

THESIS / THÈSE

MASTER EN BIOCHIMIE ET BIOLOGIE MOLÉCULAIRE ET CELLULAIRE À FINALITÉ APPROFONDIE

Étude des protéines interagissant avec la région 3' non-traduite du virus de Schmallenberg impliquées dans la régulation de la transcription et de la traduction du virus

Bogaerts, Guillaume

Award date:
2022

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Faculté des Sciences

**ÉTUDE DES PROTÉINES INTERAGISSANT AVEC LA RÉGION 3' NON-TRADUITE
DU VIRUS DE SCHMALLEMBERG IMPLIQUÉES DANS LA RÉGULATION DE LA
TRANSCRIPTION ET DE LA TRADUCTION DU VIRUS**

**Mémoire présenté pour l'obtention
du grade académique de master 120 en biochimie et biologie moléculaire et cellulaire**

Guillaume BOGAERTS

Janvier 2022

Université de Namur
FACULTE DES SCIENCES
Secrétariat du Département de Biologie
Rue de Bruxelles 61 - 5000 NAMUR
Téléphone: + 32(0)81.72.44.18 - Téléfax: + 32(0)81.72.44.20
E-mail: joelle.jonet@unamur.be - <http://www.unamur.be>

Étude des protéines interagissant avec la région 3' non-traduite du virus de Schmallenberg impliquées dans la régulation de la transcription et de la traduction du virus

BOGAERTS Guillaume

Résumé

Le virus de Schmallenberg (SBV) appartient au genre *orthobunyavirus* et au séro groupe *Simbu*. Il possède un génome à ARN simple brin tri-segmenté de polarité négative. Les orthobunyavirus ont la particularité de contenir des motifs essentiels pour leur transcription et leur traduction au niveau de leurs régions non-traduites (NTR). En effet, au sein du segment S du génome des simbuvirus, la 3'NTR est caractérisée par la présence d'une structure en tige-boucle hautement conservée, suivie d'un signal de fin de transcription ('GC'). Ensemble, ces deux régions permettent une traduction efficace de la nucléoprotéine dans un contexte poly(A) indépendant. L'établissement de mutants des motifs conservés de la 3'NTR couplé, à une méthode de capture de protéines, suivie d'une analyse en spectrométrie de masse, ont permis l'identification de protéines de liaison à l'ARN (RBP) interagissant avec le segment S de SBV. Cette stratégie vise à caractériser le mécanisme de terminaison de la transcription et d'initiation de la traduction de l'ARN viral. À cette fin, nous avons utilisé la méthode RaPID (RNA-protein interaction detection) pour identifier les partenaires protéiques potentiels qui pourraient jouer des rôles essentiels dans la transcription et la traduction du segment S. Nous avons également validé la viabilité d'un virus recombinant possédant des structures BoxB de part et d'autre des motifs ARN. Ces structures permettraient au virus de lier une biotine ligase et de biotinyler directement les protéines interagissant avec ses motifs dans un contexte d'infection virale.

Mémoire de master 120 en biochimie et biologie moléculaire et cellulaire

Janvier 2022

Promoteur: D. Coupeau

Co-promoteur: P. Renard

Remerciements

La réalisation de ce travail n'aurait pas été possible sans vous. C'est pourquoi je tiens à vous remercier, VOUS TOUS qui m'avez accompagné tout le long de ce chemin parcouru et qui avez contribué à l'élaboration de ce mémoire et à son aboutissement.

Je tiens à remercier Damien, mon promoteur, qui m'a encadré tout au long de ce mémoire et qui m'a fait partager ses brillantes intuitions. Tu as toujours pris le temps de m'expliquer en profondeur la matière abordée et de répondre à mes questions. Tu n'as cessé de me pousser à la réflexion et c'est ce qui me vaut aujourd'hui de me sentir grandi en tant que (futur) scientifique. Je tiens également à te remercier pour ta gentillesse, ta disponibilité permanente et la bienveillance dont tu as fait part durant ces 10 mois.

Je tiens à remercier Laura, doctorante à l'URVI, qui m'a également accompagné pendant ce mémoire. Tu m'auras appris la plupart de mes manipulations en veillant toujours à ce que tout soit effectué en bonne et due forme. Tu auras su trouver les mots justes pour me permettre de m'améliorer petit à petit. Nos moments partagés à discuter de notre projet commun, de protéomique, de médecine vétérinaire, de la vie, de la Colombie, de tout et de rien, ... ont rendu ces 10 mois chaleureux, plaisants et instructifs. Je ne te remercierai jamais assez de m'avoir appris tant de choses avec pédagogie et enthousiasme. Ta bonne humeur est inébranlable et c'est ce qui fait ton charme. J'ai adoré travaillé à tes côtés Laura!

Je tiens à remercier toutes et tous ceux travaillant à l'URVI, qui nous ont accueilli, mes co-mémorants et moi, avec tant de bonté et de générosité. Vous avez fait en sorte que tout se passe toujours au mieux et c'est grâce à vous tous que l'ambiance au laboratoire est excellente. Je n'oublierai pas les fous-rires, les discussions autour de la table (et d'un verre), les soirées de jeux de société, ... Et j'espère revenir très vite vous saluer. Merci à tous pour votre disponibilité sans faille et le partage de vos projets et vos connaissances lorsque j'étais curieux. Merci de m'avoir permis de passer un moment inoubliable pour la fin de mon séjour à l'UNamur.

Je tiens à remercier mes co-mémorants Valérie, Julie et Maxence, sans qui rien n'aurait été pareil. Nous avons partagé ces 10 mois ensemble, dans un même bureau la plupart du temps. Merci d'avoir partagé vos expériences, votre travail, vos idées. Valérie tu as été mon binôme sur cette fin de mémoire, tu as toujours tenté de me pousser à continuer sans lâcher même quand je perdais un peu d'espoir, je te souhaite plein de réussite. Maxence, mon binôme depuis 6 années maintenant, que ce soit en cours, au laboratoire ou dans les activités estudiantines. Merci mille fois pour ce parcours incroyable qu'on a vécu ensemble, tu fais sans l'ombre d'un doute partie de ma réussite en tant que véritable ami et soutien !

Je tiens à remercier mes professeurs de l'UNamur, qui ont su faire naître la passion pour la biologie moléculaire. Plus particulièrement, je tiens à remercier Patsy, ma co-promotrice, qui, dès ma seconde année, m'a fait découvrir avec passion la biochimie, les protéines, et tout ce que m'anime pour la suite. Merci pour vos conseils avisés, pour ces moments d'échanges et

toutes les connaissances partagées. Je tiens également à remercier, Marc, responsable MaSUn de l'UNamur, pour le temps accordé à mon projet et toute son expérience partagée.

Je tiens à remercier ma famille, qui fut d'un soutien sans faille. Ma mère, Sandrine, pour sa dévotion et son implication dans mes études. Mon frère Arnaud, pour son calme, ses avis, et sa manière de relativiser dans le bon sens du terme et qui m'a appris que le bien être est une chose essentielle pour le développement personnel. Mon frère Romain, pour ses conseils plus professionnels, son aide avec l'anglais et ses encouragements. Mon père, Marco, qui a toujours cru en moi et qui a su montrer sa fierté en tant que parent.

Je tiens à remercier tous mes amis et plus particulièrement mes colocataires Cédric, Pierre et Simon. Notre partage de domicile touche à sa fin et quelle tristesse ! Pendant ces 2 années j'ai passé un moment inoubliable avec vous. Vous faites certainement partie de ce chemin parcouru et vous m'avez aidé à devenir ce que je suis aujourd'hui. Cette « coloc » était une pure réussite, je vous aime, et à très bientôt !

À tous, pour ce chemin parcouru, tout ce qui en est sorti, et tout ce qui suivra, merci !

Table des matières

Introduction	7
1. Découverte et épidémiologie de la maladie de Schmallerberg	7
2. Vecteur du SBV	9
3. Signes cliniques de la maladie de Schmallerberg	10
4. Phylogénie du SBV	11
5. Structure et composition des orthobunyavirus.....	13
6. Cycle réplicatif des <i>Bunyavirales</i>	16
7. Transcription des <i>Bunyavirales</i>	18
8. Le segment S du virus de Schmallerberg.....	19
9. Le monde de l'ARN	21
10. Les fonctions associées aux structures de l'ARN	22
11. Utilisation de l'ARN en virologie moléculaire	23
12. Évolution des outils de détection des RBP	24
Matériel et méthodes	28
1. Lignées cellulaires et culture	28
2. Amplification virale et titrage.....	28
3. Transfections	29
Transfection des vecteurs dans des cellules BHK-21	29
Transfection pour l'obtention d'un virus recombinant BoxB par génétique inverse	29
4. Extraction d'ARN.....	31
5. Traitement poly(A) polymérase du virus recombinant BoxB.....	31
6. Purification de l'ARN par phénol : chloroforme	31
7. Reverse transcription (RT) de l'ARN en ADN complémentaire	33
8. Réaction de polymérisation en chaîne	33
Amplification des séquences d'intérêts pour l'identification des protéines.....	33
Amplification pour l'obtention du virus recombinant BoxB	33
Amplification de la 3'NTR du virus recombinant BoxB.....	33
9. Purification des produits de PCR.....	34
10. Digestion des produits de PCR.....	34
11. Ligation des produits de PCR.....	34
12. Électroporation et culture de bactéries.....	35
13. Clonage et criblage	35
14. Amplification et purification d'ADN plasmidique.....	35

15.	Identification des protéines par spectrométrie de masse.....	36
	Extraction et dosage des protéines.....	36
	Pull-down et digestion des protéines biotinylées.....	36
	Purification des protéines sur Pierce C18 Spin Tips.....	37
16.	Analyse des données.....	39
Résultats		41
1.	Sélection des mutants de la 3'NTR du segment S de SBV et construction des plasmides motif-ARN.....	41
2.	Impact de la longueur des constructions sur l'ensemble des protéines identifiées	45
3.	Étude de l'enrichissement des termes de l'ontologie génique (GO)	45
4.	Impact des mutations sur la capture et l'identification des protéines	47
	Analyse comparative des protéines identifiées en conditions non- infectées	47
	Analyse comparative des protéines identifiées en conditions d'infection	50
5.	Construction d'un segment S recombinant BoxB par mutagenèse dirigée	52
6.	Élaboration d'un virus recombinant BoxB par génétique inverse	52
7.	Évaluation de la dose infectieuse médiane du SBV recombinant BoxB.....	54
8.	Étude de la région terminale 3' du segment S du virus recombinant BoxB	54
Conclusion et perspectives		56
Références		62
Figures supplémentaires		69

Introduction

1. Découverte et épidémiologie de la maladie de Schmalleberg

Durant l'été 2011, une maladie touchant les bovidés est apparue dans le nord de l'Allemagne. Au début, cette maladie était caractérisée par de la diarrhée, une diminution de la production laitière ainsi qu'une fièvre modérée dans les élevages (Beer *et al.*, 2013 ; Hoffman *et al.*, 2012). Durant les mois qui ont suivi, plusieurs cas ont été reportés en Belgique, aux Pays-Bas et en Allemagne (**Fig. 1**), inquiétant de plus en plus les éleveurs (Claine *et al.*, 2015 ; Afonso *et al.*, 2014). Les signes cliniques se présentaient comme transitoires et non-spécifiques et pouvaient donc se rapporter à plusieurs pathogènes. Les premières analyses de diagnostic ont rapidement permis d'exclure les agents de maladies circulantes à l'époque comme le virus de la langue bleue, le virus de l'herpès bovin-1, le virus de la fièvre aphteuse et des virus exotiques comme le virus des maladies hémorragiques épizootiques (Afonso *et al.*, 2014 ; Hoffmann *et al.*, 2012). Endéans les trois mois, l'agent étiologique fut caractérisé par l'institut Friedrich-Loeffler à partir d'échantillons de sang de bovins provenant d'une ferme proche de la ville de Schmalleberg (Rhénanie du Nord-Westphalie, Allemagne) (Hoffmann *et al.*, 2012). Le nouveau pathogène a ainsi été déterminé à l'aide d'une approche métagénomique et fut nommé le virus de Schmalleberg (SBV). Les études phylogénétiques ont mis en évidence par homologie de séquences que le SBV appartient au genre *orthobunyavirus* de la famille des *Peribunyaviridae*, faisant du SBV le premier orthobunyavirus détecté en Europe (Beer *et al.*, 2013 ; Doceul *et al.*, 2013 ; Hoffmann *et al.*, 2012). Ces études ont révélé que le SBV est apparenté aux virus Shamonda, Aino et Akabane faisant partie du séro groupe *Simbu* (simbuvirus) (Hughes *et al.*, 2020) ; Yanase *et al.*, 2012. Les simbuvirus sont répertoriés en tant qu'arbovirus (arthropod-borne virus), ce qui signifie qu'ils sont transmis par un vecteur arthropode. En effet, des études ont permis la détection du virus dans des vecteurs insectes du genre *Culicoïdes*, considérés comme les vecteurs principaux du virus (De Regge *et al.*, 2012). En décembre 2011, l'infection par le SBV devint préoccupante car le virus fut détecté chez des agneaux et des veaux malformés, dont certains étaient mort-nés ou mouraient peu de temps après la naissance (van den Brom *et al.*, 2012). Il a ensuite été montré que certains des cas d'avortements induits par l'infection au SBV provenaient des premiers cas d'infection observés durant l'été 2011 et que les femelles gestantes auraient transmis le virus par voie transplacentaire lors de la phase précoce de gestation (Wernike *et al.*, 2015). Début 2012, le SBV a pour la première fois été signalé au Royaume-Uni et en France après la détection du virus dans des échantillons d'agneaux malformés (Afonso *et al.*, 2014 ; Doceul *et al.*, 2013). Par la suite, quelques mois après la découverte des effets tératogènes du SBV, le virus s'est rapidement et largement propagé en Europe. Fin avril 2012, le SBV était détecté dans 3628 troupeaux en Europe (Doceul *et al.*, 2013).

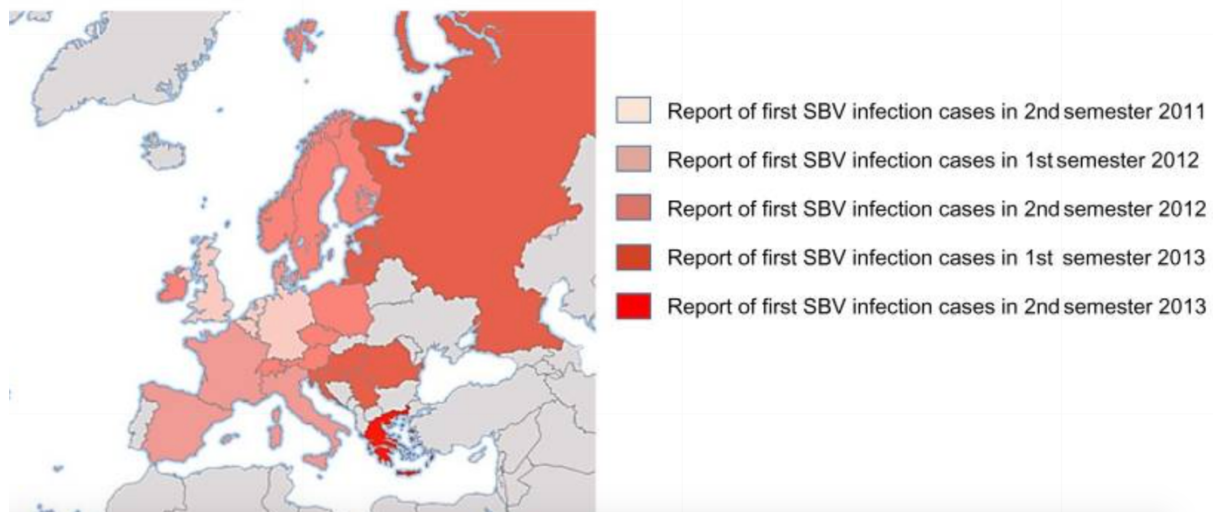


Figure 1: Carte géographique montrant la répartition du virus de Schmallenberg en Europe entre 2011 et 2013 (Claine *et al.*, 2015).



Figure 2: Photo de vecteurs arthropodes du genre *Culicoides* responsables de la maladie de Schmallenberg (Institute of Animal Health, UK)

Fin octobre 2012, l'infection par le SBV avait été confirmée par des tests sérologiques dans environ 6000 exploitations en Europe, confirmant de sa propagation vaste et rapide (Collins *et al.*, 2019 ; Doceul *et al.*, 2013 ; EFSA, 2012). Après cette première vague, de nombreuses régions ont fait face à une réémergence de la maladie de Schmallenberg. En effet, des études sérologiques et de surveillance du lait ont montré que le virus était réapparu en 2013 et 2014 dans des troupeaux de bovins et d'ovins en provenance de Belgique et d'Allemagne (Claine *et al.*, 2015 ; Wernike *et al.*, 2015). Les années suivantes, malgré un niveau grandement inférieur à celui de l'épidémie de 2012, un certain nombre de rapports ont signalé la poursuite de la circulation du virus à de faibles niveaux dans pas moins de 27 pays européens (Collins *et al.*, 2019 ; Claine *et al.*, 2015 ; Stokes *et al.*, 2016).

Aujourd'hui encore, le SBV est endémique de certaines régions d'Europe et circule à bas bruit. Il est important de continuer la surveillance car il n'est pas exclu que des flambées épizootiques surviennent dans les années à suivre. Ces flambées risquent de survenir à cause d'une diminution de l'immunité de troupeau et par le renouvellement d'une population naïve au SBV. De plus, les conditions environnementales et météorologiques pourraient avoir un impact majeur sur l'abondance et la survie des vecteurs Culicoïdes favorisant l'apparition d'une réémergence épizootique (Sailleau *et al.*, 2017 ; Gache *et al.*, 2016).

Malgré une épidémie contrôlée, la propagation vaste et rapide sur le continent européen, où le SBV est actuellement endémique, a entraîné de nombreuses pertes économiques chez les éleveurs et continue de susciter certaines inquiétudes (Wernike *et al.*, 2015)

2. Vecteur du SBV

La rapide dissémination en Europe SBV a conduit bon nombre d'études à se focaliser sur son vecteur Culicoïdes (*Diptera ceratopogonidae*) (**Fig. 2**). L'espèce *C. obsoletus* a été caractérisée comme étant le vecteur principal du SBV (Zimmer *et al.*, 2015). Les mâles sont floricoles tandis que les femelles ont besoin de repas sanguins tous les 3 à 4 jours pour la maturation de leurs œufs. Au sein du vecteur femelle, le virus se réplique dans ses glandes salivaires et se transmet aux vertébrés par voie sous-cutanée durant le repas sanguin. Les organes sensoriels olfactifs de la femelle sont capables de détecter des composés volatiles (comme le dioxyde de carbone) produits par leurs hôtes. Cette différence entre mâles et femelles implique que les femelles peuvent être retrouvées plus proches du sol et donc du bétail contrairement aux mâles qui se situent à hauteur des arbres (Zimmer *et al.*, 2015). Il a été établi que les animaux d'élevage représentaient de bonnes cibles pour les Culicoïdes. En effet, les étables sont des lieux ouverts où l'humidité et la chaleur y sont importantes en hiver, ce qui représente des lieux propices à la réplication du vecteur, au développement des larves et à la protection contre le froid hivernal. En 2013, des études de détection ont montré la présence de larves de *C. obsoletus* dans des restes de boue séchée et de matières fécales sur les murs des étables en Belgique (Zimmer *et al.*, 2014). La rapide propagation du SBV à travers l'Europe peut être expliqué par plusieurs facteurs. Premièrement, le vent constitue le principal facteur de dispersion des Culicoïdes. Il a été montré qu'ils sont capables de parcourir environ 2km de façon active durant leur vie. Cependant, des études sur le déplacement des insectes ont

démonstré que le vent pouvait accentuer leur déplacement jusqu'à 1000km (Tarlinton *et al.*, 2012), expliquant la propagation jusqu'au Royaume-Uni. Deuxièmement, les Culicoïdes présents en Europe semblent être capables de se reproduire durant les mois d'hiver où ils ne sont pas actifs (De Regge, 2017).

3. Signes cliniques de la maladie de Schmallenberg

Comme décrit précédemment, les premiers symptômes étaient transitoires et non-spécifiques chez les adultes bovins. Il s'agissait avant tout d'inappétence, d'hyperthermie, de chute de la production laitière et des diarrhées (Hoffmann *et al.*, 2012 ; Garigliany *et al.*, 2012). Hoffman et ses collègues ont confirmé ces observations par l'inoculation du SBV à des veaux sérologiquement naïfs et âgés de 9 mois (Hoffmann *et al.*, 2012). La virémie induite par le SBV dure généralement 4 à 6 jours et survient 2 à 4 jours après l'infection sous-cutanée par le vecteur. A l'heure actuelle, il n'a pas encore été déterminé où la réplication virale avait lieu, mais plusieurs infections en conditions expérimentales montrent que le matériel génétique viral est retrouvé dans les organes lymphoïdes, les poumons ainsi que dans le système reproducteur (Bayrou, 2020).

Chez les nouveau-nés les signes cliniques étaient beaucoup plus importants que chez l'adulte. Les premiers signes cliniques caractérisés étaient des signes neurologiques congénitaux et des malformations apparues chez des agneaux nouveau-nés (**Fig. 3**). Le plus souvent, des malformations musculaires étaient retrouvées à la naissance du nouveau-né (**Fig. 3A**). La cause de ces malformations est l'infection du système nerveux du fœtus, inhibant les réflexes moteurs de position et menant à des torsions du corps. Parmi celles-ci on retrouve des déformations des mâchoires, des vertèbres ou des membres. Dans certains cas, le système nerveux n'est pas capable de se développer. Outre les malformations, certains nouveau-nés morphologiquement normaux présentent des problèmes moteurs, de la cécité, de l'hyperexcitabilité, *etc.* (Garigliany *et al.*, 2012).

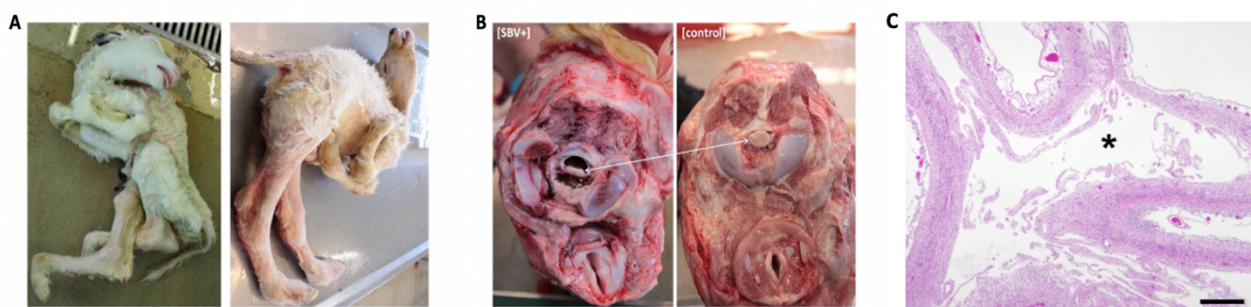


Figure 3: Signes cliniques causés par l'infection transplacentaire du SBV. A) Image d'un nouveau-né malformé résultant d'une infection par le SBV et présentant des signes d'arthrogrypose, de torticollis et de brachygnathie inférieure. B) Micromyélie dans le cas d'un agneau infecté par le SBV versus un agneau non infecté. C) Coupe du cortex cérébral avec une porencéphalie colorée à l'hématoxyline et à l'éosine provenant du cerveau d'un agneau infecté congénitalement par le SBV. La matière blanche est absente comme indiqué par (*). La matière grise adjacente est réduite en épaisseur (barre = 2 mm) (Varela *et al.*, 2013).

4. Phylogénie du SBV

Le SBV appartient à l'ordre des *Bunyavirales* du phylum *Negarnaviricota*. Ce phylum recense tous les virus à ARN monocaténaire de polarité négative. L'ordre des *Bunyavirales* contient 12 familles plus de 1500 espèces pouvant infecter des plantes, des animaux et des insectes. Malgré une grande diversité au sein des *Bunyavirales* (hôtes, vecteurs, organisation génomique, etc.), les mécanismes généraux de réplication sont conservés. En outre, cinq familles (*Arenaviridae*, *Peribunyaviridae*, *Nairoviridae*, *Phenuiviridae* et *Hantaviridae*) contiennent des agents pathogènes humains causant des maladies telles que des maladies hémorragiques des encéphalites et des maladies respiratoires (Leventhal *et al.*, 2021 ; Hughes *et al.*, 2020). Ces cinq familles partagent des caractéristiques communes dont la structure du virion, des génomes à ARN segmentés, linéaires, simple brin antisens ou ambisens et une transmission principalement via des vecteurs arthropodes. Dans l'ensemble, le nombre de segments génomiques peut varier de 2 à 5 segments. Cependant, la majorité des bunyavirus possède trois segments génomiques (Leventhal *et al.*, 2021). Les segments génomiques des *Bunyavirales* codent en règle générale pour quatre protéines structurales (la polymérase virale, deux glycoprotéines et une nucléoprotéine). De plus, certaines familles peuvent coder pour des protéines non structurales (NS). L'organisation des régions codantes de chaque segment génomique varie selon la famille (**Fig. 4**). Par exemple, les *Peribunyaviridae*, *Phenuiviridae* et *Tospoviridae* codent pour deux protéines non structurales : les protéines NSm et NSs. Le cadre de lecture de ces deux protéines peut être chevauchant avec celui des protéines structurales (*Peribunyaviridae*) ou être orienté en ambisense (*Phenuiviridae*) (**Fig. 4**) (Amroun *et al.*, 2017 ; Markleitz *et al.*, 2013). Une seconde propriété unique du génome des *bunyavirales* est la nature hautement conservée et complémentaire des extrémités 5' et 3' de chaque segment génomique, permettant la stabilité et la réplication des segments génomiques (**Table 1**) (Doceul *et al.*, 2017 ; Coupeau *et al.*, 2013 ; Walter & Barr, 2011).

La famille *Peribunyaviridae* comporte des virus enveloppés qui possèdent 3 segments distincts d'ARN simple brin à polarité négative. La famille compte 4 genres : *Orthobunyavirus*, *Herbesvirus*, *Pacuvirus* et *Shangavirus*. Le genre *Orthobunyavirus* est le plus étudié et le mieux compris de la famille et comprend des virus tels que le virus de Schmallenberg (SBV), le virus de La Crosse (LACV) et le virus Bunyamwera (BUNV). Les orthobunyavirus sont exclusivement transmis à l'hôte via un vecteur arthropode et peuvent causer des zoonoses avec des taux de mortalité élevés (Leventhal *et al.*, 2021 ; Hughes *et al.*, 2020).

