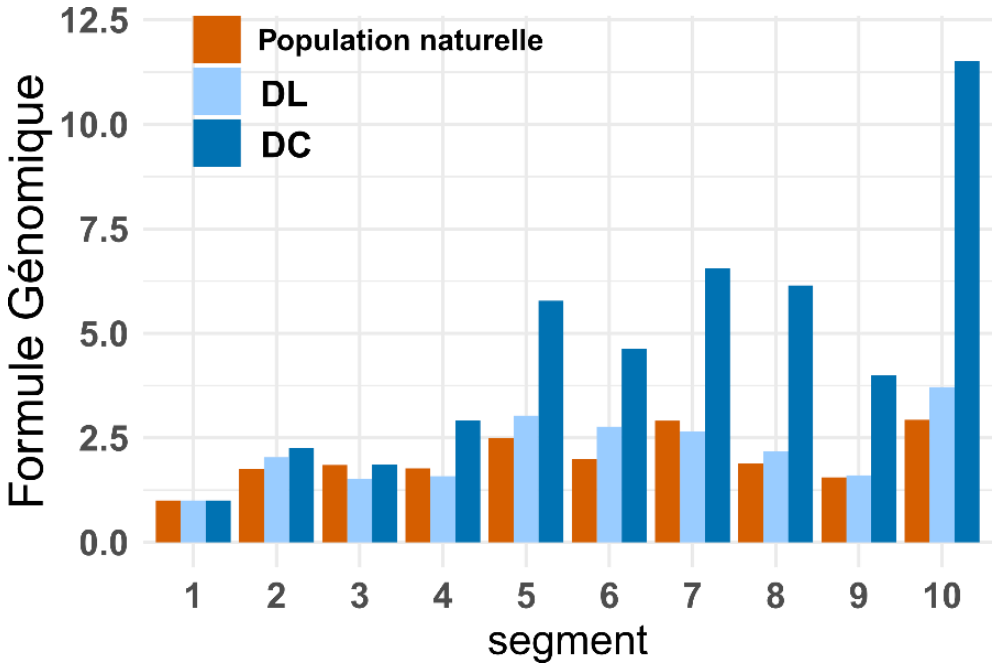


Figure 9. Comparaison des formules génomiques du virus de la FCO-4 chez son hôte *C. imicola* pour les deux états de dissémination de l'infection expérimentale (DC : dissémination complète, DL : dissémination limitée) et pour les populations virales naturelles.

DISCUSSION

La variation du nombre de copies de gènes chez le virus de la FCO-4 semble induire la convergence des segments vers des FGs dépendantes de l'hôte infecté (arthropode ou mammifère) (5). La mesure de fréquences des segments génomiques de populations sauvages de virus a permis de révéler la possibilité pour le virus de la FCO-4 à dévier d'une situation d'équilibre où chaque segment serait présent en quantité égale. Ces résultats laissent en suspens de nombreuses questions liées à la nature et à l'occurrence réelle des formules génomiques décrites. La formule génomique à l'équilibre établie pour deux espèces de virus multipartites correspond à l'état des fréquences vers lequel les segments génomiques de populations virales convergent au cours de l'infection (1, 4). L'état d'équilibre de la FG semble constituer, au moins pour le FBNSV, un avantage sélectif pour la population virale qui se traduit par une accumulation plus importante des virions (1). Au travers de notre approche nous démontrons que la formule génomique du virus de la FCO-4 peut changer au cours de l'infection d'un hôte, entre les populations virales dans l'inocula et dans les culicoïdes ayant ingéré cet inocula. Le changement dynamique de la fréquence des segments pourrait permettre la convergence vers une FG et soutiendrait l'hypothèse du caractère adaptatif de la variation du nombre de copie de gènes chez le virus de la FCO-4. Dans un second temps, nous avons testé si un lien pouvait être établi entre la fréquence des segments et la valeur sélective de la population virale. Si l'accumulation du virus ne diffère pas significativement en fonction de sa dissémination jusqu'à la tête, il existe, en revanche, des différences de FG en fonction de l'état de dissémination.

Caractère dynamique de la formule génomique

La formule génomique du virus de la FCO-4 dans l'inocula a changé lors de l'infection de son vecteur, *C. imicola*. Les fréquences des segments divergent des fréquences de l'inocula pour la majorité des segments. Les différences observées entre les formules génomiques de la souche BTV4-K1V3 initiales et 10 jours post-infection pourraient être liées à des effets non-exclusifs de goulot d'étranglement ou de sélection au cours de l'infection des cellules intestinales ou d'organes secondaire comme les corps gras.

Ces résultats ne permettent pas d'identifier des possibles différences dans les profils de fréquences au cours de l'infection car notre estimation est faite à 10 jours post-inoculation. L'infection des cellules insecte implique une voie de sortie non-lytique (18), l'infection s'établie dans les cellules et peut théoriquement durer pendant la vie de l'insecte (8). Le cycle d'infection de l'hôte culicoïde actuellement établie prévoit une dissémination du virus jusqu'aux glandes salivaires au minimum 5 jours après l'ingestion du virus de la FCO par voie orale. Ainsi les populations virales ont pu se répliquer et produire plusieurs générations de virions au cours de leur dissémination dans le compartiment corporel. Nos données ne permettent donc pas de déterminer la vitesse à laquelle la FG peut varier, cependant la convergence de nos résultats vers deux profils de FG différents entre nos conditions pourrait montrer l'effet d'une sélection à l'échelle de plusieurs générations. La vitesse de la génération de la FG sera abordée au cours du troisième chapitre de ce manuscrit.

Notre analyse ne permet pas d'explorer des différences dans les profils de fréquences entre compartiments au sein de l'insecte. Chez les moustiques infectés par le virus de la fièvre du Nil Occidental, le groupe de Grubaugh (31) a montré que les populations virales pouvaient se différencier génétiquement entre compartiments. Les critères de sélection (Ct < 31 pour les dix segments) n'ont malheureusement pas permis d'exploiter les virus présents dans la tête des culicoïdes. A l'exception des glandes salivaires, les cellules nerveuses pourraient constituer des cibles secondaires dans lesquels le virus pourrait se répliquer dans la tête (8). L'étude directe des glandes salivaires, chronophage à mettre en œuvre pour l'ensemble des individus infectés, n'aurait probablement pas prévenu le rejet des échantillons de têtes pour des critères qualitatifs (Ct < 31) à cause du nombre limité de particules virales détectable dans cet organe. Donc, il est possible que les formules génomiques soient différentes entre compartiments potentiellement sous l'effet de goulots d'étranglement mais aussi de pressions de sélections qui pourraient varier entre les compartiments.

Lien entre formule génomique et valeur sélective

Le premier facteur étudié pour tester le lien entre FG et adaptation est la capacité du virus à se disséminer et coloniser complètement l'insecte. La méthode employée pour la dissection n'élimine pas la probabilité qu'une partie des glandes salivaires se retrouve avec les compartiments thorax et abdomen lors de notre étude. La figure 1B présente la position des glandes salivaires dans les culicoïdes, elles se situent à cheval entre la tête et le thorax. Cependant les glandes salivaires n'ont jamais été décrites comme constituant une barrière à la propagation du virus (32) et la présence du virus dans la tête est souvent considérée comme une information soutenant que le virus peut être transmis (27).

Nous voulions aussi comparer la valeur sélective des populations virales disséminées ou non en fonction de leur taille de population ou charge virale. La charge virale constitue un indicateur commun de l'aptitude d'une population virale à se multiplier dans un environnement donné. Elle est généralement normalisée par l'intermédiaire d'un « gène de ménage » (i.e. gène hôte exprimé à valeur constante entre échantillons qui permet une normalisation pour tenir compte de l'efficacité de l'extraction). Nous n'avons malheureusement pas été en mesure de développer une approche permettant la quantification d'un gène de ménage chez *C. imicola*. Nous avons donc considéré le corps sans tête comme notre unité de mesure. L'absence de différence dans la charge virale entre conditions de dissémination pourrait impliquer une capacité équivalente des populations virale à se répliquer dans le corps indépendamment de leur capacité à se disséminer. Ce résultat suggère que d'autres facteurs sont susceptibles d'influencer la capacité d'une population virale à se propager jusqu'à la tête.

Variation de la formule génomique en fonction de la dissémination

Les formules génomiques sont différentes entre les états de dissémination. De plus, les profils de fréquence des populations à dissémination complète divergeaient plus de l'équimolarité, de l'inocula ou des populations naturelles. Nous n'avons pas actuellement une explication claire pour ce résultat.

La distance à la FG des populations naturelles (dGFa) plus importante quand la dissémination est complète est surprenante. Une explication possible est que les pools

de terrain pourraient être majoritairement constitués d'individus pour lesquels la dissémination n'a pas été complète. La dissémination complète du virus a pourtant lieu dans 70,4% dans les conditions d'infection expérimentale de laboratoire utilisées pour cette étude (24). Elle est peut-être un phénomène plus rare dans les conditions naturelles de circulation du virus. Plusieurs facteurs expérimentaux pourraient expliquer le succès du virus à se disséminer jusqu'à la tête. Le titre en virus de la solution utilisée pour le gorgement des culicoïdes (27) est élevé, alors qu'en condition d'infection *in vivo*, le titre décrit chez les ovins est très variable au cours de la virémie et n'atteint pas des concentrations si importantes au niveau des capillaires sanguins où viennent se gorger les culicoïdes (33). La température est aussi un facteur déterminant dans la capacité du virus à infecter le vecteur, le maintien des culicoïdes dans des enceintes climatiques à 27°C constitue des conditions favorables à la dissémination du virus de la FCO (34). Une autre approche d'infection expérimentale montre que des culiloïdes infectés transmettent le virus dans 80-100% des cas (33). Cette approche utilisait une inoculation intra-thoracique du virus concentré produit en culture cellulaire. Nous pensons que la concentration en virus de ce mode d'inoculation pourrait être associé à cette valeur extrême de taux de transmission du vecteur vers le ruminant. L'inoculation intra-thoracique revient dans ce cas à infecter indirectement les glandes salivaires en faisant abstraction des barrières et contrainte liées à la dissémination du virus dans cet hôte (8). Il serait intéressant d'étudier le succès de dissémination et le taux de transmission du virus au travers d'une approche utilisant un gorgement naturel afin de considérer l'ensemble des facteurs biologiques pouvant agir sur ces paramètres. Les variations de distance à la FG des échantillons de terrain pourraient aussi être liées aux différences dans le temps après inoculation. Nous ne disposons d'aucune information relative au temps d'infection des individus de terrain contrairement aux infections expérimentales. Finalement, l'inocula est aussi différent entre les deux populations. La souche FCO-4_K1V3 a été isolée et amplifiée en culture cellulaire, elle peut être génétiquement différente de la souche FCO-4 circulant entre hôte naturellement (35).

Nos résultats ont cependant permis de suivre les variations de fréquence des segments au niveau individuel chez le vecteur. Au total le nombre d'individus analysé devrait permettre d'estimer qu'il a une valeur représentative de la différence réelle qui existe entre les 17 échantillons à dissémination complète et les 8 échantillons à

dissémination limitée. L'ordre de variation entre segments est très proche pour le virus trouvé sur le terrain et les populations virales à dissémination incomplètes. Pour les deux conditions de dissémination le segment 1 est le moins représenté. Pour la dissémination incomplète, la variation d'un facteur 3 entre ce segment et les segments 7 et 10 suit nos observations précédentes pour l'hôte culicoïde (5). L'ordre de variation pour les populations qui ont atteint la tête varie de 4 à 6 pour les segments 5 à 9. Elle atteint un facteur 11 pour le segment 10. Ce dernier code pour les protéines non-structurelles qui permettent la sortie non lytique des cellules, elles sont essentielles à l'infection dans l'hôte culicoïdes (13, 21). Dans le cadre de l'hypothèse proposé par Sicard et al. (1), la variation du nombre de gènes pourrait participer à la régulation l'expression des gènes, la surreprésentation du segment 10 pourrait participer à la surexpression des protéines favorisant la libération des particules virales ou inhibant la production d'interféron de type I (19). La surreprésentation du segment 10 pourrait aussi favoriser l'apparition de mutation adaptative, comme une conséquence directe du nombre plus important de copies de ce segment (23). Plusieurs études portant sur la génétique du virus de la FCO indiquent que le segment 10 est l'un des segments avec la plus forte diversité génétique (35, 36). Ces mutations pourraient être aussi importantes dans le succès de dissémination du virus (37, 38). Les mécanismes de régulation de l'expression ou de génération de mutations adaptatives ne sont pas exclusifs.

Nos résultats suggèrent un rôle du segment 10 dans l'adaptation à l'hôte culicoïdes. Plus particulièrement, les protéines NS3 et NS3a dont l'expression est essentielle au succès de l'infection (13). Le nombre plus important de copies de ce segment pourrait participer à une production plus élevée de ces protéines et induire une libération importante des virions par exocytose pour poursuivre l'infection. Dans le contexte des infections persistantes qui se produisent dans les cellules d'insectes, ce système pourrait mener à une accumulation de ces protéines et permettre la sortie de groupes de virions par l'intermédiaire des lysosomes (18). Les populations virales qui infectent les cellules avec un nombre plus important de copies du segment 10, pourraient peut-être bénéficier d'une libération de la cellule plus rapide. Les résultats récents de Labadie et de ses collègues mettent en avant le rôle important d'un mode de transmission négligé entre cellules pour le virus de la FCO (39). Des vésicules extracellulaires qui portent plusieurs dizaines de particules virales sont produites au

cours de l'infection de systèmes *in vivo et in vitro*. De plus, un phénomène d'exclusion de surinfection semble limiter l'infection d'une cellule à une seule vésicule extracellulaire, tout en laissant possible l'entrée de particules virales non-enveloppées dans des vésicules (39). Par l'intermédiaire de ce mécanisme de libération, il semble possible de conserver l'information génétique complète des 10 segments au cours de l'infection d'un hôte malgré la VNCG. Cependant, la présence de telles structures au moment de la transmission entre hôtes n'est pas démontrée. Les FGs observées pour la condition de dissémination complète pourraient ne pas être identiques dans les glandes salivaires. Si une variation importante de NC était maintenue, la transmission de particules libres soumises à un effet de goulot d'étranglement semble peu probable. Il existe peut-être d'autres mécanismes favorisant la co-infection au cours de la transmission pour compenser ce coût important. Pour discuter son effet, il serait important de déterminer la nature de la VNCG chez le virus de la FCO et notamment la capacité d'une particule virale à porter un nombre différent de segments par l'intermédiaire de concaténation génomique (40), d'occupation de l'espace disponible pour des segments surnuméraires (41) ou de génération de génomes défectifs (35).

CONCLUSION

L'infection du vecteur *C. imicola*, par le virus de la FCO-4 conduit à des variations de la FG du virus après 10 jours d'infection. Ces variations semblent liées à l'état de dissémination du virus. Lorsque le virus colonise la totalité de l'insecte, les populations virales ont des variations de fréquences entre segments qui sont plus importantes. La FG montre un écart maximum d'ordre dix entre le segment 1, faiblement représenté, et le segment 10, le plus représenté. De telles variations semblent difficiles à imaginer à l'échelle de la particule virale individuelle, et impliquent probablement une modification du nombre de copies au niveau de la population. Ces résultats questionnent le potentiel de transmission de l'intégrité du génome lors des goulots d'étranglement inhérents au cycle d'un arbovirus et pourrait être compensé par des mécanismes de transmissions favorisant la co-infection ou une FG différente dans les glandes salivaires. Ces variations se traduisent par une distance aux FGs de l'inocula et à la formule génomique de terrain qui est supérieure pour l'état de

dissémination complet. Néanmoins, les charges virales des corps infectés ne changent pas en fonction de la dissémination. Ces variations pourraient donc être liées à la propagation plus rapide des populations virales pour lesquels le segment 10 est présent en quantité plus importante au sein de chaque cellule. Si la VNCG participait à la régulation de l'expression, le segment 10 qui code pour des protéines non-structurelles impliquées dans la sortie non-lytique des virions pourrait contribuer à une dissémination plus rapide de ces virions. Chez plusieurs espèces d'arbovirus, la composante génétique a été décrite comme un facteur clef du succès de l'infection (31, 37, 38, 42, 43) et la VNCG pourrait aussi accélérer la génération de mutation pour les segments les plus représentés. Les valeurs sélectives des FGs ne pouvaient pas être directement évaluées avec notre approche. Il semble qu'au terme de l'expérience, les populations virales ont atteint leur niveau d'accumulation maximale caractéristique de l'infection des culicoïdes (8). Nos résultats soutiennent le changement dynamique des FGs au cours de l'infection et un rôle de la VNCG dans le succès de l'infection du virus de la FCO-4.

REFERENCES

1. Sicard A, Yvon M, Timchenko T, Gronenborn B, Michalakis Y, Gutierrez S, Blanc S. 2013. Gene copy number is differentially regulated in a multipartite virus. Nature Communications 4:1-8.
2. Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. 2014. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. Proceedings of the National Academy of Sciences 111:16854-16859.
3. Hu Z, Zhang X, Liu W, Zhou Q, Zhang Q, Li G, Yao Q. 2016. Genome segments accumulate with different frequencies in Bombyx mori bidensovirus. J Basic Microbiol 56:1338-1343.
4. Wu B, Zwart MP, Sánchez-Navarro JA, Elena SF. 2017. Within-host Evolution of Segments Ratio for the Tripartite Genome of Alfalfa Mosaic Virus. Scientific Reports 7:1-15.
5. Moreau Y, Gil P, Exbrayat A, Rakotoarivony I, Breard E, Sailleau C, Viarouge C, Zientara S, Savini G, Goffredo M, Mancini G, Loire E, Gutierrez S. 2020. The genome segments of Bluetongue virus differ in copy number in a host-specific manner. J Virol doi:10.1128/JVI.01834-20.
6. Weaver SC, Reisen WK. 2010. Present and future arboviral threats. Antiviral Res 85:328-45.
7. Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, Moore CG, Carvalho RG, Coelho GE, Van Bortel W, Hendrickx G, Schaffner F, Elyazar IR, Teng HJ, Brady OJ, Messina JP, Pigott DM, Scott TW, Smith DL, Wint GR, Golding N, Hay SI. 2015. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife 4:e08347.
8. Mellor PS, Carpenter S, White DM. 2009. Chapter 14 - Bluetongue virus in the insect host, p 295-320. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50018-6. Academic Press, London.
9. Lee WS, Webster JA, Madzokere ET, Stephenson EB, Herrero LJ. 2019. Mosquito antiviral defense mechanisms: a delicate balance between innate immunity and persistent viral infection. Parasit Vectors 12:165.
10. Fu H, Leake CJ, Mertens PPC, Mellor PS. 1999. The barriers to bluetongue virus infection, dissemination and transmission in the vector, Culiloïdes variipennis (Diptera: Ceratopogonidae). Archives of Virology 144:747-761.
11. Hardy Jl Fau - Houk EJ, Houk Ej Fau - Kramer LD, Kramer Ld Fau - Reeves WC, Reeves WC. Intrinsic factors affecting vector competence of mosquitoes for arboviruses.
12. J. T. PAWESKA GJVaPSM. 2002. Vector competence of South African Culiloïdes species for bluetongue virus serotype 1 (BTV-1) with special reference to the effect of temperature on the rate of virus replication in C. imicola and C. bolitinos. Medical and Veterinary Entomology.
13. Feenstra F, Drolet BS, Boonstra J, van Rijn PA. 2015. Non-structural protein NS3/NS3a is required for propagation of bluetongue virus in Culiloïdes sonorensis. Parasit Vectors 8:476.
14. Ross-Smith N, Darpel KE, Monaghan P, Mertens PPC. 2009. Chapter 5 - Bluetongue virus: cell biology, p 77-99. *In* Mellor PS, Baylis M, Mertens PPC

(ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50009-5. Academic Press, London.
15. Mertens PP, Burroughs Jn Fau - Walton A, Walton A Fau - Wellby MP, Wellby Mp Fau - Fu H, Fu H Fau - O'Hara RS, O'Hara Rs Fau - Brookes SM, Brookes Sm Fau - Mellor PS, Mellor PS. Enhanced infectivity of modified bluetongue virus particles for two insect cell lines and for two Culiloïdes vector species.
16. P. Mertens JNB, A. Walton, M. P. Wellby, H. Fu, R. S. O'Hara, S. M. Brookes, and P. S. Mellor. 1996. Enhanced Infectivity of Modified Bluetongue Virus Particles for Two Insect Cell Lines and for Two Culiloïdes Vector Species. Virology.
17. Ftaich N, Ciancia C, Viarouge C, Barry G, Ratinier M, van Rijn PA, Breard E, Vitour D, Zientara S, Palmarini M, Terzian C, Arnaud F. 2015. Turnover Rate of NS3 Proteins Modulates Bluetongue Virus Replication Kinetics in a Host-Specific Manner. J Virol 89:10467-81.
18. Alex D.Hyatt YZaPR. 1993. Release of Bluetongue Virus-like Particles from Insect Cells is Mediated by BTV Nonstructural Protein NS3/NS3A. Virology.
19. Vitour D, Doceul V, Ruscanu S, Chauveau E, Schwartz-Cornil I, Zientara S. 2014. Induction and control of the type I interferon pathway by Bluetongue virus. Virus Res 182:59-70.
20. Janowicz A, Caporale M, Shaw A, Gulletta S, Di Gialleonardo L, Ratinier M, Palmarini M. 2015. Multiple genome segments determine virulence of bluetongue virus serotype 8. J Virol 89:5238-49.
21. van Gennip RG, van de Water SG, van Rijn PA. 2014. Bluetongue virus nonstructural protein NS3/NS3a is not essential for virus replication. PLoS One 9:e85788.
22. Carpi G, Holmes EC, Kitchen A. 2010. The evolutionary dynamics of bluetongue virus. J Mol Evol 70:583-92.
23. Elde NC, Child SJ, Eickbush MT, Kitzman JO, Rogers KS, Shendure J, Geballe AP, Malik HS. 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. Cell 150:831-41.
24. Federici V, Goffredo M, Mancini G, Quaglia M, Santilli A, Di Nicola F, De Ascentis M, Cabras P, Volpicelli C, De Liberato C, Satta G, Federico G, Leone A, Pisciella M, Portanti O, Pizzurro F, Teodori L, Savini G. 2019. Vector Competence of Italian Populations of Culiloïdes for Some Bluetongue Virus Strains Responsible for Recent Northern African and European Outbreaks. Viruses 11.
25. Balenghien T, Pages N, Goffredo M, Carpenter S, Augot D, Jacquier E, Talavera S, Monaco F, Depaquit J, Grillet C, Pujols J, Satta G, Kasbari M, Setier-Rio ML, Izzo F, Alkan C, Delecolle JC, Quaglia M, Charrel R, Polci A, Breard E, Federici V, Cetre-Sossah C, Garros C. 2014. The emergence of Schmallenberg virus across Culiloïdes communities and ecosystems in Europe. Prev Vet Med 116:360-9.
26. Hofmann MA, Renzullo S, Mader M, Chaignat V, Worwa G, Thuer B. 2008. Genetic characterization of toggenburg orbivirus, a new bluetongue virus, from goats, Switzerland. Emerg Infect Dis 14:1855-61.
27. Paslaru AI, Mathis A, Torgerson P, Veronesi E. 2018. Vector competence of pre-alpine Culiloïdes (Diptera: Ceratopogonidae) for bluetongue virus serotypes 1, 4 and 8. Parasit Vectors 11:466.
28. Sailleau C, Breard E, Viarouge C, Gorlier A, Quenault H, Hirchaud E, Touzain F, Blanchard Y, Vitour D, Zientara S. 2018. Complete genome sequence of

bluetongue virus serotype 4 that emerged on the French island of Corsica in December 2016. Transboundary and Emerging Diseases 65:e194-e197.
29. Feng L, Lintula S, Ho TH, Anastasina M, Paju A, Haglund C, Stenman UH, Hotakainen K, Orpana A, Kainov D, Stenman J. 2012. Technique for strand-specific gene-expression analysis and monitoring of primer-independent cDNA synthesis in reverse transcription. BioTechniques 52:263-270.
30. Rosenblad A. 2017. Multivariate Statistical Methods: A Primer (4th Edition). Journal of Statistical Software 78.
31. Grubaugh ND, Weger-Lucarelli J, Murrieta RA, Fauver JR, Garcia-Luna SM, Prasad AN, Black WCt, Ebel GD. 2016. Genetic Drift during Systemic Arbovirus Infection of Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching. Cell Host Microbe 19:481-92.
32. H. Fu CJL, P. P. C. Mertens, and P. S. Mellor. 1998. The barriers to bluetongue virus infection, dissemination and transmission in the vector, Culiloïdes variipennis (Diptera: Ceratopogonidae). Archive of Virology.
33. Baylis M, O'Connell L, Mellor PS. 2009. Chapter 15 - Rates of bluetongue virus transmission between Culiloïdes sonorensis and sheep, p 321-341. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50019-8. Academic Press, London.
34. Carpenter S, Wilson A, Barber J, Veronesi E, Mellor P, Venter G, Gubbins S. 2011. Temperature dependence of the extrinsic incubation period of orbiviruses in Culiloïdes biting midges. PLoS One 6:e27987.
35. Kopanke JH, Lee JS, Stenglein MD, Mayo CE. 2020. The Genetic Diversification of a Single Bluetongue Virus Strain Using an In Vitro Model of Alternating-Host Transmission. Viruses 12.
36. Carpi G, Holmes EC, Kitchen A. 2010. The evolutionary dynamics of bluetongue virus. Journal of Molecular Evolution 70:583-592.
37. Patterson EI, Khanipov K, Rojas MM, Kautz TF, Rockx-Brouwer D, Golovko G, Albayrak L, Fofanov Y, Forrester NL. 2018. Mosquito bottlenecks alter viral mutant swarm in a tissue and time-dependent manner with contraction and expansion of variant positions and diversity. Virus Evol 4:vey001.
38. Riegler L. 2002. Variation in African horse sickness virus and its effect on the vector competence of Culiloïdes biting midgesUniversity of Surrey.
39. Labadie T, Roy P. 2020. doi:10.1101/2020.06.11.146357.
40. Anthony SJ, Darpel KE, Belaganahalli MN, Maan N, Nomikou K, Sutton G, Attoui H, Maan S, Mertens PP. 2011. RNA segment 9 exists as a duplex concatemer in an Australian strain of epizootic haemorrhagic disease virus (EHDV): Genetic analysis and evidence for the presence of concatemers as a normal feature of orbivirus replication. Virology 420:164-71.
41. Desselberger U. 2020. What are the limits of the packaging capacity for genomic RNA in the cores of rotaviruses and of other members of the Reoviridae? Virus Res 276:197822.
42. Aaskov J, Buzacott K, Thu HM, Lowry K, Holmes EC. 2006. Long-Term Transmission of Defective RNA Viruses in Humans and Aedes Mosquitoes. Science 311:236.
43. Grubaugh ND, Smith DR, Brackney DE, Bosco-Lauth AM, Fauver JR, Campbell CL, Felix TA, Romo H, Duggal NK, Dietrich EA, Eike T, Beane JE, Bowen RA, Black WC, Brault AC, Ebel GD. 2015. Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness. PLoS Pathog 11:e1004874.

ANNEXES

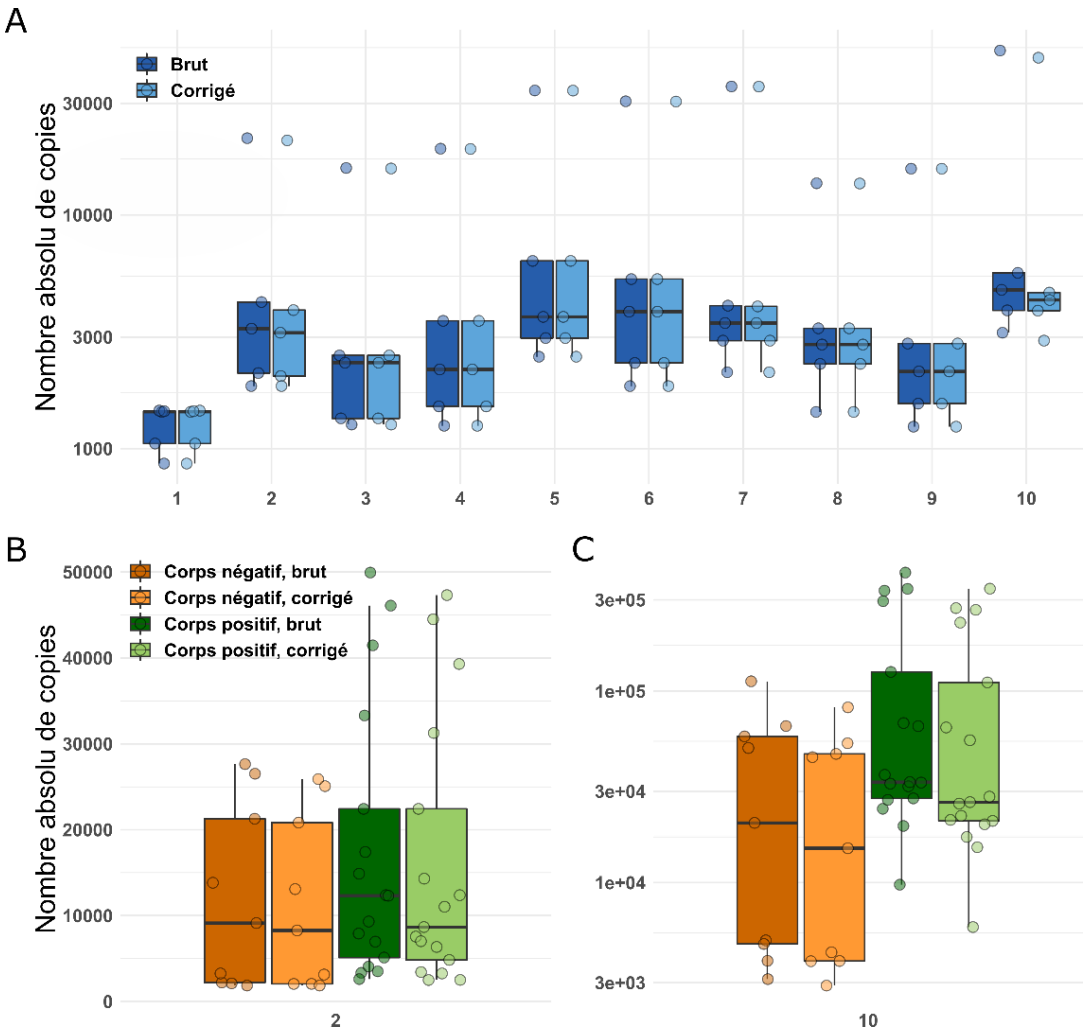


Figure S1. Effets des amplifications indépendantes de l'action des primers à l'étape de transcription inverse. A) Nombre de copies brutes ou corrigées pour chacun des dix segments. Un effet significatif est observé seulement pour les segments 2 et 10. **B)** Effet de l'amplification indépendante de l'action des primers sur l'ensemble des échantillons pour le segment 2. **C)** Effet de l'amplification indépendante de l'action des primers sur l'ensemble des échantillons pour le segment 10. L'axe des ordonnées est transformé en échelle logarithme10.

Tableau S1. A) Comparaison des données corrigées ou brutes pour l'effet d'amplification aléatoire pour les segments 2 et 10 sur le jeu de données incluant tous les échantillons. Les tests de normalité (Shapiro) et d'homocédasticité (F-test) sont validé. Les résultats des tests de Wilkoxon et de Student ne montrent pas de différence entre les jeux de données corrigés et brute pour les segments 2 et 10 (DC : dissémination complète, DL : dissémination limitée). **B)** Effet de l'amplification aléatoire sur le nombre de copies et les fréquences normalisées pour les données filtrées brutes (Ct < 31) et non corrigées.

A

Segment	statut	Wilcoxon	Student
2	**DC**	0.863	0.519
	DL	0.863	0.915
10	**DC**	0.0939	0.0847
	DL	0.0939	0.123

B

Segment	Filtre	Rapport (brute/corrigé)
2	Ct < 31	1.0889
10		1.2517

CHAPITRE 3 : Génération de la formule génomique du virus de la FCO-4

La variation du nombre de copies de gènes est un phénomène décrit chez plusieurs espèces de virus et associé à l'adaptation de nouvelles pressions de sélection. Chez le virus multipartite FBNSV, les formules génomiques changent en fonction de l'hôte plante infecté. Dans notre premier chapitre nous montrons un phénomène similaire pour le virus de la fièvre catarrhale ovine de sérotype 4 (FCO-4) où les formules génomiques sont différentes entre les hôtes ruminants et l'espèce vectrice. Au cours du deuxième chapitre nous avons établi une corrélation entre ce phénomène et la valeur sélective de la population virale lors de l'infection du vecteur *C. imicola*.

Des virus, avec différentes architectures génomiques, multipartites et segmentées, peuvent être associés au phénomène de VNCG. Chez le FBNSV, Sicard et al. ont proposé que le phénomène pourrait être lié à un échantillonnage aléatoire des segments suite à un goulot d'étranglement puis à un effet de la sélection naturelle qui favoriserait les populations avec des fréquences des segments maximisant la valeur sélective.

Au cours de ce chapitre nous questionnons la génération du phénomène de VNCG chez le virus de la FCO-4. Nous avons mis en place une expérience d'infection expérimentale en culture cellulaire dans le but de comparer deux scénarios qui pourraient conduire à la génération de la formule génomique : une génération aléatoire comme pour le virus multipartite ou une génération déterministe. Nous comparons la première génération de virions formée après 12 heures d'infection dans deux conditions de multiplicité d'infection : une condition faible où les cellules seront infectées par un seul virion et une condition de forte co-infection où la complémentation entre génomes viraux est favorisée. Les différences de formules génomiques décrites dans le Chapitre 1 nous ont mené à comparer ces conditions d'infection entre des types cellulaires correspondants aux hôtes naturels : KC pour les culicoïdes et MDBK pour les bovins.

Nos résultats soutiennent qu'un profil de fréquences spécifiques à la condition expérimentale peut être généré lors de la première génération de particules virales. Pour les conditions où la réplication est plus importante les profils de formule génomique tendent vers les FG décrites dans les hôtes naturels. Ces résultats soulignent une seconde fois la corrélation entre la FG et la valeur sélective pour les populations du virus de la FCO-4. Les coefficients de variations des dix segments sont

similaires entre les conditions d'infections suggérant qu'un effet déterministe est à l'origine de la génération de la FG pour le virus de la FCO avec une influence potentielle de facteurs d'origine virale ou cellulaire.

Ces résultats soulignent le rôle important que pourraient jouer la MOI sur la génération de la FG et sur la capacité du virus de la FCO-4 à se multiplier. Ils questionnent l'influence des particules virales défectives ou semi-infectieuses sur l'expression de la FG.

Ce chapitre présente des résultats préliminaires, il sera valorisé suite à la thèse.

Génération de la formule génomique du virus de la FCO-4

Yannis Moreau[1,2*], Etienne Loire[1,2], Patricia Gil[1,2], Serafin Gutierrez[1,2*]

[1] ASTRE, Univ Montpellier, Cirad, INRA, Montpellier, France.

[2] Cirad, UMR ASTRE, F-34398 Montpellier, France.

* Corresponding author: Yannis Moreau, UMR 117 ASTRE, Campus international de Baillarguet, 34398 Montpellier cedex 5, yannis.moreau@cirad.fr; or Serafin Gutierrez, UMR 117 ASTRE, Campus international de Baillarguet, 34398 Montpellier cedex 5, serafin.gutierrez@cirad.fr

RESUME

Le virus de la FCO-4 possède le potentiel de générer des profils non équilibrés de fréquence de ses segments génomiques. Nous avons établi que la fréquence des segments génomiques (formule génomique ou FG) était dépendante de l'hôte infecté, avec une différence marquée entre l'hôte arthropode et deux hôtes ruminants. De plus, la formule génomique semble liée à une dissémination plus rapide du virus au cours de l'infection du vecteur. Cependant, notre connaissance des mécanismes liés à la génération de la formule génomique reste très limitée. Nous questionnons ici l'origine du phénomène de la variation du nombre de copies de gène chez le virus de la FCO-4. Au travers d'une approche d'infection expérimentale sur un temps court, nous suivons l'influence de la multiplicité d'infection (MOI) et du type cellulaire bovin (MDBK) ou culicoïdes (KC) sur l'évolution de la FG pour la première génération de particule virale. Nos résultats montrent le potentiel du virus de la FCO-4 à générer des profils de fréquences spécifiques à chaque type cellulaire après un seul cycle de réplication. Ces profils sont observés même lorsqu'une faible MOI à l'inoculation limite drastiquement les chances de complémentation. Ces résultats soutiennent un effet déterministe participant à la génération de la FG.

Mots-clefs : formule génomique - Multiplicité d'infection - FCO-4 - culture cellulaire

INTRODUCTION

Le concept de formule génomique (FG) a initialement été proposé pour décrire la convergence des segments génomiques vers une fréquence dépendante de l'hôte chez le *Faba bean necrotic virus* (FBNSV) (1). Le FBNSV est un virus multipartite, chacun de ses huit segments génomiques est encapsidé individuellement. Les travaux de Sicard *et al.* ont démontré le rôle du mécanisme de variation du nombre de copies de gènes (VNCG) dans l'adaptation de ce virus à différentes plantes hôtes (1). Au cours de l'infection, toutes les cellules ne portent pas un exemplaire de chacun des segments génomiques suggérant l'existence d'un système de complémentation entre virus infectants différentes cellules : le « cycle de vie multicellulaire » (2). Plus récemment comme cela avait été originellement postulé dans leurs travaux en 2013 (1), la VNCG a été associée à une régulation de l'expression des gènes, un résultat que soutient sa fonction adaptative (R. Gallet *et al.* en préparation).

Deux hypothèses ont été proposées pour expliquer la génération de la formule génomique à l'équilibre (FGE). La FGE est la formule à laquelle les populations virales tendent au cours de l'infection dans un hôte donné (1). La FGE est supposée être la formule qui permet une expression optimale dans un hôte donné. La première hypothèse implique une génération aléatoire des formules génomiques au sein d'un hôte, suivie par une sélection des populations avec des formules proches de la formule optimale (i.e. la FGE) (1). Plus précisément, des goulots d'étranglement donneraient lieu à des populations virales avec des formules différentes lors de la colonisation de l'hôte. Ensuite, une sélection positive ferait monter en fréquence les populations avec la FG permettant un phénotype optimal. La deuxième hypothèse implique des taux de réplication différents entre segments (1). Ces taux mèneraient vers la FGE. Cependant, ce scénario déterministe pourrait impliquer la perte de segments à faible fréquence lors de goulots d'étranglement au cours de la colonisation de l'hôte, un phénomène qui n'a pas été observé chez FBNSV (1).

Nos travaux précédents explorent le potentiel des populations d'un virus à génome segmenté, le virus de la fièvre catarrhale ovine (FCO), à varier en nombre de copies de segments. Le cycle du virus de la FCO implique généralement une alternance obligatoire des cycles de réplication dans des ruminants et des insectes. Nous avons

pu montrer l'existence de profils de fréquences des segments associables au concept de formule génomique observé chez FBNSV et d'autres virus multipartites (Chapitre 1, (3)). Comme chez FBNSV, nos résultats montrent une spécificité du profil de fréquences en fonction de l'espèce hôte, avec une différence entre les populations issues des ruminants et des insectes. De plus, nos résultats suggèrent un lien potentiel entre la FG et la valeur sélective de la population virale (Chapitre 2).

Nos résultats nous ont poussé à étudier la génération des profils de fréquences chez le virus de la FCO. Les goulots d'étranglement pourraient limiter la diversité génétique du virus de la FCO lors de la transmission. La variation du nombre de copies de gènes (VNCG) pourrait induire un coût si aucun mécanisme n'existe pour compenser un potentiel déséquilibre de nombre de copies de segments au moment de la transmission. Dans le cadre où le virus de la FCO-4 dispose des éléments génomiques nécessaires à l'établissement de l'infection, le goulot d'étranglement pourrait avoir isolé des virions avec une FG particulière. La génération de la FG pourrait être aléatoire puis converger vers un état d'équilibre comme cela avait été proposé pour les virus multipartites (1). Mais elle pourrait aussi être déterminée par des facteurs viraux, comme la présence de mutations qui modifieraient l'efficacité d'encapsidation de certains segments (4). Ou de la même manière modulée par des facteurs liés à l'hôte et aux systèmes de défenses antiviraux qui peuvent varier en fonction des types cellulaires (5, 6).

Nous avons mis en place une approche expérimentale dans le but de tester les deux scénarios proposés pour expliquer la génération de la FG chez FBNSV dans notre modèle. Notre étude se place volontairement à un stade préliminaire de l'infection des cultures cellulaires. Après 12 heures d'infection, seule la population dans l'inocula a eu le temps de se répliquer et on peut limiter l'influence des pressions de sélection sur la descendance et sa formule génomique. De plus, l'approche expérimentale nous permet de définir le type cellulaire ainsi que le goulot d'étranglement lors de l'infection primaire. Dans le cadre d'un fort goulot d'étranglement, chaque cellule pourrait n'être infectée que par un seul virion. Dans ce cas, la présence des segments génomiques essentiels devrait être un prérequis à la réplication du virus. Les expériences d'assemblage montrent que les protéines du complexe de transcription (segments 1, 4 et 6) ainsi que ceux codant pour la particule core (segments 3 et 7) sont essentiels pour initier l'assemblage sans cellule (7). Aussi, l'efficacité d'encapsidation est réduite

si certains segments sont manquants, les segments de plus petite taille (segments 10, 9, 8) ont un effet plus marqué sur cette diminution de l'efficacité (7).

Dans le cadre de l'hypothèse d'une génération aléatoire de la FG, la faible multiplicité d'infection (MOI, i.e. nombre moyen de particule virale infectieuse par cellule) devrait favoriser la réplication des particules virales possédant l'ensemble des segments génomiques essentiels. Un faible MOI à l'inoculation devrait donc générer une descendance avec des profils de fréquences proche de l'équimolarité. Sous cette hypothèse, une forte MOI à l'inoculation devrait conserver le profil de l'inocula. Si l'on considère que la VNCG est générée par des facteurs déterministes liés à la génétique du virus ou de l'hôte, la génération de la FG devrait se produire rapidement, lors du premier cycle d'infection, et même à faible MOI.

MATERIELS ET METHODES

Cultures cellulaires

Les cellules utilisées pour notre expérience sont entretenues dans une salle dédiée dans les laboratoires du CIRAD (P2). Les cellules de culicoïdes KC (dérivées de cellules de l'espèce *Culicoides sonorensis*) sont maintenues dans un incubateur à 27°C dans des flasques T75 scellées, sans contrôle du taux d'hygrométrique. Pour l'entretien de cette lignée, nous utilisons le milieu *Schneider's Drosophilia medium* (SDM) supplémenté de 10% de sérum de veau fœtal (SVF). Les lignées cellulaires mammifères (MDBK, OA3 et BSR) sont cultivées en flasque T75 avec bouchon filtré et maintenues dans un incubateur à 37°C dans des conditions humides, avec 5% de CO2. Pour les cellules bovine MDBK (dérivées de cellules rénales de l'espèce *Bos taurus*) et OA3 (dérivées de cellules testiculaires de l'espèce *Ovis aries*) nous avons utilisé le milieu *Dubelcco's Modified Eagle Medium* (DMEM), supplémenté de 10% de SVF, 1% d'acides aminés non-essentiels, 1% de pyruvate de sodium et 1% de L-Glutamine (300mM). Les cellules BSR utilisées sont dérivées de cellules BHK-21 (*Baby-Hamster-Kidney*). Elles étaient entretenues avec le milieu de culture DMEM Glutamax, supplémenté de 10% de SVF et 1% de pyruvate de sodium.

Souche virale

La souche du virus de la FCO-4 utilisée a été isolée sur des cellules KC par l'unité Virologie (ANSES-INRA-EnvA, Maisons-Alfort) à partir du sang d'un mouton infecté sur l'île de Corse en 2017. Nous avons amplifié cet isola à deux reprises sur des cellules de mammifère (BSR) dans des conditions d'infection à forte MOI. La souche a été nommée FCO-4_K1B2 en conséquence de ces passages successifs. Cette souche a été titrée en utilisant la méthode de dilution limite de Spearman–Karber (8). Le titre en TCID50/mL a été transformé en pFU/mL en multipliant sa valeur par une constante de 0.69 (9) pour déterminer les concentrations finales utilisées dans nos deux conditions de MOI (*i.e.* 0.3 et 3).

Infection des cultures cellulaires

Notre approche se concentre sur le premier cycle de l'infection cellulaire. Après 12 heures, seule la première génération de particules virales issue des particules primo-infectante a été produite (7). Le temps d'expérience ne permet pas aux particules virales de première génération d'infecter de nouvelles cellules. Les populations virales étudiées sont principalement issues de cellules infectées par une seule particule virale dans la condition à faible MOI de 0.3. Dans la condition à forte MOI de 3, la population virale étudiée est issue de cellules infectées en moyenne par 3 particules virales, ce nombre peut varier en suivant une distribution gaussienne autour de cette moyenne. Nous avons adapté notre approche expérimentale afin d'infecter un nombre équivalent de cellules dans les deux conditions de MOI. Les infections à une MOI de 0.3 sont réalisées en plaque de 12 puits tandis que les infections à une MOI de 3 sont faites en plaque de 48 puits. La souche virale K1B2 a été titrée sur deux types cellulaires d'origine mammifère : MDBK et OA3 (Tableau 1). La sortie non lytique du virus de la FCO-4 ne permet pas la titration sur les cellules KC. Pour ce type cellulaire nous avons sélectionné le titre viral le plus élevé, obtenu sur les cellules OA3, afin de déterminer la concentration de virus à utiliser pour nos deux conditions de MOI. Ce choix maximise la probabilité de co-infection pour la MOI la plus forte. Avec une différence d'un log entre les conditions d'infection, le risque principal était de générer deux conditions d'infection avec un nombre de particules par cellule identique.