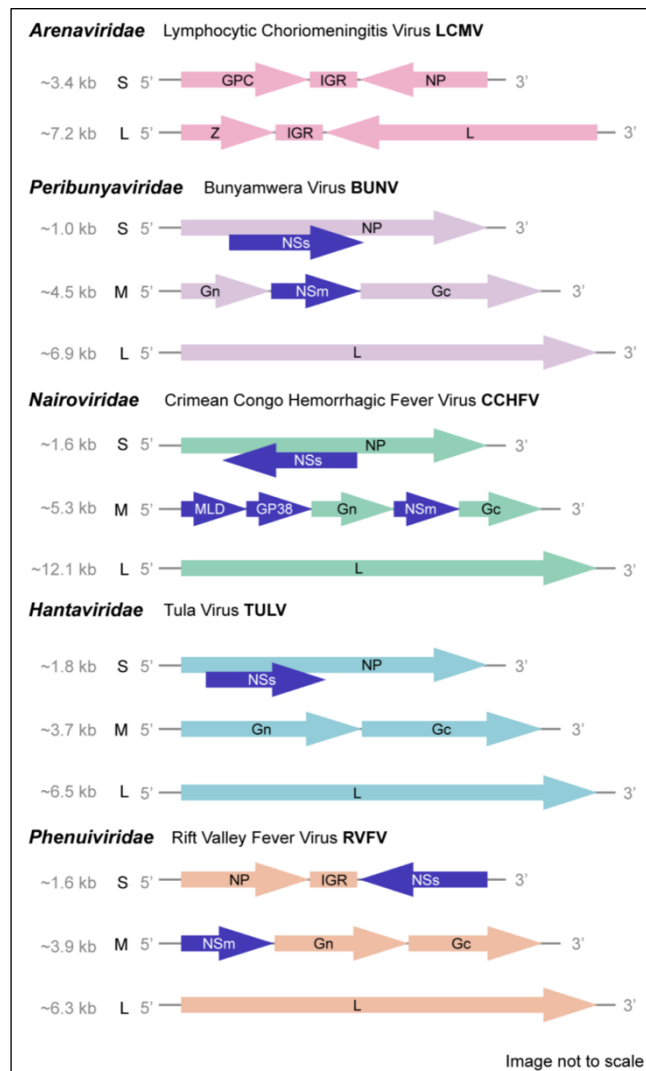


Figure 4: Organisation des génomes des *Bunyavirales*. Présentation des structures génomiques spécifiques de 5 familles de l'ordre des *Bunyavirales* : *Arenaviridae*, *Peribunyaviridae*, *Nairoviridae*, *Hantaviridae*, et *Phenuiviridae*. La souche de référence pour chaque famille est indiquée ainsi que la taille approximative de chaque segment génomique. Précurseur des glycoprotéines (GPC), nucléoprotéine (NP), région intergénique (IGR), protéine matricielle liant le zinc (Z), protéine non structurale du segment S (NSs), protéine non structurale du segment M (NSm), glycoprotéines de surface (Gn et Gc), domaine semblable à la mucine (MLD) et glycoprotéine sécrétée (GP38). Les protéines NS sont surlignées en bleu foncé et comprennent les NSs et NSm, ainsi que MLD et GP38 pour les *Nairoviridae* (Leventhal et al., 2021)

5. Structure et composition des orthobunyavirus

Les orthobunyavirus sont des virus enveloppés, de forme sphérique avec un diamètre mesurant entre 80 et 120nm. Leur bicouche lipidique est originaire de l'appareil de Golgi et est recouverte de glycoprotéines virales (Elliott, 2014) (**Fig. 6**). La particule virale comporte un génome constitué de 3 segments d'ARN monocaténares de polarité négative. Les segments génomiques sont nommés en fonction de leur taille respective : un grand segment (L), un segment de taille moyenne (M) et un petit segment (S) (**Fig. 5 et 6**) (Elliott, 2014). Chaque segment génomique est empaqueté par des oligomères de nucléoprotéines (N) apportant une stabilité et un ancrage pour l'ARN polymérase ARN-dépendante virale (RdRp), formant ensemble un complexe ribonucléoprotéique (RNP) (Ter Horst *et al.*, 2019 ; Walter & Barr, 2011). Les premières études fondamentales sur les orthobunyavirus ont été réalisées chez le virus bunyamwera (BUNV) dont la classification phylogénétique est proche du SBV. BUNV servira donc de modèle pour décrire les orthobunyavirus. Ces études montrent que chaque segment génomique code jusqu'à 3 protéines pour un total de 6 protéines codées par l'entièreté du génome de BUNV (Elliott, 1997).

Premièrement, le segment L (~6500 kb) du génome de BUNV permet l'expression de la protéine L qui a une activité RdRp (**Fig. 6**) impliquée dans la production des trois types d'ARN produits durant la réplication : l'ARN génomique de polarité négative (ARNg), l'ARN complémentaire positif (ARNc) et l'ARN messenger positif (ARNm) (Elliott, 1989). De plus, la RdRp possède un rôle majeur dans le processus de « vol de coiffe » qui permet d'amorcer la synthèse de l'ARNm viral dans le cytoplasme de la cellule hôte (Coupeau *et al.*, 2013).

Deuxièmement, le segment M (~4,5 kb) encode les deux glycoprotéines membranaires Gn et Gc ainsi qu'une protéine non structurale, la protéine NSm (**Fig. 6**) (Elliott, 2014). Les glycoprotéines Gn et Gc sont riches en cystéine et sont donc capables de réaliser de nombreux ponts disulfures intra- et intermoléculaires importants pour leur structure et leur fonction. Bien qu'elles aient des tailles nettement différentes (Gn 35kDa et Gc 110kDa), elles sont toutes deux ancrées dans la membrane et interagissent entre elles pour former des trimères d'hétérodimères présentant une extrémité carboxy-terminale cytosolique. Ce complexe trimérique permet la reconnaissance et l'attachement aux récepteurs des cellules hôtes et l'induction du processus d'endocytose. Les orthobunyavirus étant capables d'infecter de nombreux types cellulaires, cela laisse présager que le récepteur reconnu est ubiquiste (Elliott & Blakqori, 2011). En outre, la protéine Gc est habilitée à agir comme une protéine de fusion de classe II (Elliott & Blakqori, 2011). Ce type de protéine se retrouve en majorité chez les genres *alphavirus* et *flavivirus* (Gaudin, 2004). Elles ne sont pas clivées après synthèse et leur peptide de fusion est interne. Lors de l'acidification des vésicules d'endocytose, les hétérodimères modifient leurs structures et présentent leurs peptides de fusion interne induisant le relargage du matériel génétique dans le cytoplasme de la cellule hôte (Gaudin, 2004 ; Modis *et al.*, 2004). En 2016, Varela et ses collègues ont également démontré que la protéine Gc du virus de Schmallenberg jouait un rôle important dans l'inhibition de la synthèse des protéines de l'hôte en modèle murin, en induisant notamment une diminution de la réponse interféron (Varela *et al.*, 2013).

La séquence de la protéine NSm se situe entre les protéines Gc et Gn et est le fruit du clivage de la polyprotéine du segment M (Elliott, 2014). Contrairement aux 2 autres protéines du segment M, le rôle de la protéine NSm n'est pas encore bien défini. Indépendamment des protéines virales, des études ont montré sa localisation au niveau de l'appareil de Golgi. De plus, elle possède 5 domaines dont 1 permet l'interaction avec Gn et participerait à la maturation des 2 glycoprotéines (Kraatz *et al.*, 2018). La localisation de la protéine NSm suggère qu'elle est importante pour l'assemblage et la sortie du virus (Elliott, 2014). Cependant, des études ont montré que la protéine NSm n'était pas essentielle pour la réplication du virus (Kraatz *et al.*, 2018).

Dernièrement, le segment S (800-1000 nt) encode la nucléoprotéine N (25–30 kDa) mais également la seconde protéine non structurale NSs (10 kDa) (**Fig. 6**) (Elliott, 2014). Il a été montré que la nucléoprotéine est la protéine virale la plus exprimée par les cellules infectées et qu'elle comporte plusieurs fonctions essentielles aux *Bunyavirales* (Elliott, 2014 ; Jonsson *et al.*, 2010). Étant la protéine la plus exprimée lors d'infection, Boshra et ses collègues ont montré dans des modèles murins que l'immunisation contre cette unique protéine était protectrice (Boshra *et al.*, 2017).

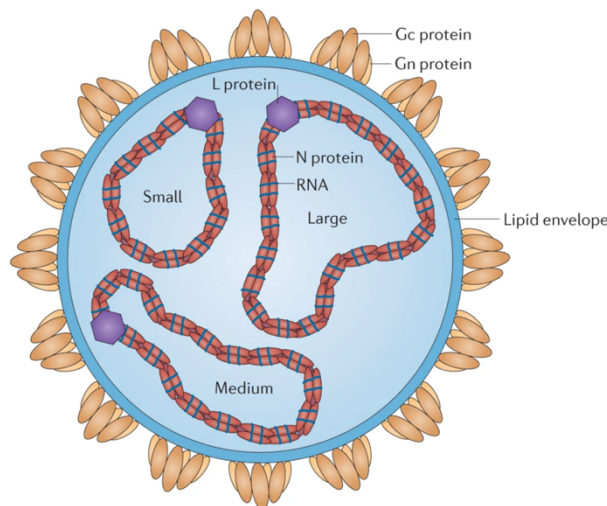


Figure 5: Représentation schématique du virion des orthobunyavirus. Les trois segments du génome (petit, moyen et grand) sont encapsidés par la protéine N (nucléocapside) pour former des complexes ribonucléoprotéiques (RNP). Les RNP s'associent à la protéine L (RdRp) et sont empaquetés dans une enveloppe lipidique modifiée par l'insertion des glycoprotéines virales Gn et Gc (Elliott, 2014).

Table 1: Familles parmi les *Bunyvirales* ainsi que la séquence de leurs extrémités terminales 3' et 5'.

Famille	Séquence
<i>Peribunyaviridae</i>	3'UCAUCAUG 5'AGUAGUGGC
<i>Hantaviridae</i>	3'AUCAUCAUCUG 5'UAGUAGUAUGC
<i>Nairoviridae</i>	3'AGAGUUUCU 5'UCUCAAGA
<i>Phénouviridae</i>	3'UGUGUUUC 5'ACCAAAG
<i>Tospoviridae</i>	3'UCUCGUUA 5'AGAGCAAU

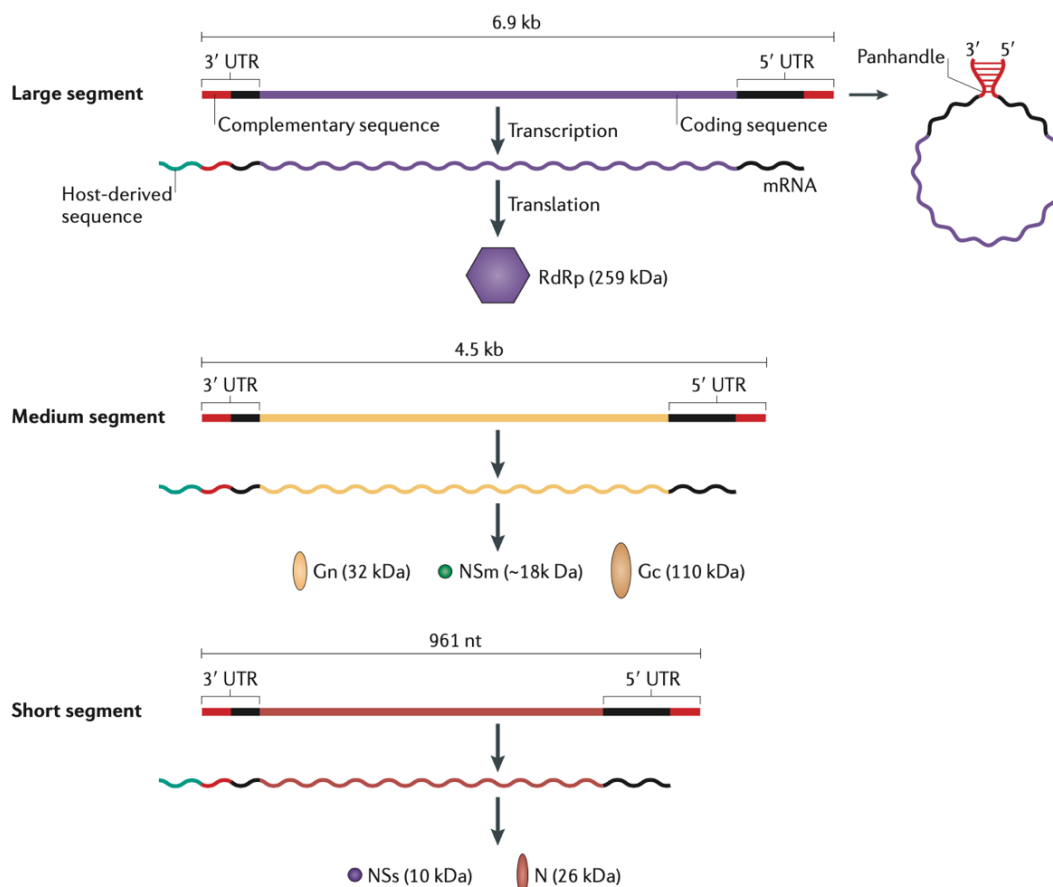


Figure 6: Stratégie de codage du génome des orthobunyavirus. Le virus Bunyamwera est utilisé comme exemple. Les petits, moyens et grands ARN génomiques sont représentés ainsi que leur longueur. Les ARN génomiques sont transcrits en ARNm, qui contiennent des séquences dérivées de l'hôte à l'extrémité 5' mais sont tronqués à l'extrémité 3' par rapport à l'ARNg. Les séquences terminales de chaque segment du génome sont complémentaires (rouge), ce qui entraîne un appariement de bases entre les extrémités 3' et 5' pour former une structure en « panhandle ». Les produits protéiques avec leur poids moléculaire estimé sont également indiqués (Elliott, 2014).

La fonction majeure de la protéine N est la formation des RNP permettant la stabilité des segments d'ARNg et d'ARNc (Kohl *et al.*, 2006). Effectivement, au moment de la formation des virions, les protéines N s'assemblent en tétramères formant un anneau qui entoure l'ARN, protégeant le génome d'une dégradation dans l'environnement cellulaire. La structure de la protéine permet également une interaction avec la RdRp ce qui permet de favoriser la transcription et la réplication (Dong *et al.*, 2013). La protéine N est également capable d'interagir avec l'extrémité C-terminal des protéines Gc et Gn, stabilisant la structure des virions. La structure de la nucléoprotéine et la modification de conformation qu'elle peut prendre, indique qu'elle est capable d'interagir avec de l'ARN ainsi qu'avec des protéines du virus (RdRp) ou de l'hôte eucaryote (Dong *et al.*, 2013).

La protéine non structurale NSs est codée par un cadre de lecture chevauchant le gène N dans le segment S (Elliott, 2014). Elle se localise principalement dans le noyau et en périphérie de celui-ci. Tout comme la protéine NSm, NSs est dispensable pour la survie et la réplication du virus dans les cellules hôtes (Amroun *et al.*, 2017 ; Varela *et al.*, 2013 ;). Néanmoins, des études montrent une atténuation de la réplication virale non négligeable chez les mutants de la protéine NSs (Δ NSs) comparé au virus sauvage (WT). Cette différence n'est cependant pas visible lors d'études dans des cellules d'insectes vecteurs infectées (Elliott, 2014 ; Eifan *et al.*, 2013). En outre, plusieurs études ont démontré que la protéine NSs était capable de réduire fortement la réponse antivirale à interféron (IFN) de l'hôte (Elliott *et al.*, 2013 ; Varela *et al.*, 2013 ; Blakqori *et al.*, 2007). Barry et ses collègues ont mis en évidence que cette protéine était capable de cibler la sous-unité RPB1 de la RNA polymérase II (RNAP_{II}) induisant un blocage de celle-ci (Barry *et al.*, 2014). Il y aurait donc une diminution de la transcription et par conséquent de la synthèse protéique *de novo*. La réduction d'IFN favorise la réplication et la propagation du SBV et contribue ainsi à sa pathogénèse. Enfin, il a été démontré que la protéine NSs pouvait induire l'apoptose par une augmentation des caspases 3 et 7 activées chez l'hôte (Aksoy & Azkur, 2018).

6. Cycle répliatif des *Bunyavirales*

Peu d'études ayant été menées sur la réplication du SBV à proprement parler, il s'agira d'un résumé sur la réplication des *Bunyavirales* (Fig. 7). Malgré quelques différences entre les diverses familles, leurs cycles répliatifs partagent de grandes similarités (Elliott, 2014).

La première étape lors de l'infection de la cellule est la reconnaissance cellulaire et l'attachement aux récepteurs. Cette étape est réalisée par les deux glycoprotéines introduites précédemment : Gc et Gn. Différents récepteurs de reconnaissance ont été identifiés comme DC-SIGN chez les phlébovirus de la famille des *Phenuiviridae* (Elliott 2014 ; Hofmann *et al.*, 2013) ou les intégrines β 1 et β 3 chez les *Hantaviridae* (Sun *et al.*, 2018 ; Jonsson *et al.*, 2010). De plus, des études portant sur l'entrée de virus très proches phylogénétiquement du SBV, tels que le La Crosse Virus (LACV) ou l'Oropouche Virus (OROV), ont démontré que l'entrée était menée par une endocytose clathrine-dépendante (Shtanko *et al.* 2014 ; Elliott 2014 ; Hollidge *et al.*, 2012). Après l'endocytose, les particules virales nécessitent d'être libérées. La fusion entre la membrane du virion et de l'endosome se produit par des changements de conformation de Gc en condition acide, libérant ainsi les RNP dans le cytoplasme (Sun *et al.*, 2018 ; Soldan *et al.*, 2010).

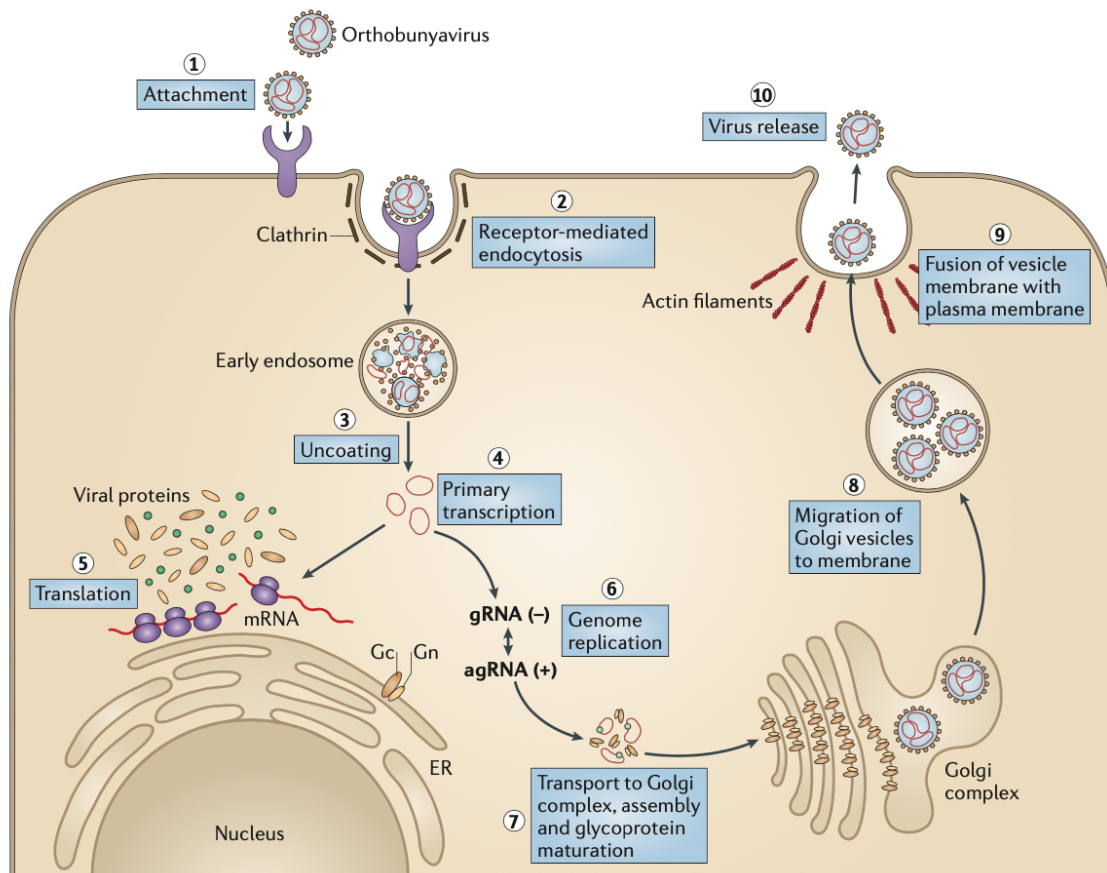


Figure 7: Schéma du cycle réplcatif des orthobunyavirus. (1) La réplication commence par l'attachement du virus aux récepteurs de l'hôte. (2) Internalisation du virus par endocytose clathrine dépendante médiée par les récepteurs. (3) Libération du génome viral dans le cytoplasme par l'acidification des vésicules endosomales et la fusion de la membrane virale avec la membrane endosomale. (4) Transcription primaire catalysée par la RdRp. (5) Traduction des ARNm viraux. Les glycoprotéines Gn et Gc se dimérisent dans le RE et se localisent au niveau de l'appareil de Golgi. (6) Les trois segments d'ARNg sont convertis en ARNc de sens positif pour la réplication du génome. (7) Les ribonucléoprotéines (RNP) sont transportées vers les membranes de l'appareil de Golgi qui ont été modifiées par l'insertion de dimères Gn-Gc. Les particules virales bourgeonnent à partir de vésicules dérivées de la membrane de Golgi. (8) Les vésicules golgiennes contenant les particules virales sont transportées vers la surface cellulaire. (9) Fusion des membranes vésiculaires avec la membrane plasmique, conduisant à la libération des virions infectieux, qui peut être facilitée par les filaments d'actine. (10) Libération des particules virales naissantes par la cellule hôte (Elliott, 2014).

Une étude de colocalisation d'un orthobunyavirus avec rab5 (GTPase régulatrice des endosomes précoces) a montré que la fusion se déroule dans les endosomes précoces (Hollidge *et al.*, 2012). Cependant, d'autres études ont montré qu'au sein des *Bunyavirales* le moment de la fusion membranaire variait et pouvait se dérouler, chez certaines espèces, aux stades endosomaux plus tardifs (Shtanko *et al.* 2014 ; Lozach *et al.* 2010). La conséquence de cette fusion est la libération des trois RNP, qui pour rappel constituent un segment d'ARNg (de polarité négative) empaqueté dans un ensemble des protéines N associées à une unique RdRp (Sun *et al.*, 2018). Lorsque les RNP sont libérées dans le cytoplasme, les segments génomiques migrent vers le réticulum endoplasmique (RE) où a lieu la transcription du génome en ARNm⁺ et en ARNc⁺ (Elliott, 2014). Les ARNm servent de brins codants pour la traduction des protéines et les ARNc servent de brins matrices pour la réplication du génome. Lorsque suffisamment de RNP sont formées et accumulées dans le cytoplasme, le virus peut commencer son empaquetage et préparer sa sortie. A l'heure actuelle, il est connu que l'empaquetage se produit au niveau de l'appareil de Golgi. La polyprotéine M y est maturée et forme l'hétéro-multimère Gn-Gc intégré dans la membrane du Golgi. Au même instant la protéine N s'accumule dans le cytoplasme reformant des RNP dirigées vers la membrane modifiée de l'appareil de Golgi. Au niveau de l'appareil de Golgi, des vésicules précurseuses vont subir une maturation en vésicules intermédiaires modifiant ainsi la sensibilité des protéines Gn et Gc à l'endoglycosidase H (Elliott, 2014). Il semblerait que les signaux d'empaquetage soient présents au niveau des régions non-traduites en 3' et 5'. Cependant, aucune étude n'a réellement pu mettre en évidence ces signaux chez les orthobunyavirus dont le but est d'avoir les trois RNP (Large, Medium et Small) dans chaque vésicule (Elliott 2014 ; Reguera *et al.*, 2013). Lorsque les virions s'accumulent dans les vésicules golgiennes, les vésicules se libèrent et sont transportées via le transport d'exocytose et le cytosquelette vers la membrane plasmique. La fusion de la bicouche lipidique d'origine golgienne des virions et de la membrane plasmique permet le relargage extracellulaire (Elliott, 2014).

7. Transcription des *Bunyavirales*

A l'exception du cadre de lecture des gènes codant pour les protéines non structurales qui peut varier, le mécanisme de transcription chez les *Bunyavirales* est partagé chez les différents genres (Amroun *et al.*, 2017 ; Elliott, 1992). La transcription commence dès que l'ARNg (polarité négative) complexé en RNP est libéré dans le cytoplasme. La transcription chez les virus à ARN de polarité négative a pour objectif de former, d'une part des ARNm (polarité positive) permettant la traduction des protéines structurales et non structurales et d'autre part, de former de l'ARNc (polarité positive) afin de reformer du génome (Elliott, 2014). En premier lieu, la RdRp associée à chaque RNP, initie la transcription primaire pour donner naissance aux ARNm des segments S, M et L. Les RNP des *Bunyavirales* ont la particularité de former une structure dite en « panhandle ». En effet, les extrémités 3'NTR et 5'NTR (régions non-traduites) des segments génomiques sont complémentaires sur 25 à 35 nucléotides (**Table 1**) (Doceul *et al.*, 2017). La complémentarité des extrémités permet la stabilité de la RNP sous une forme double-brin (Sun *et al.*, 2018 ; Mir *et al.* 2006). Ce court segment double-brin de la « panhandle » sert d'amorce pour l'initiation de la transcription, produisant des ARNm coiffés

en 5' (Coupeau *et al.*, 2013 ; Walter & Elliott, 2011). Cependant, il a été démontré que la polymérase des *Bunyavirales* ne possède pas la capacité de produire cette coiffe. Comme tous les bunyavirus, le SBV utilise un mécanisme de « vol de coiffe » (Coupeau *et al.*, 2013). Ce mécanisme débute par un clivage des 10 à 18 premiers nucléotides en position 5' des ARNm de l'hôte par une activité endonucléase de la RdRp virale. Les fragments nucléotidiques ainsi « volés » à l'hôte sont utilisés pour amorcer la traduction ARNm (Sun *et al.*, 2018 ; Coupeau *et al.*, 2013). De plus, chez les orthobunyavirus, il a été établi que les ARNm étaient dépourvus de queues poly(A) et ne nécessitaient pas la présence de protéines de liaison à la queue poly(A) (PABP) (Coupeau *et al.*, 2013 ; Vera-Otarola *et al.*, 2010). Cela suggère un mécanisme alternatif poly(A) indépendant pour l'initiation de la traduction chez le SBV.

C'est précisément l'interrogation sur un mécanisme alternatif qui a défini la suite et les objectifs de ce mémoire : comprendre comment certaines des protéines sont hautement traduites par un mécanisme poly(A) indépendant, et comment le SBV définit un transcrit d'ARNm plutôt qu'un ARNc. Dernièrement, des études de séquence et de structure du segment S codant pour la protéine N ont mis en exergue des particularités pouvant être reliées à la transcription et la traduction du SBV (Coupeau *et al.*, 2013).

8. Le segment S du virus de Schmallenberg

Il a été démontré que les protéines L et M sont faiblement traduites en comparaison avec la protéine N codée par le segment S. En effet, Jonsson et ses collègues ont montré que lors d'une infection virale par des hantavirus, la protéine N est une des protéines les plus exprimées par la cellule hôte (Jonsson *et al.*, 2010). Toutes ces indications amènent à s'interroger sur la régulation du segment S et sa capacité à former un ARNm hautement traduit. Bien que les trois ARNm (S, M et L) soient dépourvus d'une queue poly(A), les sites de clivage des extrémités 3' diffèrent. En effet, contrairement aux segments L et M où aucun clivage clair d'ARNm n'a pu être observé, le segment S est spécifiquement clivé en ARNm (Coupeau *et al.*, 2013). Il existe deux hypothèses pour expliquer ce clivage entre l'ARNm et l'ARNc du segment S. La première s'expliquerait par un décrochage prématuré de la RdRp sur le complexe RNP en forme de « panhandle », conduisant à un brin tronqué servant d'ARNm. La seconde se base sur une séquence spécifique de clivage, permettant un clivage post-transcriptionnel de l'ARNc pour former un ARNm (Coupeau *et al.*, 2013). A l'heure actuelle, il n'a pas encore été déterminé dans quelles conditions la RdRp peut se décrocher ou de quelle manière le brin d'ARNc est clivé.

L'étude d'alignement de séquences et des profils de terminaison du segment S des simbuivirus a montré la présence d'une structure en tige-boucle et d'un signal de terminaison 5'-GCN₁-₃GC-3' ('GC') dans la région 3'NTR du SBV (**Fig. 8**) (Coupeau *et al.*, 2013). Cette étude a été réalisée sur plusieurs simbuivirus et tous possédaient les deux motifs (Coupeau *et al.*, 2013). Le fait que ces deux motifs soient conservés au sein des 3'NTR de ce séro groupe (**Fig. 9**) suggère qu'ils sont importants pour la régulation de la transcription et de la traduction du segment S.

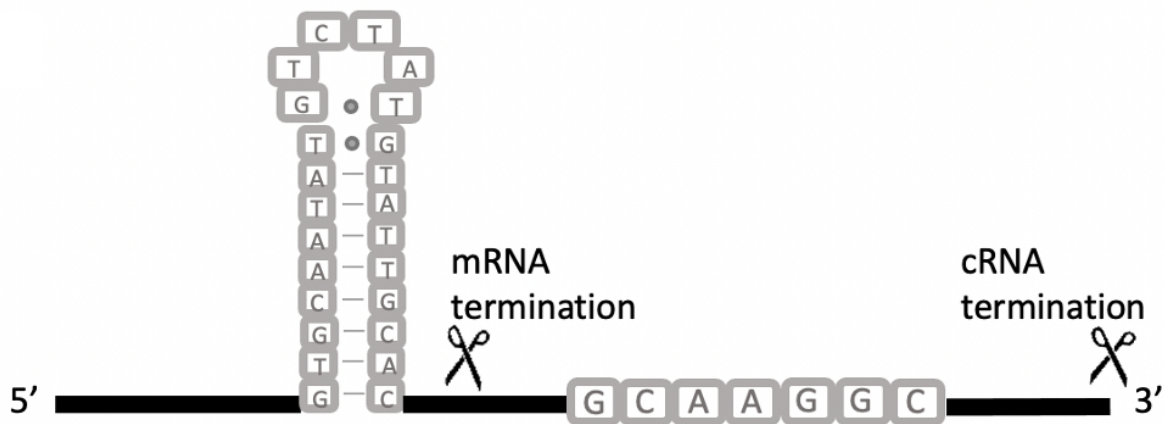


Figure 8: Schéma des motifs ARN de la 3’NTR du SBV. La séquence correspond à la structure en tige-boucle de SBV suivie du signal ‘GC’. Les ciseaux représentent les sites de clivages de l’ARNm et de l’ARNc.

SBV:	5'	CC- <u>GUGCAAUAUGUCUAUGUAUUGCAC</u> - <u>ACCAUUAUACU</u> -GCAAGGC-UUGA 3'
SHAV:	5'	CC- <u>GUGCAAUAUGUCUAUGUAUUGCAC</u> - <u>ACCAUUAUACU</u> -GCAAGGC-UUGA 3'
SATV:	5'	CC- <u>GUGCAAUAUGUCUAUGUAUUGCAC</u> - <u>ACCAUUAUACU</u> -GCAAGGC-UUGA 3'
DOUV:	5'	CC- <u>GUGCAAUAUGUCUAUGUAUUGCAC</u> - <u>ACCAUUAUACU</u> -GCAAAGC-UUGA 3'
AKAV:	5'	AA- <u>GUGCACCAUGUCUAUGUGGUGCAC</u> - <u>ACCUUUGUUCA</u> -GCU--GC-UUGG 3'
SIMV:	5'	UA- <u>GUACAACAUGUCUAUGUGUUGUAC</u> - <u>GCCAUUU--CA</u> -GCA--GC-UUGU 3'
SABOV:	5'	UU- <u>GUGUAUCAUGUCUAUGUGAUACAC</u> - <u>GCCAAUU--CU</u> -GCU--GC-UUGU 3'
PEAV:	5'	UA- <u>GU AUGGUGUUC CACAACACCAUAC</u> - <u>GCCAUUU--CU</u> -GCA--GC-AUUG 3'
AINOV:	5'	CA- <u>GU AUGGUGUUC CAAAACACCAUAC</u> - <u>GCCAUUC--CU</u> -GCA--GC-AUUG 3'
SANV:	5'	CA- <u>GU AUGGUGUUC CAAAACACCAUAC</u> - <u>GCCAUUU--CU</u> -GCA--GC-AUUG 3'

Figure 9: Alignement de séquences nucléotidiques de l’extrémité 3’NTR du segment S au sein des simbuvirus. L’encadré de gauche représente la séquence de la structure en tige-boucle. Les nucléotides soulignés forment les bras complémentaires de la tige-boucle. Les nucléotides surlignés représentent les différences entre les virus du séro groupe Simbu. La séquence encadrée de droite représente le ‘GC’ signal. La structure tige-boucle et le ‘GC’ signal sont conservés au sein des simbuvirus et séparé par 9-11nt représentant le signal de terminaison de l’ARNm (encadré central).

DOUV : virus Douglas ; AKAV : virus Akabane ; SIMV : virus Simbu ; SABOV : Sabo virus ; PEAV : Peaton virus ; AINOV : Aino virus ; SANV : Sango virus (Coupeau et al., 2013).

Enfin, ces observations amènent à deux interrogations : (1) La traduction efficace de la protéine N est-elle due à la présence de ces motifs ? (2) Le 'GC' signal est-il un site de clivage qui permet la formation d'un ARNm à partir d'un ARNc ?

Bien que la majorité des ARNm soient polyadénylés, des mécanismes indépendants de la queue poly(A) ainsi que des PABPs ont déjà été observés. L'exemple le plus connu est décrit par Marzluff et ses collègues, où ils ont montré que des niveaux élevés de traduction se produisent parmi les ARNm d'histones par le biais d'une tige-boucle localisée dans une 3'UTR non-polyadénylée (Marzluff *et al.*, 2008). De plus, les structures stables de type tige-boucle à l'extrémité 3' des petits ARNm sont connues pour stabiliser les transcrits en protégeant contre l'action des exonucléases (Ford & Wilusz, 1999). En ce qui concerne le signal 'GC', sa présence est uniquement retrouvée dans les ARNg et l'ARNc. Cela suggère que le signal 'GC' est un signal permettant le clivage en amont de son site (Coupeau *et al.*, 2013).

Étant donné la localisation et la conservation du signal 'GC' et de la tige-boucle, nous avons émis l'hypothèse que les 2 motifs pourraient jouer des rôles cellulaires ou viraux clés, impliqués dans la terminaison de la transcription ainsi que dans l'initiation de la traduction en recrutant des protéines.

9. Le monde de l'ARN

Les prochains paragraphes présentent l'ARN sous toutes ses formes afin de mieux discerner l'évolution des connaissances sur cette molécule, étant donné qu'elle se situe au centre de notre hypothèse sur la régulation de la réplication du SBV.

Alors que le dogme central en biologie du vivant définit un modèle linéaire dans lequel l'ADN précède l'ARN et la protéine, plusieurs théories s'accordent aujourd'hui sur le fait que l'ARN est apparu avant l'ADN (Higgs *et al.*, 2015 ; Atkins *et al.*, 2011). L'hypothèse selon laquelle, l'ARN serait l'élément majeur du dogme central se base sur le fait qu'il possède entre autres, certaines propriétés catalytiques intrinsèques et que la synthèse des protéines peut se produire en absence d'ADN mais pas de l'ARN. Dès lors, un mode de vie primaire basé sur l'ARN et non l'ADN a été étudié en profondeur. Ces études suggèrent que l'ADN aurait été sélectionné en tant qu'archive de l'information génétique dans les cellules afin de créer un système plus stable que les anciens systèmes ARN-protéines. (Higgs *et al.*, 2015 ; Dworkin *et al.*, 2003 ; Freeland *et al.*, 1999 ; Gesteland *et al.*, 1993 ; Lazcano *et al.*, 1992 ; Lazcano *et al.*, 1988).

L'ARN sous ces différentes formes, tel que décrit aujourd'hui, est une des molécules centrales de l'expression génique dans toutes les formes de vie existantes, servant de messager et de régulateur (Atkins *et al.*, 2011). De plus, il est connu que l'ARN peut faire office de génome chez de nombreux virus en plus de leur fonctions catalytiques. En effet, les virus à ARN possèdent comme matériel génétique un ARN simple- ou double-brin et sont capables de se répliquer en synthétisant le brin complémentaire par appariement de bases. Cela démontre que l'ARN peut supplanter l'ADN comme support de l'information génétique (Flores *et al.*, 2014).

10. Les fonctions associées aux structures de l'ARN

L'ARN est le plus souvent représenté comme une molécule « linéaire », ciblée par des complexes de maturation des ARN et de la traduction en protéines. L'ARN est pourtant bien plus que cela. En effet, la découverte des ribozymes (Nissen *et al.* 2000 ; Guerrier-Takada *et al.*, 1983) a montré que l'ARN peut également exercer une fonction biochimique de catalyse qui est dépendante de sa structure dynamique, elle-même régie par des facteurs tels que la température, l'état énergétique cellulaire, les hélicases ARN, les protéines chaperonnes et d'autres protéines interagissant avec l'ARN (RBP) (Lewis *et al.*, 2017). Pour démontrer cela, une étude a su mettre en évidence, via transfert d'énergie par résonance de type Förster (FRET), que l'ARN présent dans les ribosomes était dynamique et que son activité était dépendante de son état structural pouvant être modifié par l'association avec une protéine du ribosome (Kim *et al.*, 2014). Ils ont également établi que la protéine ribosomique pouvait se lier à certaines régions flexibles de l'ARN uniquement dans une conformation spatiale spécifique. Une protéine ne se liera qu'à un ARN dont la structure est compatible à son domaine de liaison, permettant la formation d'un complexe. Lors de la formation de ce complexe, la structure de la protéine ainsi que de l'ARN sont modifiées pour que chacune s'adapte au site de liaison. Finalement, Kim et ses collègues ont montré que la fonction catalytique des ribosomes était attribuée à ces modifications induites (Hall, 2014 ; Kim *et al.*, 2014).

En plus d'une activité catalytique dépendante de sa conformation spatiale induite, des études ont mis en exergue la capacité des ARN à former des structures secondaires et tertiaires (Atkins *et al.*, 2011). Tous les types d'ARN (codants ou non-codants) présentent des structures leur permettant d'exercer leur maturation et leur fonction. Ces structures tridimensionnelles peuvent permettre l'interaction avec des protéines formant ainsi des complexes ribonucléoprotéiques (RNP) stables. Les RNP sont connus pour être impliqués dans la régulation transcriptionnelle et traductionnelle (Majumder *et al.*, 2021 ; Lewis *et al.*, 2017).

Des éléments de structure secondaire tels que des tige-boucles, des hélices, des boucles internes, *etc.* peuvent être prédits lorsque les nucléotides interagissent dans l'espace deux à deux (Lewis *et al.*, 2017). Dans ce travail, la structure en tige-boucle sera développée car elle apparaît comme exerçant un rôle majeur dans la réplication de certain virus. En effet, certains virus à ARN de polarité positive ou négative ont la capacité de se répliquer directement dans le cytoplasme en complexant, grâce à leur tige-boucle, des RBP de la cellule hôte (Choi *et al.*, 2021 ; Bujalowski *et al.*, 2017). Les tige-boucles sont constituées d'une suite de nucléotides simple-brin s'appariant entre eux, via les duplexes nucléotidiques A-U et G-C ainsi que des paires en appariement « bancal » comme les duplexes G-U (Majumder *et al.*, 2021 ; Parisien & Major, 2008).

Un exemple bien décrit de la relation entre la structure de l'ARN et la réplication a été étudié chez les *Flaviviridae*. En effet, il a été démontré que les flavivirus possèdent des éléments structurés dans leur régions 3' NTR et 5'NTR, qui mènent à des interactions ARN-ARN permettant des liaisons entre les deux extrémités du génome. Dans cet exemple, plusieurs tige-boucles, conservées au sein des flavivirus, ont été caractérisées comme interagissant ensemble et servant de promoteur pour la réplication virale. En outre, une série de mutations dans les

structures ont pu montrer que l'activité était liée à la structure même de ces régions et non à leur séquence. Cela montre l'importance d'un mécanisme conservé chez les flavivirus pour l'initiation de leur réplication (Choi *et al.*, 2021 ; Bujalowski *et al.*, 2017 ; Davis *et al.* 2007 ; Yu & Markoff, 2005). D'autres exemples chez les virus ont été mis en avant par l'étude des complexes ARN-protéine comme chez les *Picornaviridae*, le virus de l'Hépatite C (HCV), *etc.* (Yi *et al.*, 2005 ; Andino *et al.*, 1990). Ces diverses études ont souligné l'importance de la structure des ARN dans l'interaction avec d'autres ARN et/ou protéines.

11. Utilisation de l'ARN en virologie moléculaire

L'importance des RNP dans les processus cellulaires aussi bien viraux, eucaryotes ou procaryotes a permis le développement de nouvelles techniques permettant leur caractérisation au sein de leur environnement cellulaire.

Plusieurs approches ont été utilisées afin de détecter les potentiels partenaires protéiques des ARN viraux. L'objectif étant généralement de déterminer les RBP responsables de la réplication des virus. Par exemple, une étude a utilisé un système de capture d'ARN par affinité, c'est-à-dire de courtes sondes d'ADN (39nt) biotinylées et complémentaires à la 3'NTR du virus de l'hépatite C (HCV). Cette région contient des motifs conservés au sein de ces virus qui semblent être impliqués dans leur réplication et/ou leur pathogénicité (Gerber, 2021 ; Kim *et al.*, 2007 ; Harris *et al.*, 2006). Ces sondes servent d'appât moléculaire afin de capturer les protéines de l'hôte interagissant avec cette région de l'HCV. Ils ont ensuite capturé les sondes en déposant un lysat cellulaire sur une colonne contenant de la streptavidine grâce à l'interaction forte de celle-ci pour la biotine. Cela a permis d'isoler les RBP associées à la séquence d'intérêt (Gerber, 2021). Une des limitations majeures de cette technique reste la non-spécificité de liaison des RBP à la sonde d'ADN ou au complexe ADN-ARN, due au lysat protéique appliqué *in vitro* sur les sondes ainsi qu'aux lavages qui peuvent influencer les RNP capturées. Une approche intéressante pour pallier à cette limitation est la méthode PAIR (identification des RBP assistée par peptide et acide nucléique) qui utilise des sondes d'acide nucléique peptidique (ANP) couplées à un peptide pouvant pénétrer dans les cellules (CPP). L'ANP employé est complémentaire à l'ARN viral cible et peut donc interagir dans le cytosol, réduisant ainsi la non-spécificité liée à l'utilisation de sondes *in vitro* (Majumder *et al.*, 2021 ; Upadhyay *et al.*, 2013). Étant donné que les ANP sont des « imitations » de l'ADN, ils ne sont probablement pas reconnus par les protéines cellulaires et sont, par conséquent, très stables (Gerber, 2021).

Chez le virus de la dengue (DENV), une autre approche a été exploitée pour purifier les complexes ARN-protéine (Phillips *et al.*, 2016). Cette méthode consiste à utiliser une dizaine d'ASO (antisense oligonucleotides), pouvant s'hybrider à des régions différentes de l'ARN viral, et notamment aux motifs conservés dans les NTR de DENV. Les ASO, préalablement biotinylées en leur extrémité 5', ont été liés de manière covalente aux ARN viraux non-polyadénylés par rayonnement ultraviolet (UV) permettant la capture des complexes (Phillips *et al.*, 2016). Cette méthode de liaison par UV est souvent utilisée en biologie cellulaire car elle permet, sous conditions dénaturantes, d'isoler uniquement ce qui est lié de manière covalente à l'ARN. L'avantage de cette méthode est de directement lier les RBP à l'ARN, pouvant donc

être purifiés dans leurs états natifs. De plus, cette technique est spécifique aux interactions ARN-protéines. Cependant, la liaison par UV comporte plusieurs limitations comme la préférence pour les séquences riches en U et les ARN simple brin, une efficacité qui varie en fonction des acides aminés et une liaison irréversible entre l'ARN et les protéines (Majumder *et al.*, 2021 ; Ramanathan *et al.*, 2019). Pour pallier à certaines de ces limitations, des agents chimiques comme le formaldéhyde sont également utilisés afin de purifier les complexes ARN-protéines à une distance de 2Å. Cette technique est plus efficace que les UV et le lien est réversible. Cependant, cette méthode inclut des liaisons protéines-protéines et ADN-protéines qui sont indésirables dans le cas de l'étude des RBP (Majumder *et al.*, 2021 ; Ramanathan *et al.*, 2019). Malgré beaucoup de similarités entre les deux techniques, elles comportent certains avantages et désavantages qui sont à évaluer en fonction de la question biologique établie.

12. Évolution des outils de détection des RBP

Ces dernières années, le domaine de l'identification des RBP s'est fortement développé, compte-tenu de l'enjeu majeur qu'il représente (Gerber, 2021). Dès lors, de nouvelles techniques de détection de protéines liées à l'ARN ont été développées. Ces nouvelles méthodes se basent, non pas sur la capture par affinité ou sur la création de liaisons covalentes, mais sur des aptamères d'ARN et un marquage de proximité *in vivo* (Trinkle-Mulcahy, 2019). En effet, ces méthodes visent à utiliser des ARN structurés localisés à proximité des motifs ARN à étudier et possédant un domaine d'interaction avec des protéines de marquage spécifiques (Sternburg *et al.*, 2020 ; Trinkle-Mulcahy, 2019 ; Ramanathan *et al.*, 2019). Dans ce travail seul un marquage de proximité par la biotine ligase a été utilisé et sera donc décrit. Le principe se base sur la conversion par la biotine ligase de la biotine en biotinyl-5'-AMP intermédiaire capable de lier de façon covalente les lysines situées à proximité. Dans un environnement réduit, la biotinyl-5'-AMP devient très rapidement inactive et ce, sur une distance maximale de 20nm (Ramanathan *et al.*, 2019). Les deux premières techniques centrées sur l'ARN étaient le BioID (proximity-dependent biotinylation identification) et l'APEX (engineered ascorbate peroxidase) à partir desquelles de nouvelles applications ont été réalisées augmentant l'efficacité et la spécificité de ces techniques d'identification de RBP. La première génération de BioID (**Fig. 10**) a été réalisée par ingénierie génétique d'une biotine ligase d'*E.coli* mutée R118G (BirA*) générant un site catalytique qui diminue la rétention de la molécule active (biotinoyl-5'-AMP). La protéine BirA* est couplée à un peptide qui est reconnu et se lie à un aptamère à proximité du site d'intérêt et peut ainsi biotinyler les RBP qui pourront ensuite être capturées grâce à des billes de streptavidine (Trinkle-Mulcahy, 2019). Le principe d'APEX est semblable à celui de BioID mais diffère par l'utilisation d'une enzyme différente. Dans ce cas, l'enzyme utilisée est une ascorbate peroxydase qui catalyse l'oxydation de biotine-phénol en radical biotine-phénoxy de courte durée (<1 ms) en présence de H₂O₂ (Nguyen *et al.*, 2019).

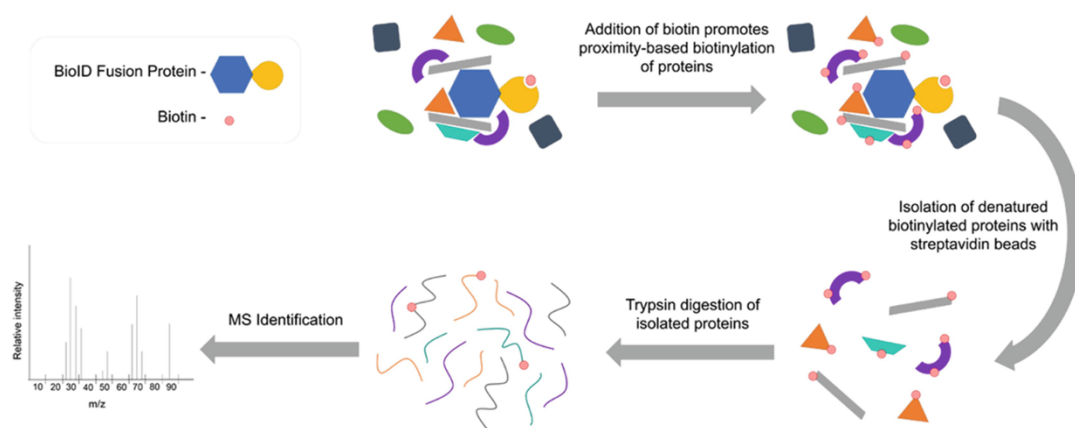


Figure 10: Schéma de l'utilisation de la biotine ligase pour un marquage de proximité. Une protéine reconnaissant une séquence spécifique est fusionnée à une biotine ligase. En présence de la cible et de biotine, la biotine ligase peut marquer les protéines à proximité. Une fois marquées, les protéines peuvent être isolées par billes de streptavidine. Les protéines purifiées peuvent être digérées et analysées par spectrométrie de masse (Sears *et al.*, 2019).

Table 2: Liste de différentes approches de marquage de proximité par la biotine ligase ou l'ascorbate peroxydase (Trinkle-Mulcahy, 2019).

Tag	Enzyme Activity	Source	Mutations	Size	Reference
BirA*	Biotin Ligase (promiscuous)	<i>E. Coli</i>	R118G	35 kDa	1
TurboID	Biotin Ligase (promiscuous)	<i>E. Coli</i>	Q65P, I87V, R118S, E140K, Q141R, S150G, L151P, V160A, T192A, K194I, M209V, M241T, S263P, I305V	35 kDa	30
MiniTurbo	Biotin Ligase (promiscuous)	<i>E. Coli</i>	N-terminal aa1-63 deleted (DNA binding domain); Q65P, I87V, R118S, E140K, Q141R, S150G, L151P, V160A, T192A, K194I, M209V, I305V	28 kDa	30
BioID2	Biotin Ligase (promiscuous)	<i>A. Aeolicus</i>	R40G	27 kDa	28
BASU	Biotin Ligase (promiscuous)	<i>B. Subtilis</i>	N-terminal aa1-65 deleted (DNA binding domain); R124G, E323S, G325R	28 kDa	29
APEX	Ascorbate peroxidase	Pea (synthetic)	K14D, W41F, E112K	28 kDa	9-10
APEX2	Ascorbate peroxidase	Soybean (synthetic)	K14D, W41F, E112K, A134P	28 kDa	15

Dans un premier temps, ces deux méthodes étaient destinées à marquer les interactions protéines-protéines dans les cellules vivantes, mais les avancées actuelles ont permis l'étude de complexes ARN-protéine tout en augmentant la spécificité et l'efficacité du système. Effectivement, l'ajout d'aptamères et d'une biotine ligase couplée à un peptide reconnu par l'aptamère a été une révolution pour la découverte des RBP (Trinkle-Mulcahy, 2019). Un problème souvent rencontré lors de l'utilisation de ces techniques est le temps nécessaire pour la biotinylation des RBP. Dans ce sens, les méthodes ont évolué pour pallier cette activité « lente » des biotine ligases. Les chercheurs ont donc développé des variantes pour améliorer les techniques comme TurboID, BioID2, APEX2, BASU, *etc.* (**Table 2**) (Trinkle-Mulcahy, 2019). Par exemple, le système TurboID est identique à celui de BioID mais la BirA* utilisée contient 14 mutations qui augmentent considérablement l'efficacité de marquage. En effet, alors qu'un marquage classique BioID par biotine dure jusqu'à 18h, le TurboID ne prend que 10 min pour obtenir des résultats similaires (Branon *et al.*, 2018).

Les avantages de ces techniques de marquage par proximité en comparaison des méthodes de capture par affinité sont : (1) Aucun composé toxique n'est nécessaire ; (2) La capacité à capturer les interactions faibles et/ou transitoires ; (3) La force de l'association de la biotine avec la streptavidine permet des extractions et des lavages plus rigoureux, diminuant les contaminants de fond (Trinkle-Mulcahy, 2019). Ces techniques par marquage de proximité peuvent être associées avec des approches de protéomiques quantitatives permettant l'identification et la comparaison des RBP (Majumder *et al.*, 2021).

Dans le cadre de ce travail, la nouvelle technique RaPID (RNA protein interaction detection) a été utilisée afin de détecter et d'identifier les RBP se liant aux motifs de la région 3'NTR du segment S de SBV (**Fig. 11**). Analogue des systèmes décrits précédemment, cette méthode comporte un marquage protéique dans les cellules vivantes qui est dépendant de la proximité est basé sur la biotine ligase. RaPID implique l'utilisation de deux composants : un aptamère (BoxB) et une protéine de fusion capable d'interagir spécifiquement avec les BoxB (**Fig. 11**). Les aptamères BoxB sont des courts ARN (15nt) de structure en forme de tige-boucle provenant du phage λ qui encadrent de part et d'autre (5' et 3') les motifs de l'ARN à étudier. La protéine de fusion est constituée d'une biotine ligase dérivée de *Bacillus subtilis* (BASU) fusionnée à un peptide N du phage λ se liant spécifiquement aux BoxB. Une fois les motifs de l'ARN entourés par les BoxB, l'expression de BASU permet la biotinylation des protéines se trouvant à une distance de ≈ 66 nt en présence de biotine. Une fois les RBP biotinylées, elles sont récupérées avec des billes couplées à la streptavidine dans des conditions strictes (Ramanathan *et al.*, 2018). En outre, la technique RaPID a été associée avec une approche de protéomique quantitative : la spectrométrie de masse en tandem par chromatographie en phase liquide à ionisation par électrobulisisation (LC/ESI-MS/MS). Un spectromètre de masse est un instrument analytique qui va aider à identifier et à quantifier des molécules présentes dans un échantillon en mesurant le rapport masse sur charge (m/z) et l'abondance des ions en phase gazeuse (**Fig.12**). Au même titre que la capture et la purification des RBP, l'identification des protéines est essentielle pour déterminer les interactants de l'ARN d'un virus. C'est une méthode quantitative ce qui permet la comparaison des RBP entre plusieurs conditions/échantillons donnés.