Tableau 1. **Conditions d'infection expérimentale pour les deux MOI et types cellulaires.**

Cellule	Titration sur	Titre (pFU/ml)	Plaque	Nombre initial de cellules	MOI
KC	OA3	1,44E+06	P12	8,00E+05	0,3
			P48	2,00E+05	3
MDBK	MDBK	5,63E+05	P12	3,20E+05	0,3
			P48	8,00E+04	3
OA3	OA3	1,44E+06	P12	3,20E+05	0,3
			P48	8,00E+04	3

La veille des infections, les cellules sont semées à la concentration appropriée afin d'obtenir un tapis à approximativement 70% de confluence (Tableau 1). Le tapis

cellulaire est lavé à 3 reprises avec du PBS 1X avant d'être infecté avec 500 µl de solution virale diluée dans un milieu de culture non supplémenté correspondant au type cellulaire. Les plaques sont agitées lentement pendant 1 minute puis placées dans un incubateur pendant 2 heures pour permettre l'adhésion et l'entrée du virus. Le surnageant est ensuite enlevé et le tapis rincé de nouveau 3 fois au PBS 1X. Un volume de 500 µl de milieu non supplémenté est ajouté et les cellules sont replacées dans l'incubateur pour 10 heures supplémentaires. Les tapis cellulaires étaient cultivés dans un volume de 500 µl de milieu SDM pour les cellules KC et de milieu DMEM pour les cellules MDBK. A terme, les plaques sont scellées et placées dans un congélateur à -80°C jusqu'à l'extraction ARN. Chaque condition d'infection a été répétée dans 6 réplicats reproduits sur deux journées différentes. Un contrôle négatif a été ajouté pour chaque condition et répétition.

Extraction des ARN viraux

Les solutions de cellules lysées par congélation ont été aliquotées individuellement et 150 µl de chaque puit ont été utilisés pour procéder à l'extraction. L'extraction est faite avec le kit 'MagMAX CORE Nucleic Acid Purification Kit' (Applied Biosystem). La lyse est faite en présence d'acrylamide linéaire comme carrier. Les étapes suivantes d'extraction sont faites en suivant les instructions du fournisseur. L'élution des ARN est faite avec de l'eau sans nucléase.

Quantification des segments génomiques par RT-qPCR

Nous avons utilisé une approche de RT-qPCR spécifique pour quantifier les segments génomiques du virus de la FCO en ciblant les brins négatifs de l'ARNdb (Chapitre 1). Cette méthode a fait l'objet de validation y compris à partir de virus produit en culture cellulaire (3). Brièvement, les ARNdb sont dénaturés 3 minutes à 95°C avec 20% de DMSO avant l'étape de RT. Elle est réalisée avec le kit RevertAid (FicherScientific), séparément pour chacune des 10 cibles génomiques. A cette étape des primers ciblant spécifiquement le brin négatif sont utilisés, ainsi seuls les brins génomiques présents au sein des particules virales sont quantifiés. Nous utilisons la méthode de fluorescence SYBRGREEN pour nos réactions de qPCR. La robustesse

de cette approche a été établie sur différents modèles viraux pour l'étude de la variation du nombre de copies de gènes (1, 10). Pour contrôler la spécificité nous avons validé l'absence d'amplification de nos primers sur les types cellulaires étudiés et chaque réaction a été vérifiée par analyse de la courbe de dissociation. Chaque réaction est réalisée en triplicata. Notre méthode de quantification est absolue, les nombres de copies sont calculés grâce à un plasmide pGEM-T qui porte un insert contenant chacune des 10 cibles de qPCR en un exemplaire (3). Ce plasmide est utilisé comme point de gamme à concentration constante pour chacune de nos réactions et donc il est inclus dans chaque plaque de qPCR. Cette approche nous permet également de valider la précision de notre méthode en vérifiant que chaque cible est bien détectée en quantité équimolaire pour ce contrôle.

Effet de l'amplification indépendante de l'action des primers à l'étape de transcription inverse

Les segments 2 et 10 ont été testés pour détecter un effet d'amplification indépendant de l'action des primers à l'étape de transcription inverse (RT) (11). Nous avons précédemment décrit son influence sur les résultats de RT-qPCR au cours de notre étude de la variation du nombre de copies virales dans les populations sauvages du virus de la FCO-4 (3). Pour contrôler son effet, une étape de RT est réalisée sans ajout de primer pour chaque échantillon. Les brins d'ADNc formés sont quantifiés avec l'approche de qPCR développée pour chaque segment. Les produits de PCR résultent donc d'une amplification indépendante de l'action des primers à l'étape de RT. Nous utilisons ensuite deux jeux de données pour qualifier l'influence de l'effet. Le jeu de données brutes présente les valeurs non modifiées du nombre de copies (NC). Le jeu de données corrigées est généré par soustraction de la valeur du nombre de copies finales obtenue sans primer à l'étape de RT à la valeur brute du nombre de copies pour chaque segment. Nous avons suivi l'influence de l'effet d'amplification, indépendamment de l'action des primers spécifiquement sur les segments 2 et 10, en comparant l'effet de la correction sur nos 47 échantillons.

Statistiques

Nos résultats ont été analysé avec le logiciel R (R Développement Core Team, 2011, 239 version 1.2.5033). Les comparaisons entre condition ont été faites selon les résultats des tests de normalité des distributions et d'homocédasticité. Le test de Wilcoxon est utilisé si la normalité des distributions est rejetée (Shapiro), le test de Welch si l'homocédasticité est rejetée (var. Test) et le test de Student si toutes les conditions précédentes sont validées. Les fréquences sont comparées séparément pour chaque segment par un test de Wilcoxon. Les indices de distance ont été calculés comme suit :

(1) distance à l'équimolarité : dEQ = $\sum_i^{10}|p\mathrm{i} - 0.1|/2$

(2) distance à la formule génomique de l'inocula : dIN = $\sum_i^{10}|pi - pINi|/2$

(3) distance à la formule génomique dans des culicoïdes sauvages : dGFa = $\sum_i^{10}|pi - pGFi|/2$

(4) distance à la formule génomique dans des vaches : dGFb = $\sum_i^{10}|pi - pGFi|/2$

suivant l'équation de Manly (12) où *i* est le segment, *p* la fréquence relative du segment dans l'échantillon, *0.1* la fréquence relative à l'équimolarité, *pIN* la fréquence de l'inocula, *pGF* la fréquence de la formule génomique. La formule génomique a été calculée séparément en fonction des conditions d'infection en divisant la valeur médiane de fréquence d'un segment donné par la valeur médiane de fréquence du segment la plus faiblement représenté.

RESULTATS

Effet d'amplification indépendant de l'action des primers à l'étape de RT

Les résultats de nos études précédentes ont montré un effet de l'amplification indépendant de l'action des primers à l'étape de RT (AIAPET) pour les segments 2 et 10 (3). L'ordre de variation lié à une amplification non spécifique est comparable à nos résultats précédents. Pour nos résultats de qPCR sur les virus infectant des cultures de cellule, son effet compte en moyenne pour 31,35% du nombre de copies totales quantifiées pour le segment 10. L'effet est plus faible pour le segment 2 avec 14,5% de variation du nombre de copies en moyenne. Le nombre de copies ne différait pas significativement entre les jeux de données corrigées et brutes pour les segments 2 et 10 (Wilcoxon, p-value > 0.05, Figure 1). Afin de surveiller un possible impact sur nos résultats de ce phénomène, l'ensemble des tests statistiques a été fait en parallèle pour les deux jeux de données. Les résultats présentés ci-dessous correspondent au jeu de données corrigées et toute différence de résultat entre les jeux est indiquée dans le test.

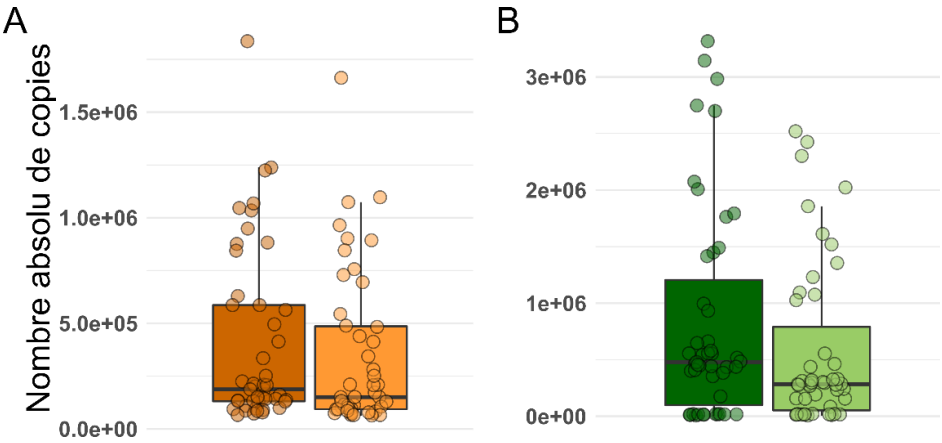


Figure 1. Comparaison du nombre absolu de copies entre les jeux de données brutes et corrigées. Pour chacune des figures le jeu de données brutes est présenté dans la couleur foncée et le jeu de données corrigées dans la couleur plus claire. A) Comparaison des traitements pour le segment 2, aucune différence significative de nombre de copies absolues n'est observée entre les jeux de données (Wilcoxon, p-value = 0.20). B) Comparaison des jeux de données pour le segment 10. Il n'y a pas de différence significative entre les deux traitements (Wilcoxon, p-value = 0.058).

Fréquences des segments après un cycle d'infection

Notre approche expérimentale *in vitro* devait initialement inclure un type cellulaire correspondant à chaque espèce d'hôte étudié sur le terrain. Nous avons donc réalisé les infections en parallèle sur trois types cellulaires (bovin, ovin et culicoïdes). Cependant, les résultats de qPCR pour le type cellulaire OA3 (cellules de mouton, espèce *Ovis aries*) se sont avérés inexploitables et nous sommes dans l'impossibilité de reproduire l'expérience par contrainte de temps. Les résultats sur le type OA3 ne sont donc pas présentés.

La figure 2 présente les fréquences des segments dans les différentes conditions expérimentales. Les fréquences de la majorité des segments déviaient significativement de l'équimolarité dans les deux conditions de MOI (Figure 2, Annexes Tableau 1). A faible MOI, une fréquence significativement inférieure à 0,1 a été observée pour les segments 1, 2, 3 et 4 (Welch unilatéral, p-values < 0.05) et pour les segments 1, 3 et 10 à forte MOI. Une fréquence significativement supérieure à 0,1 est observée pour les segments 5, 6, 7 et 10 à faible MOI (Welch unilatéral, p-values < 0.05) et pour les segments 6, 7 et 8 à forte MOI (Welch unilatéral, p-values < 0.05).

Nous avons comparé les fréquences pour chaque segment à celles dans l'inocula. Pour la majorité des segments, leurs fréquences étaient significativement différentes entre l'inocula et les quatre conditions d'infection (Figure 2, Annexes Tableau 2). Une fréquence significativement plus basse était observée pour les segments 1, 2 et 3 (Welch unilatéral, p-values < 0.05) a faible MOI, et pour le segment 1 à forte MOI. Une fréquence significativement plus haute était observée pour les segments 6, 7, 9 et 10 à faible MOI (Welch unilatéral, p-values < 0.05), et pour les segments 6, 7, 8 et 9 à forte MOI (Welch unilatéral, p-values < 0.05).

Les fréquences différaient aussi entre les deux conditions de MOI pour chaque type cellulaire (Figure 2). Dans certains cas, les distributions non gaussiennes de fréquences ne permettent pas d'utiliser une ANOVA. C'est le cas pour le segment 3 chez les virus produits sur KC à MOI 0.3 (Shapiro, p-value = 9.32E-03) et pour le segment 10 sur MDBK dans les deux conditions de MOI (Shapiro, MOI 0.3 p-value = 3.05E-03, MOI 3 p-value = 2.40E-02). De même, les conditions d'homogénéité des variances sont rejetées pour les segments 6 et 8 (Fligner-Killeen, p-values < 0.01). Le test de Wilcoxon a donc été utilisé pour l'ensemble des segments et donne des

résultats de même significativité que les tests post-hoc de Tukey sur les ANOVA des segments où les prérequis du test étaient valides. Nos résultats témoignent d'une situation complexe, avec des variations de fréquences qui diffèrent entre segments en fonction de nos conditions d'infection. Les résultats statistiques des comparaisons par segment entre conditions sont présentés dans le Tableau 3 en Annexes. Nous distinguons quatre comportements différents des fréquences des segments en fonction des conditions de MOI :

(i) La fréquence des segments varie de la même façon entre les conditions de MOI pour les deux types cellulaires. La forte MOI induit une augmentation de la fréquence pour les segments 1, 2 et 3, mais son effet est inverse pour le segment 10 (Wilcoxon, KC p-values < 0,05, MDBK p-values < 0,05).

(ii) La MOI peut aussi avoir un effet opposé sur les fréquences en fonction du type cellulaire. Pour le segment 8 en particulier : la forte MOI induit une diminution de la fréquence pour les cellules insectes mais une augmentation dans les cellules mammifères (Wilcoxon, KC p-value = 4.44E-05, MDBK p-value = 7.40E-07).

(iii) Dans certains cas, la MOI n'affecte les fréquences que pour un type cellulaire. Dans les cellules MDBK, la forte MOI implique une augmentation de la fréquence pour le segment 4, elle induit en revanche une diminution de la fréquence pour le segment 5 (Wilcoxon, MDBK p-values < 0,05). De la même manière, pour la forte MOI, une diminution de la fréquence est observée pour le segment 6 dans les cellules KC (Wilcoxon, KC p-value = 1,48E-06).

(iv) Les fréquences de certains segments ne sont pas affectées par les différentes conditions. Pour les segments 7 et 9 les fréquences ne diffèrent pas entre les conditions de MOI et les types cellulaires (Wilcoxon, p-values > 0,05 pour les cellules KC et MDBK).

Pour les différences entre les types cellulaires, dans la condition d'infection à faible MOI, seuls les segments 1 et 2 possèdent des fréquences significativement supérieures dans le type cellulaire MDBK (Wilcoxon, MOI 0.3, p-values < 0,05). Dans les conditions d'infection à forte MOI, les segments 3, 4 et 8 sont plus représentés dans le type cellulaire MDBK (Wilcoxon, MOI 3, p-values < 0,05). En revanche, les segments 1, 5 et 10 ont des fréquences plus élevées dans le type cellulaire insecte, KC (Wilcoxon, MOI 3, p-values < 0,05).

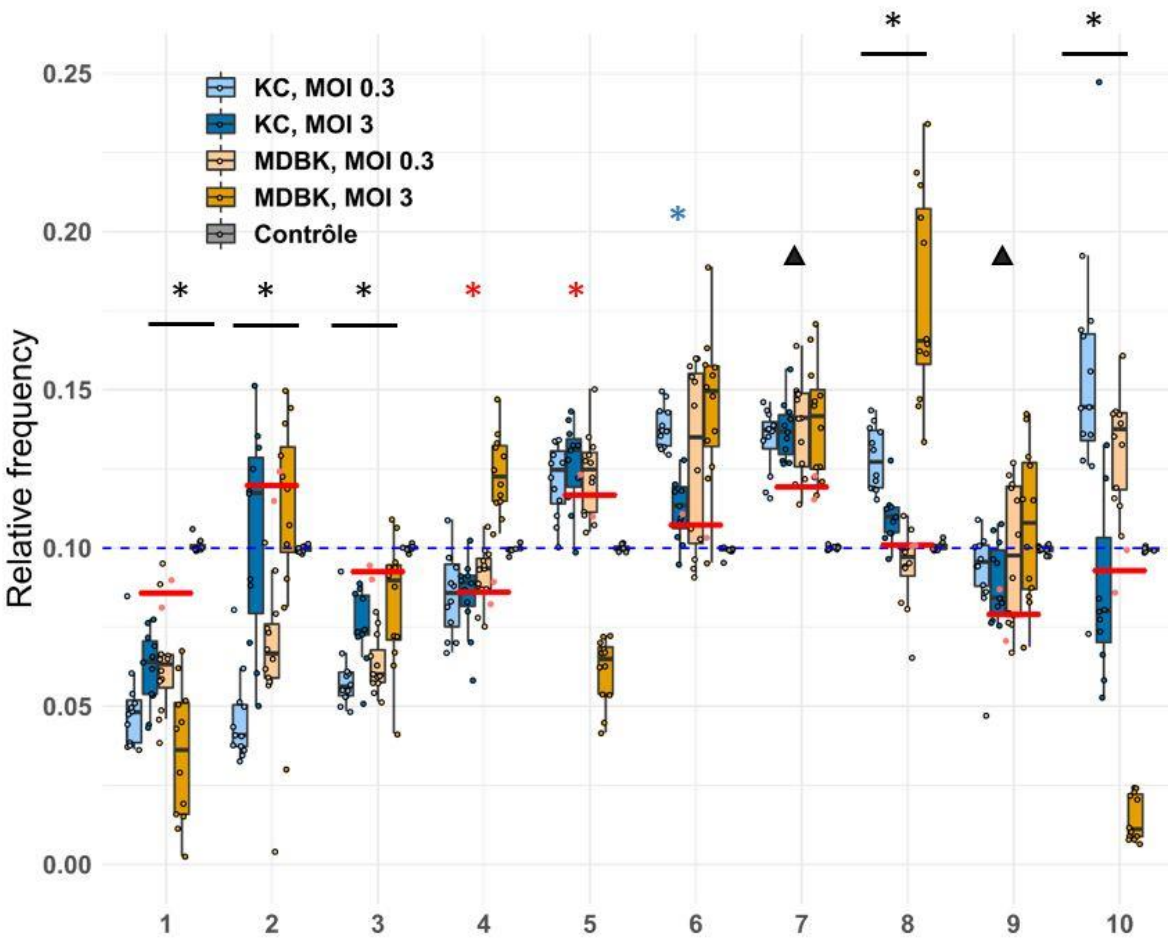


Figure 2. Influence de la MOI sur les fréquences relatives des segments génomiques du virus de la FCO-4 en fonction de la MOI à l'inoculation. La ligne bleue en pointillé indique la fréquence théorique attendue (0,1) pour les segments s'ils étaient tous présents en quantité équimolaire. Les points rouges représentent les valeurs de fréquences de l'inocula tandis que les droites de la même couleur correspondent à leurs valeurs moyennes. Les fréquences du contrôle plasmide ont été inclues pour chaque segment. Les différences significatives ont été déterminées séparément pour chaque segment avec un test de Wilcoxon (p-values < 0.05). Les fréquences pour lesquels une différence entre les conditions de MOI sont observées pour les deux types cellulaires sont marquées par un astérisque noir. Si une différence significative n'est observée que pour un seul type cellulaire, les segments ont été annotés par un astérisque rouge (MDBK) ou bleu (KC). Finalement, les segments pour lesquels aucune différence significative n'est trouvée entre les conditions sont marqués d'un triangle noir (segments 7 et 9).

Influence des conditions d'infection sur la variance des fréquences

Nous avons exploré si la variance des fréquences changeait en fonction de la MOI. Augmenter la MOI devrait faciliter l'apparition de la FG optimale sous le scénario de « génération aléatoire » et donc réduire la variance. Au contraire, sous le scénario « déterministe », augmenter la MOI devrait soit augmenter la variance, soit ne pas avoir une influence majeure. Les coefficients de variations des fréquences ont été estimés pour chaque segment et comparés en fonction des conditions d'infection (Figure 3). Nos résultats soutiennent des variations de la distribution des fréquences équivalentes pour les dix segments pour les différentes conditions d'infection.

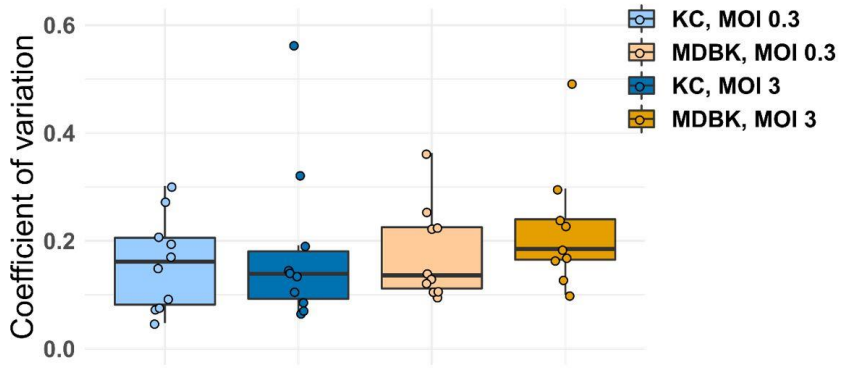


Figure 3. Coefficients de variation calculés pour chaque segment dans nos différentes conditions d'infection. Les coefficients de variation ne présentent pas de variation significative entre les conditions de MOI et de type cellulaire (Welch, p-value > 0,05).

Distance des fréquences par rapport à l'équimolarité et à l'inocula

Nous avons estimé un indice de distance des fréquences de chaque échantillon et l'inocula à la fréquence attendue en situation d'équimolarité (dEQ). Cet indice diffère significativement entre nos traitements et sa valeur moyenne d'origine dans l'inocula (test de Student, unilatéral, p-value < 0.05, Figure 4A). Des différences sont observées dans chaque type cellulaire en fonction de la MOI (Figure 4A). Pour les cellules KC, la distance à l'équimolarité est significativement plus élevée pour la condition d'infection à forte MOI (Welch, p-value = 4.37E-04). Au contraire, les populations virales issues de l'infection des cellules MDBK ont une distance à l'équimolarité plus importante pour la condition de MOI forte (Welch, p-value = 2.22E-04).

Nous avons également calculé un indice de distance à l'inocula pour chaque échantillon (dIN). La distance à l'inocula différait significativement entre les conditions de MOI (Figure 4B). Les variations de cet indice sont identiques à celles de l'indice de distance à l'équimolarité : pour les infections sur cellules KC, la distance à l'inocula est significativement plus faible pour la condition d'infection à forte MOI (Welch, p-value = 6.16E-04). Tandis que pour les virus issus de l'infection des cellules MDBK la distance à l'inocula est plus faible pour la condition de MOI faible (Welch, p-value = 7.15E-04). Les résultats présentés dans la figure 4 montrent que les distances à l'équimolarité et à l'inocula changent dans le même sens pour un type cellulaire donné. Cependant, le sens du changement est inversé entre types cellulaires : les distances diminuent avec la MOI pour les cellules KC mais elles augmentent pour les cellules MDBK.