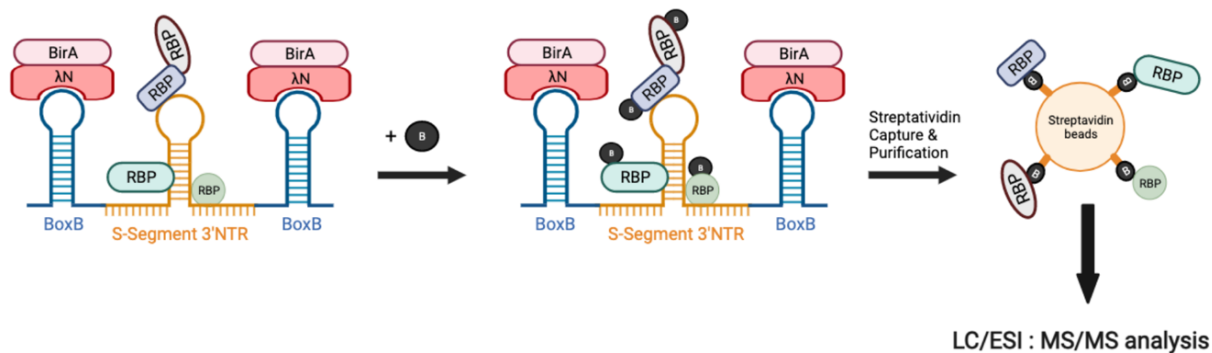


Figure 11: Schéma du principe de la technique RaPID appliqué au segment S de SBV. Utilisation de BoxB du phage λ (bleu) qui encadrent de part et d'autre les motifs ARN (jaune) dans la région 3'NTR. La biotine ligase BirA (rose) est fusionnée à un peptide (N) (rouge) du phage λ. Le peptide N reconnaît la BoxB et en présence de biotine peut marquer les protéines qui interagissent avec le motif ARN. Une fois marquées, les protéines sont capturées par des billes de streptavidine, purifiées et analysées par spectrométrie de masse en tandem (LC/ESI : MS/MS analysis).

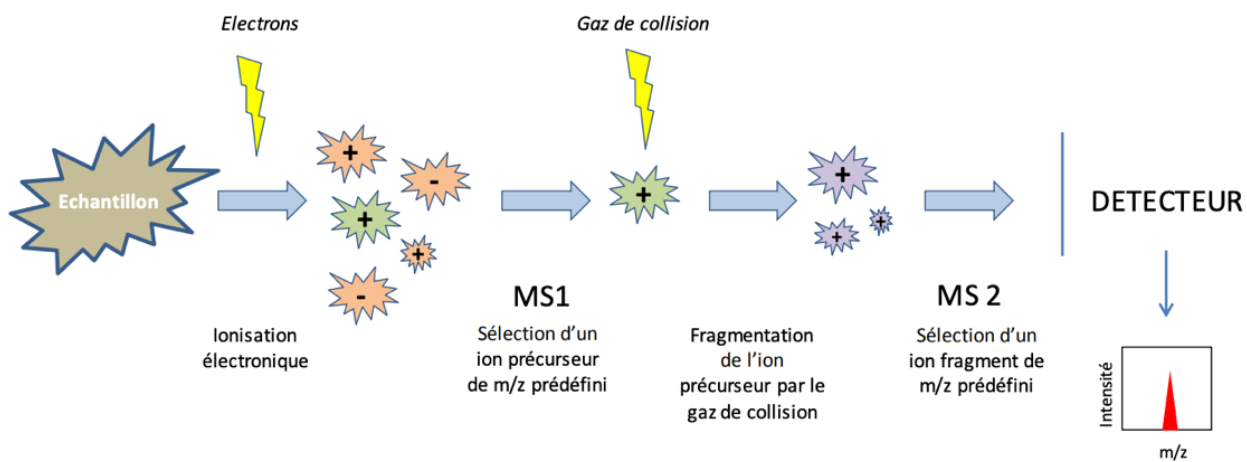


Figure 12: Principe de la fragmentation des composés par un spectromètre de masse en tandem (Willemin, 2014).

Matériel et méthodes

1. Lignées cellulaires et culture

Les cellules BHK-21 (Baby Hamster Kidney Cells) (ATCC® CCL-10™) sont des fibroblastes adhérents issus de rein de bébé hamster. Cette lignée cellulaire a été utilisée parce qu'elle est déficiente de la réponse à l'interféron de type 1, entraînant une grande sensibilité à de nombreux virus, dont le virus de Schmallenberg (Elliott *et al.*, 2013 ; Truant *et al.*, 1997). Le milieu de croissance utilisé est constitué du Glasgow Minimum Essential Medium (GMEM) complété par du tryptose phosphate broth (2,95g/L) (TPB), de la pénicilline (100 µg/ml), de la streptomycine (100µg/ml) et de 10% sérum bovin fœtal (FBS). Les cellules ont été incubées à 37°C dans une atmosphère contenant 5% de CO₂. La procédure de maintien de la lignée cellulaire consiste à retirer le milieu lorsque les cellules sont à confluence. Par la suite, du PBS-EDTA (phosphate buffered saline - Ethylene Diamine Tetra Acetic), un chélateur d'ions bivalents, a été ajouté pour laver l'excès de protéines sériques et affaiblir les jonctions intercellulaires dépendantes du calcium (4~10mL pendant 1 min.). Ensuite, le PBS-EDTA est retiré et la trypsine-EDTA est ajoutée. La trypsine inhibe l'adhésion intercellulaire en clivant les protéines d'adhésion et en chélatant les ions bivalents comme le calcium. Une fois les cellules détachées, du GMEM complet est ajouté pour diluer et inhiber l'effet de la trypsine afin de remettre les cellules en suspension dans de nouvelles flasques.

Les cellules BSR-T7 sont des cellules dérivées des BHK-21 exprimant de manière stable l'ARN polymérase T7 (Buchholz *et al.*, 1999). L'expression de la polymérase permet la production de virus recombinant par génétique inverse. (Elliot *et al.*, 2013). Les BSR-T7 ont été cultivées dans du GMEM complété par du TPB, 10 % de FBS, de la pénicilline (100 µg/ml), de la streptomycine (100µg/ml) et du G418 (200µg/ml).

2. Amplification virale et titrage

Trois flasques de 175 cm² de cellules de BHK-21 à 80% de confluence ont été infectées par le SBV avec une MOI de 0,01 (Multiplicity of infection) dans du milieu d'infection : Eagle's minimal essential medium (EMEM) supplémenté par 1% de FBS et 1% de PS. La faible quantité de FBS dans le EMEM rend ce milieu pauvre, ce qui facilite l'infection des cellules. Deux heures après infection, le EMEM a été remplacé par du milieu riche, le GMEM complet. 48h après infection, le surnageant a été récupéré et centrifugé à 4°C pendant 20min à 3000g afin de retirer les débris cellulaires. Finalement, le surnageant a été aliquoté et stocké à -80°C. Cinq échantillons de surnageant contenant du SBV ont été choisis aléatoirement et dégelés sur glace afin de calculer le titre viral du stock. Pour ce faire, des cellules BHK-21 ont été ensemencées sur des plaques 96 puits (0,34 cm²/puits) jusqu'à atteindre 70% de confluence. Les aliquots du surnageant contenant le SBV ont été dilués de 10⁻¹ jusque 10⁻⁸ pour infecter chaque ligne de la plaque avec 50µL de milieu d'infection. Après 2h d'incubation, 150µL de GMEM complet ont été ajoutés dans tous les puits pour une durée d'incubation de 72h. 72h après l'infection, les cellules ont été colorées au crystal violet permettant la distinction entre cellules mortes et les cellules vivantes. 50µL d'une solution de 4% de PFA (paraformaldéhyde)

a été ajoutée dans chaque puits pendant 30min pour fixer les cellules au support. Après 30min le PFA a été remplacé par 100µL d'une solution de crystal violet dans chaque puits. Le milieu est ensuite jeté et les cellules sont séchées pendant 24h. Pour finir, les virus ont été titrés par la méthode de dose infectieuse standard de 50% (TCID₅₀) permettant de déterminer la moyenne sur les 5 échantillons de SBV (Reed & Muench, 1938).

3. Transfections

Transfection des vecteurs dans des cellules BHK-21

Conformément à la technique RaPID (Ramanathan *et al.*, 2019), deux types de plasmide ont été nécessaires. Le premier constitue le backbone de clonage, appelé plasmide motif-ARN (Addgene plasmid #107253). Ce plasmide a permis la transcription des deux motifs en région 3'NTR à étudier sous le contrôle du promoteur fort du cytomégalovirus (CMV). Ce plasmide contient 3 BoxB (15nt) "GCCCTGAAAAAGGGC" qui encadrent de part et d'autre les deux motifs et le gène codant pour l'eGFP. Le second plasmide nommé RaPID (Addgene plasmid #107250) permet l'expression de la biotine ligase BirA modifiée de *Bacillus subtilis* (BASU) et de la protéine mCherry également sous le contrôle du promoteur CMV (Annexe 1.B). L'expression des deux gènes codant pour les protéines fluorescentes (eGFP et mCherry) ont permis de contrôler l'efficacité des transfections (**Fig. 13**).

Les cellules BHK-21 ont étéensemencées la veille dans des plaques 6 puits (9,6 cm²/puits) pour atteindre une confluence de 80%. Les cellules ont été transfectées avec un mélange contenant 8µg de vecteur, 1,5µg de plasmide RaPID et 18µL de LipofectamineTM 2000 (InvitrogenTM) suivant le protocole du producteur. Le milieu Opti-MEMTM (Thermo ScientificTM) a été utilisé pour permettre une transfection efficace. 24h après la transfection, la biotine a été ajoutée au milieu pour atteindre une concentration finale de 2µM et incubée pendant 18h.

Deux séries de triplicats ont subi une étape d'infection supplémentaire 20h après transfection. Les BHK-21 ont été infectées avec une MOI de 1 de SBV pendant 2h dans du milieu d'infection. Après les 2h d'incubation, le milieu a été remplacé par du GMEM complet et de la biotine en concentration finale de 2µM pendant 12h. Chaque condition a été réalisée en triplicat afin de permettre une étude statistique valide.

Transfection pour l'obtention d'un virus recombinant BoxB par génétique inverse

Les cellules BSR-T7 ont étéensemencées dans des plaques 6 puits (9,6 cm²/puits) pour atteindre une confluence de 70%. Les cellules ont été transfectées avec un mélange contenant 4µg de vecteur pUC57 comprenant le segment S recombinant, 4µg de segment L, 4µg de segment M et 18µL de LipofectamineTM 2000 (InvitrogenTM) suivant le protocole du producteur. 5h après transfection l'Opti-MEMTM a été remplacé par du GMEM complet pour une incubation de 72h à 37°C et 5% de CO₂.

Afin d'étudier la viabilité et la réplication du SBV recombinant, le surnageant de transfection en BSR-T7 contenant des potentiels virions a été utilisé pour infecter des BHK-21 qui permettent une meilleure réplication virale et des effets cytopathiques plus marqués.

Pour contrôler l'efficacité des transfections, les cellules ont été observées sous un microscope fluorescent à contraste de phase (Zeiss-Axiovert 40 CFL).

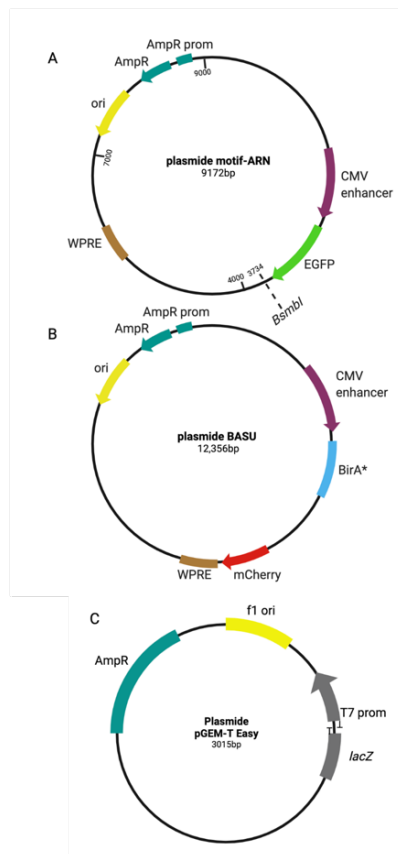


Figure 13: Schéma des différents plasmides utilisés. A) Plasmide motif-ARN comprenant le site de restriction *BsmBI*. B) Plasmide BASU comprenant la biotine ligase modifiée de *Bacillus subtilis*. C) Plasmide pGEM-T Easy comprenant le promoteur T7 dans le gène *lacZ*.

Amp : gène de résistance à l'ampicilline ; prom : promoteur ; CMV : Cytomégalovirus ; WPRE : l'élément de régulation post-transcriptionnelle du virus de l'hépatite de Woodchuck (WHV); eGFP : gène de la protéine verte fluorescente modifiée ; mCherry : gène de la protéine fluorescentes rouge monomérique ; *lacZ* : gène de l'opéron lactose.

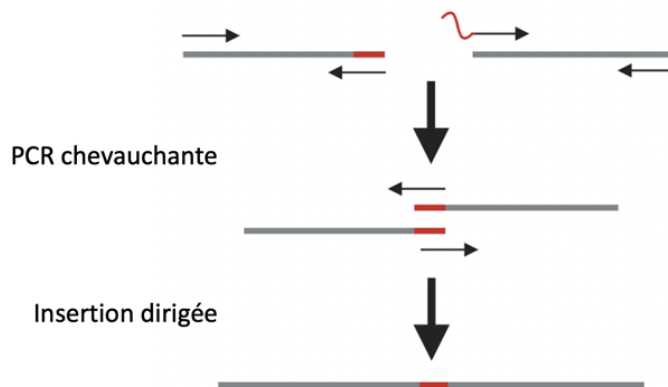


Figure 14: Schéma de mutagenèse dirigée. Les séquences sont représentées en gris et les nucléotides complémentaires en rouge. Les petites flèches désignent les primers.

4. Extraction d'ARN

L'ARN viral a été extrait suivant le protocole du fabricant "Macherey-Nagel : Viral RNA isolation ». Dans un premier temps, 150µL d'échantillon ont été mélangés avec 600 µL de RAV1 afin de lyser les virus pendant 5 min à 70°C. L'ARN a été mélangé à 600µL d'éthanol pur et fixé sur la membrane de silice par centrifugation 1min à 8000g. La membrane a ensuite été rincée à l'aide des buffers RAW et RAV3 et deux étapes de centrifugation de 1min à 8000g. L'échantillon a été finalement nettoyé et séché par l'ajout de 200µL de RAV3 centrifugé 4min à 11000g. L'ARN a ensuite été élué dans 50µL d'eau sans RNase.

Une fois l'ARN isolé, deux traitements successifs à la DNase I[®] (BioLabs[®]Inc.) ont été appliqués dans le but de retirer le maximum de plasmides résiduels suite à la transfection. Dans un volume final de 100µL, de l'eau traitée par du pyrocarbonate de diéthyle (DEPC), 10µL du buffer de DNase 10X (BioLabs[®]Inc.), 10µg d'ARN, 10U de DNase I[®] ont été incubés ensemble 20 min à 37°C.

5. Traitement poly(A) polymérase du virus recombinant BoxB

Lors l'amplification de la 3'NTR du virus recombinant BoxB, un traitement par l'*E. coli* poly(A) polymérase (BioLabs[®]Inc.) a été effectuée en suivant le protocole du fabricant (BioLabs : NEB# M0276). La solution de réaction comprend 5µg d'ARN isolé à partir du surnageant des infections en BHK-21, 2µL *E. coli* Poly(A) Polymérase Buffer 10X (BioLabs[®]Inc.), 1mM d'ATP, 5U d'*E. coli* Poly(A) polymérase et de l'eau pure (DEPC) dans un volume final de 20µL.

6. Purification de l'ARN par phénol : chloroforme

Après traitement DNase (BioLabs[®]Inc.) et poly-adénylation (BioLabs[®]Inc.), une purification phénol : chloroforme suivie d'une précipitation à l'éthanol des transcrits d'ARN sont réalisées pour éliminer les enzymes et les nucléotides libres restant. Le volume de réaction a été ajusté à 300µL avec de l'eau exempte de nucléase et 300µL de phénol : chloroforme : 3-méthylbutan-1-ol (ratio 24,5 : 24,5 : 1) ont été ajoutés. Après avoir bien mixé et centrifugé la solution pendant 15min à 4°C, la phase du dessus (aqueuse) a été récupérée. Le même volume de chloroforme : 3-méthylbutan-1-ol (ratio 49 : 1) a ensuite été ajouté à la phase recueillie et centrifugé à pendant 15min à 4°C. Par la suite, la phase aqueuse a été précipitée dans 2 fois son volume en éthanol pur et de l'acétate de sodium 0,3M final. La solution a été incubée pendant 1h à -80°C avant centrifugation de 15000g pendant 15min à 4°C. Une fois le surnageant retiré, la solution a été lavée dans 1mL d'éthanol 70% pour finalement être séchée et resuspendue dans de l'eau pure (DEPC).

Table 3: PCR et liste des amplicons et des primers associés. Chaque PCR et primers associés sont représentés par une séquence forward et une séquence reverse. Les lettres minuscules représentent un adaptateur ajouté aux primers. Les lettres majuscules représentent la séquence complémentaire à l'amplicon. Les lettres en gras représentent la séquence d'une BoxB.

PCR	Séquence primer Forward (5'-3')	Séquence primer Reverse (5'-3')	Taille (pb)
3'NTR S-segment	agctaagcttggagacgACCTCTTCAT CACAGATCTTCA	agctegtctctagctctAGTAGTGTTC ACTTATTAACATC	154
Screen pNTRs	CAAAGACCCCAACGAGAAGC	AGTAGTGTTCCTCACTTATTAA CTATCTTA	293
<u>Segment S recombinant</u>			
Segment A	TGTA AAAACGACGGCCAGTG	TTGAAGATCTGTGATGAAGAGG TTT	822
Segment B	tcttcacacagatcttaa GCCCTGAA AAAGGGCAAGCTTTTTCCGT	ctccacttattaactatcttaa GCCCTTTTTTC AGGGCAAGAGCTCAGAA	120
Segment C	TTAAGATAGTTAATAAGTGG AGAAC	GGAAACAGCTATGACCATG	235
Segment B-C	tcttcacacagatcttaa GCCCTGAA AAAGGGCAAGCTTTTTCCGT	GGAAACAGCTATGACCATG	355
Segment A-B-C	TGTA AAAACGACGGCCAGTG	GGAAACAGCTATGACCATG	1177
Screen A-B-C	TGTA AAAACGACGGCCAGTG	ctccacttattaactatcttaa GCCCTTTTTTC AGGGCAAGAGCTCAGAA	810
RACE-PCR	GCTAGAACCTTCTTGCAGCA ATTCG	GCTGTCAACGATACGCTACGTA ACG	
Nested PCR	TCAACATCTAAACCTCTTCAT CACAG	CGCTACGTAACGGCATGACAGT G	114-154
Screen RACE-PCR	TGTA AAAACGACGGCCAGTG	GGAAACAGCTATGACCATG	334-374
RT RACE Oligo d(T) RACE spécifique	GCTGTCAACGATACGCTACGTAACGGCATGACAGTG(T) ₂₄		60+
RT (BoxB post infection)	AGTAGTGTTCCTCACTTATTAACATCTTA		

7. Reverse transcription (RT) de l'ARN en ADN complémentaire

Les ARN obtenus après purification ont été rétro-transcrits en ADNc par la SuperScript IV Reverse Transcriptase (SSIV™, Invitrogen™). Une première RT pour valider la séquence du virus recombinant BoxB a été réalisée en utilisant un primer complémentaire au segment S du SBV (**Table 3**). Une seconde RT a été réalisée avec un oligo(d)T spécifique du kit GeneRacer™ pour l'étude de la 3'NTR (**Table 3**).

L'ARN a été mélangé avec 2,5µM d'oligo(d)T ou 0,1µM de primer, 1µM de dNTP, d et de l'eau DEPC pour « annealer » le primer et la solution d'ARN. Ce mélange a été incubé 5min à 65°C. Ensuite, 4µL de buffer SSIV 5X (Invitrogen™), 5mM de DTT, 40 U de RNaseOUT™ et 200 U de SuperScript IV™ Reverse Transcriptase, ont été ajoutés au mélange et incubés 10 min à 55°C. La réaction a ensuite été inhibée par une incubation de 10 min à 80°C.

8. Réaction de polymérisation en chaîne

Amplification des séquences d'intérêts pour l'identification des protéines

La PCR a été appliquée pour amplifier la séquence complète (120nt) de l'ADNc de la région 3'NTR du segment S à partir des plasmides pUC57. Deux amorces spécifiques ont été utilisées dans une concentration finale de 0,5µM (**Table 3**). Ces amorces comprennent des queues avec le site de restriction de l'enzyme pour la future digestion *BsmBI* (BioLabs®Inc.). L'ADN polymérase d'haute-fidélité Q5® (BioLabs®Inc.) a été choisie en raison de son haut niveau de fidélité. 35 cycles ont été effectués dans les conditions de température (NEB tcalculator) déterminées par la séquence des primers (**Table 3**).

Amplification pour l'obtention du virus recombinant BoxB

Trois segments différents (A, B et C) ont été amplifiés par l'ADN polymérase de haute-fidélité Q5® (BioLabs) afin de réaliser un virus recombinant à partir d'un segment S recombinant dans le vecteur pUC57. Les PCR ont été effectuées séparément dans des conditions de température propres à la séquence de leurs primers (**Table 3**). Une fois chaque segment amplifié, une seconde PCR a été réalisée pour hybrider les fragments entre eux, formant un segment A-B-C correspondant au segment S recombinant souhaité (**Fig. 14**). Pour ce faire, 10 cycles ont été effectués sans primers, permettant l'hybridation des séquences complémentaires entre elles, suivis de 25 cycles contenant les primers adéquats à l'amplification du segment complet (**Table 3**).

Amplification de la 3'NTR du virus recombinant BoxB

L'ADNc obtenu après RT du surnageant des infections en BHK-21 été amplifié par le protocole de RACE-PCR (rapid amplification of cDNA-ends by polymerase chain reaction) disponible chez le fabricant (Invitrogen™, GeneRacer™ Kit). Pour l'amplification, le protocole de l'ADN polymérase GoTaq® (Promega) a été utilisé. La PCR a été réalisée en utilisant un primer reverse capable de s'hybrider sur l'adaptateur de l'oligo d(T) spécifique du kit GeneRacer™ (**Table 3**) dans une concentration finale de 0,6M et un primer forward spécifique au segment S en concentration finale de 0,2M (**Table 3**).

Pour augmenter la spécificité de l'amplification, une seconde PCR a été réalisée directement sur le produit de PCR issu de la RACE-PCR. La PCR nichée suit le même protocole que la RACE-PCR mais contient des primers internes à l'amplicon précédemment amplifié. Le primer reverse spécifique de cette PCR se trouvait dans une concentration finale de 0,2M. Le primer forward était également spécifique au segment S et la concentration finale était de 0,2M (**Table 3**).

Tous les produits de PCR ont migré sur gel d'agarose (1,5 à 2%) pendant 45min à 100V pour vérifier l'amplification spécifique des inserts.

9. Purification des produits de PCR

L'étape de purification des amplicons a été réalisée directement à partir du produit PCR en suivant le protocole du fabricant "Macherey-Nagel : PCR clean-up and Gel extraction kit". Tout d'abord, le tampon NT1 a été ajouté à l'échantillon PCR (200µL de NT1 pour 100µL d'échantillon d'ADN). Le mélange obtenu a été transféré sur colonne contenant une membrane de silice par une étape de centrifugation de 1min à 11000g pour fixer l'ADN à la membrane. Par la suite, deux lavages avec 650µL de NT3 ont été effectués, chacun suivi d'une étape de centrifugation (11000g, 1min). Une nouvelle étape de centrifugation a été appliquée pour sécher la colonne (11000g, 2min). La dernière étape consistait à éluer l'ADN de la membrane de silice en ajoutant 18µL d'eau (DEPC) suivie d'une dernière étape de centrifugation 1min à 11000g.

10. Digestion des produits de PCR

La digestion des vecteurs a été réalisée sur les inserts en utilisant l'enzyme de restriction *BsmBI* afin de générer des extrémités compatibles avec le plasmide motif-ARN. A cette fin, la totalité de l'amplicon a été mélangée avec 2,5U d'enzyme de restriction *BsmBI*, 5µL de buffer de digestion (r3.1) 10x (NEBuffer™) et de l'eau (DEPC) pour atteindre un volume final de 50µL/échantillon. Une fois les inserts digérés, ils ont été purifiés en appliquant le même protocole de purification de l'ADN mentionné précédemment.

Le segment S recombinant a été digéré par les enzymes *KpnI* ET *BamHI* (BioLabs® Inc.) et le buffer de digestion 10x (NEBuffer r1.1), suivant le protocole du fabricant (BioLabs® Inc.).

11. Ligation des produits de PCR

La ligation des inserts dans le plasmide motif-ARN a été réalisée en utilisant l'ADN ligase T4 (Promega). 3µL d'inserts ont été mélangés avec 100ng de plasmide motif-ARN, 5µL de tampon 2X et 2U d'enzyme ADN ligase T4. Les produits de ligation ont ensuite été incubés pendant 12h à 12°C.

Le segment S recombinant a été ligué dans un vecteur pUC57 suivant le protocole de l'ADN ligase T4 (Promega).

Le produit purifié de la PCR nichée a directement été ligué dans un vecteur pGEM®-T Easy (Promega) par la méthode de « clonage TA » pour l'étude de l'impact de la BoxB sur la réplication virale.

12. Électroporation et culture de bactéries

Tous les produits de ligation ont été mélangés avec des bactéries *E. coli* TG1 compétentes (ratio 1:50) afin d'y introduire les plasmides contenant les inserts. Pour ce faire, les bactéries TG1 et les produits de ligation ont été placés dans une cuvette de 2mm pré-refroidie et ont été électroporés par un électroporateur ECM 319 BTX à 2500V, 25 μ F durant 5ms, perméabilisant la membrane des bactéries. Instantanément, 150 μ L de LB ont été ajoutés pour permettre la récupération des bactéries. Enfin, les *E. coli* transformées ont été cultivées pendant la nuit dans du LB-agar supplémenté en ampicilline (100 μ g/mL).

13. Clonage et criblage

Afin de sélectionner les bactéries qui possèdent le plasmide contenant l'insert, un criblage par PCR a été réalisé. Des colonies suffisamment isolées l'une de l'autre ont été prélevées au hasard sur boîtes de pétri et mises en suspension dans 25 μ L de mélange PCR préparé selon le protocole du fabricant (Promega). Il comprend 5 μ L de tampon vert 5x, une concentration finale de 0,2mM de dNTPs, 0,625U d'ADN polymérase GoTaq[®] (Promega) et 0,2 μ M de primer forward et de primer reverse (**Table 3**).

Lors de la sélection des vecteurs pour l'identification des protéines, le primer forward choisi était spécifique au plasmide et le primer reverse spécifique à l'insert, afin de vérifier l'orientation de l'insert dans le vecteur.

En ce qui concerne le clonage du segment segment S recombinant, il a été réalisé en utilisant 2 primers propres à la séquence du segment S du SBV (**Table 3**).

14. Amplification et purification d'ADN plasmidique

Les clones bactériens contenant les vecteurs d'intérêts ont été amplifiés pendant une nuit à 37°C dans 200mL de LB-ampicilline (100 μ g/mL) sous agitation constante à 180rpm.

Le protocole de purification a été réalisé par midiprep suivant le protocole du fabricant (Macherey-Nagel : « NucleoBond Xtra Midi kit »).

La première étape consiste en une centrifugation de la culture (4500g, 20min à 4°C). Le surnageant a été jeté et le culot a été remis en suspension dans 8mL de tampon RES. Ensuite, 8mL de tampon de lyse sont ajoutés pour perméabiliser les membranes. Après 5min à température ambiante, le tampon de lyse a été neutralisé par du tampon NEU. Pendant ce temps d'incubation, la colonne a été équilibrée en appliquant 12mL de tampon EQU. Par après, la solution neutralisée a été versée sur le bord de la colonne en gardant le précipité contenant l'ADN génomique dans le récipient. La membrane de silice contenant l'ADN plasmidique a été lavée une première fois en appliquant 5mL de tampon EQU et une seconde fois en utilisant 8mL de tampon WASH. Après élution des plasmides de la membrane de silice par le tampon ELU, 3.5mL d'isopropanol ont été ajoutés et mélangés soigneusement avant une étape de centrifugation (15000g, 30min à 4°C) pour culoter l'ADN plasmidique. Enfin, de l'éthanol (75°) a été utilisé pour laver le culot. Une dernière étape de centrifugation à sec (7500g, 5min à T° ambiante) a été réalisée et l'échantillon a été resuspendu dans 100 μ L d'eau DEPC. La qualité

et la concentration des plasmides ont ensuite été analysées par un spectrophotomètre NanoDrop (Thermo Scientific™) pour évaluer la qualité et la concentration des plasmides.

Pour vérifier les séquences des inserts, les échantillons ont été envoyés au séquençage (Eurofins Genomics) et alignés avec le logiciel *Geneious* prime, en les comparant avec la séquence de référence disponible du segment S (numéro d'accèsion HE649914).

Une fois les séquences vérifiées, les constructions destinées à l'identification des protéines ont été appelées pNTR_L-WT, pNTR_L-5P/3P, pNTR_L-5P, pNTR_L-GC/TT.

15. Identification des protéines par spectrométrie de masse

Extraction et dosage des protéines

Trois heures avant la fin de la transfection des vecteurs dans les cellules BHK-21 en présence de biotine, le milieu a été renouvelé avec du milieu complet (GMEM) pour éliminer l'excès de biotine. Les cellules ont été récoltées à l'aide de PBS-EDTA, puis ont été lavées deux fois avec du PBS 1X froid. Les cellules ont ensuite été lysées en appliquant 200µL de tampon de lyse (**Table 3**), 4µL de cocktail d'inhibiteurs de protéase 50x (cOmplete™, Roche) et 20µL de solution détergente Triton™X-100 Surfact-Amps™ 25% (Thermo Scientific™). Une étape de sonication (Amplitude 50, 10s) a été réalisée et 220µL de Tris froid pH7 (4°C) ont été ajoutés au lysat cellulaire, suivi d'une seconde étape de sonication (Amplitude 70, 10s). Le mélange a ensuite été centrifugé à 16000g pendant 10min à 4°C pour éliminer les débris cellulaires. Le surnageant a été transféré dans un nouveau tube collecteur et stocké à -80°C.

Le kit Pierce™ BCA Protein Assay Kit (Thermo Scientific™) a été utilisé afin de doser les extraits protéiques. C'est une méthode hautement sensible qui permet de déterminer la concentration en protéines d'une solution par un dosage colorimétrique grâce à la réduction du cuivre en présence des protéines dans un milieu alcalin. La droite de calibration a été effectuée sur des concentrations finales de BSA (bovine serum albumine) comprise entre 20 et 2000 µg/ml.

Les échantillons ont été mélangés (ratio 1:20) avec le « Working reagent » préparé selon les modalités du kit. Les extraits protéiques ont ensuite été incubés en absence de lumière pendant 30 min à 37°C. Les mesures colorimétriques ont été effectuées à l'aide d'un spectrophotomètre Multiskan® EX (Thermo Scientific™) à une longueur d'onde de 550nm.

Pull-down et digestion des protéines biotinylées

Le pull-down des protéines a été réalisé en utilisant les Dynabeads™ M-280 Streptavidin d'Invitrogen (Thermo Scientific™). Tout d'abord, les Dynabeads (100µL/échantillon) ont été lavées 2 fois avec du tampon de lyse (**Table 4**), suivies de 4 lavages à froid avec du PBS-50. Les billes lavées ont été remises en suspension dans 100µL d'échantillon de protéines et incubées pendant une nuit à 4°C sur roue. Le jour suivant, plusieurs étapes de lavages sur rack suivi d'incubations de 8 min sur roue à 24°C ont été réalisées : deux lavages dans du buffer 1 suivi d'un lavage dans du buffer 2, d'un lavage dans du buffer 3 et de deux lavages dans le buffer 4 (**Table 5**). Par la suite, les protéines biotinylées liées aux billes ont été digérées pendant une nuit dans une solution 20µL de buffer 4 contenant 500ng de trypsine à 450rpm 24°C. Ensuite, le surnageant contenant les protéines digérées a été transféré dans un nouveau tube

collecteur pour une seconde digestion dans une solution contenant 250ng de trypsine pendant 3h. Cette étape a été suivie de deux passages sur rack magnétique dans le but d'enlever la solution de billes résiduelle.

Purification des protéines sur Pierce C18 Spin Tips

Après avoir retiré toutes les billes, les protéines ont été purifiées à l'aide de Pierce™ C18 spin tips (Thermo Scientific™) pour l'analyse par spectrométrie de masse (MS). Avant d'effectuer la purification, le pH des échantillons doit être ajusté avec 2,5% d'acide trifluoracétique (TFA) pour obtenir un pH <4. Tout d'abord, les pointes C18 ont été placées dans des tubes de centrifugation et prétraitées en ajoutant 20µL d'une solution de 0,1% de TFA dans 80% d'acétonitrile (ACN) à 1000g pendant 1min. Les tips ont ensuite été rincés avec 20µL de TFA à 0.1% et centrifugés à 1000g pendant 30s. Après avoir transféré le tips C18 dans un nouveau tube collecteur, la solution comprenant les peptides d'intérêts (25µL) a été transférée sur tips et centrifugée deux fois (en réutilisant le flux) à 1000g pendant 1min. Les spin tips C18 ont ensuite été lavés deux fois en utilisant 20µL de TFA 0.1% et centrifugés à 1000g pendant 1min. Pour éluer les protéines, les tips ont été placés sur de nouveaux tubes en appliquant deux fois 10µL de TFA 0.1% dans 80% ACN pour une étape de centrifugation à 1000g pendant 1min. Un speed vacuum (ScanVac ScanSpeed40) a été utilisé pendant 12min à 45°C jusqu'à quasi-sécheresse des échantillons qui ont été resuspendus dans 0.1% TFA et 2% ACN:eau. Finalement, les peptides résultants ont été analysés par chromatographie en phase liquide à nanocapillaires couplée à l'ionisation en électrospray et à la spectrométrie de masse en tandem (LC-ESI-MS/MS : MaSUN-URBC).

Table 4: Composition du tampon de lyse stocké à T° ambiante.

Name	Final Concentration
NaCl	0.50M
Tris pH 4	0.05M
SDS	0.20%
DTT	1mM
H ₂ O	/

Table 5: Composition des tampons Wash stockés à T° ambiante.

<u>Wash Buffer1</u>	
Name	Final Concentration
SDS	2%
H ₂ O	/
<u>Wash Buffer2</u>	
Name	Final Concentration
Na-DOC	0.10%
Triton	1%
NaCl	0.5M
HEPES	0.05M
EDTA	1mM
H ₂ O	/
<u>Wash Buffer3</u>	
Name	Final Concentration
Na-DOC	0.50%
LiCl	0.25M
NP-40	0.50%
Tris pH4	0.01M
EDTA	1mM
H ₂ O	/
<u>Wash Buffer4</u>	
Name	Final Concentration
Tris pH 4	0.05M
H ₂ O	/

16. Analyse des données

Après la LC-MS/MS, les triplicats ont été regroupés par catégories de réplicats contenant les différentes conditions. Le fichier de données ainsi obtenu a été analysé à l'aide des logiciels tels que Scaffold 5.0 et STRING 11.0.

L'identification des protéines a été effectuée sur le logiciel Scaffold (Scaffold 5.0) en appliquant comme paramètres de filtrage un maximum de 1% de FDR (False Discovery Rate) et un minimum de 2 peptides associées par protéine. Les banques de données utilisées par le logiciel pour l'identification des protéines étaient *Bos spp.*, *Mesocricetus auratus* et le virus de Schmallerberg. La lecture des spectres des protéines identifiées s'est réalisée en comparant les valeurs quantitatives du nombre de spectres totaux normalisés sur l'ensemble des spectres identifiés. L'analyse statistique s'est effectuée par des tests de Student entre 2 catégories d'échantillons et par ANOVA (analysis of variance) entre les différentes catégories. La sélection des candidats protéiques a été établie lorsque les valeurs des variances entre catégories d'échantillons étaient validées par les tests et correspondaient à une p-value <0,05 et un \log_2 Fold-change <-2 ou >2.

Afin de détecter les interactants non représentatifs des techniques appliquées, le logiciel STRING 11.0 a été utilisé pour prédire les interactions protéine-protéine entre les protéines qui ont été sélectionnées après analyse statistique. La prédiction d'une association physique (directe) entre les acteurs a été appliquée avec un score d'interaction minimum requis de 0,700, représentant une confiance élevée.

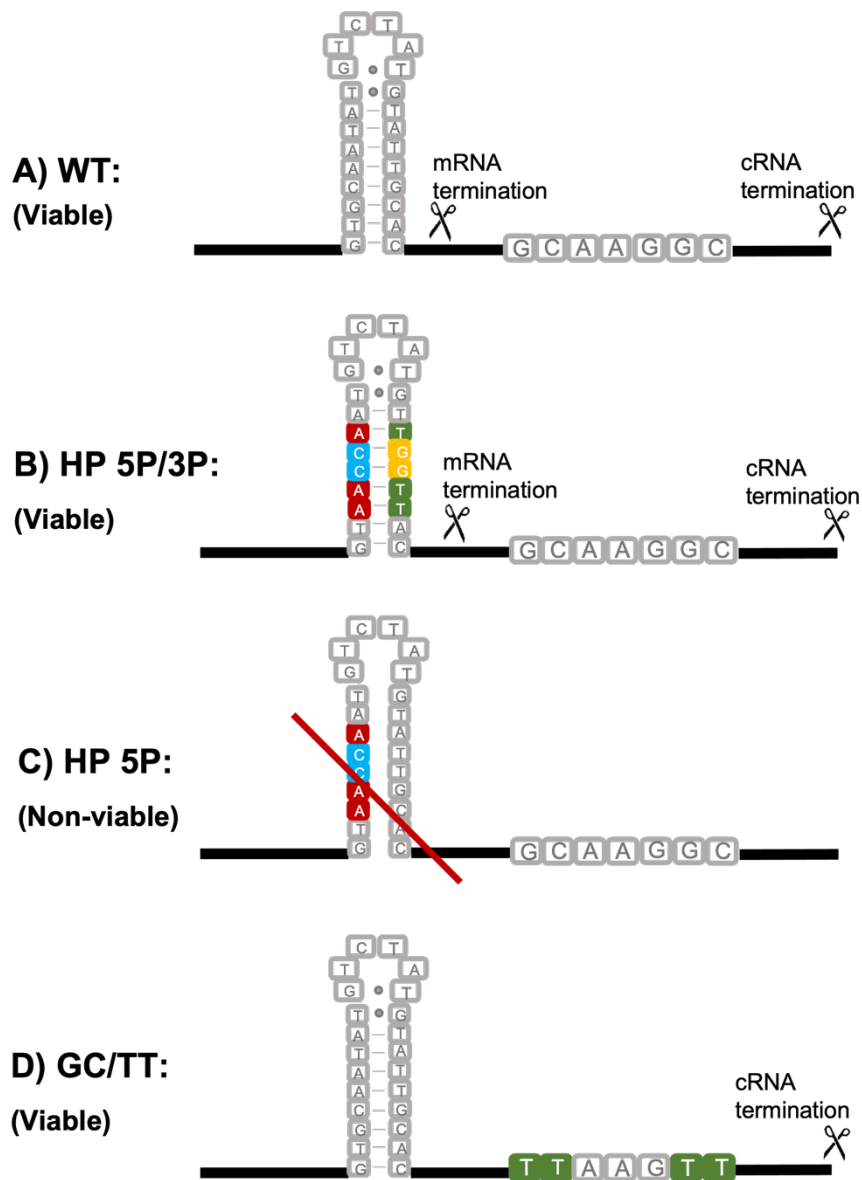


Figure 15: Schéma des motifs ARN en région 3'NTR de SBV WT et des mutants sélectionnés. Les terminaisons 3'NTR présentent en orientation 5'-3' une structure tige-boucle et un signal 'GC'. Les sites de clivage des ARNc et des ARNm sont indiqués par les ciseaux. Les mutants HP 5P/3P et GC/TT sont viables et le mutant HP 5P est non viable.

Résultats

1. Sélection des mutants de la 3'NTR du segment S de SBV et construction des plasmides motif-ARN

Afin d'étudier les RBP se liant aux motifs de la région non-traduite en 3' du segment S du SBV, une série de vecteurs ont été construits. Ces vecteurs comprennent la séquence sauvage (WT) du SBV ou des séquences mutées de ces motifs. Les différentes mutations (**Fig. 15**) ont été réalisées avec comme objectif de déstabiliser ou non la structure de la tige-boucle ou la séquence des motifs ARN.

Le choix des mutants et la validation de ceux-ci ont pu être réalisés grâce aux diverses études qui sont actuellement menées dans le laboratoire URVI au sein de l'Université de Namur. Un ensemble de mutants de la séquence 3'NTR ont été réalisés en effectuant des substitutions dans les deux motifs ARN (Bonil, données non publiées). L'objectif de ces mutations est de déterminer leurs impacts sur la cinétique de croissance virale. Par conséquent, des mutations ont été effectuées dans la séquence des bras ou de la boucle de la structure tige-boucle et dans le signal 'GC' (Bonil, données non publiées). Cet ensemble de virus mutés a servi de base pour réaliser la sélection des mutants utilisés pour la reconnaissance des protéines aux motifs ARN du SBV. Le choix final des mutants s'est basé sur des critères de modification ou d'inhibition de la séquence de la tige-boucle, d'une modification du signal 'GC' et d'une différence significative de la croissance virale entre les mutants et le virus WT (Bonil, données non publiées).

Le virus HP 5P/3P contient 5 mutations complémentaires sur chaque bras de la tige-boucle. Les mutations sont compensatoires et mènent à une structure intacte dont la séquence est modifiée. Les données du laboratoire indiquent des courbes de croissance, une dose infectieuse médiane et un phénotype similaires entre le WT et HP 5P/3P. Cela montre qu'une modification de la séquence nucléotidique mais une conservation de la structure de la tige-boucle produit un phénotype semblable (Bonil, données non publiées). Ce mutant a donc été sélectionné afin d'observer si le recrutement des RBP est structure dépendant ou non, par comparaison avec le WT (**Fig. 15.B**).

Le virus HP 5P présente 5 mutations sur l'un des deux bras, inhibant la formation de la tige-boucle. Les études menées à l'URVI ont montré que ce mutant n'était pas capable de se répliquer et qu'il était par conséquent non-viable. Les résultats observés sur les virus WT et HP 5P/3P en comparaison avec le virus HP 5P, indiquent que le rôle potentiel de la structure tige-boucle est structure dépendant et non séquence dépendant (Bonil, données non publiées). Ce mutant a donc pour objectif la discrimination des RBP se liant à la structure tige-boucle (**Fig. 15.C**).

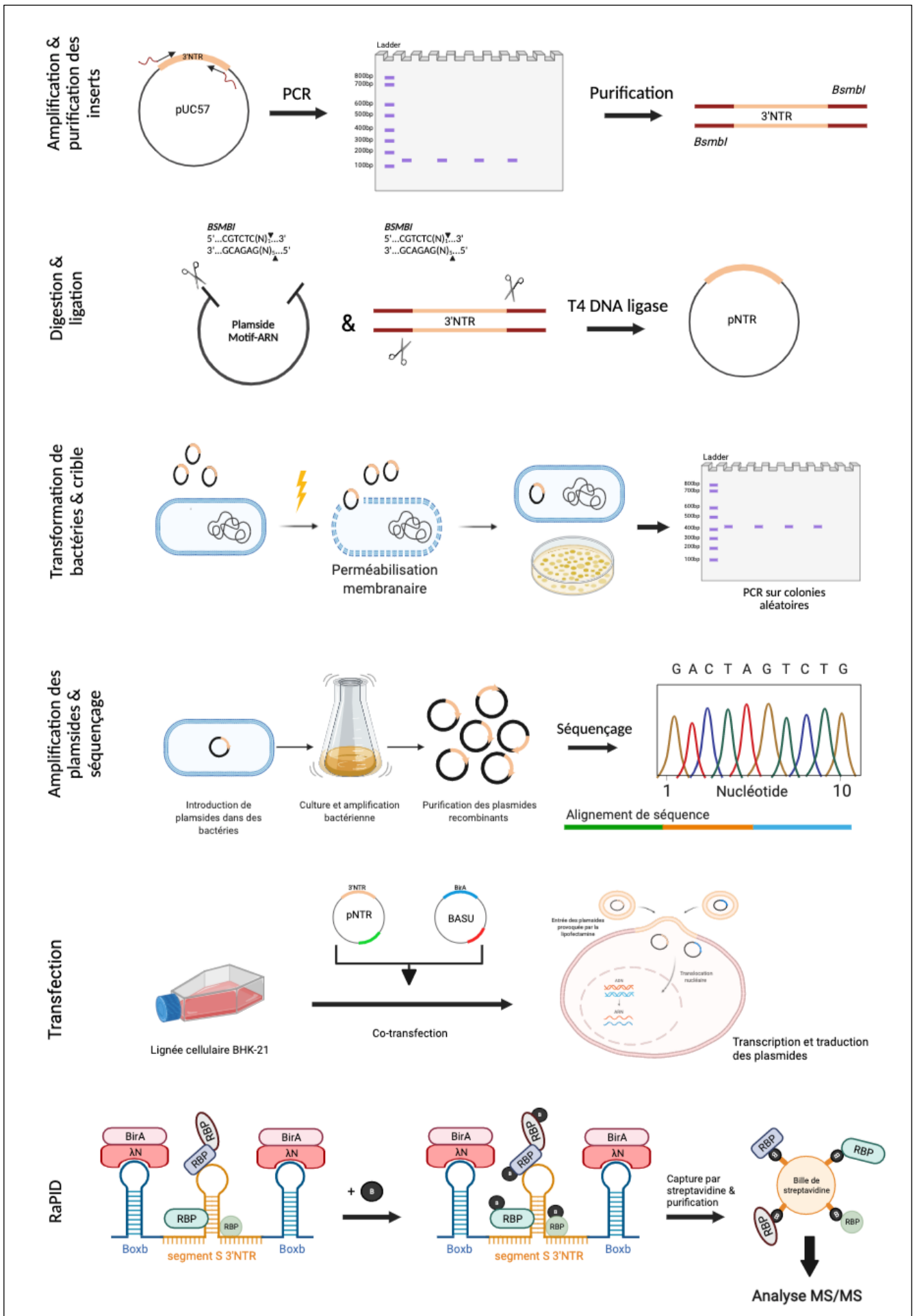


Figure 16: Vue d'ensemble schématique du déroulement des étapes, depuis l'amplification du segment S de la 3' NTR, jusqu'à la détection des interactions ARN-protéines (RaPID) et l'analyse des protéines.

Le virus GC/TT contient 4 substitutions dans le signal 'GC' (les 'GC' des extrémités sont substitués en 'TT'). Les données du laboratoire montre que ce mutant présente un phénotype atténué associé à une inhibition de la production d'ARNm. Par comparaison avec d'autres mutants de la séquence du signal 'GC', les études réalisées en URVI ont pu mettre en évidence que le rôle du signal était séquence dépendant (Bonil, données non publiées). Le mutant GC/TT a été sélectionné pour investiguer les RBP permettant le clivage potentiel de l'ARNc en ARNm (**Fig. 15.D**).

Une fois les mutants sélectionnés, nous avons réalisé de nouvelles banques plasmidiques comprenant l'entièreté de la 3'NTR (120nt) avec comme objectif la comparaison des RBP capturées par diverses constructions (pNTR_L). Les constructions ainsi générées (pNTR_L) seront comparées aux constructions déjà créées au laboratoire de l'URVI qui contiennent une séquence d'intérêt de 60nt entourant les motifs ARN (pNTR_{Sh}). Néanmoins, il nous a semblé judicieux de réaliser les pNTR_L pour éviter un encombrement stérique sur les séquences plus courtes.

Tout d'abord, chaque mutant a été amplifié par PCR à partir de plasmides pUC57 contenant les séquences d'intérêts (WT et mutants). Après l'étape de purification des inserts, le vecteur et chaque insert ont été digérés par *BsmBI* et ligaturés ensemble. Une fois les plasmides électroporés dans des bactéries compétentes, un crible par PCR sur des colonies choisies au hasard a été appliqué dans le but de confirmer la présence de l'insert dans le vecteur (**Fig. 16**). Une fois les séquences d'intérêt (WT et mutants) clonées dans les plasmides motif-ARN, les constructions ont été nommées pNTR_L (ex : pNTR_L-WT, pNTR-5P). Les pNTR_L ont été transfectés conformément à la méthode RaPID pour le marquage des RBP. La lignée cellulaire BHK-21, qui supporte des niveaux élevés de réplication du SBV, a été co-transfectée par deux plasmides : les pNTR et le plasmide BASU (**Fig. 13**). Un contrôle BirA seul a également été réalisé en transfectant uniquement le vecteur BASU afin d'identifier les protéines marquées uniquement grâce à l'expression de la biotine ligase. L'expression d'eGFP et de mCherry montre que les constructions pNTR et BASU, respectivement, sont bien transfectées au sein des cellules.

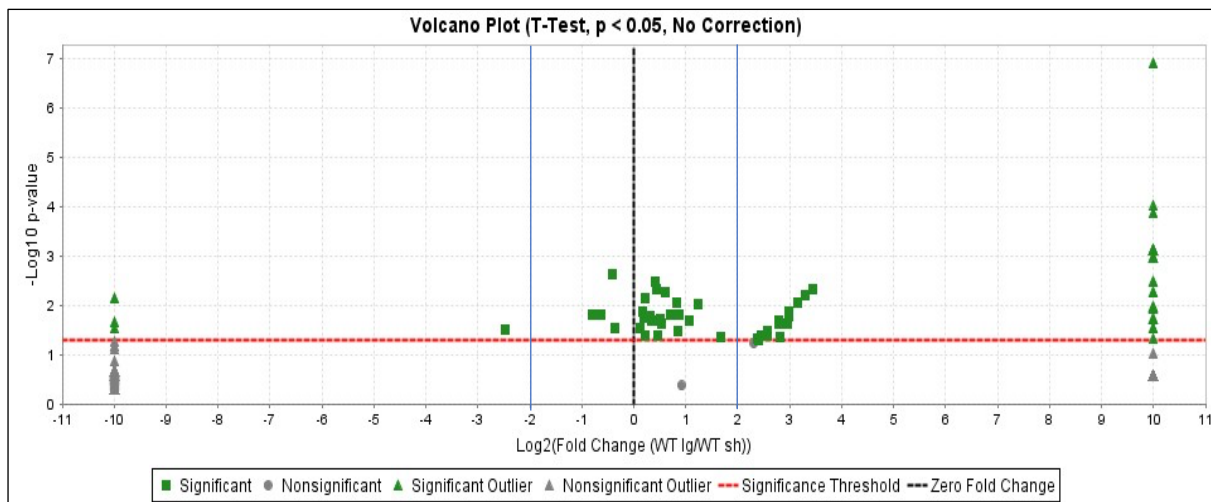


Figure 17: Volcanoplot comparant les protéines différentiellement exprimées identifiées entre WT_L et WT_{Sh} . Les lignes bleues correspondent à la limite minimum désignée (valeur absolue) du $\log_2 F_c$ pour la comparaison des protéines.

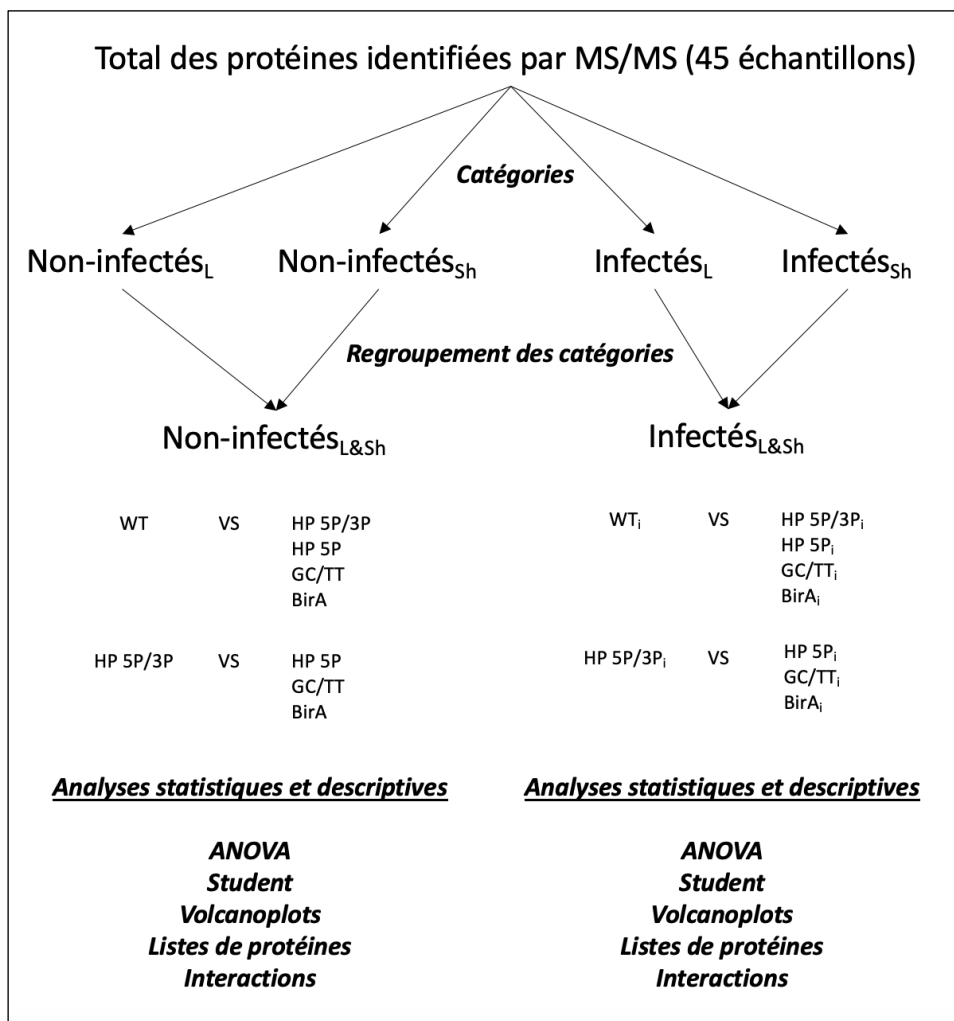


Figure 18: Schéma récapitulatif des différentes étapes pour l'analyse du RBPome interagissant avec les motifs ARN.

2. Impact de la longueur des constructions sur l'ensemble des protéines identifiées

Avant d'analyser l'impact des mutants sur l'interaction de partenaires protéiques, nous avons comparé les vecteurs longs (pNTR_L: 120nt) et les vecteurs courts (pNTR_{Sh}: 60nt) sur l'ensemble des protéines identifiées. Chaque condition a été réalisée en triplicat à l'exception de quelques conditions où l'identification des protéines n'a pas été possible sur un réplicat. Dans un premier temps, la quantité totale de protéines identifiées par le logiciel Scaffold était de 13564 protéines comprises dans 3377 clusters.