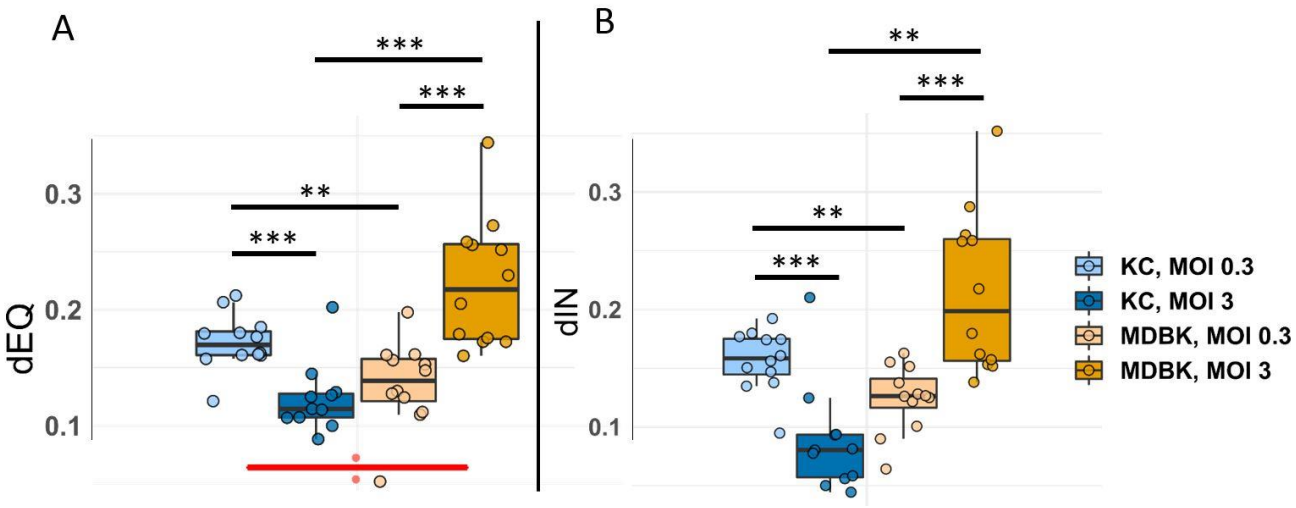


Figure 4. (A) Indice de distance à l'équimolarité (dEQ) calculé par échantillon pour chacune de nos conditions par rapport à l'attendu théorique de fréquence en situation d'équimolarité entre segments. Les astérisques indiquent les différences significatives observées entre les conditions de MOI (test de Welch, p-value < 0.05). La MOI a une influence significative sur l'indice de la distance à l'équimolarité. Les points rouges correspondent à l'indice de distance calculé pour l'inocula, la barre de la même couleur correspond à leur valeur moyenne. **(B)** Indice de distance à l'inocula (dIN) défini par rapport aux fréquences moyennes des segments du virus utilisé pour l'infection et représenté pour chaque condition. La MOI a un effet significatif sur la distance à l'inocula (Welch, p-value < 0.05), il est variable en fonction des types cellulaires. Les astérisques marquent la significativité des différences observées par type

Distance entre les populations expérimentales et naturelles

Les conditions d'infection *in vitro* diffèrent singulièrement des conditions naturelles d'infection notamment car elles permettent de faire abstraction de nombreuses pressions de sélections. Nous étions curieux de comparer l'état des fréquences et l'effet des MOI sur ces dernières par rapport aux formules génomiques observées en conditions naturelles d'infection (Figure 5). Nous avons calculé l'indice de distance à la formule génomique des populations naturelles en utilisant les fréquences médianes des segments viraux trouvées dans les populations issues soit de vaches pour les cellules MDBK, soit de culicoïdes pour les KC. Entre nos conditions de MOI, des différences significatives sont trouvées pour les cellules KC et MDBK (Welch, p-value = 4.80E-02 et 9.08E-04 respectivement). Pour les cellules KC, les indices de distance à la formule génomique sont significativement différents de la dGFa des populations naturelles à faible et à forte MOI (Welch, p-value = 5.80E-06 et 2.22E-02 respectivement). Les mêmes différences sont observées pour les cellules MDBK à faible et à forte MOI (Welch, p-value = 2.74E-05 et 5.30E-06 respectivement). Nous avons constaté que les indices de distance à la formule génomique ne variaient pas seulement en fonction du type cellulaire (Annexes Tableau 4).

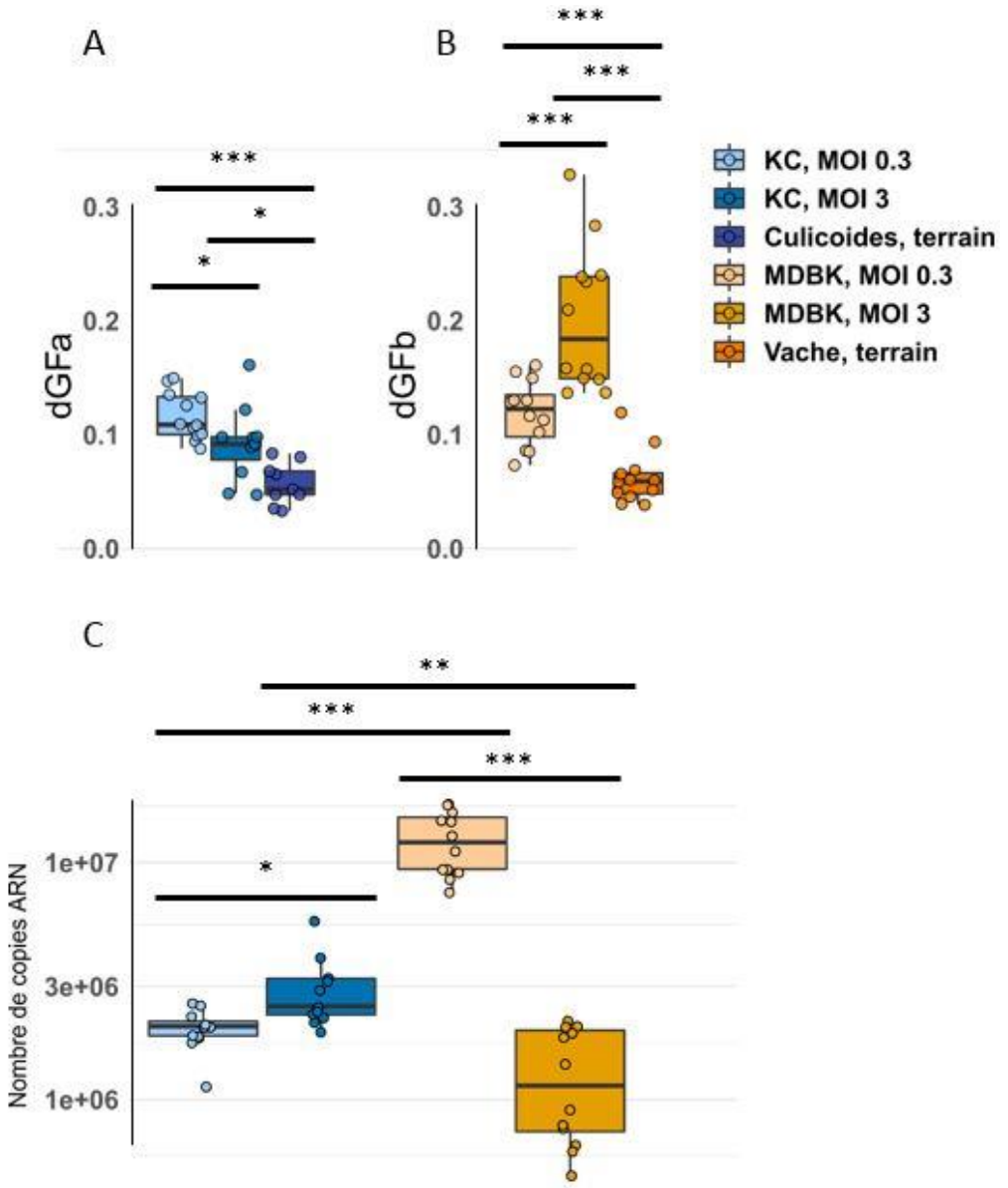


Figure 5. Indices de distance à la formule génomique calculé par rapport aux valeurs médianes de fréquence des formules génomiques des virus décrits dans des populations naturelles, soit des **(A)** hôtes arthropodes (dGFa), soit **(B)** des bovins (dGFb). Les espèces hôtes correspondent aux types cellulaires testés dans notre approche. Les indices des populations virales de terrain ont été ajoutés afin de rendre compte de la distance à la formule génomique existant en conditions naturelles. Les astérisques indiquent les différences significatives entre chaque condition en fonction du groupe étudié (Welch, p-value < 0.05). **(C)** Nombre cumulé de copies des segments génomiques pour les différentes conditions d'infection. Pour chaque échantillon les nombres de copies des 10 segment sont additionnés. Entre les conditions de MOI, les différences significatives de charge virale sont indiquées par des astérisques (test de Welch, p-value < 0.05).

La convergence vers une FG à l'origine de l'augmentation de la charge virale

Nos résultats ci-dessus suggèrent fortement que chaque lignée cellulaire représente un environnement distinct des autres lignées pour le profil de fréquences. Cette observation, similaire à celle identifiée chez les populations naturelles (Chapitre 1), pourrait impliquer que certains profils permettent d'optimiser la multiplication, et donc la charge virale, dans un environnement ou un type cellulaire donné. Nous avons donc cherché à estimer la charge virale (*i.e.* le nombre de copies de tous les segments viraux dans un échantillon) dans chaque condition expérimentale. Cette variable peut être utilisée comme proxy de la valeur sélective des populations virales en culture cellulaire, cependant la concentration en particule virale infectieuse lui est généralement préférée. Nos résultats précédents indiquent cependant que le nombre de copies cumulé d'ARN semble constituer un proxy assez fiable du titre viral (Annexes Figures 6A et 6B).

Nous avons estimé et comparé la charge virale entre les différentes conditions de notre expérience. La MOI à l'inoculation a eu un effet significatif sur la charge virale pour les deux types cellulaires (Figure 5C). Pour les cellules d'insecte, l'infection à faible MOI a induit une charge virale qui est significativement plus faible que pour la condition à forte MOI (Welch, p-value = 0,017). Le nombre moyen de copies cumulées est de 2,00E06 et de 2,97E06 respectivement pour les MOI faible et forte, soit un ordre de différence de 1,48 entre les conditions de MOI. La tendance inverse est observée pour les cellules MDBK. La charge virale est significativement supérieure en condition de faible MOI (Welch, p-value = 3,62E-07). Pour cette condition, le nombre de copies cumulées est en moyenne de 1,26E07 contre 1,31E06 copies à forte MOI. Un ordre 10 de différence est observé pour la charge virale entre nos conditions d'infection pour le type cellulaire mammifère. Donc, les conditions de MOI qui ont induit une charge virale supérieure étaient aussi celles plus proches des populations naturelles et ceci pour les deux lignées cellulaires (Figure 5). Si la charge virale est estimée par segment, le nombre de copies quantifiées par segment conserve l'ordre des variations observées entre les conditions d'infection pour huit des dix segments génomiques (Annexe, Figure 5)

L'effet opposé de la MOI sur la charge virale en fonction de l'origine des lignées cellulaires avait été observé lors d'expériences préliminaires pour la mise en place des infections expérimentales. Ces expériences ont été réalisées avec les lignées KC et

VERO. La lignée VERO est la lignée d'origine vertébré susceptible à l'infection par le virus de la FCO mais elle a une origine non-ruminant (primate). Ces expériences préliminaires n'ont pas été faites avec les mêmes conditions (MOI plus faibles et temps d'infection plus longs) mais elles ont aussi montré le même effet opposé de la MOI sur la charge virale entre les lignées insecte et vertébré (Annexes Figure 6A). De plus, l'effet opposé a été observé non seulement pour la charge virale estimée par RT-qPCR mais aussi lors des titrations en culture cellulaire (Annexes, Figure 6B).

Nous avons exploré davantage le possible lien entre la charge virale et les distances aux trois profils de fréquences étudiés ci-dessus : les populations naturelles, l'équimolarité et l'inocula. Nous avons comparé la charge virale à chacune de distances dans les échantillons de chaque condition expérimentale par des régressions linéaires. Nous n'avons pas détecté de corrélation significative entre la charge virale et les distances sauf pour la condition « forte MOI/cellules MBDK » (Annexes, Figures 1, 2, 3 et 4). Dans cette condition, les trois distances corrélaient avec la charge virale avec un pente positive (R^2 > 0,72 ; p-values < 1E-04). Ces significativités pour la condition de forte MOI sur les cellules MBDK doivent cependant être relativisées car les deux expériences réplicas ont un effet fort sur la significativité de l'interaction et nous n'avons pas détecté des différences significatives quand chaque expérience réplica a été analysée séparément (régression linéaire, R^2 < 0.13 ; p-values > 0,477).

DISCUSSION

Nos travaux visent à préciser le processus qui conduit à la génération d'une fréquence spécifique des segments génomiques du virus de la FCO-4 dans ces différents hôtes et au cours de l'infection. Pour déterminer quels facteurs pouvaient influencer le phénomène nous avons étudié le premier cycle de l'infection. Après 12 heures d'infection le virus de la FCO-4 n'a pu effectuer qu'un seul cycle de réplication (i.e. seules les particules virales de l'inocula ont pu se répliquer). Les conditions d'infection *in* vitro limitent l'influence des pressions de sélection, les fréquences de segment générées devraient être dépendantes de la capacité des génomes présents au sein de chaque cellule à produire une descendance plus ou moins importante. Notre approche permet en particulier de contrôler l'influence de la MOI sur le phénomène de VNCG, avec l'effet potentiel de la complémentation entre les génomes viraux lors de la co-infection cellulaire sur l'émergence de profils particuliers. Nos résultats soutiennent la capacité du virus de la FCO à générer une infection à faible MOI. Dans cette condition, les FG des populations virales changent peu entre les deux types cellulaires. En revanche, à forte MOI, différents profils de fréquences sont observés. Dans ce système d'infection nous montrons que les différents types cellulaires peuvent conduire à l'expression d'une FG dépendante de l'environnement, cependant les profils semblent aussi très dépendants de l'efficacité de réplication du virus. Aussi, nos résultats soutiennent plutôt le scénario déterministe. Les coefficients de variation ne semblent pas être influencés par l'effet du goulot d'étranglement à l'inoculation.

Biais possibles

La mise en place de cette expérience a soulevé différentes questions quant aux facteurs qui pourraient influencer la nature de nos résultats. Le titre obtenu sur les cellules OA3 a été utilisé pour calculer la MOI pour le type cellulaire KC. Nous avons sélectionné le titre le plus élevé pour augmenter la probabilité d'avoir une condition d'infection avec un MOI supérieure à 1. Plusieurs scénarios sont donc envisageables. Le titre pourrait être surévalué, avec une MOI qui pourrait être inférieure à 1 dans nos deux conditions d'infection. Dans ce cas, il y aurait rarement plus d'un génome par cellule et nous observerions une variation de charge virale due au nombre plus

important de cellules infectées pour la condition où la MOI est plus forte. Ce scénario n'est pas soutenu par les différences que nous observons entre les deux conditions de MOI. Une expérience préliminaire menée au cours de ce projet montrait l'absence de variation entre les profils de FGs entre deux conditions de MOI faibles (< 0,1) pour le type cellulaire KC après 48 heures d'infection (Annexes figure 7). Nos résultats montrent bien une différence de fréquences des segments dépendante de la MOI pour KC (figure 2). A l'inverse, si le titre est plus élevé que le titre de référence utilisé, les deux conditions de MOI pourraient permettre la co-infection à l'inoculation. Des expériences additionnelles sont requises pour clarifier ce point.

L'effet de la MOI dans nos conditions d'infection est complexe mais d'autres études montrent des résultats similaires. Les travaux de Moreno et al. (13) témoignent de l'effet complexe de ce facteur sur le cycle de réplication pour le virus de la chorioméningite lymphocytaire (LCMV), le virus de la stomatite vésiculaire (VSV) et le virus de la fièvre aphteuse (FMDV). Pour le LCMV le titre (pfu/ml) de la descendance, après 48 heures d'infection, est supérieur dans une condition d'infection à faible MOI (= 0.01) par rapport à la condition de forte MOI (= 10). Ces différences se traduisent par une diminution du nombre de copies ARN viraux détectées pour la condition de forte MOI ainsi qu'un ratio du nombre de particules infectieuses par nombre de copies ARN beaucoup plus faible pour cette condition. Pour le VSV, ces deux conditions d'infection n'affectent pas le titre de la descendance virale produite. En revanche, pour le FMDV, le titre de la descendance est plus élevé pour une MOI de 1 que pour une MOI faible (= 0.001). Cette étude souligne également l'importance que la diversité génétique peut avoir sur la génération de particules virales infectieuses. Pour le LMDV, en présence d'un agent mutagène, le titre, le nombre de copies et le ratio de virions infectieux sont plus forts pour la condition de forte MOI.

Nos résultats pourraient être influencés par des spécificités de la souche FCO-4_K1B2 utilisée comme inocula. FCO-4_K1B2 provient d'une population virale isolée à partir de sang d'un mouton infecté. Les étapes d'isolation et l'amplification de cette souche en culture cellulaire ont pu mener à l'apparition de nouvelles (14, 15). Kopanke et al. montrent dans ce sens des mutations probablement spécifiques à l'infection en culture cellulaire avec le virus de la FCO-17. Ces auteurs ont fait dix passages du virus en cultures cellulaires soit de vache (BPAEC), soit de culicoïdes (CuVaW3), soit en alternance dans les deux types cellulaires. Leurs résultats montrent la fixation de

mutations spécifiques communes aux trois conditions (16). En prenant un modèle plus proche du notre, le passage du virus de la FCO-3 entre KC et BSR diminue la diversité génétique de la population virale, elle peut cependant être rétablie après plusieurs amplifications dans les cellules BSR (15). Les auteurs soulignent aussi une diminution de la virulence d'un virus amplifié sur BSR sur les cellules insectes KC. La souche FCO-4_K1B2 et son historique de passage pourraient conduire à des différences de virulence entre les types cellulaires KC et MDBK. L'influence potentielle de l'historique de passage n'enlève rien à nos observations sur la génération des profils des fréquences ou sur les différences entre les conditions expérimentales.

Quel mode de génération de la formule génomique ?

La fréquence des segments génomique peut dévier de l'inocula après seulement un cycle de réplication. Ces variations sont similaires pour huit segments dans la condition de faible MOI. En revanche, à forte MOI, seulement cinq segments ont des variations similaires entre les deux types cellulaires. La complémentation possible à forte MOI a peut-être une influence sur la capacité des populations virales à diverger de l'état des fréquences initiales. Cependant les populations virales qui se sont répliquées sur les cellules KC à forte MOI présentent l'indice de distance à l'inocula le plus faible. Ces différences sont directement liées aux variations plus fortes des fréquences pour les cellules MDBK à forte MOI. La complémentation ne constitue pas le seul facteur à l'origine du changement de FG.

A faible MOI, la présence d'un seul virion par cellule devrait permettre la réplication de particules portant les gènes indispensables à l'assemblage de nouvelles particules. Lors d'une expérience d'assemblage dans un système sans cellule, il a été montré que le virus de la FCO pouvait générer des particules virales à partir des protéines constitutives du complexe de transcription,de la capside core et de certains segments génomiques (7). L'absence d'un segment diminue l'efficacité d'encapsidation pour le segment 6. Les segments ont une influence différente sur l'efficacité d'encapsidation, avec une diminution plus importante lorsqu'un des segments de plus petite taille est absent. En système d'infection expérimentale, les résultats sont plus mitigés, il semble par exemple que le virus de la FCO peut se répliquer en l'absence des protéines codées par le segment 10 dans les cellules mammifères et insectes mais avec une

virulence diminuée (17). Les segments 2 et 10 jouent un rôle central dans la virulence de la FCO-8, mais les segments 1, 4, 5, 6, 7 et 9 participent aussi (18). A faible MOI, l'absence de convergence des populations virales vers une fréquence à l'équilibre pourrait suggérer qu'il existe des mécanismes autres que la co-infection par des virions libres permettant la complémentation. L'agrégation des particules virales, principalement présente sous une forme enveloppée en culture cellulaire, pourrait permettre le regroupement de plusieurs particules virales au sein de vésicules extracellulaires pour le virus de la FCO (19). De plus, le potentiel de ce virus à transférer ses ARN ou ses protéines entre les cellules adjacentes constituent une autre piste à explorer. Ce phénomène permet par exemple au FBNSV de compléter son cycle de réplication au sein d'une cellule malgré l'absence de certains segments génomiques dans la cellule (2).

La génération aléatoire de la formule génomique devrait induire des différences entre nos coefficients de variation pour les différentes conditions d'infection. Les coefficients de variations sont identiques y compris pour la condition d'infection à forte MOI ou la complémentation des génomes devrait permettre une variation plus importante. Nos résultats soutiennent donc le scénario déterministe qui implique des facteurs liés à la génétique du virus de la FCO-4 dans la génération des FG. Les fréquences observées après un cycle d'infection présentent cependant une distribution relativement large. Cette variabilité implique que le scénario de génération de la FG n'est probablement pas exclusivement déterministe.

Quelle influence de la MOI sur les profils des fréquences ?

La MOI a une influence sur les profils des fréquences à plusieurs niveaux. Pour les deux types cellulaire l'augmentation de la MOI se traduit par une augmentation de la fréquence des segments 2 et 3, ainsi qu'une diminution de la fréquence du segment 10. Les gènes des segments 2 et 10 sont impliqués dans des facteurs de virulence du virus de la FCO-8 (18) mais nous ne possédons pas une explication claire quant aux variations de fréquences conservées pour ces trois segments. Néanmoins aux deux valeurs de MOI testées, la fréquence des segments dévie de l'équimolarité et de l'inocula pour une majorité des segments. La comparaison des indices de distance à

l'équilibre, à l'inocula et de la valeur de nombres cumulés de copies ARN montre que leurs variations sont dépendantes des MOI différentes pour chaque type cellulaire.

Nos résultats suggèrent aussi une influence de la MOI dépendante de la lignée cellulaire. Le nombre des segments qui ont des fréquences différentes entre les conditions de MOI est proche pour MDBK et KC (7 vs 6). Cependant, les variations de fréquences ne sont pas toujours conservées (segments 4, 5, 6) et le sens de variation peut être différent entre lignées (segment 8). Il serait intéressant de pouvoir valider la conservation de ce phénomène avec d'autres lignées de cellules insecte et vertébré, mais également avec d'autres souches du virus de la FCO. Les résultats pourraient être différents au moins chez les cellules de mammifères qui peuvent mener à la génération d'une diversité génétique virale différentes entre BSR et MDBK (16, 20). Nos données suggèrent que la charge virale finale pourrait permettre de généraliser nos observations pour les deux lignées : les conditions d'infection qui permettent la plus forte charge virale sont celles qui génèrent les profils de séquence les plus proche des populations naturelles, indépendamment de l'hôte d'origine. L'utilisation d'un gradient plus important de valeurs de MOI au cours de l'infection pourrait permettre de clarifier l'interaction étroite entre la MOI, la charge virale et son effet dépendant du type cellulaire pour la souche FCO-4_K1B2.