Afin de comparer l'efficacité de la biotinylation en fonction de la taille des constructions pNTR_L et pNTR_{Sh}, des volcanoplots ont été réalisés à partir des données extraites de Scaffold. Les volcanoplots sont basés sur un test statistique de Student et le Fold change (Fc) entre deux conditions sélectionnées pour le test. **La figure 17** montre peu de variation entre les catégories WT_L et WT_{Sh}. En effet, la majorité des protéines identifiées ont une valeur de log₂Fc comprise entre -2 et 2 et ont une p-value supérieure à 0,05 (encadré marron). Les valeurs aux extrémités du volcanoplot représentent les valeurs pour lesquelles les protéines sont présentes dans une catégorie mais absentes dans l'autre. La comparaison WT_L et WT_{Sh} montre que certaines protéines sont absentes en WT_L (**Fig.17**). Cependant, les valeurs des p-values et la lecture des spectres associés à ces protéines, nous ont permis de les exclure de la comparaison. La même analyse a été appliquée à toutes les autres catégories (pNTR_L vs pNTR_{Sh}) (données non montrées).

Après analyse comparative entre les vecteurs pNTR_L et pNTR_{Sh}, nous avons estimé qu'ils étaient similaires. Il a donc été décidé de regrouper ces catégories ensemble afin d'augmenter le nombre de réplicats par catégories et par conséquent, la puissance des tests statistiques. Les nouvelles catégories correspondent donc à aux deux vecteurs pNTR_{L+Sh} (pNTR) (**Fig. 18**).

3. Étude de l'enrichissement des termes de l'ontologie génique (GO)

Avant d'analyser et de comparer les sets de protéines identifiées, nous avons analysé quels clusters de fonctions cellulaires étaient enrichis lors de l'utilisation de la technique RaPID. Très peu d'annotations fonctionnelles existent pour les gènes de *Mesocricetus auratus*. C'est pourquoi, à l'aide du logiciel DAVID GO, nous avons converti les identifiants géniques en identifiants appartenant à *homo sapiens* et *Mus musculus*. Les sets de gènes étant semblables chez l'homme et la souris, nous avons continué l'étude des GO avec les identifiants de *Homo sapiens*. Une fois les listes de gènes d'*Homo sapiens* formées, nous avons comparé les catégories WT, BirA et 5P afin de déterminer quels clusters étaient enrichis lors de l'utilisation de la méthode RaPID (**Fig. 19**). Nous pouvons remarquer que le profil d'enrichissement du vecteur seul (BirA) est semblable à celui de WT. Les trois domaines comportant les scores d'enrichissement les plus élevés sont des clusters de fonction protéique impliqués dans les interactions cellulaires (~80), la maturation de l'ARN (~21) et les interactions avec l'ARN (~19).

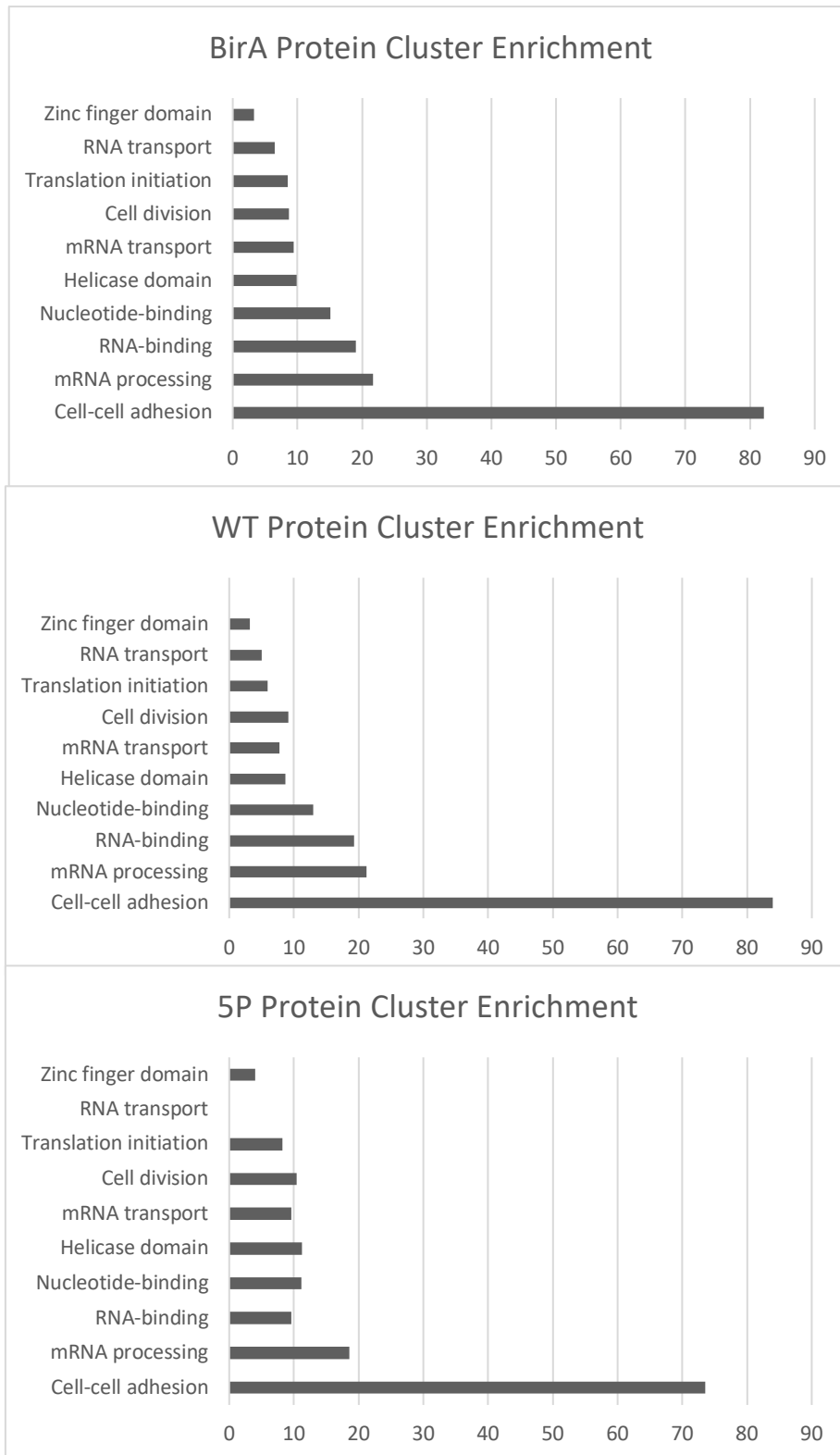


Figure 19: Histogramme de l'enrichissement des domaines et clusters protéiques (DAVID GO).

Pour la condition HP 5P, c'est-à-dire, le mutant inhibant la structure tige-boucle, le graphique montre un score d'enrichissement plus faible pour les protéines de liaison à l'ARN (9,62) (RBP). En ce qui concerne les protéines impliquées dans le transport de l'ARN, aucun enrichissement n'a été déterminé par le logiciel DAVID pour la condition HP 5P (**Fig. 19**).

4. Impact des mutations sur la capture et l'identification des protéines

L'objectif étant de comparer les RBP interagissant avec les constructions WT et mutantes, la première étape a été de vérifier si les catégories étaient comparables entre elles. Pour ce faire, un histogramme a été réalisé, mettant en avant le nombre de protéines identifiées par catégorie (**Fig. 20**). Le nombre de protéines identifiées varie entre 2773 (WTi) et 4470 (5P/3P) avec en moyenne 3531 protéines identifiées par catégorie (**Fig. 20**). A partir de ce résultat, nous avons estimé que nous pouvions poursuivre l'investigation sur le rôle des motifs ARN par comparaison de protéines entre les catégories.

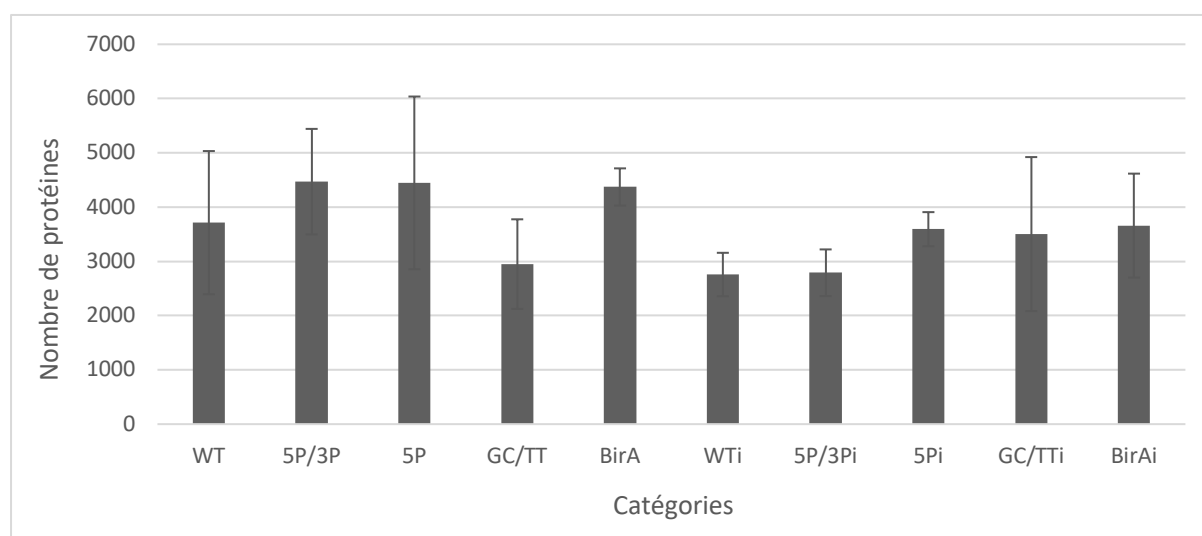


Figure 20: Histogramme représentant le nombre moyen de protéines identifiées par catégorie d'échantillons. Les barres d'erreurs sont associées aux écarts type.

Analyse comparative des protéines identifiées en conditions non-infectées

Chaque analyse comparative a été réalisée entre deux catégories sélectionnées et par un test statistique de Student. Les protéines différenciellement identifiées entre les catégories correspondent à celles validées par les test T ainsi qu'exprimant un $\text{Log}_2(\text{Fc}) < -2$ ou $> +2$. Les volcanoplots issus de Scaffold montrent les variations statistiques entre les protéines identifiées dans la catégorie WT en comparaison des catégories HP 5P/3P, HP 5P GC/TT et BirA (**Fig. 21**). De manière générale, les volcanoplots montrent une grande similitude entre les différentes constructions (**Fig. 21**). Le vecteur BirA ne comprend que l'expression de la biotine ligase sans la présence de construction et représente un contrôle d'autobiotinylation du système. Seulement 2 protéines ont été identifiées de manière significativement supérieure dans la catégorie WT comparée à la catégorie BirA : les protéines Helz et MBNL1 (**Table 6**). L'analyse statistique comparative entre la catégorie WT et la catégorie GC/TT, a mis en

évidence 14 protéines surexprimées en WT dont les protéines Helz et MBNL1 (**Table 6**). La comparaison entre les catégories WT et 5P a uniquement affiché une seule protéine surexprimée en WT : la protéine Dsc1.

L'analyse a également été réalisée entre la catégorie 5P/3P et les catégorie de mutants (**Fig. supplémentaire 1**).

Table 6: Gènes correspondant aux protéines différenciellement identifiées par comparaison de la construction WT(i) et des mutants.

Catégories comparées	Gènes ID
WT vs BirA	Helz, MBNL1
WT vs 5P	Dab2, Pspc1, Nelfe
WT vs GC/TT	Pgd, Srsf2, Esyt1, Rfx5, Nhlrc2, Helz, Micall2, MBNL1, Exosc9, Eya3, H671_3g9162, Tjap1, Rab3gap2
WTi vs 5Pi	Dsc1
WTi vs GC/TTi	Sarnp, TAF15, Dsc1, Mustn1

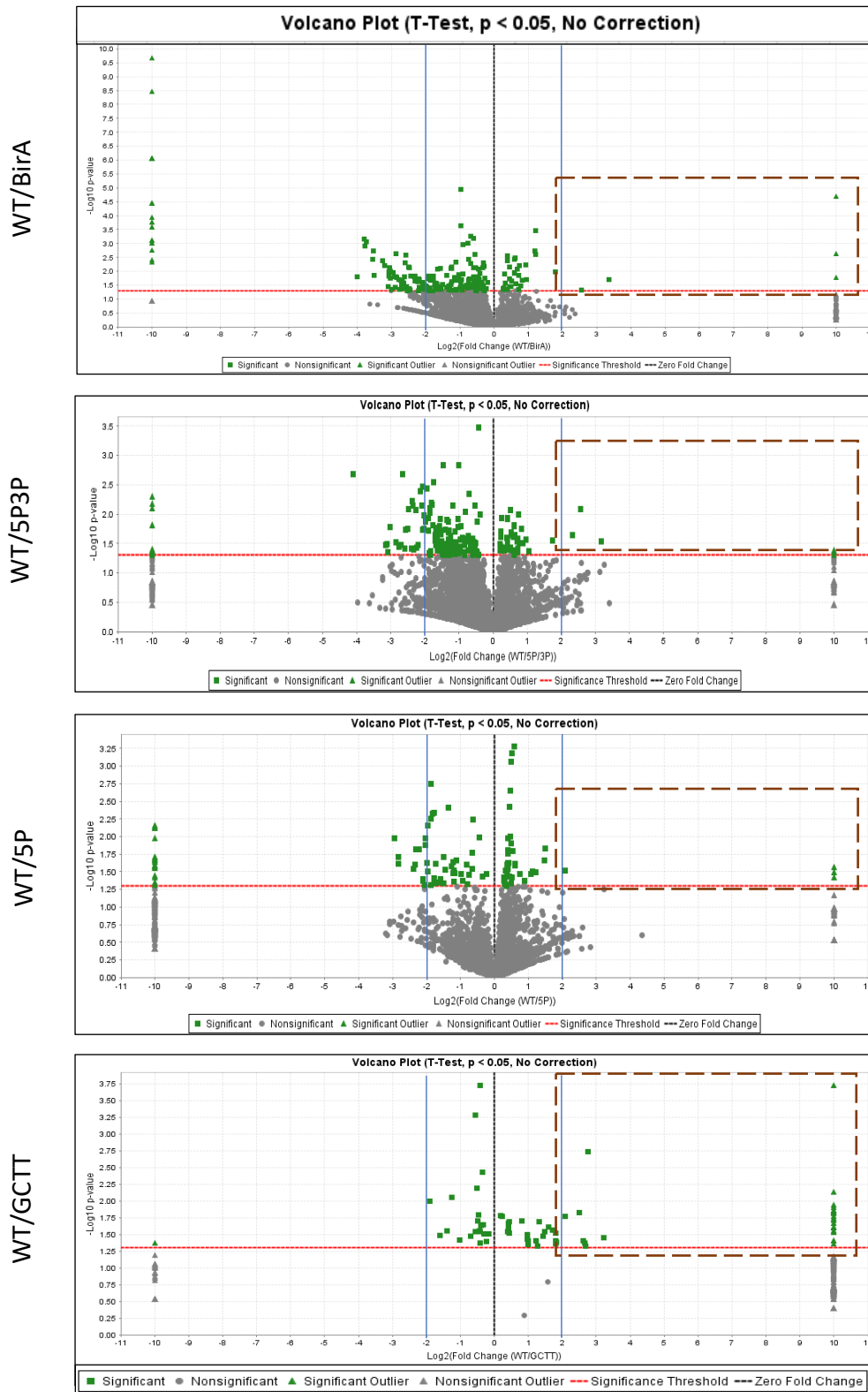


Figure 21: Volcanoplots des protéines différemment exprimées entre les catégories. La ligne rouge correspond au seuil de p-value au-dessus duquel les protéines sont sélectionnées. Les lignes bleues correspondent à la limite minimum désignée (valeur absolue) du $\log_2 Fc$ pour la comparaison des protéines. L'encadré en pointillés marron représente les protéines sélectionnées lors des comparaisons entre la condition WT et les conditions mutantes.

Analyse comparative des protéines identifiées en conditions d'infection

Pour identifier les RBP qui interagissent avec nos constructions, nous avons également réalisé la même étude sur un modèle de BHK-21 infecté par le SBV en parallèle du protocole expérimental. En effet, les cellules ont été infectées 2 heures avant l'ajout de la biotine pour une durée totale de 12 heures. Toutes les étapes décrites précédemment ont également été effectuées sur les catégories infectées, menant à un regroupement des pNTR_L et pNTR_{Sh} infectés (i) (**Fig. 22**). Bien que les volcanoplots (**Fig. 22**) affichent des tendances similaires lors de l'utilisation d'un modèle infecté et non-infecté, les protéines différenciellement identifiées ne sont pas identiques (**Table 6**). En effet, aucune protéine n'a été statistiquement validée lors de la comparaison des catégories WT_i et BirA_i. La comparaison entre les catégories WT_i et 5P_i a uniquement affiché une seule protéine surexprimée en WT : la protéine Dsc1. Cette même protéine a également été retrouvée aux côtés des protéines Sarnp, TAF15, Mustn1 après l'analyse comparative entre les catégories WT_i et GCTT_i (**Table 6**).

L'analyse a également été réalisée entre la catégorie 5P3P_i et les catégorie de mutants (**Fig. supplémentaire 1**).

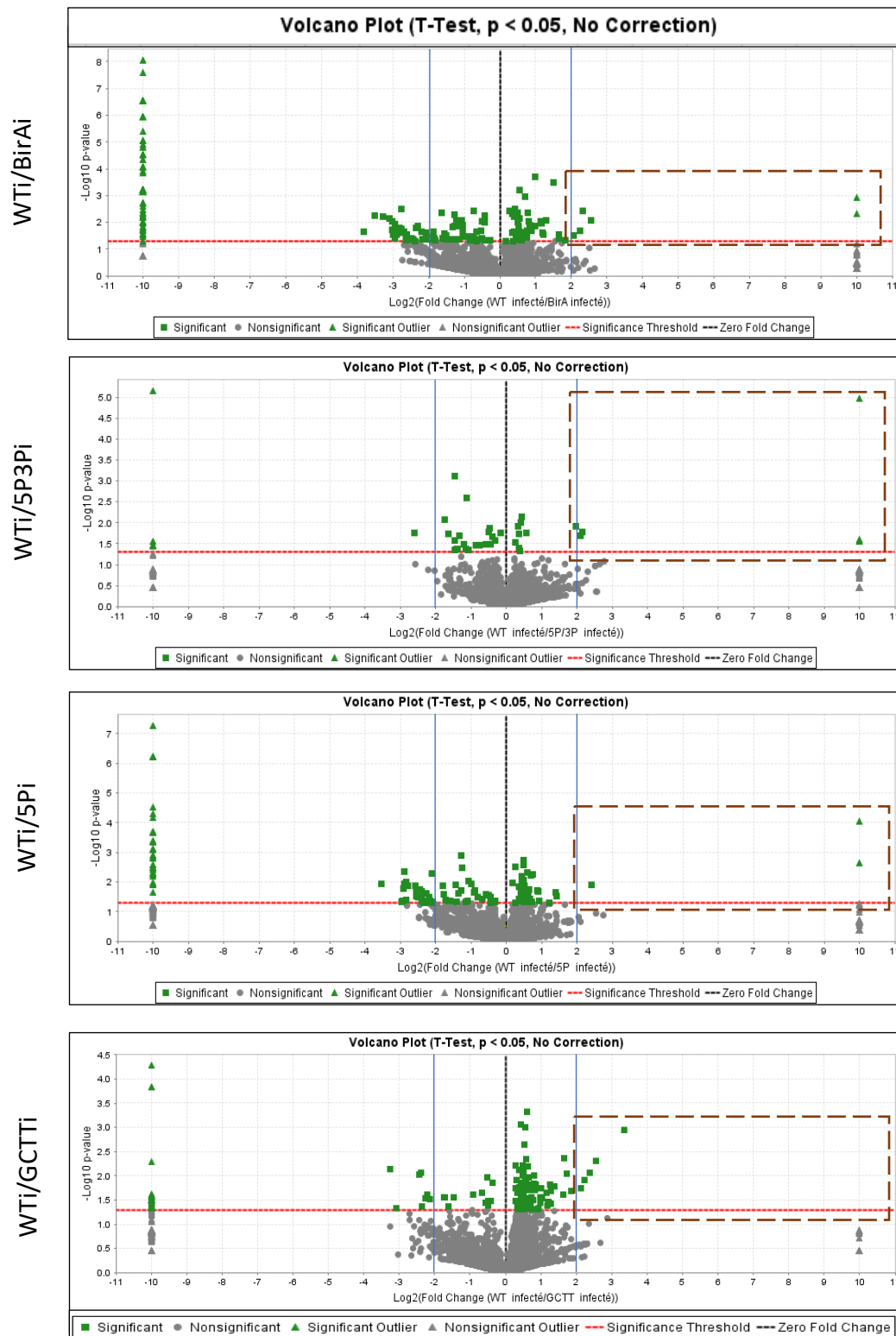


Figure 22: Volcanoplots des protéines différentiellement exprimées entre les catégories en condition d'infection. La ligne rouge correspond au seuil de p-value au-dessus duquel les protéines sont sélectionnées. Les lignes bleues correspondent à la limite minimum désignée (valeur absolue) du $\log_2\text{Fc}$ pour la comparaison des protéines. L'encadré en pointillés marron représente les protéines sélectionnées lors des comparaison entre la condition WT et les conditions mutantes.

5. Construction d'un segment S recombinant BoxB par mutagenèse dirigée

Après analyse des catégories infectées et non-infectées, peu de différence entre les protéines ont été observées. Nous avons tenté une approche alternative pour la technique RaPID afin de coupler la technique de marquage de proximité à une infection virale en conditions normales d'infection. Pour ce faire un virus recombinant BoxB a été réalisé par génétique inverse. L'objectif étant d'obtenir un segment S recombinant qui possède une unique BoxB de part et d'autre des deux motifs à étudier en région 3'NTR. La création du mutant s'est réalisée à partir de deux vecteurs, le pUC57 et le pNTR_{Sh}-WT (**Fig. 23**). Dans un premier temps, plusieurs PCR séquentielles et chevauchantes ont été effectuées sur 3 séquences distinctes, nommées Bloc A, B et C. La première séquence amplifiée à partir d'un vecteur pUC57 (Bloc A : 822nt) comprend la séquence codante du segment S ainsi que le promoteur T7 (**Fig. 23**). La seconde séquence, issue de pNTR_{Sh} (Bloc B : 150nt), correspond aux deux motifs ARN dans la 3'NTR, encadrés par une seule structure BoxB de part et d'autre de cette région (**Fig. 23**). Le Bloc B a été amplifié en utilisant des adaptateurs spécifiques complémentaires avec l'extrémité du 3' du Bloc A et l'extrémité 5' du Bloc C. La dernière séquence (Bloc C : 256nt) correspond à la fin de la région 3'NTR du segment S et la portion du vecteur pUC57 comprenant la séquence du ribozyme (**Fig. 23**).

Après amplification et migration sur gel de chaque Bloc (A, B et C), les produits de PCR ont été utilisés pour effectuer des PCR chevauchantes afin de liguer les Blocs les uns aux autres. Dans ce sens, deux PCR distinctes ont été effectuées, une première comportant les Blocs A et B et une seconde comportant les Blocs B et C. Les adaptateurs présents sur la séquence du Bloc B ont permis l'hybridation et l'amplification du Bloc A-B et du Bloc B-C (**Fig. 23**). Une fois les deux fragments amplifiés, une dernière PCR du Bloc A et du Bloc B-C a été réalisée afin d'amplifier la séquence du nouveau Bloc A-B-C par mutagenèse dirigée. Nous avons décidé de réaliser la dernière PCR à partir du fragment B-C car la bande d'amplification était beaucoup plus intense. Le segment S recombinant ainsi généré a été cloné et vérifié par alignement de séquences, confirmant la présence d'une BoxB de part et d'autre des motifs ARN.

6. Élaboration d'un virus recombinant BoxB par génétique inverse

Le virus recombinant BoxB a été réalisé par génétique inverse (**Fig 24**). Pour ce faire, les segments génomiques L (WT) et M (WT) et S (WT ou recombinant), chacun compris dans un plasmide pUC57, ont été co-transfectés dans des cellules BSR-T7 (**Fig. 24.A**). La lignée BSR-T7 est utilisée car elle permet l'expression constitutive de l'ARN polymérase T7.

Le virus a été amplifié dans des cellules BHK-21 en utilisant le surnageant de la transfection des cellules BSR-T7. Après l'infection, la lignée BHK-21 affichait des effets cytopathiques (données non montrées), démontrant la capacité répliquative du virus recombinant BoxB. L'ARN viral a été extrait et séquencé afin de s'assurer de la présence des BoxB. Après la réalisation d'un stock du virus recombinant, nous avons titré ce virus pour de futures manipulations à dose virale contrôlée (**Fig. 24.C**).

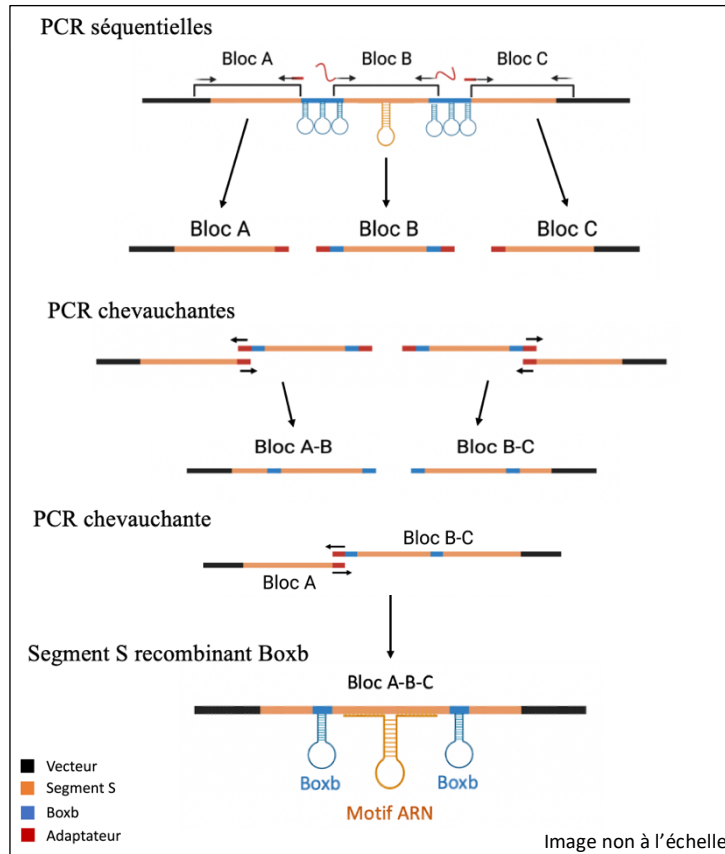


Figure 23: Schéma des différentes étapes pour la construction du segment S recombinant BoxB.