Implications pour l'infection in vivo

Nos résultats suggèrent des formules génomiques différentes en fonction de la lignée cellulaire et de la MOI (Figure 7A), avec des différences d'abondance entre les segments pouvant dépasser un facteur 10x (Figure 7B). Pour les cellules KC, le segment 7 est le plus représenté pour les deux conditions de MOI. Les profils de FG sont assez proches pour ce type cellulaire, à l'exception du segment 10 qui est trois fois plus présent à faible MOI. L'ordre de variation maximal est de 3.5 entre les segments 2 et 7 pour la MOI faible tandis qu'il est d'environ 2 entre le segment 1 et 7 pour la MOI forte. Pour les cellules MDBK, les conditions de MOI ont un effet majeur sur l'ordre de différence entre les segments. A faible MOI, les variations sont proches des FGs observées sur KC dans la même condition (Figure 7A). En revanche, pour la MOI forte, l'ordre de variation entre les segments 8 et 10 est de 17,5x. Il est possible que le faible nombre de copies ARN viraux détecté soit associé au déséquilibre plus

important des fréquences entre les segments pour cette condition d'infection. Cette différence pourrait être associée à une sur-représentation de génomes défectifs au sein de la population virale (21). Elle pourrait conduire à un effet délétère pour la réplication du virus comme cela a été décrit chez différentes espèces de virus (22). En analysant les résultats d'une infection expérimentale antérieure, nous avons estimé le nombre de particules virales infectieuses par rapport au nombre de copies ARN quantifiées (Annexe, Figure 7C). En partant de l'hypothèse qu'il y'a en moyenne dix segments par particule virale, nos estimations prévoient que 1 à 10% des particules virales sont infectieuses à l'issue d'une infection sur cellules insectes. En revanche leur fréquence est inférieure à 1 % à l'issue de l'infection sur cellules VERO. Ces observations pourraient expliquer le phénomène particulier observé à forte MOI dans les cellules MDBK, les génomes défectifs pourraient se répliquer plus rapidement et entrer en compétition avec les segments génomiques complets (23). Ces estimations s'approchent de celles obtenues pour la grippe A ou une partie importante de la population est incapable d'exprimer au moins un gène viral (24, 25).

Si des variations similaires se produisaient *in vivo*, elles impliqueraient que la FG pourrait être influencée non seulement par l'hôte mais aussi par des changements démographiques au cours de l'infection. Les goulots d'étranglement lors de l'infection primaire d'un organe ou d'un hôte pourraient donner lieu à des MOI faibles comme dans notre design (26, 27). À l'inverse, la MOI pourrait être très élevée lors du pic de virémie. Ces dynamiques démographiques et leur lien avec la MOI ont été mis en évidence chez certains virus (28). Nous ne disposons malheureusement pas de données relatives à la MOI au cours de l'infection naturelle chez le virus de la FCO. Il serait intéressant par la suite d'étudier la dynamique temporelle de la variation de la FG au cours de l'infection *in vivo* et *in vitro*. Il semble aussi important de déterminer l'effet des goulots d'étranglement au cours de l'infection.

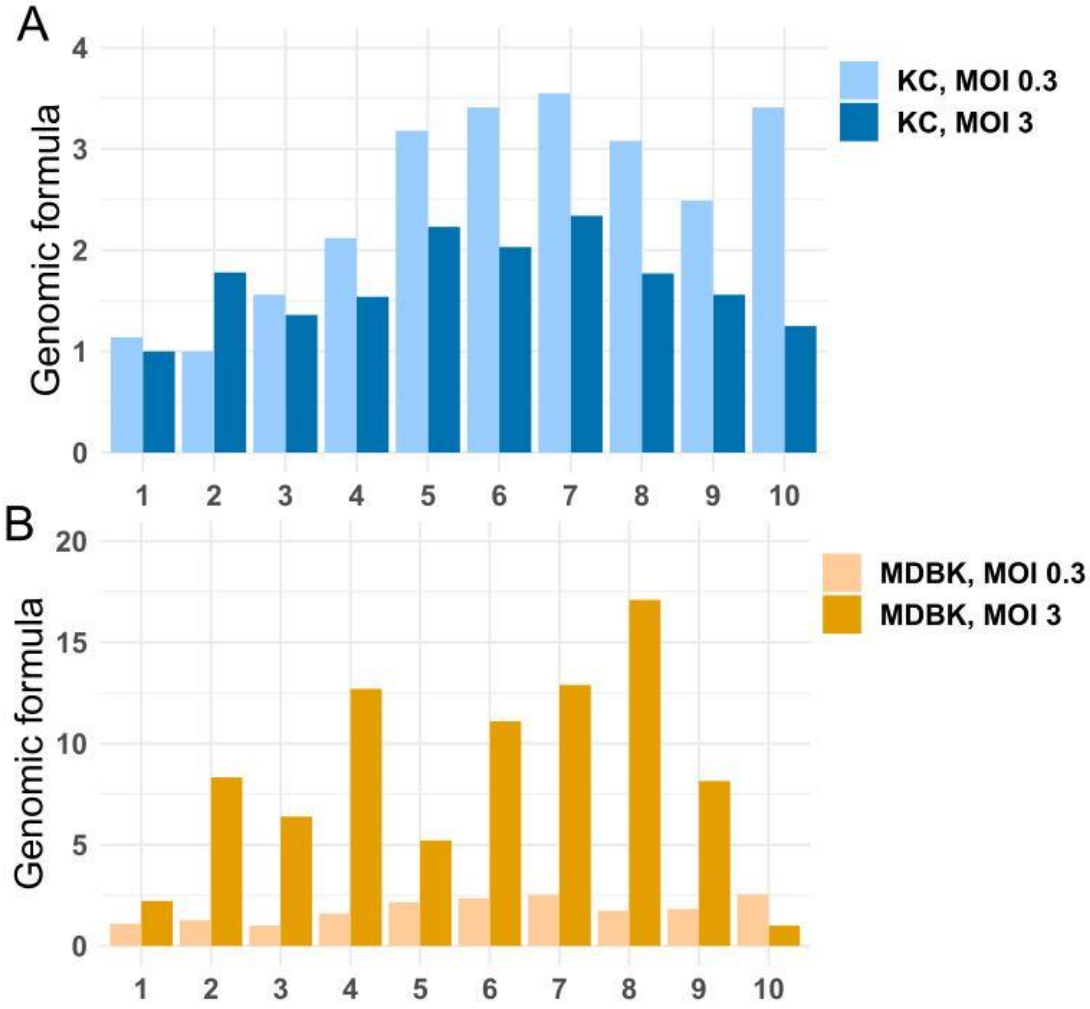


Figure 6. Formule génomique en fonction des conditions d'infection. Les MOI faibles sont présentées dans une couleur plus claire. (A) Formule génomique des virus pour les infections sur cellules KC. (B) Formules génomiques des populations virales à l'issue de l'infection sur MDBK.

CONCLUSION

La variation du nombre de copies de gènes est un mécanisme qui peut être observé aux stades préliminaires de l'infection d'une cellule. Son amplitude, en termes de variations entre les différents segments du virus de la FCO-4, est influencée par le nombre de particules virales initialement présent au sein des cellules infectées. Bien que les conditions de cultures cellulaires limitent l'interférence de facteurs hôtes et donc les pressions de sélection, nous observons potentiellement un effet du passage du virus dans un type cellulaire particulier avant l'infection. Les différences de formule

génomique que nous observons sont dépendantes du type cellulaire et de la MOI. Des travaux complémentaires seraient nécessaires pour déterminer le mécanisme moléculaire derrière le phénomène de VNCG chez ce virus. La capacité du virus de la FCO-4 à se multiplier (i.e. charge virale) semble dépendante de la formule génomique et du type cellulaire. Ainsi pour les deux conditions où la réplication est plus importante nous observons des profils de formule génomique proches des FGs dans les hôtes naturels. Ces résultats soulignent le lien potentiel entre la FG et la valeur sélective pour les populations du virus de la FCO-4. L'accumulation du virus, plus faible dans une condition d'infection à forte MOI sur les cellules de mammifères, est peut-être liée à la présence de particules virales défectives (DIPs). Nous avons estimé que les DIPs pouvaient constituer plus de 99% de la population virale pour ce virus dans les cellules VERO, elles pourraient être à l'origine de cette diminution de réplication virale. Ces résultats soulignent le rôle important que pourraient jouer la MOI sur la génération de la FG et donc sur la capacité du virus de la FCO-4 à se multiplier. Nos résultats invitent à l'étude de l'impact de la MOI au cours des infections naturelles et à l'implication de mécanismes favorisant la complémentation des génomes chez le virus de la FCO.

REFERENCES

1. Sicard A, Yvon M, Timchenko T, Gronenborn B, Michalakis Y, Gutierrez S, Blanc S. 2013. Gene copy number is differentially regulated in a multipartite virus. Nature Communications 4:1-8.
2. Sicard A, Pirolles E, Gallet R, Vernerey MS, Yvon M, Urbino C, Peterschmitt M, Gutierrez S, Michalakis Y, Blanc S. 2019. A multicellular way of life for a multipartite virus. Elife 8.
3. Moreau Y, Gil P, Exbrayat A, Rakotoarivony I, Breard E, Sailleau C, Viarouge C, Zientara S, Savini G, Goffredo M, Mancini G, Loire E, Gutierrez S. 2020. The genome segments of Bluetongue virus differ in copy number in a host-specific manner. J Virol doi:10.1128/JVI.01834-20.
4. Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. 2014. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. Proc Natl Acad Sci U S A 111:16854-9.
5. Schnettler E, Ratinier M, Watson M, Shaw AE, McFarlane M, Varela M, Elliott RM, Palmarini M, Kohl A. 2013. RNA interference targets arbovirus replication in Culicoides cells. J Virol 87:2441-54.
6. Vitour D, Doceul V, Ruscanu S, Chauveau E, Schwartz-Cornil I, Zientara S. 2014. Induction and control of the type I interferon pathway by Bluetongue virus. Virus Res 182:59-70.
7. Sung PY, Roy P. 2014. Sequential packaging of RNA genomic segments during the assembly of Bluetongue virus. Nucleic Acids Res 42:13824-38.
8. Freeman KP, Castro AE, Kautz CE. 1986. Unusual characteristics of a parvovirus isolated from a clinically ill steer. Veterinary Microbiology 11:61-68.
9. King L. 2012. The baculovirus expression system: a laboratory guide. Springer Science & Business Media.
10. Hu Z, Zhang X, Liu W, Zhou Q, Zhang Q, Li G, Yao Q. 2016. Genome segments accumulate with different frequencies in Bombyx mori bidensovirus. J Basic Microbiol 56:1338-1343.
11. Feng L, Lintula S, Ho TH, Anastasina M, Paju A, Haglund C, Stenman UH, Hotakainen K, Orpana A, Kainov D, Stenman J. 2012. Technique for strand-specific gene-expression analysis and monitoring of primer-independent cDNA synthesis in reverse transcription. BioTechniques 52:263-270.
12. Rosenblad A. 2017. Multivariate Statistical Methods: A Primer (4th Edition). Journal of Statistical Software 78.
13. Moreno H, Tejero H, de la Torre JC, Domingo E, Martin V. 2012. Mutagenesis-mediated virus extinction: virus-dependent effect of viral load on sensitivity to lethal defection. PLoS One 7:e32550.
14. Sailleau C, Breard E, Viarouge C, Gorlier A, Quenault H, Hirchaud E, Touzain F, Blanchard Y, Vitour D, Zientara S. 2018. Complete genome sequence of bluetongue virus serotype 4 that emerged on the French island of Corsica in December 2016. Transboundary and Emerging Diseases 65:e194-e197.
15. Lean FZX, Neave MJ, White JR, Payne J, Eastwood T, Bergfeld J, Di Rubbo A, Stevens V, Davies KR, Devlin J, Williams DT, Bingham J. 2019. Attenuation of Bluetongue Virus (BTV) in an in ovo Model Is Related to the Changes of Viral Genetic Diversity of Cell-Culture Passaged BTV. Viruses 11.

16. Kopanke JH, Lee JS, Stenglein MD, Mayo CE. 2020. The Genetic Diversification of a Single Bluetongue Virus Strain Using an In Vitro Model of Alternating-Host Transmission. Viruses 12.
17. van Gennip RG, van de Water SG, van Rijn PA. 2014. Bluetongue virus nonstructural protein NS3/NS3a is not essential for virus replication. PLoS One 9:e85788.
18. Janowicz A, Caporale M, Shaw A, Gulletta S, Di Gialleonardo L, Ratinier M, Palmarini M. 2015. Multiple genome segments determine virulence of bluetongue virus serotype 8. J Virol 89:5238-49.
19. Labadie T, Roy P. 2020. A non-enveloped arbovirus released in lysosome-derived extracellular vesicles induces super-infection exclusion. PLoS Pathog 16:e1009015.
20. Caporale M, Di Gialleonorado L, Janowicz A, Wilkie G, Shaw A, Savini G, Van Rijn PA, Mertens P, Di Ventura M, Palmarini M. 2014. Virus and host factors affecting the clinical outcome of bluetongue virus infection. J Virol 88:10399-411.
21. Vignuzzi M, Lopez CB. 2019. Defective viral genomes are key drivers of the virus-host interaction. Nat Microbiol 4:1075-1087.
22. Ana Grande-Perez ELz, Pedro Lowenstein, Esteban Domingo, and Susanna C. Manrubia. 2005. Suppression of viral infectivity through lethal defection. PNAS.
23. Brooke CB. 2017. Population Diversity and Collective Interactions during Influenza Virus Infection. J Virol 91.
24. Russell AB, Trapnell C, Bloom JD. 2018. Extreme heterogeneity of influenza virus infection in single cells. Elife 7.
25. Brooke CB, Ince WL, Wrammert J, Ahmed R, Wilson PC, Bennink JR, Yewdell JW. 2013. Most influenza a virions fail to express at least one essential viral protein. J Virol 87:3155-62.
26. Grubaugh ND, Ruckert C, Armstrong PM, Bransfield A, Anderson JF, Ebel GD, Brackney DE. 2016. Transmission bottlenecks and RNAi collectively influence tick-borne flavivirus evolution. Virus Evol 2:vew033.
27. Mellor PS, Carpenter S, White DM. 2009. Chapter 14 - Bluetongue virus in the insect host, p 295-320. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50018-6. Academic Press, London.
28. Gutierrez S, Michalakis Y, Blanc S. 2012. Virus population bottlenecks during within-host progression and host-to-host transmission. Curr Opin Virol 2:546-55.

ANNEXES

Tableau S1. Différence des fréquences observées à l'attendu théorique de 0.1 pour les différents segments dans les conditions expérimentales. Le tableau présente les p-values des résultats du test de Welch one-sided appliqué séparément pour chaque segment et corrigé pour le taux de découverte de faux positifs par la méthode Benjamin-Hopkins.

Segment	MDBK				KC			
	0,3		3		0,3		3	
	LESS	GREATER	LESS	GREATER	LESS	GREATER	LESS	GREATER
1	3,74E-06	1	4,58E-07	1	1,02E-07	1	6,26E-07	1
2	5,16E-04	1	0,87	0,519	6,34E-08	1	0,848	0,728
3	3,92E-08	1	6,45E-03	1	1,15E-07	1	1,65E-05	1
4	4,96E-03	0,999	1	7,79E-05	2,25E-03	0,999	2,25E-03	0,999
5	1	4,89E-05	1,21E-07	1	1	4,89E-05	1	4,89E-05
6	1	2,70E-03	1	4,46E-05	1	1,05E-09	1	1,23E-03
7	1	1,62E-06	1	4,66E-06	1	1,31E-07	1	1,31E-07
8	0,253	0,937	1	4,64E-06	1	1,69E-06	1	1,39E-03
9	0,526	0,991	0,832	0,674	0,116	0,991	3,66E-02	0,991
10	1	4,75E-05	5,18E-13	1	1	3,03E-04	9,02e- 1	0,732

Tableau S2. Différence des fréquences observées à l'inocula pour les différents segments dans les conditions expérimentales. Le tableau présente les p-values des résultats du test de Welch one-sided appliqué séparément pour chaque segment et corrigé pour le taux de découverte de faux positifs par la méthode Benjamin-Hopkins.

Segment	MDBK				KC			
	0,3		3		0,3		3	
	LESS	GREATER	LESS	GREATER	LESS	GREATER	LESS	GREATER
1	2,69E-04	1	5,19E-05	1	2,94E-06	1	4,07E-05	1
2	9,12E-06	1	0,192	1	2,30E-09	1	9,02E-02	1
3	4,16E-07	1	6,27E-02	1	9,13E-07	1	3,08E-04	1
4	1	5,04E-02	1	9,02E-07	0,86	0,588	0,86	0,588
5	0,98	0,122	2,59E-09	1	0,98	0,123	0,98	8,03E-02
6	1	1,61E-02	1	1,93E-04	1	9,10E-09	1	4,84E-02
7	1	6,45E-04	1	9,33E-04	1	2,65E-04	1	1,18E-04
8	1	0,96	1	5,24E-06	1	2,41E-06	1	2,74E-03
9	0,999	7,88E-03	0,999	4,41E-03	0,999	7,88E-03	0,999	7,88E-03
10	1	7,34E-06	1,36E-12	1	1	8,91E-05	1	0,51

Tableau S4. Comparaison des fréquences des 10 segments génomiques. Les deux colonnes de type cellulaire présentent les résultats (p-values) du test de Wilcoxon entre la MOI de 0.3 et de 3 dans chaque lignée de cellule. Les colonnes MOI présentent les résultats du test de Wilcoxon (p-values) de comparaison entre les cellules pour chacune des deux MOI.

segment	Type cellulaire		MOI	
	KC	MDBK	0,3	3
1	2,25E-02	4,51E-03	1,45E-02	2,18E-03
2	4,44E-05	6,56E-04	5,56E-03	0,525
3	4,49E-03	6,81E-03	0,198	0,26
4	0,928	1,48E-06	0,242	1,48E-06
5	0,316	7,40E-07	0,843	1,48E-06
6	1,48E-06	0,291	0,932	4,01E-04
7	0,786	0,932	0,378	0,976
8	4,44E-05	7,40E-07	0,74	1,48E-06
9	0,347	0,319	0,671	5,95E-02
10	4,49E-03	7,40E-07	0,101	1,48E-06

Tableau S4. Valeurs médianes des indices de distance à la formule génomique du virus de la FCO-4 pour l'hôte arthropode (dGFa) et bovin (dGFb).

	KC		MDBK		Culicoides	Vaches
MOI	0,3	3	0,3	3	NA	NA
dGFa	0,109	0,0918	0,0858	0,218	0,0521	0,0877
dGFb	0,145	0,0767	0,123	0,184	0,0958	0,0591

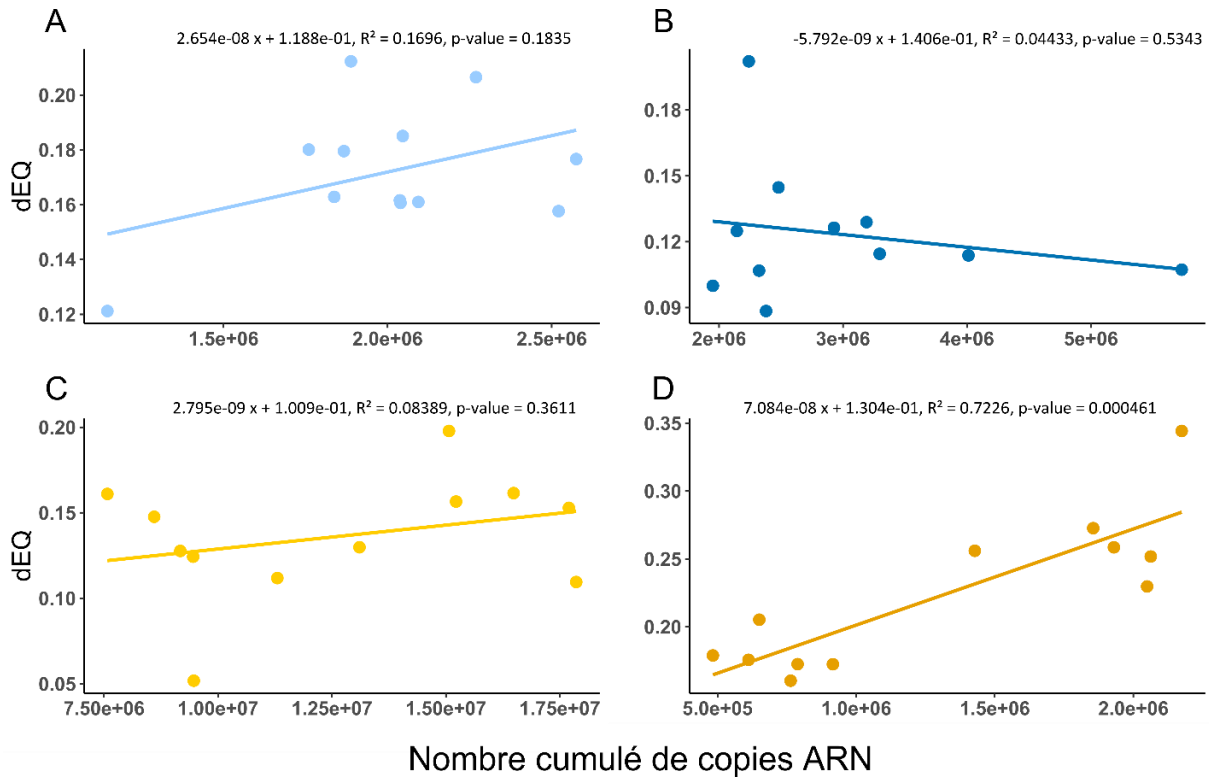


Figure S1. Régression linéaire de la distance à l'équimolarité par la charge virale. Les équations des droites, les coefficients de corrélation et la significativité de l'interaction sont indiqués pour chaque condition. Les distances à l'équimolarité sont représentées en bleu pour les virus produits sur cellules de culicoïdes KC, en clair pour l'infection à faible MOI (A) et en foncé pour celle à forte MOI (B). Pour les cellules de vache MDBK, l'infection à faible MOI est représentée en jaune clair (C) et en jaune foncé pour la forte MOI (D).