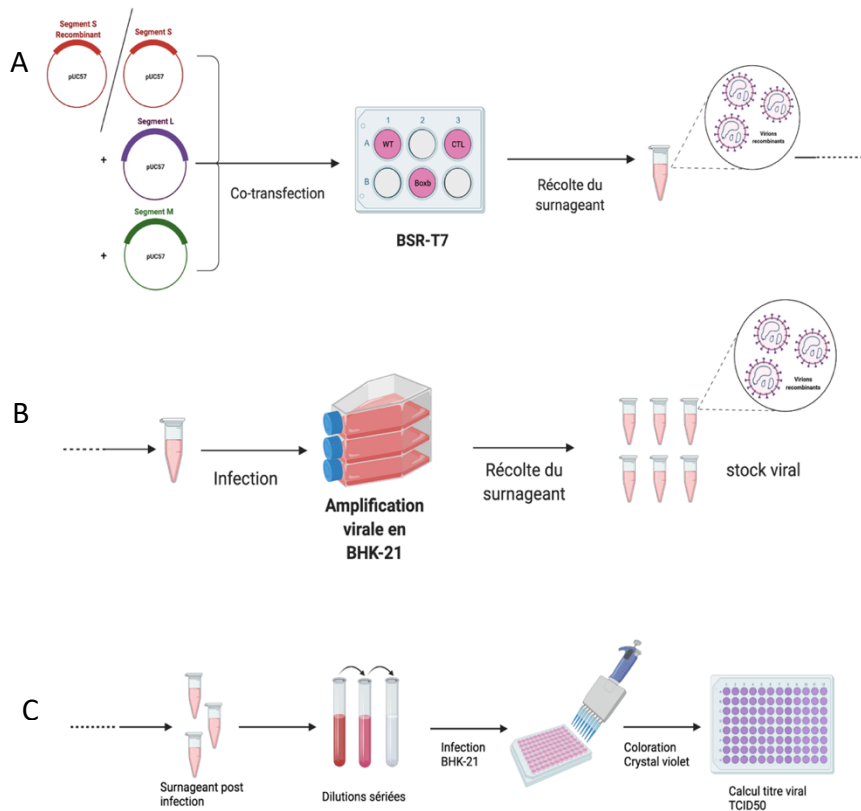


Figure 24: Vue d'ensemble schématique du déroulement de la création d'un virus recombinant par génétique inverse jusqu'au calcul du titre viral par la méthode TCID₅₀. A) Génétique inverse par co-transfection de trois plasmides contenant chacun un segment génomique du virus. B) Amplification virale. C) Méthode de titrage par TCID₅₀.

7. Évaluation de la dose infectieuse médiane du SBV recombinant BoxB

Le dosage de la charge virale du virus recombinant BoxB a été calculée par la méthode de dose infectieuse standard de 50% (TCID₅₀) de Reed et Muench. Les puits ont été ensemencés avec des cellules BHK-21 qui ont ensuite été infectées par des virions issus du stock après l'amplification virale (**Fig. 24.C et Fig. 25**) Après coloration et rinçage, chaque puits non coloré correspond à un puits où les cellules sont mortes ce qui correspond à une capacité répliquative du virus. Nous constatons de la lyse cellulaire jusqu'aux dilutions 10⁻⁴ de l'inoculum viral pour les plaques A et B. (**Fig. 25**). Le contrôle sans virus affiche uniquement des puits colorés, reflétant l'absence de virus. Le titre viral a été calculé pour l'ensemble du stock sur 5 aliquots pris aléatoirement et s'élève à 9,55 x 10⁴ TCID₅₀/mL. Le virus recombinant BoxB est donc viable mais atténué lorsqu'on compare son TCID₅₀/mL à celui du SBV WT qui atteint en moyenne 3 x 10⁶ TCID₅₀/mL

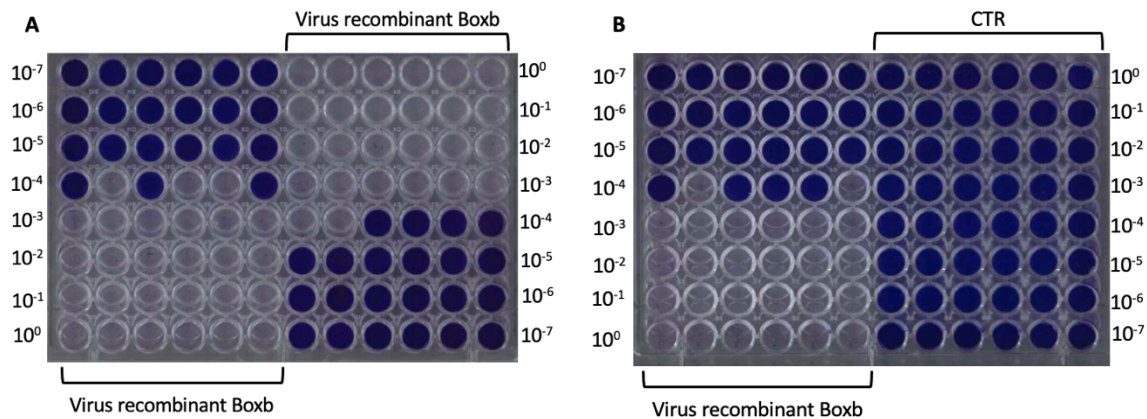


Figure 25: Méthode TCID₅₀ des SBV recombinant BoxB sur plaque 96 puits. A) Dilutions sériées de deux stocks de virus. B) Dilutions sériées d'un stock viral et du contrôle négatif. Les puits transparents représentent les cellules mortes et les puits colorés représentent les puits où les cellules sont vivantes fixées au support.

8. Étude de la région terminale 3' du segment S du virus recombinant BoxB

Après avoir calculé le titre viral du virus recombinant BoxB, nous avons étudié l'impact de l'ajout des BoxB dans la 3'NTR sur la terminaison de transcription. L'analyse des pools d'ARN a été réalisée par une RACE-PCR permettant l'identification de l'extrémité des ARN. Après purification et séquençage du produit de la PCR nichée (**Fig. 26**), les séquences ont été alignées sur une séquence de référence pour une étude comparative de la terminaison de transcription des clones (**Fig. 27**).

La **figure 26** montre que sur un pool de 30 clones issus du SBV WT, 10 possèdent un ARN non-tronqué (33%). 17 clones WT terminent leur transcription en amont du signal 'GC' sur un site de clivage 'CA', ce qui représente 56% des clones WT (**Fig. 26**). 23 clones ont été séquencés afin de déterminer les extrémités des mutants BoxB. 10 clones présentent une

transcription du segment S non-tronquée (ARNc) (43%). 13 clones (56%) ont une fin de transcription prématurée. Ces terminaisons se localisent aléatoirement dans la région 3’NTR du segment S recombinant et ne correspondent pas au clivage observé dans les ARNm du virus WT.

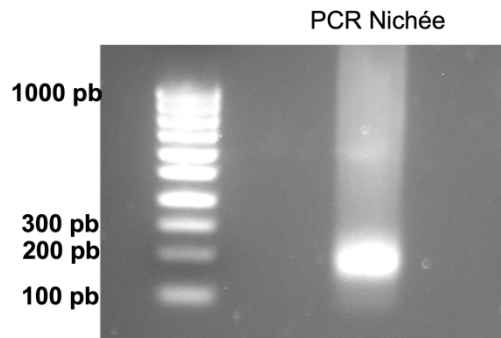


Figure 26: Gel de PCR nichée. Amplification de la 3’NTR du virus recombinant BoxB.

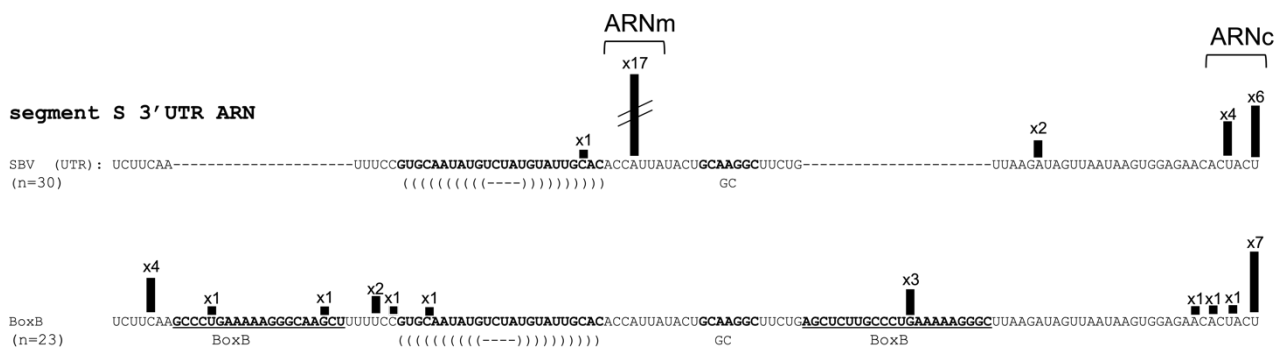


Figure 27: Étude de l'extrémité 3'UTR par RACE-PCR. Comparaison des profils de pool d'ARN entre le SBV WT et le SBV recombinant BoxB. Les barres noires indiquent les nucléotides terminaux par nombre de clones donnés. Les séquences en gras correspondent au signal 'GC', la structure en tige-boucle et les BoxB. Les parenthèses désignent les nucléotides complémentaires formant les bras de la structure en tige-boucle. Les séquences correspondant aux BoxB sont soulignées.

Conclusion et perspectives

Dans cette étude nous avons mis en œuvre différents protocoles expérimentaux afin de mettre en évidence les mécanismes régulant la transcription et la traduction du virus de Schmallenberg. Le SBV est un virus possédant un génome à ARN tri-segmenté de polarité négative (Elliott, 2014). Lors de la réplication virale, les segments génomiques sont transcrits en ARNc et ARNm de polarité positive (Elliott, 2014). L'ARNc est utilisé comme brin matrice pour reformer du génome alors que l'ARNm sert de brin codant pour la traduction des protéines structurales et non-structurales. Les ARNm du SBV sont coiffés en 5' mais ne possèdent pas de queue poly(A) à l'extrémité 3' (Coupeau *et al.*, 2013 ; Walter & Elliott, 2011 ; Vera-Otarola *et al.*, 2010). Le segment S du SBV code pour la nucléoprotéine N. Lors de l'infection par les bunyavirus, la protéine N est une des protéines les plus exprimées par la cellule hôte (Jonsson *et al.*, 2010). Elle est essentielle pour la réplication du SBV, la formation des virions et la stabilité génomique (Kohl *et al.*, 2006). À l'heure actuelle, peu de choses sont connues sur les mécanismes cellulaires utilisés par le segment S pour réguler son expression, mais des motifs importants et conservés ont été identifiés (Coupeau *et al.*, 2013).

Dans un premier temps, une structure en tige-boucle à l'extrémité 3'NTR des ARNm a été mise en évidence. Les ARNm du SBV présentent une extrémité 3' écourtée en comparaison avec les ARNc. Deuxièmement, il a été démontré qu'un signal de clivage ('GC'), responsable de la production de ces ARNm, se localisait 10nt en aval de la structure tige-boucle. Ainsi, l'extrémité 3'NTR des ARNm présenterait une structure tige-boucle terminale qui servirait au mécanisme poly(A) indépendant. Par conséquent, étant donné la localisation et la conservation du signal 'GC' et de la structure tige-boucle, nous avons émis l'hypothèse que les 2 motifs à ARN jouent des rôles clés dans la terminaison de la transcription ainsi que dans l'initiation de la traduction par le recrutement de protéines formant des complexes avec l'ARN. Cette hypothèse se base sur des études qui ont identifié des motifs à ARN similaires en tant que facteurs essentiels de mécanismes de régulation de la transcription et traduction chez les virus mais également chez les eucaryotes.

Plusieurs mécanismes de traduction poly(A) indépendants viraux et eucaryotes ont été mis en évidence (Blakqori *et al.*, 2009 ; Marzluff *et al.*, 2008). En effet, plusieurs virus possèdent des ARN structurés, localisés dans leurs régions codantes ou non-traduites. Par ailleurs, certaines structures tige-boucles ont été décrites comme augmentant l'activité traductionnelle des ARNm viraux non-polyadénylés (Delmas *et al.*, 2019). Par exemple, Blakqori et ses collègues ont montré que la 3'UTR de l'ARN du segment S du virus Bunyamwera (BUNV) stimule efficacement la synthèse des protéines virales indépendamment d'une queue poly(A) ou de PABP (Blakqori *et al.*, 2009). Chez un autre membre des *Bunyavirales*, l'Andes hantavirus (ANDV), il a été rapporté que les protéines virales n'étaient pas nécessaires pour la stimulation de la traduction cap-dépendante, renforçant l'hypothèse selon laquelle les protéines de l'hôte sont les acteurs de ce mécanisme (Vera-Otarola *et al.*, 2010). En fait, des complexes RNP entre des RBP de l'hôte et des structures d'ARN virales ont déjà été décrits chez les *Flaviviridae*, notamment chez le virus de la Dengue (DENV) et le virus du Nil occidental (WNV) (Dong *et*

al., 2015 ; Holden *et al.*, 2004). Les flavivirus sont des virus à ARN simple brin de polarité positive qui possèdent des éléments structurés dans leur génome. En effet, Dong et ses collègues ont montré que DENV et WNV possèdent des éléments structurés dans leurs extrémités 3' UTR et 5' UTR capables d'interagir entre elles, initiant la réplication virale par le recrutement de protéines de la famille Sm et Sm-like (Dong *et al.*, 2015). De plus, chez les eucaryotes, les histones, protéines cellulaires très abondantes, sont connues pour être pourvues d'un ARNm coiffé et non-polyadénylé (Marzluff *et al.*, 2008). De manière intrigante, l'extrémité 3' UTR du pré-ARNm de l'histone ressemble légèrement à celle de l'ARN du SBV. En effet, elle comprend un élément en aval de l'histone (HDE), une région riche en purines conservée et localisée 15nt en aval du site de clivage de l'ARNm, permettant des interactions ARN-ARN (Marzluff *et al.*, 2008). Le même type d'interaction pourrait être retrouvé avec le signal 'GC' pour favoriser le clivage de l'ARNc en ARNm, mais nous ne pouvons pas exclure l'interaction avec des protéines eucaryotes. La comparaison des protéines entre GC/TT(i) et WT(i) pourrait permettre de valider ou non cette possibilité. De plus, en amont de l'HDE, une structure en tige-boucle hautement conservée est retrouvée, ce qui rappelle la structure présente à l'extrémité 3' de l'ARN du SBV. Cela soulève la possibilité que les bunyavirus utilisent à leurs propres fins un processus de traduction cellulaire indépendant de la queue poly(A) (Williams & Marzluff, 1995 ; Marzluff *et al.*, 2008). S'appuyer sur ces modèles est une première étape pour identifier des candidats protéiques similaires pour la transcription et la traduction du SBV.

Partant de notre hypothèse, nous avons utilisé la méthode RaPID pour capturer et identifier les RBP qui interagissent avec la tige-boucle et le signal 'GC' dans le 3'NTR du segment S du SBV. Cette méthode a été choisie car elle s'est avérée fructueuse dans la découverte de RBP de l'hôte, essentielles à la réplication de virus. En effet, des RBP interagissant avec la 3'UTR de l'ARN du virus Zika (ZIKV) ont été identifiées par cette approche. ZIKV est un flavivirus non-polyadénylé qui contient des structures tige-boucles conservées dans ses régions UTR (Ramanathan *et al.*, 2019 ; Wang *et al.*, 2017). En outre, Ramanathan et ses collègues ont démontré l'efficacité de cette technique en testant des motifs ARN interagissant avec des protéines connues, afin de vérifier si la méthode permettait l'identification de trois RBP. Après une analyse MS/MS, les trois protéines étaient effectivement identifiées. L'avantage majeur de la méthode RaPID est la biotinylation des cibles dans des cellules vivantes. En effet, la liaison covalente de la biotine sur les protéines liant les motifs ARN permet une capture efficace des protéines et un lavage très rigoureux réduisant le bruit de fond. En outre, l'utilisation d'une biotine ligase modifiée de *Bacillus subtilis* (BASU), permet une biotinylation plus efficace des protéines que les méthodes utilisant une biotine ligase d'*E.coli*. (Trinkle-Mulcahy, 2019 ; Ramanathan *et al.*, 2019).

Conformément aux ressources disponibles sur la technique RaPID, cette méthode est efficace sur des ARN ne dépassant pas 66 nucléotides de long. Effectivement, l'enzyme BASU n'est pas capable de biotinyler les protéines au-delà d'un intervalle de 10nm de la BoxB (Ramanathan *et al.*, 2019). Toutefois, Ramanathan et ses collègues ont réussi à identifier des RBP différenciellement exprimées entre la 3'UTR complète de ZIKV (428 nt) et des mutants sélectionnés (Ramanathan *et al.*, 2019). Dès lors, nous avons décidé d'utiliser deux types de

constructions, des pNTR_{Sh} (60nt) comprenant les motifs ARN et des pNTR_L (120nt) comprenant l'entièreté de la 3'NTR. Le choix de pNTR_L nous a semblé judicieux pour éviter l'encombrement stérique dû aux nombreuses structures de l'ARN, à savoir : BoxB (3x), tige-boucle, 'GC', BoxB (3x). Bien que la taille des constructions diffère, nous avons pu montrer dans notre cas, que la longueur n'était pas discriminante pour l'identification des protéines. En effet, la position des 3 BoxB de part et d'autre du fragment analysé permet d'augmenter la distance de biotinylation jusqu'à 140nt (Majumder *et al.*, 2021). Par conséquent, nous avons rassemblé les vecteurs longs et courts ensemble pour augmenter le nombre d'échantillons par catégorie et donc la puissance des tests statistiques.

Lorsque nous avons comparé les catégories ensemble, seules quelques protéines différenciellement exprimées sont ressorties. L'utilisation d'un contrôle vecteur BirA seul nous a semblé importante pour avoir un point de référence sur l'autobiotinylation du système.

Les analyses statistiques par rapport au contrôle BirA ont montré que seules les protéines Helz et MBNL1 ne pourraient être utilisées pour la suite des analyses comparatives entre les conditions. Cela pourrait s'expliquer par des concentrations trop élevées du vecteur BASU lors des transfections, menant à une saturation du système. Les protéines MBNL1 et Helz sont des protéines de liaison à l'ARN importantes dans les étapes d'expression et de maturation de celui-ci. Investiguer leurs rôles par étude de co-localisation et d'interférence par ARN pourrait être intéressant pour la suite des travaux.

De manière générale, la comparaison des catégories a montré des résultats bruts similaires, c'est-à-dire peu de protéines significativement différentes entre le WT et les mutants. Des profils similaires de comparaison sont obtenus à travers les volcanoplots. Cette observation n'est pas surprenante étant donné que peu de protéines étaient attendues comme RBP spécifiques des motifs. Cependant, les protéines différenciellement exprimées après analyses statistiques pourraient être biaisées par le nombre de spectres comptabilisés lors de la comparaison des catégories. En effet, les protéines sortantes affichaient des taux de lecture de 2 à 6 spectres attribués par protéine, ce qui peut relever de la variabilité biologique et/ou technique. L'utilisation du logiciel CRAPome (Contaminant Repository for Affinity Purification) pourrait permettre d'augmenter la couverture et d'améliorer la caractérisation du fond associé à notre protocole expérimental. De plus, même si la méthode RaPID semble très attrayante pour les études sur des cellules vivantes, elle présente néanmoins plusieurs limites. Par exemple, toutes les cellules présentent des protéines biotinylées endogènes qui pourraient supplanter la faible abondance de certaines RBP lors de l'analyse en tandem. Dans notre cas, moins de 70µg de protéines par échantillon ont été envoyés en MS, ce qui accentue encore fortement cet aspect limitant. La technique repose également sur la biotinylation des résidus lysine, ce qui implique que toutes les protéines ne seront pas détectées de manière semblable (Ramanathan *et al.*, 2019). En outre, lors de l'utilisation de la technique RaPID, l'expression des constructions est dépendante d'une transfection. Cela constitue également une probable limitation de la méthode. En effet, le cycle répliatif du SBV est exclusivement situé dans le cytoplasme (Elliot, 2014), mais les transfections nécessitent un passage par le noyau. Par conséquent, nous pouvons supposer que cela augmente le bruit de fond, c'est-à-dire les protéines non-spécifiques capturées et identifiées. Pour pallier à cet inconvénient, nous avons

utilisé la méthode RaPID en condition d'infection par le SBV. Cette étape permet de simuler l'infection et donc de se rapprocher des conditions répliquatives du virus dans son milieu. Malgré l'infection des cellules par le SBV, aucune protéine n'était significativement plus exprimée en WTi par rapport au contrôle BirAi seule. Au total, deux protéines semblent intéressantes avec le modèle d'infection, Sarnp et TAF15. Ces deux protéines s'associent à l'ARN et sont impliquées dans sa maturation, elles pourraient, au même titre que MBNL1 et Helz, être investiguées lors de futures recherches. A l'heure actuelle, il n'a pas encore été possible de déterminer avec exactitude un complexe RNP responsable de la régulation de la transcription et de la traduction du SBV. Cependant, nous pouvons tout de même appuyer que la technique enrichit le pool de protéines de liaison à l'ARN et de maturation de l'ARN. Cela montre qu'un ajustement du protocole expérimental pourrait permettre l'identification des RBP responsables de la réplication du SBV. Par ailleurs, lors des transfections, les constructions n'expriment pas les motifs ARN dans leurs conditions natives (virus/vecteur, localisation, concentration, signal poly(A) localisé après les motifs, cadre de lecture différent, absence de la région 5'NTR de SBV, encombrement stérique), modifiant possiblement le RBPome d'interaction avec les deux motifs.

Étant donné l'étendue des conditions utilisées, nous avons tout de même voulu aller plus loin et essayer de déterminer les interactions potentielles entre l'ensemble des protéines sélectionnées en condition WT, WTi, HP 5P/3P et HP 5P/3Pi comparées au reste des catégories (STRING software) (**Fig. 28**). Nous pouvons constater que plusieurs nœuds d'interaction entre les protéines sont prédits. L'étude de l'ontologie de ces gènes montre également un enrichissement dans des processus de maturation, de transport et d'interaction avec les ARN, concordant avec notre modèle.

De plus, pour pallier les limitations dues à l'expression des constructions, nous avons réalisé un SBV recombinant BoxB viable. L'objectif de ce virus recombinant est de posséder des BoxB de part et d'autre des motifs ARN dans la 3'NTR du segment S. Une tentative avait précédemment été réalisée à l'URVI de l'UNamur, mais le virus comportant les 3 BoxB n'était pas viable. Il a donc été décidé d'ajouter une unique BoxB de part et d'autre des motifs. Ce virus s'est montré viable et très prometteur pour la suite des recherches. En effet, les données générées par la RACE-PCR chez le SBV recombinant, confirment que le segment S présente une variabilité de son site de clivage au sein de son extrémité. Par la méthode de RACE-PCR nous avons identifié les fréquences entre les différents pools d'ARN. Outre la présence de formes tronquées, il a été constaté que la majorité des virus recombinant BoxB présentait la séquence complète (ARNc). Dès lors, comme le SBV recombinant est capable de se répliquer et de produire des ARNg, il pourrait faire office de vecteur BoxB à la place des vecteurs pNTR. L'utilisation directe d'un virus viable et non d'un plasmide pourrait permettre l'identification spécifique des protéines interagissant avec le SBV ce qui serait intéressant pour limiter les perturbations structurales et par conséquent l'interaction avec les RBP. Ce protocole nécessiterait uniquement une transfection du vecteur BASU exprimant la biotine ligase de *Bacillus subtilis* fusionnée au peptide N et d'une infection par le SBV recombinant BoxB. L'ajout de biotine dans le milieu ferait de notre virus un virus capable d'autobiotinyler les RBP d'intérêt. Pour finaliser ce protocole, il serait judicieux d'évaluer la capacité d'une BoxB à

recruter la biotine ligase et à biotinyler les protéines par test d'activité enzymatique ou encore de tester la viabilité d'un virus avec d'autres aptamères comme les MS2, PP7 ou Csy4 (Gemmill *et al.*, 2020).

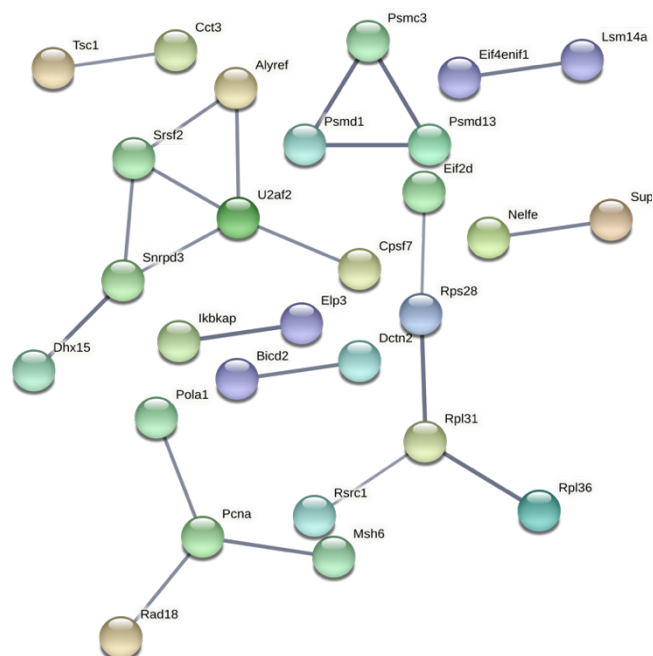


Figure 28: Prédiction de nœuds d'interaction entre protéines après analyses MS/MS. Seules les protéines ayant une prédiction d'interaction ont été conservées. Les lignes grises représentent une interaction prédite entre deux protéines avec une confiance élevée (=0,700). Chaque sphère colorée représente une protéine (STRING Software).

Il est également intéressant de noter que de nouveaux systèmes de biotinylation, ciblant des ARN natifs et sans l'ajout d'aptamères, ont été développés (Weissinger *et al.*, 2021). Par exemple, le système CRUIS (CRISPR-based RNA-United Interacting System), qui utilise la dCas13a modifiée et le PUP-IT (pupylation-based interaction tagging), est capable de cibler directement des séquences d'ARN spécifiques tout en étant fusionné à l'enzyme PafA qui marque les RBP environnantes. En outre, la distribution de CRUIS est principalement dans le cytoplasme en raison de la localisation de PafA (Weissinger *et al.*, 2021 ; Zhang *et al.*, 2020), ce qui correspond à la localisation du cycle répliatif du SBV. Cependant, cette technique reste difficile à mettre en œuvre, notamment la création d'un ARNguide, spécifique de la région des motifs ARN, n'interférant pas avec les éléments structurés. Une autre technique pourrait également être utilisée pour l'identification des RBP dans le cas du SBV. En effet, le ChiRP-MS (comprehensive identification of RNA-binding proteins by mass spectrometry) a été utilisé dans plusieurs études récentes pour identifier des protéines interagissant avec le génome de virus à ARN et de longs ARN non-codants (Flynn *et al.*, 2021 ; Lan *et al.*, 2020 ; Ooi *et al.*, 2019). Le ChiRP-MS se base sur la création de liaisons covalentes au niveau d'interactions ARN-protéines *in vivo*. Les complexes sont ensuite purifiés à l'aide d'oligonucléotides antisens biotinylés, dirigés contre l'ARN d'intérêt. Les protéines co-précipitées sont éluées doucement et identifiées par spectrométrie de masse. Cette technique se veut intéressante car elle n'implique pas l'utilisation de vecteur et seul un modèle d'infection par le SBV serait suffisant pour purifier les RNP.