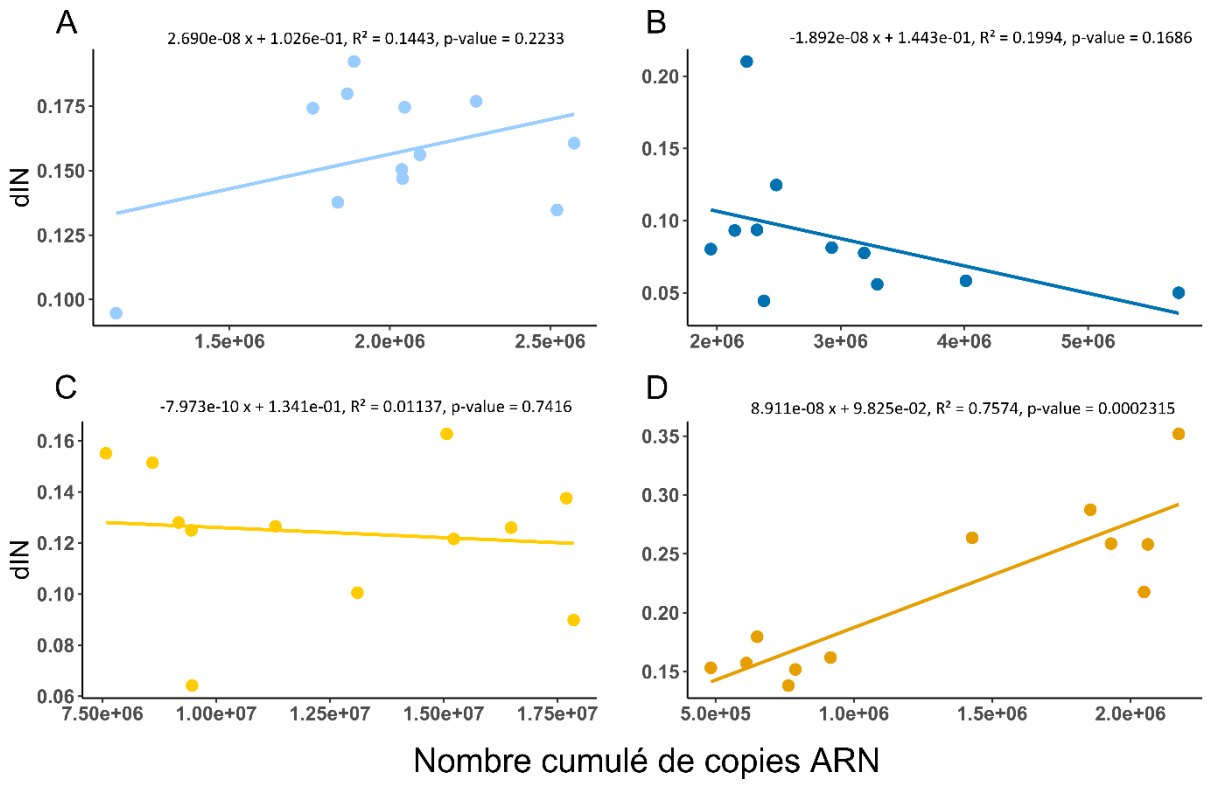


Figure S2. Droites de régression linéaire de la distance à l'inocula par la charge virale. Les équations des droites, les coefficients de corrélation et la significativité de l'interaction sont indiqués pour chaque figure. Les distances à l'équimolarité sont représentées en bleu pour les cellules de culicoïdes KC, (A) en clair pour l'infection à faible MOI et (B) en foncé pour celle à forte MOI. Les distances à l'équimolarité sont représentées en jaune pour les cellules de vache MDBK, (C) en clair pour l'infection à faible MOI et (D) en foncé pour celle à forte MOI.

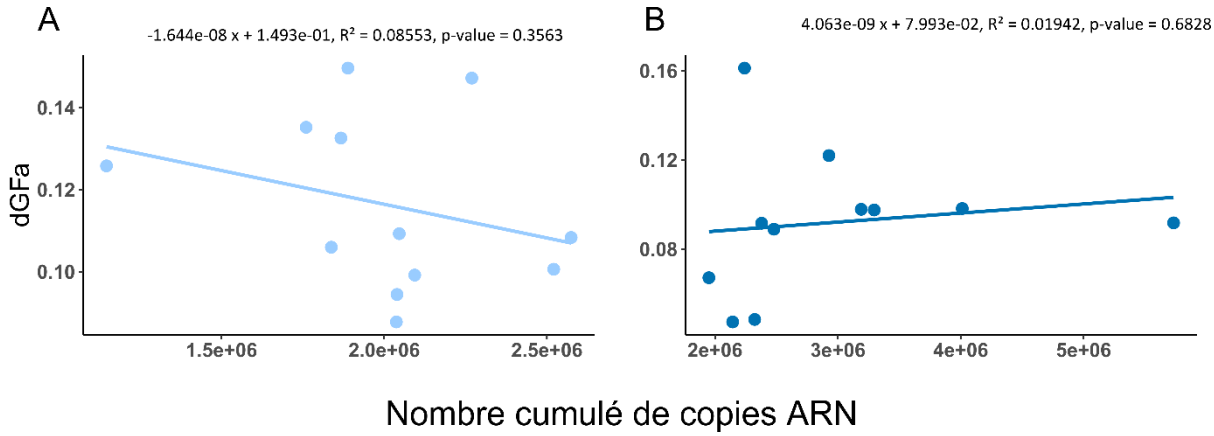


Figure S3. Droites de régression linéaire de l'indice pour la distance à la formule génomique de l'hôte culicoïdes par le nombre cumulé de copies d'ARN viral pour les cellules d'insecte aux MOI de 0.3 (A) et de 3 (B).

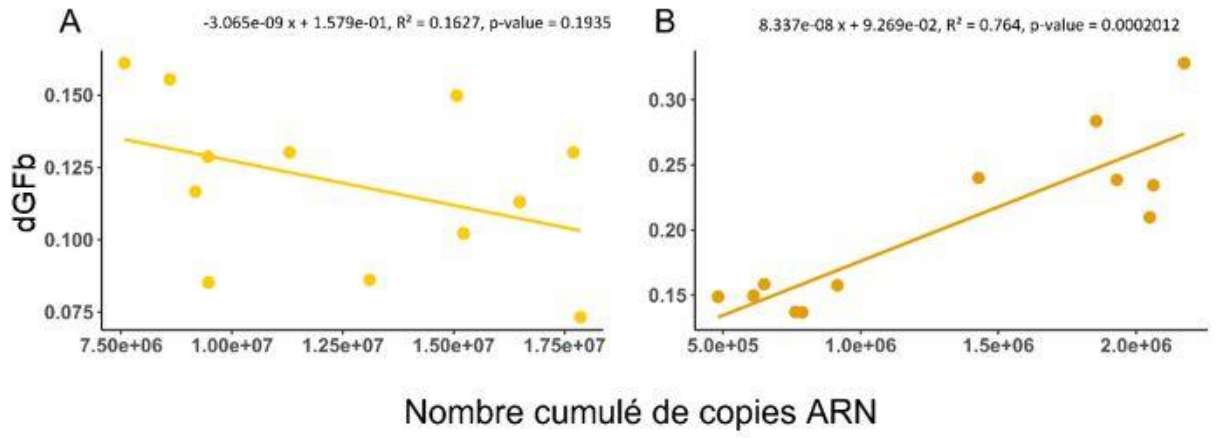


Figure S4. Droites de régression linéaire de l'indice pour la distance à la formule génomique de l'hôte vache par le nombre cumulé de copies d'ARN viral pour les cellules de vache aux MOI de 0.3 (A) et de 3 (B).

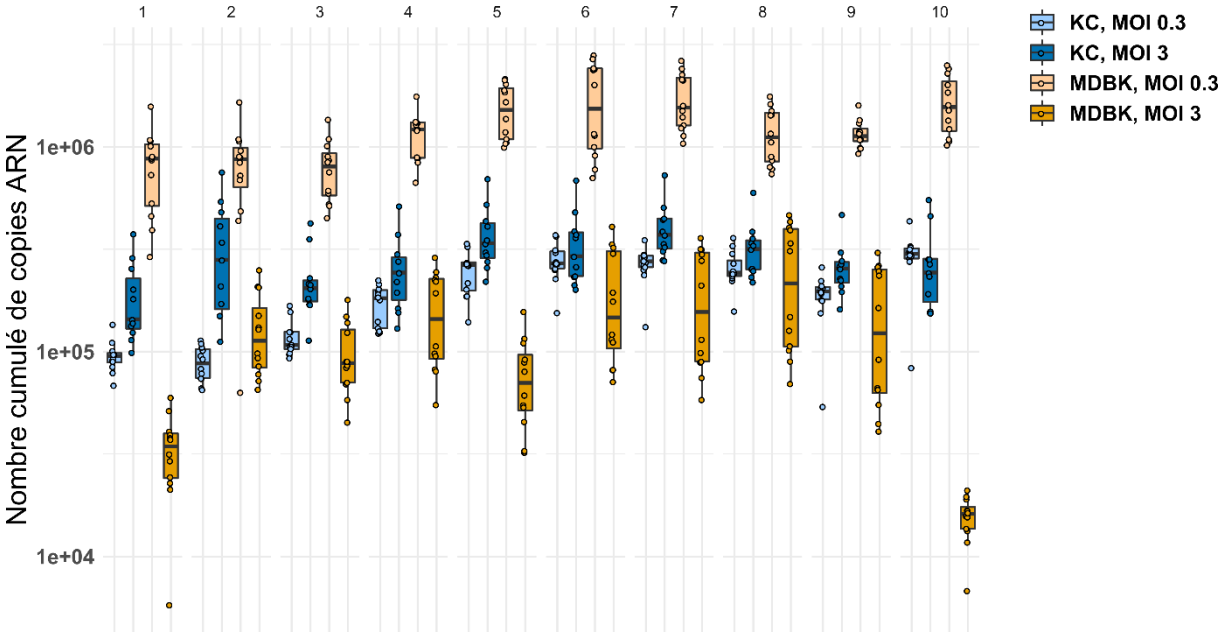


Figure S5. Nombre de copies quantifiées pour chacun des dix segments génomiques entre chaque condition d'infection. Les segments sont présentés en haut de la figure. L'échelle des ordonnées a été transformée en échelle logarithmique (10). A l'exception des segments 2 et 10, l'ordre de variation est conservé entre les conditions par rapport à l'ordre cumulé des copies pour les dix segments présentés dans la figure 5.

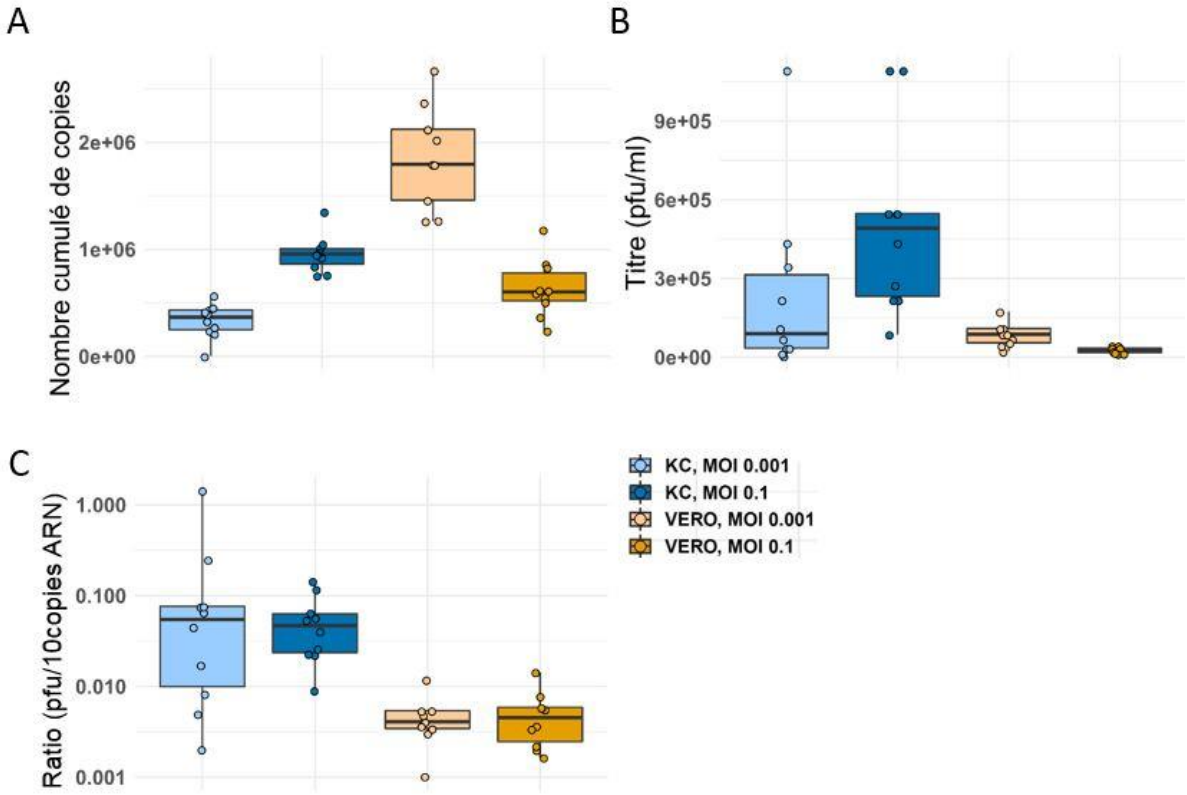


Figure S6. Résultats d'infection expérimentale après 48 heures d'infection sur des cellules KC et VERO à deux MOI inférieures à 1. Chaque condition comprenait 10 réplicas biologiques. **(A)** Nombre cumulé de copies des segments génomiques pour les différentes conditions d'infection. Pour chaque échantillon, les nombres de copies des 10 segments sont additionnés. **(B)** Titre estimé à partir du type cellulaire VERO pour chacune des conditions d'infection. **(C)** Estimation de la proportion de particules virales infectieuses par rapport au nombre de particules virales totales, en considérant que chaque virion contient 10 segments génomiques en moyenne.

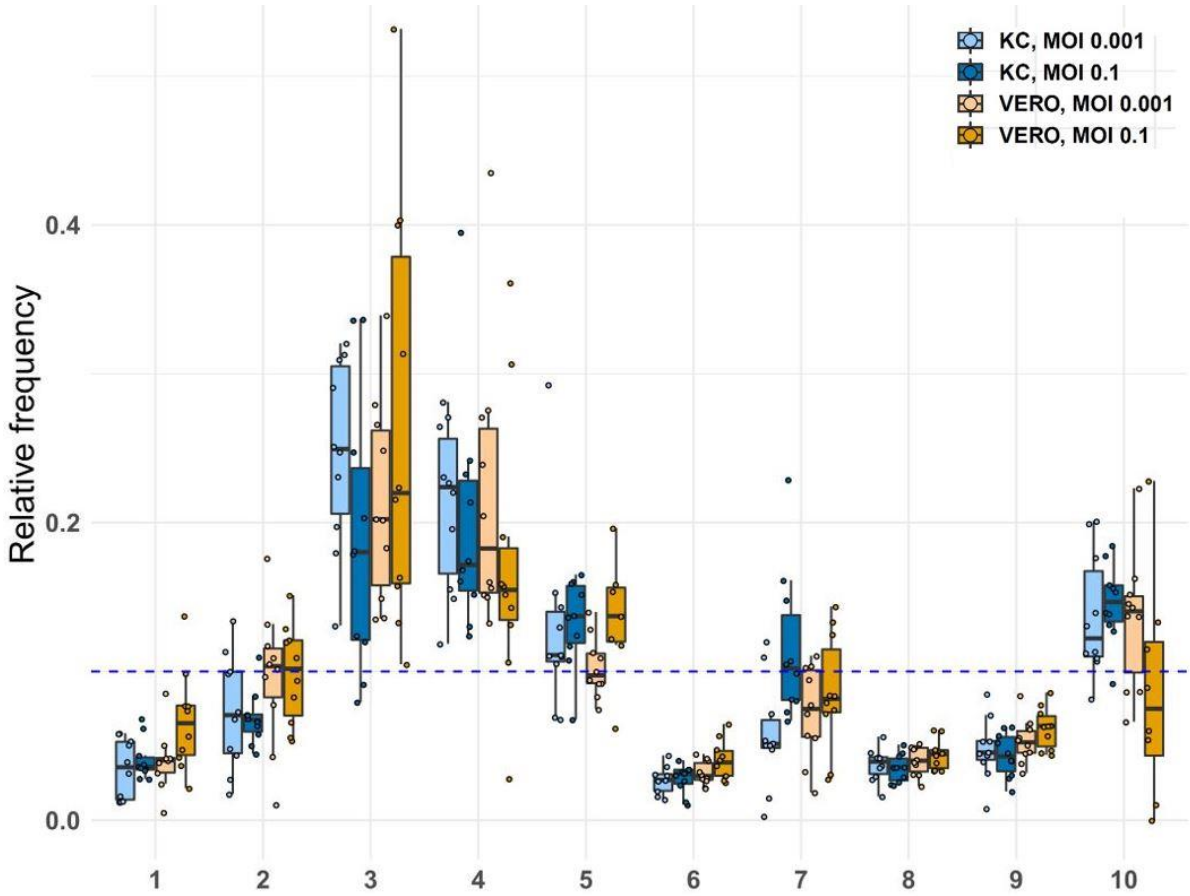


Figure S7. Influence de la MOI sur les fréquences relatives des segments génomiques du virus de la FCO-4 en fonction du type cellulaire. La ligne en pointillé bleue indique la fréquence théorique attendue (0,1) pour les segments s'ils étaient tous présents en quantité équimolaire. Cette figure présente les fréquences des segments quantifiés après 48 heures d'infection sur des cellules KC et VERO infectées à deux MOI inférieures à 1. Chaque condition comprenait 10 réplicas biologiques. La souche virale utilisée a été isolée sur KC puis amplifiée une fois sur cellule VERO. Les couples de primers utilisés et les conditions de réaction de RT-qPCR ne sont pas identiques à ceux utilisés pour l'approche présentée dans ce chapitre.

DISCUSSION GENERALE

Le cycle caractéristique des arbovirus implique une alternance obligatoire entre un vecteur et un hôte végétal ou animal. L'adaptation à ces cycles s'est produite pour des espèces virales de familles différentes représentant une variété d'architectures génomiques, et une large gamme d'hôtes (1). Si l'infection conduit à une prolifération du virus au sein de l'hôte, elle peut ultimement mener à sa transmission lors de l'interaction entre hôte et vecteur, qui est vitale pour ce dernier. Pour le virus, ce cycle implique l'entrée, la réplication et la sortie dans des cellules d'espèces avec une génétique et, donc, un métabolisme différent. La capacité à se multiplier dans des hôtes très distants est d'autant plus surprenante que le génome des arbovirus a une taille réduite et ne contient généralement qu'un faible nombre de gènes (1). De plus, les évènements de transmission et la dissémination du virus au cours de l'infection d'un hôte peuvent conduire à des réductions drastiques de diversité génétique, au moins temporaire, sous l'effet de goulot d'étranglement (2-4). L'alternance d'hôtes implique également un régime de pression de sélection variable qui peut se traduire par un effet de sélection purifiante pour les arbovirus (5-7). Ainsi, les séquences consensus des génomes d'arbovirus semblent assez conservé dans la nature (8).

Pour maintenir ce cycle en alternance, les arbovirus animaux ont probablement développé des adaptations qui permettent une expression génique adaptée aux différents hôtes (9). Nous avons proposé deux hypothèses pour expliquer le maintien de ces cycles associés à l'alternance d'hôte. Plus spécifiquement, ce projet visait l'étude de deux mécanismes qui pourraient potentiellement faciliter l'alternance d'hôte chez le virus de la fièvre catarrhale ovine (FCO) :

i) La variation du nombre de copies de gènes (VNCG) au niveau de la population virale pourrait agir comme un système de régulation de l'expression des gènes. Cette hypothèse fait écho aux découvertes récentes du rôle potentiel de la VNCG chez un nanovirus multipartite de plante, le *Faba bean necrotic stun virus* (FBNSV). Chez FBNSV, la fréquence des segments génomiques varie en fonction des hôtes plantes (10). Aucun mécanisme de ce genre n'avait été décrit chez un arbovirus

avec une gamme d'hôtes animaux. Le but est de déterminer si un tel mécanisme existe et s'il joue un rôle dans l'adaptation du virus de la FCO-4 à l'alternance d'hôte.

ii) Nous voulions déterminer si une adaptation à l'hôte après chaque évènement de transmission par génération de mutations adaptatives était possible. Les virus à ARN, comme la plupart des arbovirus, génèrent des erreurs de réplication et un taux de réplication élevé qui favorisent l'apparition de variations génétiques (11). Il est possible que les variations de pression de sélection favorisent certains génotypes viraux en fonction de l'hôte. Nous souhaitions déterminer si des mutations spécifiques à l'hôte peuvent être générées au cours du cycle d'infection et le cas échéant si ces génotypes peuvent se rétablir à chaque étape du cycle**.** Cette dernière hypothèse n'a malheureusement été approchée que tardivement au cours du projet et les résultats associés ne sont pas présenté dans ce manuscrit.

Nous avons choisi le virus de la FCO, un arbovirus infectant des ruminants et des moucherons culicoïdes, comme modèle pour explorer ces hypothèses. Ce virus possède un génome ARN double brin divisé en dix segments génomiques. Ces segments de taille variable (de 0,8 à 3,95 kb), codent, pour sept d'entre eux, une protéine unique qui a une fonction structurelle ou fonctionnelle pour le (12). Le virus de la FCO, avec son fort niveau de segmentation, constitue un modèle idéal à l'étude du phénomène de variation du nombre de copies de gènes. L'évolution vers une forme génomique segmentée pourrait constituer un avantage sélectif en permettant une régulation de l'expression des gènes en fonction du nombre de copies (10, 13). Dans ce contexte nous nous sommes, en premier lieu, intéressés à l'existence de ce phénomène et à l'importance qu'il pouvait revêtir en fonction des hôtes impliqués dans le cycle d'infection du virus de la FCO-4.

1. La variation du nombre de copies des segments chez le virus de la FCO-4, un rôle potentiel dans l'adaptation au cycle en alternance ?

Nous avons étudié le phénomène de VNCG dans des populations naturelles du virus de la FCO-4 circulant sur les îles de Corse et de Sardaigne en 2017. L'étude de populations naturelles présente l'avantage d'éliminer les biais relatifs à l'isolement et au passage du virus en culture cellulaire. Il permet aussi de rendre compte d'un état

naturel du virus ; les infections sont d'origine vectorielles et les effets sélectifs correspondent à la réalité biologique du cycle en alternance. Par contre, la durée des infections et la valeur sélective des souches étudiées sont inconnus. La nature des échantillons donc limite les interprétations liées à la cinétique de l'infection ou au caractère adaptatif de nos observations.

Nous avons démontré que les segments génomiques ne sont pas présents en quantité égale à l'échelle des populations virales étudiées. Nos résultats montrent que quatre segments dévient significativement de l'équimolarité pour les trois hôtes. Certains segments, en revanche, dévient de l'état d'équilibre en fonction des hôtes (Chapitre 1). Les profils de fréquence, ou formules génomiques (FGs), que nous avons observé sont différents entre certains hôtes. Plus précisément, les hôtes ruminants, plus proches d'un point de vue phylogénétique, présentent des FGs qui sont très proches. Au contraire, la FG du virus présent dans le vecteur arthropode est différente de celle des deux hôtes ruminants. Ces différences soutiennent un lien entre la VNCG et l'adaptation à l'alternance d'hôtes chez le virus de la FCO-4.

La VNCG est un processus qui existe dans la nature, son lien avec la régulation de l'expression est établi dans des systèmes non-viraux (14). Chez les virus, l'hypothèse du rôle du nombre de copies dans la régulation de l'expression des gènes a été initialement proposé par Sicard *et al.* (10). Dans le cadre de cette hypothèse, les résultats obtenus pourraient souligner l'importance de certains gènes en fonction des hôtes. Le segment 2 code pour la protéine de la capside externe à l'origine de la principale réponse immunitaire chez les mammifères (15). La souche de la FCO-4 est particulièrement pathogène chez les ovins ; chez cette espèce, le segment 2 est surreprésenté. Le segment 7, à l'opposé, est plus présent chez les culicoïdes. Il code pour la protéine externe de la particule core (dépourvue des protéines de la capside externe VP2 et VP5) : VP7, qui agit comme principal ligand au récepteur cellulaire, chez les culicoïdes (16). Chez le vecteur, l'infectivité de la particule virale est augmenté en l'absence de la protéine VP2. L 'infection par des particules core est possible, avec une infectivité équivalente à celle de la particule virale complète dans les cellules insectes. En revanche, les particules core ont une infectivité extrêmement faible dans les cellules mammifères (17). Le segment 10 présente la plus forte variation entre hôtes mammifères et insectes. Chez les premiers, son rôle est principalement associé à la modulation de la réponse Interféron de type I (18). Pour tous les hôtes, les

protéines NS3 et NS3a, codées par le segment 10, sont associées aux voies de sortie par perméabilisation de la membrane cellulaire ou par l'intermédiaire de lysosomes produits au niveau du réticulum endoplasmique (9, 19). Ces voies sont actives chez les mammifères au début de l'infection de la cellule et semblent être les seules voies de sortie existantes dans les cellules insectes.

Nous avons voulu tester un possible lien entre la VNCG et l'expression des gènes viraux. Notre approche consistait à quantifier les brins ARN positifs du BTV-4 pour comparer le nombre de copies génomiques aux nombres de copies des ARN messagers dans les mêmes échantillons. Malheureusement, notre méthode de quantification s'est avérée moins robuste pour la quantification des brins positifs. Nous avons rencontré des problèmes de spécificité lors du contrôle des courbes de dissociation pour la moitié des segments. Des problèmes similaires avaient prolongé le développement de notre première approche. De plus, le nombre de copies des brins positifs était inférieur à celui des brins négatifs chez les segments pour lesquels la spécificité était valide. La conservation de ces échantillons à moyen terme ou leur manipulation occasionnelle a pu mener à cette diminution de la concentration en ARN viraux.

Nos échantillons de terrain ne nous permettaient pas de caractériser la convergence vers une FG spécifique pour chaque hôte. L'effet des FGs sur la valeur sélective des populations virales n'a pas pu être étudié, non plus. Pour répondre à ces questions nous avons profité d'un dispositif expérimental mis en place à l'Istituto Zooprofilattico Sperimentale de Teramo (20) (Chapitre 2). L'infection par voie orale de culicoïdes de l'espèce *Culicoides imicola* avec le virus de la FCO-4 permettait de reproduire les infections du vecteur en condition contrôlées. Ces infections expérimentales permettaient donc de suivre la variation potentielle de formule génomique entre l'inocula et la population virale au terme de l'infection. De plus, l'analyse de la présence du virus dans la tête nous a permis de différencier les populations virales selon leur dissémination dix jours après inoculation. Les échantillons ont été classé en deux catégories : dissémination complète ou dissémination incomplète. Au-delà de fournir des informations sur la vitesse de dissémination, la colonisation de la tête est considérée comme un indice de la capacité du virus à infecter les glandes salivaires et donc à être transmis (20, 21). Nous avons donc considéré la colonisation de la tête

comme un proxy de la valeur sélective de la population virale. Nous avons comparé les deux niveaux de dissémination à l'échelle du corps des individus.

La FG changeait au cours de l'infection et en fonction du niveau de dissémination du virus. Dans nos conditions expérimentales, une distance plus importante à la FG de l'inocula est corrélée à la dissémination complète. Cependant, la charge virale était équivalente entre les deux niveaux de dissémination. Le virus pourrait avoir atteint son niveau d'accumulation maximal dans les deux conditions (21-23). L'hypothèse de la VNCG comme mécanisme adaptatif favorisant l'alternance d'hôtes n'est donc pas rejetée par nos observations. Malheureusement, nos observations, à elles seules, ne permettent pas d'établir un lien entre la FG et la valeur sélective. Pour déterminer ce lien, il faudrait développer un système permettant de reproduire les FGs décrites pour les deux états de dissémination. Le développement de méthode permettant l'infection par des virions possédant les FGs décrites ne semble pas possible à mettre en œuvre. Par contre, l'étude au niveau cellulaire par transfection des ARNm aux ratios correspondants pourrait permettre d'établir si ces différences mènent à une infection plus rapide.

Les variations de fréquences que nous décrivons pourraient avoir une implication importante sur le succès de la transmission du virus à l'hôte vertébré, notamment si des goulots d'étranglement sévères ont lieu lors de la (2, 24). En effet, des FGs divergentes de l'équimolarité pourraient impliquer la perte d'information génétique lors des goulots d'étranglement pendant la transmission. Cette situation soulève la question de la conservation de la FG dans les glandes salivaires. Malheureusement, les charges virales trop limitées dans la tête n'ont pas permis une estimation fiable de la FG dans ce compartiment. Chez le culicoïde, les différents organes sont compartimentés (22) et leur infection pourrait induire des effets de goulots d'étranglement. Il est donc possible que la FG soit différentes entre les compartiments ou entre les organes pour un même individu. Par exemple, chez le virus de la fièvre du Nil occidental, un arbovirus monopartite de la famille des *Flaviviridae*, la diversité génétique des populations virales change au cours de l'infection entre les organes du vecteur sous des effets répétés de goulot d'étranglement (4). Ainsi la FG décrite dans le corps des individus lorsque la dissémination est complète ne correspondrait pas nécessairement à la FG dans les glandes salivaires. Une alternative à l'existence de

FGs différentes dans les glandes salivaires est l'existence des mécanismes putatifs favorisant la co-infection au moment de la transmission. Par exemple la salive des moustiques contient différentes molécules qui peuvent affecter des facteurs hôtes comme la contraction des capillaires sanguins, l'inflammation des tissus, la coagulation et les défenses immunitaires (25). Certaines de ces molécules peuvent avoir un effet direct en augmentant l'infectivité des arbovirus, comme cela a été démontré pour le virus du Nil occidental (26) et le virus de la Dengue (27). Pour le virus de la FCO aussi, la salive augmente l'infectivité, en participant à l'inhibition des voies de défense de l'hôte mouton (28). Nous pensons qu'il serait important, aussi, de déterminer le rôle que pourraient jouer certaines de ces molécules dans l'agrégation de particules virales. Chez le virus de la FCO, un autre mécanisme qui pourrait favoriser la co-infection serait sa transmission sous la forme de vésicule extracellulaire rassemblant plusieurs particules virales (29).

Ce premier ensemble de résultats établit la preuve que la VNCG peut exister chez des populations naturelles d'un virus à génome segmenté. Il reste encore à élucider si le phénomène de VNCG facilite l'adaptation du virus de la FCO à travers la régulation de l'expression. Nous avons également établi une corrélation entre la FG et la valeur sélective de la population virale. La VNCG pourrait jouer un rôle dans la capacité de dissémination du virus de la FCO-4 au sein de l'hôte arthropode par l'intermédiaire de mécanismes qui restent à caractériser.

2. Génération de la VNCG

Nous avons testé l'effet de la taille de population initiale et du type cellulaire sur la VNCG pour la génération de virions formés après un seul cycle d'infection (i.e. 12 heures d'infection). Nous avions deux scénarios pour expliquer la génération de la VNCG. i) La FG pourrait être généré aléatoirement puis sélectionnée pour son état optimal en fonction de l'hôte. ii) La FG pourrait aussi être déterminée par des facteurs génétiques et moléculaires liés au virus et/ou à l'hôte. Nos résultats montrent qu'à partir de la même population virale, la VNCG mène à des profils différents au sein de chaque type cellulaire au cours du premier cycle de réplication. Ces résultats suggèrent un effet déterministe de génération du profil de FG qui serait alors lié à des facteurs hôte ou à la génétique virale.

Quel est donc le mécanisme moléculaire derrière la génération de FGs spécifiques à chaque hôte ? La perte d'un segment lors de l'encapsidation, la duplication des gènes par concaténation (30), l'encapsidation de segments surnuméraires (31) ou la présence de délétions (30, 32) pourraient expliquer les variations que nous observons entre les différents segments. Cependant, ces phénomènes n'ont pas été observés dans des populations naturelles du virus de la FCO. Récemment, la capacité d'encapsidation limite des *Reoviridae* a été questionnée. Certains membres de la famille des *Reoviridae* peuvent encapsider jusqu'à 12 segments. Pour le virus de la FCO, les sites de liaisons pour le complexe de transcription, au sein de la pro-capside virale (VP3 et VP7), sont disponibles et pourraient théoriquement permettre à deux segments surnuméraires d'être encapsidés. De plus, chez les *Reoviridae,* les complexes de transcription associés aux segments ne sont pas spécifiques (33). Une évaluation récente de l'espace disponible au sein des particules virales soutien la possibilité en théorie pour l'encapsidation de segments surnuméraires. Chez le virus de la FCO et son proche cousin EHDV, la formation de concatémères de segment peut se produire (30) résultant en une encapsidation complète ou partielle d'un même segment en dupliqua. Finalement comme pour la majorité des espèces virales connues, des délétions dans les parties centrales des segments peuvent se produire (30, 32) et être favorisées dans le contexte de forte concentration des particules virales notamment. Les particules virales produites sont alors qualifiées de défectives. La présence de ces particules a été décrite pour la première fois dans les années 70 chez les *Reoviridae* (22) et plus récemment chez les sérotypes 1, 8 et 17 (24, 47). Les VNCG mesurées pourraient associées de manière non-exclusives à cet ensemble de phénomènes pour le virus de la FCO-4.

L'encapsidation du virus de la FCO implique un système d'interaction entre segments qui permet l'appariement entre eux dans un ordre spécifique (34-36). Une vision consensuelle prévoie que chaque particule virale contient une copie de chaque segment (12). Néanmoins, le mécanisme d'encapsidation et, notamment les étapes menant à l'encapsidation de ces différents segments au sein de la pro-capside virale ne sont pas connus.

Le niveau de segmentation est généralement associé à la spécificité des mécanismes d'encapsidation. Il est moins coûteux pour une population avec un faible niveau de segmentation (trois segments ou moins) de posséder un processus d'encapsidation flexible, voire non spécifique : la probabilité d'infection d'une cellule par l'ensemble du génome permettant la génération de particules virales complètes est plus importante. Par exemple l'encapsidation des trois segments génomiques pourrait ne pas être spécifique pour le virus de la Fièvre de la vallée du Rift (RVFV), un arbovirus, de l'Ordre des *Bunyavirales* (37). Des résultats en cours de publication démontrent que le nombre de particules virales possédant un génome complet varie en fonction du type cellulaire, mammifère ou insecte, chez le RVFV et le virus de Schmallenberg (appartenant aussi à l'Ordre des *Bunyavirales*) (38). Pour le type cellulaire insecte, le nombre de particules virales complètes est plus important. A l'opposé, dans la famille des *Cystoviridae*, pour les Bactériophages λ 6 qui possèdent un niveau de segmentation équivalent, l'encapsidation des trois segments génomiques est régulée par un système sélectif de reconnaissance par les pro-capsides virales qui changent consécutivement de conformation suite aux liaisons successives avec les segments viraux (39). Le niveau de segmentation n'est donc pas le seul critère régissant la spécificité de l'encapsidation et le cycle biologique du virus pourrait participer à la sélection du mécanisme d'encapsidation. Pour les virus à fort degré de segmentation des familles *Orthomyxoviridae* et *Reoviridae,* les mécanismes d'encapsidation ne sont pas entièrement décrits. Pour les deux familles, des motifs conservés aux extrémités des segments et, dans une moindre mesure, distribués sous forme de motifs conservés dans les séquences génomiques, participent aux interactions nécessaire à la production de particules virales dans des populations clonales étudiées *in vitro* (12, 34-36). Chez les *Reovirus* et notamment chez le virus de la FCO, la présence de tous les segments favorise l'encapsidation dans un système d'assemblage artificiel en présence des protéines du complexe de transcription (VP1, VP4, VP6) et de la particule core (VP3, VP7)) (35). Cependant, l'absence d'un segment se traduit par une efficacité d'encapsidation grandement réduite du segment 6, dans ce cas particulier) (35). Les segments de plus petite taille semblent jouer un rôle central dans le processus d'association entre segments ; le segment 10 réduit l'efficacité d'encapsidation du segment 6 de 90%, à côté, l'absence du segment 1 ou 3 réduit cette efficacité d'environ 50%) (35). L'interaction entre segments est étudiée dans des expériences permettant de faire abstraction des facteurs cellulaires ou viraux. De mon

point de vue, si ces approches démontrent avec une grande précision les interactions possibles au niveau génomique et protéique, elles masquent l'influence potentielle de plusieurs autres mécanismes qui pourraient participer à la dynamique de ces interactions (40). En travaillant avec des populations clonales, le rôle de la diversité génétique ou de mutations spécifiques peut être masqué. Il a cependant été démontré le rôle majeur que pouvait avoir une mutation dans la régulation des interactions entre segments pour la souche PR8 (H1N1) chez IAV (41). Pour cette souche, une mutation sur une position unique du gène de la neuraminidase induit une diminution de l'encapsidation du segment codant pour l'hémagglutinine. Chez plusieurs souches du IAV, l'absence de certaines protéines ou segments au cours de l'infection d'une cellule a été décrit (42, 43). Pour la souche H3N2, la probabilité moyenne de réplication d'un segment est de 58 %. Les auteurs ont estimé qu'au cours de l'infection environ 3,6 virions devaient infecter la cellule pour que l'ensemble du génome soit présent. Chez IAV, le ratio de particules infectieuses par rapport au nombre total de particules est dépendant des souches (42, 43), mais des travaux montrent qu'une proportion importante de cellules infectées, allant jusqu'à 90%, ne produisent pas de particules virales infectieuses (41, 44). Chez le virus de la FCO-4, nous avons estimé le ratio de particule infectieuses / particules portant dix copies ARN comme étant inférieur à 10% pour une souche sauvage, à partir d'une infection en culture cellulaire. Nos estimations sont directement dépendantes de la présupposition que chaque particule virale possède dix copies ARN encapsidé, elles sont donc à considérer avec grande précaution. Aucune étude n'a évalué ce ratio chez le virus de la FCO à notre connaissance.

Les ordres de variations supérieurs à dix que nous décrivons au cours de notre approche pour la condition de MOI forte dans les cellules de vaches, pourrait être causé par la génération de segments avec des délétions, avec des taux différents entre segments. Des éléments défectifs peuvent être générés au cours du cycle d'infection du virus de la FCO en culture cellulaire (32, 44). Les segments portant des délétions pourraient permettre les interactions entre segments requis pour l'encapsidation car leurs extrémités sont conservées (30, 32).

Nos résultats questionnent notre connaissance sur la composition du génome au sein des particules virales du virus de la FCO. Chez le virus de la FCO, les méthodes de

visualisation par FISH ou de quantification à l'échelle du virion ou de la cellule n'ont jamais été appliqué, et pourraientt permettre de décrire l'effet de la VNCG à une échelle plus fine (38, 45).

3. Potentielle influence sur l'évolution du virus de la FCO

Les RdRp virales qui ne possèdent pas de système de correction des erreurs de réplication sont répandues dans la majorité des familles virales ARN (11). L'augmentation du nombre de copies d'un segment dans une population virale pourrait donc augmenter sa diversité génétique et favoriser la sélection de mutants avec une meilleure valeur sélective. Ce phénomène putatif ressemblerait ainsi à l'adaptation par sélection des copies de gènes chez des organismes avec des génomes permettant la duplication de gènes. Par exemple, un mécanisme d'extension et contraction du nombre de copies d'un gène a été décrit pour la souche ΔE3L du poxvirus vaccinal, un virus avec un long génome à ADN double brin (46). Sous l'effet d'une augmentation de la pression de sélection sur le gène K3L impliqué dans les facteurs d'adaptation à la gamme d'hôte, ce dernier augmente son nombre de copies sur le génome au cours du temps. Cette accumulation plus importante du nombre de copies favorise la génération d'une mutation adaptative qui se répand rapidement. La fixation de la mutation adaptative conduit à une diminution du nombre de copies des versions non mutantes du gène (46). Nos résultats montrent en ce sens que pour le virus de la FCO-4, les gènes les plus importants pour les facteurs de virulence (segments 2 et 10) (47) sont présents en quantité plus importante à l'échelle des populations virales. Pour la diversité génétique plus élevée et les pressions de sélection plus importantes qui leurs sont associés (32, 48, 49), nous ne pouvons distinguer la cause de la conséquence. Ces différences suggèrent une influence potentielle de la VNCG sur l'évolution des segments.

D'un point de vue évolutif, une FG divergente de l'équimolarité peut générer un coût associé à la perte des gènes lors des possibles goulots d'étranglement subis au cours du cycle d'infection. Ce coût pourrait être compensé par l'existence de mécanismes adaptatif favorisant la co-infection. Le coût induit par le maintien de segments modifiant le nombre de copies (défectifs, dupliqués...) pourrait être contrebalancé par la capacité de réassortiment du virus de la FCO et par un effet de sélection naturelle sur les FG

adaptatives. Les résultats récents de T. Labadie mettent en avant un mécanisme avec un rôle potentiel dans le maintien de l'intégrité génomique du virus de la FCO. Pour ce virus, il existe une forme d'unité infectieuse enveloppée regroupant plusieurs génomes viraux (29). Au sein de chaque hôte, des profils divergents pourrait être favorisé par l'intermédiaire de ces vésicules extracellulaires. Néanmoins, leur présence, au moment de la transmission, n'est pas démontrée. La complémentation entre génomes viraux est peut-être un prérequis à l'établissement d'une infection par le virus de la FCO. Chez le virus de la FCO, le phénomène de réassortiment témoigne d'une situation de co-infection récurrente entre différents sérotypes (49-51). Il est donc fortement probable que la co-infection se produise régulièrement en condition naturelle d'infection. Au-delà de la co-infection, un mécanisme, récemment décrit, pourrait aussi limiter le coût associé à une FG divergente de l'équimolarité ; le FBNSV peut assembler des particules virales complètes sans posséder l'intégralité de son génome au sein de la cellule infectée. Les ARNm et les protéines virales pourraient circuler entre les cellules adjacentes permettant cette complémentation par un mécanisme baptisé « vie multicellulaire » (45). Son existence chez d'autres forme virale à génome segmenté, comme le virus de la FCO, compenserait, au moins en partie, le coût associé au maintien de l'intégrité génétique lors de la transmission.

4. Limitations potentielles liées aux approches et aux méthodes

L'interprétation de nos quantifications des segments génomiques par RT-qPCR est soumise à des précautions. Pour chaque segment, une section spécifique du génome viral est ciblée. La présence de génomes défectifs, dans la population virale, pourrait mener à une quantification de segments qui ne sont pas entiers, sans possibilité de les dissocier d'une version complète du gène. De la même façon, l'absence d'hybridation d'un primer ne signifie pas, à proprement parlé, l'absence du segment dans son ensemble. La conception des primers à partir de la séquence du même virus de la FCO-4 circulant sur l'île de Corse permet de se soustraire au biais potentiel de différences génétiques qui pourraient fausser nos quantifications. Mais la présence d'éléments défectifs et la position des délétions sur les segments pourraient certainement influencer nos mesures. La solution aurait pu consister à construire des couples de primers supplémentaires pour chaque segment. Cependant, au vu des

difficultés rencontrées pour quantifier certains segments, ce travail aurait demandé un développement incompatible avec les contraintes de temps du projet. Par contre, les différences reproductibles que nous observons entre les hôtes ou les conditions expérimentales, minorent l'impact de cette limitation.

Le cycle de réplication du virus de la FCO prédit que la synthèse du brin génomique complémentaire (ARN négatif) se produit au sein de la capside virale, comme pour la majorité des virus ARNdb (52). Ce mécanisme est supposé limiter une réponse ARN interférant activée par la présence d'ARNdb virale non-encapsidé (53). Cependant, la démonstration formelle de cette réplication au sein de la particule virale n'est pas établie pour le virus de la FCO. Notre méthode se base sur la quantification du brin négatif afin d'associer précisément le nombre de copies, à la présence de brins génomiques au sein des particules virales. Pour le virus de la FCO, il existe des méthodes de purification par ultracentrifugation qui permettent de purifier les fractions de virions en fonction de la composition de leur capside (il existe trois formes différentes de particules non enveloppées chez le virus de la FCO). Cette méthode est chronophage et ne permet pas de se soustraire complètement à l'influence des ARN simple brin qui possèdent une affinité pour certaines protéines de la capside. Une alternative aurait pu consister à utiliser une méthode sélective d'extraction, comme l'isolation des ARNdb au lithium chloride après extraction. Encore une fois, j'y opposerai l'argument de temps de réalisation lié à la quantité d'échantillons traités au cours de nos approches. La faible quantité de virus présente dans les échantillons est aussi contraignante pour la réalisation d'étapes de purification supplémentaires. Notre approche théorique visant à quantifier le brin négatif constitue pour moi un élément robuste soutenant la quantification exclusive de brins génomiques encapsidés.

En amont de ce projet un effet potentiel de VNCG chez le sérotype 8 du virus de la FCO avait été observé dans notre groupe. Néanmoins l'ensemble des approches, développées ici, se concentre sur le sérotype 4 du virus de la FCO et le phénomène que nous décrivons pourrait aussi constituer une exception. Certains groupes ont déjà des méthodes de quantification segments-spécifiques qui montrent des variations temporelles de nombre de copies correspondant aux brins génomiques du virus de la FCO (54). Il serait rapide pour eux de tester si le potentiel adaptatif de la VNCG peut être étendu à d'autres sérotypes. J'espère, en ce sens, que nos résultats pousseront

d’autres groupes à considérer l’éventualité du phénomène de la variation du nombre de copie de gènes chez le virus de la FCO.