Pour aller plus loin, plusieurs expériences peuvent être réalisées afin d'améliorer la compréhension des principaux acteurs derrière les processus de transcription et de traduction du SBV. Tout d'abord, il est nécessaire de répéter cette expérience plusieurs fois avec un plus faible nombre d'échantillons afin d'ajuster le protocole expérimental dans des conditions adéquates pour la lecture des RBP par le MS. Les résultats peuvent également être consolidés en comparant plusieurs méthodes basées sur le marquage de proximité comme CRISPR et/ou les liaisons covalentes *in vivo* par ChIRP-MS. Deuxièmement, le SBV se réplique fortement dans les tissus nerveux, changer la lignée cellulaire nous permettrait donc d'obtenir un modèle plus réaliste des partenaires de l'ARN du SBV. Par exemple, l'utilisation des cellules CPT-Tert (Arnaud *et al.*, 2010), une lignée cellulaire immortalisée de plexus choroïdes de mouton, pourrait être intéressante. De plus, le SBV étant un arbovirus, il serait également fascinant d'étudier son cycle répliatif dans des cellules d'insectes. En effet, la stratégie répliatif ne peut pas être similaire à celle des cellules de l'hôte car le virus doit se répliquer très faiblement afin de ne pas tuer son vecteur. Par exemple, une lignée cellulaire KC pourrait convenir à notre objectif car elles sont développées à partir d'embryons du moucheron piqueur *Culicoides variipennis* (Wechsler *et al.* 1989). Il serait donc intéressant de comparer les niveaux de RBP interagissant avec les motifs 3'NTR entre les lignées cellulaires BHK21, CPT-Tert et KC. Enfin, nous pourrions expérimenter des études directes sur les protéines sélectionnées en interférant avec les ARNm codant pour les protéines cibles afin d'évaluer l'impact sur la répliation dans un contexte biologique. La réalisation de ces différentes expériences pourrait finalement mettre en évidence les protéines essentielles pour la transcription et traduction hautement efficace de la protéine N du SBV et de mieux comprendre les systèmes poly(A) indépendants.

Références

- Afonso, A., Abrahantes, J. C., Conraths, F., Veldhuis, A., Elbers, A., Roberts, H., ... & Richardson, J. (2014). The Schmallenberg virus epidemic in Europe—2011–2013. *Preventive veterinary medicine*, 116(4), 391-403.
- Aksoy, E., & Azkur, A. K. (2018). Schmallenberg virus induces apoptosis in Vero cell line via extrinsic and intrinsic pathways in a time and dose dependent manner. *Journal of Veterinary Medical Science*, 18-0582.
- Amroun, A., Priet, S., de Lamballerie, X., & Quérat, G. (2017). Bunyaviridae RdRps: structure, motifs, and RNA synthesis machinery. *Critical reviews in microbiology*, 43(6), 753-778.
- Arnaud, F., Black, S. G., Murphy, L., Griffiths, D. J., Neil, S. J., Spencer, T. E., & Palmarini, M. (2010). Interplay between ovine bone marrow stromal cell antigen 2/tetherin and endogenous retroviruses. *Journal of virology*, 84(9), 4415-4425.
- Atkins, J. F., Gesteland, R. F., & Cech, T. (Eds.). (2011). *RNA worlds: from life's origins to diversity in gene regulation*.
- Barry, G., Varela, M., Ratinier, M., Blomström, A. L., Caporale, M., Seehusen, F., ... & Palmarini, M. (2014). NSs protein of Schmallenberg virus counteracts the antiviral response of the cell by inhibiting its transcriptional machinery. *The Journal of general virology*, 95(Pt 8), 1640.
- Bayrou, C. (2020). *Contribution à la biologie du virus de Schmallenberg chez les ruminants (Doctoral dissertation, Université de Liège, Liège, Belgique)*.
- Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., ... & Ting, A. Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nature biotechnology*, 36(9), 880-887.
- Beer, M., Conraths, F. J., & Van der Poel, W. H. M. (2013). 'Schmallenberg virus'—a novel orthobunyavirus emerging in Europe. *Epidemiology & Infection*, 141(1), 1-8.
- Blakqori, G., Delhaye, S., Habjan, M., Blair, C. D., Sánchez-Vargas, I., Olson, K. E., ... & Weber, F. (2007). La Crosse bunyavirus nonstructural protein NSs serves to suppress the type I interferon system of mammalian hosts. *Journal of virology*, 81(10), 4991-4999.
- Blakqori, G., van Knippenberg, I., & Elliott, R. M. (2009). Bunyamwera orthobunyavirus S-segment untranslated regions mediate poly (A) tail-independent translation. *Journal of virology*, 83(8), 3637.
- Boshra, H. Y., Charro, D., Lorenzo, G., Sanchez, I., Lazaro, B., Brun, A., & Abrescia, N. G. (2017). DNA vaccination regimes against Schmallenberg virus infection in IFNAR^{-/-} mice suggest two targets for immunization. *Antiviral research*, 141, 107-115.
- Bujalowski, P. J., Bujalowski, W., & Choi, K. H. (2017). Interactions between the dengue virus polymerase NS5 and stem-loop A. *Journal of virology*, 91(11), e00047-17.
- Choi, K. H. (2021). The Role of the Stem-Loop A RNA Promoter in Flavivirus Replication. *Viruses*, 13(6), 1107.
- Claine, F., Coupeau, D., Wiggers, L., Muylkens, B., & Kirschvink, N. (2015). Schmallenberg virus infection of ruminants: challenges and opportunities for veterinarians. *Veterinary Medicine: Research and Reports*, 6, 261.
- Collins, Á. B., Doherty, M. L., Barrett, D. J., & Mee, J. F. (2019). Schmallenberg virus: a systematic international literature review (2011-2019) from an Irish perspective. *Irish veterinary journal*, 72(1), 1-22.
- Cotter, R. J. (Ed.). (1993). *Time-of-flight mass spectrometry*. American Chemical Society.

- Coupeau, D., Claine, F., Wiggers, L., Martin, B., Kirschvink, N., & Muylkens, B. (2013). Characterization of messenger RNA termini in Schmallenberg virus and related Simbuviruses. *Journal of General Virology*, 94(11), 2399-2405.
- Davis, W. G., Blackwell, J. L., Shi, P. Y., & Brinton, M. A. (2007). Interaction between the cellular protein eEF1A and the 3'-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. *Journal of virology*, 81(18), 10172-10187.
- Delmas, B., Attoui, H., Ghosh, S., Malik, Y. S., Mundt, E., & Vakharia, V. N. (2019). ICTV virus taxonomy profile: Birnaviridae. *Journal of General Virology*, 100(1), 5-6.
- De Regge, N., Deblauwe, I., De Deken, R., Vantieghem, P., Madder, M., Geysen, D., ... & Cay, A. B. (2012). Detection of Schmallenberg virus in different *Culicoides* spp. by real-time RT-PCR. *Transboundary and emerging diseases*, 59(6), 471-475.
- De Regge, N. (2017). Akabane, Aino and Schmallenberg virus—where do we stand and what do we know about the role of domestic ruminant hosts and *Culicoides* vectors in virus transmission and overwintering?. *Current opinion in virology*, 27, 15-30.
- Doceul, V., Lara, E., Sailleau, C., Belbis, G., Richardson, J., Bréard, E., ... & Zientara, S. (2013). Epidemiology, molecular virology and diagnostics of Schmallenberg virus, an emerging orthobunyavirus in Europe. *Veterinary Research*, 44(1), 1-13.
- Doceul, V., Wernike, K., Vitour, D., and Laloy, E. (2017). Schmallenberg Virus. In *Emerging and Re-Emerging Infectious Diseases of Livestock*, J. Bayry, ed. (Cham: Springer International Publishing), pp. 99–119.
- Dong, H., Li, P., Elliott, R. M., & Dong, C. (2013). Structure of Schmallenberg orthobunyavirus nucleoprotein suggests a novel mechanism of genome encapsidation. *Journal of virology*, 87(10), 5593-5601.
- Dong, Y., Yang, J., Ye, W., Wang, Y., Miao, Y., Ding, T., ... & Xu, Z. (2015). LSm1 binds to the Dengue virus RNA 3'UTR and is a positive regulator of Dengue virus replication. *International journal of molecular medicine*, 35(6), 1683-1689.
- Dworkin, J. P., Lazcano, A., & Miller, S. L. (2003). The roads to and from the RNA world. *Journal of Theoretical Biology*, 222(1), 127-134.
- Eifan, S., Schnettler, E., Dietrich, I., Kohl, A., & Blomström, A. L. (2013). Non-structural proteins of arthropod-borne bunyaviruses: roles and functions. *Viruses*, 5(10), 2447-2468.
- Elliott, R. M. (2014). Orthobunyaviruses: recent genetic and structural insights. *Nature Reviews Microbiology*, 12(10), 673-685.
- Elliott, R. M. (1989). Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera virus, the prototype of the family Bunyaviridae. *Virology*, 173(2), 426-436
- Elliott, R. M. (1997). Emerging viruses: the Bunyaviridae. *Molecular medicine*, 3(9), 572-577.
- Elliott, R. M., & Blakqori, G. (2011). Molecular biology of orthobunyaviruses.
- Elliott, R. M., Blakqori, G., van Knippenberg, I. C., Koudriakova, E., Li, P., McLees, A., ... & Szemiel, A. M. (2013). Establishment of a reverse genetics system for Schmallenberg virus, a newly emerged orthobunyavirus in Europe. *The Journal of general virology*, 94(Pt 4), 851.

- Eloy, J. F. (1986). Quantitative microanalysis by time-of-flight laser probe mass spectrometry. *Scanning Electron Microscopy*, 1986(4), 1.
- European Food Safety Authority. (2012). “Schmallenberg” virus: analysis of the epidemiological data and assessment of impact. *EFSA Journal*, 10(6), 2768. .
- Flores, R., Gago-Zachert, S., Serra, P., Sanjuán, R., & Elena, S. F. (2014). Viroids: survivors from the RNA world?. *Annual review of microbiology*, 68, 395-414.
- Flynn, R. A., Belk, J. A., Qi, Y., Yasumoto, Y., Wei, J., Alfajaro, M. M., ... & Satpathy, A. T. (2021). Discovery and functional interrogation of SARS-CoV-2 RNA-host protein interactions. *Cell*, 184(9), 2394-2411.
- Ford, L. P., & Wilusz, J. (1999). 3'-Terminal RNA structures and poly (U) tracts inhibit initiation by a 3'-->5'exonuclease in vitro. *Nucleic acids research*, 27(4), 1159.
- Freeland, S. J., Knight, R. D., & Landweber, L. F. (1999). Do proteins predate DNA? *Science*, 286(5440), 690-692.
- Gache, K., Hosteing, S., Perrin, J. B., Zientara, S., Bournez, L., & Touratier, A. (2016). Surveillance du virus Schmallenberg en France: une circulation peu intense en 2014.
- Garigliany, M. M., Bayrou, C., Kleijnen, D., Cassart, D., Jolly, S., Linden, A., & Desmecht, D. (2012). Schmallenberg virus: a new Shamonda/Sathuperi-like virus on the rise in Europe. *Antiviral research*, 95(2), 82-87.
- Gaudin, Y. (2004). Structure des glycoprotéines de fusion virale de classe II dans leur conformation fusogène. *Virologie*, 8(2), 157-158.
- Gemmill, D., D'souza, S., Meier-Stephenson, V., & Patel, T. R. (2020). Current approaches for RNA-labelling to identify RNA-binding proteins. *Biochemistry and Cell Biology*, 98(1), 31-41.
- Gerber, A. P. (2021). RNA-Centric Approaches to Profile the RNA–Protein Interaction Landscape on Selected RNAs. *Non-coding RNA*, 7(1), 11.
- Gesteland, R. F. (1993). The RNA world: the nature of modern RNA suggests a prebiotic RNA world (No. 575.2 RNA).
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35(3), 849-857.
- Hall, K. B. (2014). Protein binding cannot subdue a lively RNA. *Nature*, 506(7488), 303-304.
- Harris, D., Zhang, Z., Chaubey, B., & Pandey, V. N. (2006). Identification of cellular factors associated with the 3'-nontranslated region of the hepatitis C virus genome. *Molecular & cellular proteomics*, 5(6), 1006-1018.
- Hoffmann, B., Scheuch, M., Höper, D., Jungblut, R., Holsteg, M., Schirmmeier, H., ... & Beer, M. (2012). Novel orthobunyavirus in cattle, Europe, 2011. *Emerging infectious diseases*, 18(3), 469.
- Hofmann, H., Li, X., Zhang, X., Liu, W., Kühl, A., Kaup, F., ... & Pöhlmann, S. (2013). Severe fever with thrombocytopenia virus glycoproteins are targeted by neutralizing antibodies and can use DC-SIGN as a receptor for pH-dependent entry into human and animal cell lines. *Journal of virology*, 87(8), 4384-4394.
- Holden, K. L., & Harris, E. (2004). Enhancement of dengue virus translation: role of the 3' untranslated region and the terminal 3' stem-loop domain. *Virology*, 329(1), 119-133.

- Hollidge, B. S., Nedelsky, N. B., Salzano, M. V., Fraser, J. W., González-Scarano, F., & Soldan, S. S. (2012). Orthobunyavirus entry into neurons and other mammalian cells occurs via clathrin-mediated endocytosis and requires trafficking into early endosomes. *Journal of virology*, 86(15), 7988-8001.
- Hughes, H. R., Adkins, S., Alkhovskiy, S., Beer, M., Blair, C., Calisher, C. H., ... & Consortium, I. R. (2020). ICTV virus taxonomy profile: Peribunyaviridae. *The Journal of general virology*, 101(1), 1.
- Jonsson, C. B., Figueiredo, L. T. M., & Vapalahti, O. (2010). A global perspective on hantavirus ecology, epidemiology, and disease. *Clinical microbiology reviews*, 23(2), 412-441.
- Kim, C. S., Seol, S. K., Song, O. K., Park, J. H., & Jang, S. K. (2007). An RNA-binding protein, hnRNP A1, and a scaffold protein, septin 6, facilitate hepatitis C virus replication. *Journal of virology*, 81(8), 3852-3865.
- Kim, H., Abeysirigunawardena, S. C., Chen, K., Mayerle, M., Ragunathan, K., Luthey-Schulten, Z., ... & Woodson, S. A. (2014). Protein-guided RNA dynamics during early ribosome assembly. *Nature*, 506(7488), 334-338
- Kohl, A., Lowen, A. C., Leonard, V. H., & Elliott, R. M. (2006). Genetic elements regulating packaging of the Bunyamwera orthobunyavirus genome. *Journal of General Virology*, 87(1), 177-187.
- Kraatz, F., Wernike, K., Reiche, S., Aebischer, A., Reimann, I., & Beer, M. (2018). Schmallenberg virus non-structural protein NSm: Intracellular distribution and role of non-hydrophobic domains. *Virology*, 516, 46-54.
- Lan, Z., Yao, X., Sun, K., Li, A., Liu, S., & Wang, X. (2020). The interaction between lncRNA SNHG6 and hnRNPA1 contributes to the growth of colorectal cancer by enhancing aerobic glycolysis through the regulation of alternative splicing of PKM. *Frontiers in oncology*, 10, 363.
- Leventhal, S. S., Wilson, D., Feldmann, H., & Hawman, D. W. (2021). A Look into Bunyavirales Genomes: Functions of Non-Structural (NS) Proteins. *Viruses*, 13(2), 314.
- Lewis, C. J., Pan, T., & Kalsotra, A. (2017). RNA modifications and structures cooperate to guide RNA-protein interactions. *Nature reviews Molecular cell biology*, 18(3), 202-210.
- Lozach, P. Y., Mancini, R., Bitto, D., Meier, R., Oestereich, L., Överby, A. K., ... & Helenius, A. (2010). Entry of bunyaviruses into mammalian cells. *Cell host & microbe*, 7(6), 488-499.
- Majumder, M., & Palanisamy, V. (2021). Compendium of Methods to Uncover RNA-Protein Interactions In Vivo. *Methods and Protocols*, 4(1), 22.
- Marklewitz, M., Zirkel, F., Rwego, I. B., Heidemann, H., Trippner, P., Kurth, A., ... & Junglen, S. (2013). Discovery of a unique novel clade of mosquito-associated bunyaviruses. *Journal of virology*, 87(23), 12850-12865.
- Marzluff, W. F., Wagner, E. J., & Duronio, R. J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly (A) tail. *Nature Reviews Genetics*, 9(11), 843-854.
- Mir, M. A., & Panganiban, A. T. (2006). The bunyavirus nucleocapsid protein is an RNA chaperone: possible roles in viral RNA panhandle formation and genome replication. *Rna*, 12(2), 272-282.
- Modis, Y., Ogata, S., Clements, D., & Harrison, S. C. (2004). Structure of the dengue virus envelope protein after membrane fusion. *Nature*, 427(6972), 313-319.
- Nguyen, T. M. T., Kim, J., Doan, T. T., Lee, M. W., & Lee, M. (2019). APEX proximity labeling as a versatile tool for biological research. *Biochemistry*, 59(3), 260-269.

- Ooi, Y. S., Majzoub, K., Flynn, R. A., Mata, M. A., Diep, J., Li, J. K., ... & Carette, J. E. (2019). An RNA-centric dissection of host complexes controlling flavivirus infection. *Nature microbiology*, 4(12), 2369-2382.
- Parisien, M., & Major, F. (2008). The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature*, 452(7183), 51-55.
- Petrović, M., Hernando, M. D., Díaz-Cruz, M. S., & Barceló, D. (2005). Liquid chromatography–tandem mass spectrometry for the analysis of pharmaceutical residues in environmental samples: a review. *Journal of Chromatography A*, 1067(1-2), 1-14
- Phillips, S. L., Soderblom, E. J., Bradrick, S. S., & Garcia-Blanco, M. A. (2016). Identification of proteins bound to dengue viral RNA in vivo reveals new host proteins important for virus replication. *MBio*, 7(1), e01865-15.
- Ramanathan, M., Porter, D. F., & Khavari, P. A. (2019). Methods to study RNA–protein interactions. *Nature methods*, 16(3), 225-234.
- Ramanathan, M., Majzoub, K., Rao, D. S., Neela, P. H., Zarnegar, B. J., Mondal, S., ... & Khavari, P. A. (2018). RNA–protein interaction detection in living cells. *Nature methods*, 15(3), 207-212.
- Reed, L. J., & Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *American journal of epidemiology*, 27(3), 493-497.
- Reguera, J., Weber, F., & Cusack, S. (2010). Bunyaviridae RNA polymerases (L-protein) have an N-terminal, influenza-like endonuclease domain, essential for viral cap-dependent transcription. *PLoS pathogens*, 6(9), e1001101.
- Reguera, J., Malet, H., Weber, F., & Cusack, S. (2013). Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proceedings of the National Academy of Sciences*, 110(18), 7246-7251.
- Sailleau, C., Breard, E., Viarouge, C., Vitour, D., Romey, A., Garnier, A., ... & Zientara, S. (2017). Re-emergence of bluetongue virus serotype 8 in France, 2015. *Transboundary and emerging diseases*, 64(3), 998-1000.
- Sears, R. M., May, D. G., & Roux, K. J. (2019). BioID as a tool for protein-proximity labeling in living cells. In *Enzyme-Mediated Ligation Methods* (pp. 299-313). Humana, New York, NY.
- Shtanko, O., Nikitina, R. A., Altuntas, C. Z., Chepurnov, A. A., & Davey, R. A. (2014). Crimean-Congo hemorrhagic fever virus entry into host cells occurs through the multivesicular body and requires ESCRT regulators. *PLoS pathogens*, 10(9), e1004390.
- Soldan, S. S., Hollidge, B. S., Wagner, V., Weber, F., & González-Scarano, F. (2010). La Crosse virus (LACV) Gc fusion peptide mutants have impaired growth and fusion phenotypes, but remain neurotoxic. *Virology*, 404(2), 139-147.
- Sternburg, E. L., & Karginov, F. V. (2020). Global approaches in studying RNA-binding protein interaction networks. *Trends in Biochemical Sciences*, 45(7), 593-603.
- Stokes, J. E., Baylis, M., & Duncan, J. S. (2016). A freedom from disease study: Schmallenberg virus in the south of England in 2015. *Veterinary Record*, 179(17), 435-435.
- Tarlinton, R., Daly, J., Dunham, S., & Kydd, J. (2012). The challenge of Schmallenberg virus emergence in Europe. *The Veterinary Journal*, 194(1), 10-18.
- Ter Horst, S., Conceição-Neto, N., Neyts, J., & Rocha-Pereira, J. (2019). Structural and functional similarities in bunyaviruses: Perspectives for pan-bunya antivirals. *Reviews in medical virology*, 29(3), e2039.

- Trinkle-Mulcahy, L. (2019). Recent advances in proximity-based labeling methods for interactome mapping. *F1000Research*, 8.
- Upadhyay, A., Dixit, U., Manvar, D., Chaturvedi, N., & Pandey, V. N. (2013). Affinity capture and identification of host cell factors associated with hepatitis C virus (+) strand subgenomic RNA. *Molecular & Cellular Proteomics*, 12(6), 1539-1552.
- Van den Brom, R., Lutikholt, S. J., Lievaart-Peterson, K., Peperkamp, N. H., Mars, M. H., Van Der Poel, W. H., & Vellema, P. (2012). Epizootic of ovine congenital malformations associated with Schmallenberg virus infection. *Tijdschrift voor diergeneeskunde*, 137(2), 106-111.
- Varela, M., Schnettler, E., Caporale, M., Murgia, C., Barry, G., McFarlane, M., ... & Palmarini, M. (2013). Schmallenberg virus pathogenesis, tropism and interaction with the innate immune system of the host. *PLoS pathogens*, 9(1), e1003133/
- Vera-Otarola, J., Soto-Rifo, R., Ricci, E. P., Ohlmann, T., Darlix, J. L., & López-Lastra, M. (2010). The 3' untranslated region of the Andes hantavirus small mRNA functionally replaces the poly (A) tail and stimulates cap-dependent translation initiation from the viral mRNA. *Journal of virology*, 84(19), 10420-10424.
- Walter, C. T., & Barr, J. N. (2011). Recent advances in the molecular and cellular biology of bunyaviruses. *Journal of General Virology*, 92(11), 2467-2484.
- Wang, A., Thurmond, S., Islas, L., Hui, K., & Hai, R. (2017). Zika virus genome biology and molecular pathogenesis. *Emerging microbes & infections*, 6(1), 1-6.
- Wechsler, S. J., McHolland, L. E., & Tabachnick, W. J. (1989). Cell lines from *Culicoides variipennis* (Diptera: Ceratopogonidae) support replication of bluetongue virus. *Journal of invertebrate pathology*, 54(3), 385-393.
- Weissinger, R., Heinold, L., Akram, S., Jansen, R. P., & Hermesh, O. (2021). RNA Proximity Labeling: A New Detection Tool for RNA–Protein Interactions. *Molecules*, 26(8), 2270.
- Wernike, K., Hoffmann, B., Conraths, F. J., & Beer, M. (2015). Schmallenberg Virus Recurrence, Germany, 2014. *Emerging Infectious Diseases*, 21(7), 1202-1204.
- Willemin, M. É. (2014). Modélisation de la toxicocinétique des isomères cis et trans de la perméthrine et de ses métabolites chez le rat et de leur métabolisme sur hépatocytes humains (Doctoral dissertation, Université de Technologie de Compiègne).
- Williams, A. S., & Marzluff, W. F. (1995). The sequence of the stem and flanking sequences at the 3' end of histone mRNA are critical determinants for the binding of the stem-loop binding protein. *Nucleic acids research*, 23(4), 654-662.
- Yanase, T., Kato, T., Aizawa, M., Shuto, Y., Shirafuji, H., Yamakawa, M., & Tsuda, T. (2012). Genetic reassortment between Sathuperi and Shamonda viruses of the genus *Orthobunyavirus* in nature: implications for their genetic relationship to Schmallenberg virus. *Archives of virology*, 157(8), 1611-1616.
- Yi, M., & Lemon, S. M. (2003). Structure–function analysis of the 3' stem-loop of hepatitis C virus genomic RNA and its role in viral RNA replication. *Rna*, 9(3), 331-345.
- Yu, L., & Markoff, L. (2005). The topology of bulges in the long stem of the flavivirus 3' stem-loop is a major determinant of RNA replication competence. *Journal of virology*, 79(4), 2309-2324.
- Zhang, Z., Sun, W., Shi, T., Lu, P., Zhuang, M., & Liu, J. L. (2020). Capturing RNA–protein interaction via CRUIS. *Nucleic acids research*, 48(9), e52-e52.

Zimmer, J. Y., Brostaux, Y., Haubruge, E., & Francis, F. (2014). Larval development sites of the main *Culicoides* species (Diptera: Ceratopogonidae) in northern Europe and distribution of coprophilic species larvae in Belgian pastures. *Veterinary Parasitology*, 205(3-4), 676-686.

Zimmer, J. Y., Verheggen, F. J., Haubruge, E., & Francis, F. (2015). Orientation behaviour of *Culicoides obsoletus* (Diptera: Ceratopogonidae), a relevant virus vector in northern Europe, toward host-associated odorant cues. *Veterinary parasitology*, 211(3-4), 274-282.

Figures supplémentaires

1: Table des gènes correspondant aux protéines différentiellement identifiées par comparaison de la construction 5P/3P(i) et des mutants.

Catégories comparées	Gènes ID
5P/3P vs 5P	Crtc3, H671_5g15095, Rsrc1, Hspb1, Txndc9, Eif2d, Cpsf7, H671_7g18194, Glrx3
5P/3P vs GC/TT	Parp4, RPL31, I79_020377, A6R68_11509, Psm13, Eif4enif1, Fam98a, I79_006633,, Cpsf7, A6R68_20798, CgPICR_018114, Scrib, Hsd17b4, Phactr4, Alyref, Tbce, Crtc3,, CgPICR_002358, Dock1, Pygb, Shc1, Msh6, H671_1g1562, CgPICR_007501, Rsrc1, Kif13b, Psm1, DHX15, Supt5h, Fermt2, CgPICR_016165, Srsf2, Hspb1, Irf2bpl, Helz, CgPICR_007964, A6R68_06547, Ppfibp1, Ndr1, Zmynd8, Abcf3, Bbx, Kri1, Polr1e, Gak, M91_11412, Pcn1, Cdr2l, Ap3d1, Kif1c, CgPICR_016903, LOC101840342, U2AF2,, I79_020820, Nek7, A6R68_20623, H671_5g14943, Dennd4c, Rps28, Dctn2, Cct3, LOC101842215, LOC101838805, LOC102928996, A6R68_20243, Psmc3, Tjap1, Gb1, Pola1, MBNL1, Fhod1, Kidins220, Paxbp1, A6R68_06374, Nab2, Kif2a, Rnf20, Bcd2, Elp3, Eif2d, Sars, Rp9, Ptpn23, Mcm4, SELENBP1, A6R68_06293,, H671_4g12082, Hars, Adk, Unc45a, Glrx3, M91_03498, Arhgef1, Thop1, Elp1, Adss, Rad18, Prune1, Nupl2, Wdr46
5P/3Pi vs 5Pi	Arl6ip4, Prpf4b, Rab27a
5P/3Pi vs GC/TTi	Pum2, Celf2, Prpf4b, Sumo1, Edf1, CgPICR_012931, Fabp5, Nkap, LZ, Rpl36, Mcrip2, Rab27a