5. Perspectives

Nos résultats soulèvent beaucoup de nouvelles questions. Notre volonté de tester la possibilité d’existence de la VNCG chez le virus de la FCO et les premiers résultats positifs nous ont conduit à une étude plus avancée de ce phénomène jusqu’alors non considéré chez un arbovirus. L’absence de travaux portant sur ce sujet est aussi expliquée par les contraintes qu’imposerait un tel phénomène pour un virus à fort niveau de segmentation génomique et par l’existence de mécanismes permettant une interaction spécifique entre segments favorisant la production de particules virales théoriquement toutes infectieuses et portant l’ensemble de l’information génétique. Cette vision peut être mise en relief suite aux descriptions d’évènement de duplication génomique, d’existence de génomes défectifs ou de potentielles particules virales pour l’encapsidation de segments surnuméraires (30-32), et bien entendu par nos résultats. Il sera intéressant d’établir dans quelle mesure ces mécanismes influencent le phénomène que nous observons aujourd’hui en étudiant les cellules au niveau individuelle afin d’établir i) la proportion de particules virales infectieuses, ii) la probabilité de réplication de chaque segment au cours de l’infection, et iii) le rôle de la complémentation virale dans le succès de l’infection. Il semble, en tout cas, certain que le phénomène de VNCG est soumis à plusieurs conditions qui devraient être évaluées, en priorité. La taille des populations virales qui initient l’infection au moment de la transmission me semble être le paramètre le plus critique à déterminer. Le ratio de particules infectieuses sur le nombre de particules totales, constitue un élément central à la caractérisation de la composition génétique des virions, il faudrait étendre son évaluation pour le virus de la FCO. A cet effet, des méthodes de quantifications ciblant les protéines de capsides pourraient évaluer le nombre total de particules virales de manière plus précise. Les résultats récents de T. Labadie (29) fournissent des pistes de réflexion qui méritent d’être explorées dans le cadre de la transmission collective, combinée à une potentielle détermination génétique de la FG qui induirait une diminution drastique du coût lié à la transmission du virus.

Il serait particulièrement intéressant de pouvoir déterminer, sur un même groupe d'échantillons, le nombre de copies génomique, le nombre de transcrits et la quantité de protéines, afin de tester le lien fonctionnel de la VNCG avec l'expression chez le virus de la FCO. Le caractère adaptatif de la VNCG et notamment les profils de FG que nous avons décrit, n'a pas été suffisamment exploré. Il serait nécessaire de mettre en place un système pour tester l'influence des FGs sur la valeur sélective des populations virales. Par exemple, les FGs observées dans le corps des culicoïdes en fonction de l'état de dissémination, mènent-elles à des phénotypes d'infections reproductibles ?

Enfin, nous questionnons le phénomène de VNCG et son extension à l'ensemble des sérotypes du virus de la FCO. Peut-il être dépendant des espèces impliquées dans le système, comme le laissent à penser les différences de diversité génétique et de taux de réassortiment variable entre sérotypes en fonction de la géographie (49, 50) ? Ce phénomène de variation du nombre de copies de gênes est-il associé aux différences observées entre cellules insectes et mammifères pour le virus de la fièvre de la vallée du Rift ou d'autres virus à génome segmenté ? Plus largement, le rôle des génomes défectifs au sein des populations virales est questionné, y compris chez les virus monopartites où les délétions peuvent être bénéfiques (55, 56)…

CONCLUSION GENERALE

Le virus de la FCO-4 avec son génome divisé en dix segments possède le potentiel de s'adapter au cycle en alternance d'hôtes sous l'effet de variations de fréquences de ses segments. Les variations, que nous décrivons entre segments, montrent qu'il existe des profils ou formules génomiques qui dépendent, non seulement, des espèces hôtes, mais aussi de la capacité du virus à se disséminer pour être transmissible. Des différences de FGs sont observées entre les espèces hôtes mammifères (*Ovis aries* et *Bos taurus*) et l'espèce vectrice (*Culicoides imicola*). Chez l'espèce vectrice, les FGs des virus présents dans le corps changent en fonction de la capacité du virus à se disséminer jusqu'à la tête. Les segments impliqués dans ces variations, qu'elles soient dépendantes de l'hôte ou de la capacité de dissémination virale, présentent des fonctions particulières en fonction de l'hôte et pourraient être

associé au succès de l'infection. Plusieurs études ont montré qu'il existe des variations de la diversité génétique entre les segments du virus de la FCO. Ces différences peuvent varier en fonction des sérotypes. Néanmoins, pour les segments invariablement décrit comme possédant une diversité génétique plus importante (segments 2 et 10), nos résultats montrent que leur fréquence relative est plus importante en condition naturelle d'infection. Comme cela avait été suggéré pour les virus multipartites, nous supposons que ces variations pourraient être associées à la régulation de l'expression des gènes viraux. Cependant des études complémentaires sont encore nécessaires pour la validation de cette hypothèse. Les populations virales de la FCO-4 peuvent générer des FGs différentes en fonction du type cellulaire insecte (KC) ou mammifère (MDBK), mais aussi en fonction de complémentation entre génomes ou non. La génération de la FG pourrait être déterminé par des facteurs génétiques d'origine virale ou hôte. Pour le maintien du cycle en alternance, la taille de population au moment de la transmission est un facteur clef. L'étude de mécanismes favorisant la co-infection et la probabilité de complémentation en conditions naturelles d'infection sont donc nécessaires pour comprendre comment ce phénomène pourrait se maintenir dans le contexte d'un cycle en alternance d'hôte. Nos résultats suggèrent un rôle de la VNCG dans l'adaptation à l'alternance d'hôte chez le virus de la FCO-4. Nos résultats invitent à l'étude de ce phénomène dans d'autres sérotypes du virus de la FCO, et soutiennent son rôle potentiel dans l'adaptation aux changements d'hôtes dans d'autres espèces virales.

REFERENCES

1. Hubalek Z, Rudolf I, Nowotny N. 2014. Arboviruses pathogenic for domestic and wild animals. Adv Virus Res 89:201-75.
2. Gutierrez S, Michalakis Y, Blanc S. 2012. Virus population bottlenecks during within-host progression and host-to-host transmission. Curr Opin Virol 2:546-55.
3. Grubaugh ND, Fauver JR, Ruckert C, Weger-Lucarelli J, Garcia-Luna S, Murrieta RA, Gendernalik A, Smith DR, Brackney DE, Ebel GD. 2017. Mosquitoes Transmit Unique West Nile Virus Populations during Each Feeding Episode. Cell Rep 19:709-718.
4. Patterson EI, Khanipov K, Rojas MM, Kautz TF, Rockx-Brouwer D, Golovko G, Albayrak L, Fofanov Y, Forrester NL. 2018. Mosquito bottlenecks alter viral mutant swarm in a tissue and time-dependent manner with contraction and expansion of variant positions and diversity. Virus Evol 4:vey001.
5. Grubaugh ND, Smith DR, Brackney DE, Bosco-Lauth AM, Fauver JR, Campbell CL, Felix TA, Romo H, Duggal NK, Dietrich EA, Eike T, Beane JE, Bowen RA, Black WC, Brault AC, Ebel GD. 2015. Experimental evolution of an RNA virus in wild birds: evidence for host-dependent impacts on population structure and competitive fitness. PLoS Pathog 11:e1004874.
6. Coffey LL, Vignuzzi M. 2011. Host alternation of chikungunya virus increases fitness while restricting population diversity and adaptability to novel selective pressures. J Virol 85:1025-35.
7. Caporale M, Di Gialleonorado L, Janowicz A, Wilkie G, Shaw A, Savini G, Van Rijn PA, Mertens P, Di Ventura M, Palmarini M. 2014. Virus and host factors affecting the clinical outcome of bluetongue virus infection. J Virol 88:10399-411.
8. Ciota AT, Kramer LD. 2010. Insights into arbovirus evolution and adaptation from experimental studies. Viruses 2:2594-617.
9. Alex D.Hyatt YZaPR. 1993. Release of Bluetongue Virus-like Particles from Insect Cells is Mediated by BTV Nonstructural Protein NS3/NS3A. Virology.
10. Sicard A, Yvon M, Timchenko T, Gronenborn B, Michalakis Y, Gutierrez S, Blanc S. 2013. Gene copy number is differentially regulated in a multipartite virus. Nature Communications 4:1-8.
11. Elena SF, Sanjuan R. 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J Virol 79:11555-8.
12. Roy P. 2017. Bluetongue virus structure and assembly. Curr Opin Virol 24:115-123.
13. Holmes EC. 2009. The Evolution and Emergence of RNA Viruses. Oxford University Press: New York, NY, USA.
14. Yuriy Mileykoa RIJ, and Joshua S. Weitza. 2008. Small-scale copy number variation and large-scale changes in gene expression. PNAS.
15. P. Roy TU, A. A. Van Dijk and B. J. Erasmus. 1990. Recombinant Virus Vaccine for Bluetongue Disease in Sheep. Journal of virology.
16. Tan BH, Nason E, Staeuber N, Jiang W, Monastryrskaya K, Roy P. 2001. RGD tripeptide of bluetongue virus VP7 protein is responsible for core attachment to Culicoides cells. J Virol 75:3937-47.

17. P. P. C. Mertens JNB, A. Walton, M. P. Wellby, H. Fu, R. S. O'Hara, S. M. Brookes, and P. S. Mellor. 1996. Enhanced Infectivity of Modified Bluetongue Virus Particles for Two Insect Cell Lines and for Two Culicoides Vector Species. Virology.
18. Vitour D, Doceul V, Ruscanu S, Chauveau E, Schwartz-Cornil I, Zientara S. 2014. Induction and control of the type I interferon pathway by Bluetongue virus. Virus Res 182:59-70.
19. Han Z, Harty RN. 2004. The NS3 protein of bluetongue virus exhibits viroporin-like properties. J Biol Chem 279:43092-7.
20. Federici V, Goffredo M, Mancini G, Quaglia M, Santilli A, Di Nicola F, De Ascentis M, Cabras P, Volpicelli C, De Liberato C, Satta G, Federico G, Leone A, Pisciella M, Portanti O, Pizzurro F, Teodori L, Savini G. 2019. Vector Competence of Italian Populations of Culicoides for Some Bluetongue Virus Strains Responsible for Recent Northern African and European Outbreaks. Viruses 11.
21. Fu H, Leake CJ, Mertens PPC, Mellor PS. 1999. The barriers to bluetongue virus infection, dissemination and transmission in the vector, Culicoides variipennis (Diptera: Ceratopogonidae). Archives of Virology 144:747-761.
22. Mellor PS, Carpenter S, White DM. 2009. Chapter 14 - Bluetongue virus in the insect host, p 295-320. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50018-6. Academic Press, London.
23. Chandler Lj Fau - Ballinger ME, Ballinger Me Fau - Jones RH, Jones Rh Fau - Beaty BJ, Beaty BJ. The virogenesis of bluetongue virus in Culicoides variipennis.
24. Grubaugh ND, Ruckert C, Armstrong PM, Bransfield A, Anderson JF, Ebel GD, Brackney DE. 2016. Transmission bottlenecks and RNAi collectively influence tick-borne flavivirus evolution. Virus Evol 2:vew033.
25. Calvo E, Sanchez-Vargas I, Favreau AJ, Barbian KD, Pham VM, Olson KE, Ribeiro JM. 2010. An insight into the sialotranscriptome of the West Nile mosquito vector, Culex tarsalis. BMC Genomics 11:51.
26. Colpitts TM, Conway MJ, Montgomery RR, Fikrig E. 2012. West Nile Virus: biology, transmission, and human infection. Clin Microbiol Rev 25:635-48.
27. Conway MJ, Watson AM, Colpitts TM, Dragovic SM, Li Z, Wang P, Feitosa F, Shepherd DT, Ryman KD, Klimstra WB, Anderson JF, Fikrig E. 2014. Mosquito saliva serine protease enhances dissemination of dengue virus into the mammalian host. J Virol 88:164-75.
28. Pages N, Breard E, Urien C, Talavera S, Viarouge C, Lorca-Oro C, Jouneau L, Charley B, Zientara S, Bensaid A, Solanes D, Pujols J, Schwartz-Cornil I. 2014. Culicoides midge bites modulate the host response and impact on bluetongue virus infection in sheep. PLoS One 9:e83683.
29. Labadie T, Roy P. 2020. A non-enveloped arbovirus released in lysosome-derived extracellular vesicles induces super-infection exclusion. doi:10.1101/2020.06.11.146357.
30. Anthony SJ, Darpel KE, Belaganahalli MN, Maan N, Nomikou K, Sutton G, Attoui H, Maan S, Mertens PP. 2011. RNA segment 9 exists as a duplex concatemer in an Australian strain of epizootic haemorrhagic disease virus (EHDV): Genetic analysis and evidence for the presence of concatemers as a normal feature of orbivirus replication. Virology 420:164-71.

31. Desselberger U. 2019. What are the limits of the packaging capacity for genomic RNA in the cores of rotaviruses and of other members of the Reoviridae? Virus Res 276:197822.
32. Kopanke JH, Lee JS, Stenglein MD, Mayo CE. 2020. The Genetic Diversification of a Single Bluetongue Virus Strain Using an In Vitro Model of Alternating-Host Transmission. Viruses 12.
33. Noad R, Roy P. 2009. Chapter 4 - Bluetongue virus replication and assembly, p 53-76. *In* Mellor PS, Baylis M, Mertens PPC (ed), Bluetongue doi:https://doi.org/10.1016/B978-012369368-6.50008-3. Academic Press, London.
34. Burkhardt C, Sung PY, Celma CC, Roy P. 2014. Structural constraints in the packaging of bluetongue virus genomic segments. J Gen Virol 95:2240-2250.
35. Sung PY, Roy P. 2014. Sequential packaging of RNA genomic segments during the assembly of Bluetongue virus. Nucleic Acids Res 42:13824-38.
36. Boyce M, McCrae MA, Boyce P, Kim JT. 2016. Inter-segment complementarity in orbiviruses: a driver for co-ordinated genome packaging in the Reoviridae? J Gen Virol 97:1145-1157.
37. Wichgers Schreur PJ, Kortekaas J. 2016. Single-Molecule FISH Reveals Non-selective Packaging of Rift Valley Fever Virus Genome Segments. PLoS Pathog 12:e1005800.
38. Erick Bermúdez-Méndez EAK, Cindy M. Spruit, Jeroen Kortekaas, Paul J. Wichgers Schreur. 2020. Visualizing the RNP content of single bunyavirus virions reveals more 2 efficient genome packaging in the arthropod host. Nature Communications _ Preprint.
39. Xueying Qiao GC, Jian Qiao, and Leonard Mindich. 1995. In Vitro Packaging of Individual Genomic Segments of Bacteriophage □6 RNA: Serial Dependence Relationships

Journal of virology.
40. Mohl BP, Kerviel A, Labadie T, Matsuo E, Roy P. 2020. Differential Localization of Structural and Non-Structural Proteins during the Bluetongue Virus Replication Cycle. Viruses 12.
41. Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. 2014. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. Proc Natl Acad Sci U S A 111:16854-9.
42. Brooke CB, Ince WL, Wrammert J, Ahmed R, Wilson PC, Bennink JR, Yewdell JW. 2013. Most influenza a virions fail to express at least one essential viral protein. J Virol 87:3155-62.
43. Jacobs NT, Onuoha NO, Antia A, Steel J, Antia R, Lowen AC. 2019. Incomplete influenza A virus genomes occur frequently but are readily complemented during localized viral spread. Nat Commun 10:3526.
44. M. Nonoyama YW, and A. F. Graham. 1970. Defective Virions of Reovirus. Journal of Virology.
45. Sicard A, Pirolles E, Gallet R, Vernerey MS, Yvon M, Urbino C, Peterschmitt M, Gutierrez S, Michalakis Y, Blanc S. 2019. A multicellular way of life for a multipartite virus. Elife 8.
46. Elde NC, Child SJ, Eickbush MT, Kitzman JO, Rogers KS, Shendure J, Geballe AP, Malik HS. 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. Cell 150:831-841.

47. Janowicz A, Caporale M, Shaw A, Gulletta S, Di Gialleonardo L, Ratinier M, Palmarini M. 2015. Multiple genome segments determine virulence of bluetongue virus serotype 8. J Virol 89:5238-49.
48. Carpi G, Holmes EC, Kitchen A. 2010. The evolutionary dynamics of bluetongue virus. J Mol Evol 70:583-92.
49. Nomikou K, Hughes J, Wash R, Kellam P, Breard E, Zientara S, Palmarini M, Biek R, Mertens P. 2015. Widespread Reassortment Shapes the Evolution and Epidemiology of Bluetongue Virus following European Invasion. PLoS Pathogens 11:1-23.
50. Jacquot M, Rao PP, Yadav S, Nomikou K, Maan S, Jyothi YK, Reddy N, Putty K, Hemadri D, Singh KP, Maan NS, Hegde NR, Mertens P, Biek R. 2019. Contrasting selective patterns across the segmented genome of bluetongue virus in a global reassortment hotspot. Virus Evolution 5:1-14.
51. Shaw AE, Ratinier M, Nunes SF, Nomikou K, Caporale M, Golder M, Allan K, Hamers C, Hudelet P, Zientara S, Breard E, Mertens P, Palmarini M. 2013. Reassortment between two serologically unrelated bluetongue virus strains is flexible and can involve any genome segment. J Virol 87:543-57.
52. Patton JT, Spencer E. 2000. Genome replication and packaging of segmented double-stranded RNA viruses. Virology 277:217-25.
53. Schnettler E, Ratinier M, Watson M, Shaw AE, McFarlane M, Varela M, Elliott RM, Palmarini M, Kohl A. 2013. RNA interference targets arbovirus replication in Culicoides cells. J Virol 87:2441-54.
54. Pullinger GD, Guimera Busquets M, Nomikou K, Boyce M, Attoui H, Mertens PP. 2016. Identification of the Genome Segments of Bluetongue Virus Serotype 26 (Isolate KUW2010/02) that Restrict Replication in a Culicoides sonorensis Cell Line (KC Cells). PLoS One 11:e0149709.
55. Lopez-Ferber M, Simon O, Williams T, Caballero P. 2003. Defective or effective? Mutualistic interactions between virus genotypes. Proc Biol Sci 270:2249-55.
56. Vignuzzi M, Lopez CB. 2019. Defective viral genomes are key drivers of the virus-host interaction. Nat Microbiol 4:1075-1087.

RESUME

Les arbovirus forment un groupe polyphylétique d'espèces virales. Elles partagent la capacité d'infecter des arthropodes pour être transmise à différentes espèces d'hôte englobant les règnes animaux et végétaux. La convergence de virus aux architectures structurelles et génomiques si différentes vers ce mode de vie implique probablement des adaptations qui favorisent le maintien de ce cycle. Parmi elles, La variation du nombre de copie de gènes (VNCG) est un mécanisme candidat qui a récemment été mise en évidence chez un *nanovirus* de plante. Chez ce virus multipartite - dont les segments génomiques sont encapsidés individuellement - les fréquences de chacun de ces segments convergent vers une valeur spécifique dépendante de l'espèce hôte infectée. Nous questionnons ici le potentiel de la VNCG à jouer un rôle dans l'adaptation d'un arbovirus à génome segmenté (i.e. les segments sont séparés mais encapsidé au sein d'un même particule virale) : le virus de la fièvre catarrhale ovine (FCO). Le virus de la FCO est un virus de la famille des *Reoviridae* dont le génome se divise en dix segment. Nous avons développé une approche de RT-qPCR pour quantifier chaque segment génomique. Cette approche a été utilisée pour analyser des populations du virus de la FCO de sérotype 4 (FCO-4) issues d'un évènement épizootique naturel. Nous avons étudié les variations de fréquence de chacun de ses dix segments chez deux espèces de ruminant et chez le diptère *Culicoides imicola*, un des vecteurs principaux du virus de la FCO. Ces résultats présentent pour la première fois l'existence du phénomène de VNCG dans des populations naturelles d'un virus animal. De plus, des profils de fréquences de segments (formule génomique) dépendent de l'hôte infecté : ils sont similaires entre les hôtes ruminants et diffèrent entre ruminants et diptère. Ces résultats soutiennent un rôle potentiel de la VNCG dans l'adaptation à l'alternance d'hôte chez le virus de la FCO. Une approche d'infection expérimentale sur *C. imicola* nous a ensuite permit de démontrer que la formule génomique change lors de l'infection d'un nouvel hôte. Nous avons mis en évidence que les populations virales ayant une dissémination complète dans l'insecte ont une formule génomique différente de celles à dissémination incomplète. Ce résultat suggère un possible lien entre la VNCG et la valeur sélective de la population virale. Nos expérimentations *in vitro* montrent également que des formules génomiques spécifiques à un environnement donné peuvent être générées du premier cycle de réplication. Ce résultat suggère que la génération de la formule génomique implique des mécanismes déterministes encore inconnus. La VNCG constitue donc un phénomène potentiellement associé à l'adaptation au cycle en alternance du virus de la FCO-4. Son rôle plus large dans l'alternance d'hôte chez le virus de la FCO, voire chez d'autres arbovirus segmentés, mériterait d'être exploré.

Mots-clés : virus de la fièvre catarrhale ovine, variation du nombre de copies de gènes, arbovirus, alternance d'hôtes.

Adaptive mechanisms facilitating host alternation in Bluetongue virus

Arboviruses form a polyphyletic group of viral species. They share the ability to infect arthropods for transmission to different host species encompassing the animal and plant kingdoms. The convergence of viruses with such different structural and genomic architectures towards this way of life probably implies adaptations that favor its maintenance. Among them, variation in gene copy number (GCNV) is a candidate mechanism recently identified in a plant *nanovirus*. In this multipartite virus - whose genomic segments are individually packaged - the frequencies of each of these segments converge towards a specific value depending on the infected host species. We question here the potential of GCNV to play a role in the adaptation of an arbovirus with a segmented genome (i.e. the segments are separated but packaged within the same viral particle): the bluetongue virus (BTV). BTV is a virus of the Reoviridae family, whose genome is divided into ten segments. We have developed an RT-qPCR approach quantifying each genomic segment. This approach has been used to analyze populations of the serotype 4 BTV (BTV-4) resulting from a natural epizootic event. We studied the frequency variations of each of its ten segments in two ruminant species and in the dipteran midge Culicoides imicola, one of the main vectors of the BTV. These results present for the first time the existence of the phenomenon of GCNV in natural populations of an animal virus. In addition, segment frequency profiles (genomic formula) depend on the infected host: they are similar between ruminant hosts and differ between ruminant and diptera. These results support a potential role of GCNV in adaptation to host alternation in BTV. An experimental infection approach with *C. imicola* then allowed us to demonstrate that the genomic formula changes when a new host is infected. We showed that virus populations with complete dissemination in the insect have a different genomic formula than those with incomplete dissemination. This result suggests a possible link between GCNV and the selective value of the virus population. Our in vitro experiments also show that genomic formulas specific to a given environment can be generated from the first replication cycle. This result suggests that the generation of the genomic formula involves yet unknown deterministic mechanisms. GCNV is therefore a phenomenon potentially associated with adaptation to the alternating cycle of the BTV-4 virus. Its broader role in host alternation in BTV virus, and even in other segmented arboviruses, is worth exploring.

Key-words: Bluetongue virus, gene copy number variation, arbovirus, host alternation