Diversité et évolution de l'Ostreid Herpesvirus type 1

Camille Pelletier

▶ **To cite this version:**

Camille Pelletier. Diversité et évolution de l'Ostreid Herpesvirus type 1. Sciences agricoles. Université de La Rochelle, 2023. Français. NNT : 2023LAROS030 . tel-04549017

HAL Id: tel-04549017
https://theses.hal.science/tel-04549017

Submitted on 16 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE LA ROCHELLE

ÉCOLE DOCTORALE N°618 - EUCLIDE

Unité Adaptation Santé des Invertébrés Marins (PDG-RBE-ASIM), IFREMER, La Tremblade

THÉSE

présentée par :

Camille PELLETIER

Soutenue le 24 Novembre 2023

pour l'obtention du grade de Docteur de l'Université de La Rochelle

Discipline : Biologie de l'environnement, des populations, écologie

Diversité et évolution de l'*Ostreid herpesvirus type 1*

Jury :

Fabrice Pernet	Chargé de Recherche, IFREMER Brest, Rapporteur
Nicolas Berthet	Chargé de Recherche, Institut Pasteur Paris, Rapporteur
Christine Paillard	Directrice de Recherche, CNRS Brest, Examinatrice
Tristan Renault	Directeur de Recherche, IFREMER Nantes, Examinateur
Elisabeth Herniou	Directrice de Recherche, CNRS Tours, Examinatrice
Benjamin Morga	Chargé de Recherche, IFREMER La Tremblade, Directeur de thèse

Membres invités :

Laetitia Canini	Chargée de Recherche, ANSES Maisons-Alfort, Examinatrice
Claire Martenot	Chargée de Recherche, ANSES Ploufagran, Examinatrice

A mes parents

et mes grands parents

Remerciements

Je souhaite tout d'abord remercier les membres de ce jury de thèse qui ont évalué mon travail. Je remercie le Docteur Fabrice Pernet et le Docteur Nicolas Berthet d'avoir accepté d'être les rapporteurs de ce travail. Merci également au Docteur Christine Paillard, au Docteur Elisabeth Herniou et au Docteur Tristan Renault d'avoir accepté d'être les examinateurs de ce manuscrit. Enfin, merci au Docteur Laëtitia Canini et au Docteur Claire Martenot d'avoir participer à l'évaluation de mon travail en temps que membres invités.

Merci également aux membres de mon comité de thèse, Docteur Tristan Renault, Docteur Julien Thézé et Docteur Sonia Burrel. Merci pour vos conseils avisés qui m'ont permis d'avancer efficacement et d'atteindre mes objectifs de recherche avec succès.

J'exprime ma profonde gratitude envers mon directeur de thèse, le Docteur Benjamin Morga, ainsi que mes deux encadrants scientifiques, le Docteur Germain Chevignon et le Docteur Maude Jacquot. Merci pour toutes les corrections et améliorations que vous avez pu apporter à ce manuscrit. Votre précieuse contribution à cette étape cruciale qu'est la thèse ne peut être sous-estimée. Vous m'avez non seulement guidé dans la réussite de ce projet, mais vous avez aussi joué un rôle essentiel dans ma croissance personnelle et professionnelle. Votre soutien m'a incité à développer mes compétences, à renforcer ma confiance en moi, et finalement à prendre mon envol dans ma carrière.

Benjamin, je te remercie chaleureusement. Il y a près de sept ans, tu m'as ouvert les portes de l'Ifremer lors de mon stage de Licence, une opportunité exceptionnelle qui m'a permis de plonger dans le monde de la recherche. Ce moment a été un tournant dans ma vie, m'inspirant à suivre cette voie avec détermination. Aujourd'hui, sept années plus tard, je te suis reconnaissante pour ta confiance en moi depuis le début, pour avoir cru en mes capacités, et pour m'avoir guidée vers un sujet de thèse aussi captivant. Je te remercie également pour ton inlassable dévouement à trouver des financements qui ont rendu mes expérimentations possibles.

Germain, tu n'étais pas impliqué initialement dans ma thèse, mais le destin a fait que ton recrutement à l'Ifremer est l'une des choses les plus bénéfiques qui me soit arrivée durant ces quatre années. Ton arrivée à Ifremer a été une source d'inspiration pour moi. Tes conseils avisés m'ont été d'une aide précieuse et m'ont permis d'apprendre tout au long de ces trois ans. Ta franchise, bien que parfois un peu déconcertante, a toujours été appréciée, car elle m'a permis de grandir professionnellement. Je tiens également à te remercier pour ton implication minutieuse dans le design de chacune des figures et présentations que j'ai pu créer. Ton expertise en matière de design a toujours été d'une grande aide

pour améliorer mes figures parfois « horriblement moches » ou « trop fluos ». Je suis reconnaissante pour le temps que tu as passé à corriger chaque présentation avec tant de soin. Merci pour la découverte de Washington. Et bien sûr, je ne peux pas oublier les nombreuses sessions burger du vendredi midi.

Maude, tout comme Germain, je suis profondément reconnaissante et je ne saurais trouver les mots pour exprimer toute ma gratitude envers toi. Ton enseignement en phylodynamique, génétique et épidémiologie a été incroyablement enrichissant, et grâce à toi, j'ai acquis de nouvelles et solides compétences *dans ces domaines. Ta patience et ta disponibilité pour répondre à mes demandes de "soutien scolaire" ont été inestimables, et tu as toujours su trouver les mots réconfortants pour m'encourager dans les moments de doute. Je suis reconnaissante d'avoir eu la chance de bénéficier de ton expertise et de ta bienveillance. Tu as été bien plus qu'une encadrante, tu as été une véritable épaule maternelle sur laquelle j'ai pu m'appuyer ces derniers mois. Merci pour ton écoute inestimable, ton soutien indéfectible, et pour avoir apaisé mes questionnements, remises en question et doutes. Grâce à toi, j'ai appris que même dans les moments les plus compliqués, il existe toujours une lueur au bout du tunnel. Tu es cette bonne étoile qui a illuminé ma thèse lorsque j'en avais le plus besoin. Merci encore pour tout ce que tu as fait pour moi. Je garderai en souvenir ta générosité et ta bonne humeur.*

Nicole, merci pour tout ce que tu m'as enseigné. Depuis ma licence 3, tu m'as formé en biologie moléculaire. Tu m'as appris à être précise et rigoureuse, ce qui m'a été d'une aide inestimable pour mener à bien mes travaux de thèse. Ta bonne humeur constante a rendu notre travail ensemble non seulement instructif, mais aussi agréable. Alors que tu entames cette nouvelle étape de ta vie en tant que retraitée, je tiens à te remercier encore une fois pour ton dévouement et ta contribution à mon développement. Profite bien de ta retraite bien méritée.

Serge, merci pour tes précieux enseignements en bio-informatique et en génétique, qui ont été essentiels pour mes débuts sur Datarmor. Les solides bases que tu m'as transmises ont été d'une importance cruciale pour l'analyse de l'ensemble des données que j'ai accumulées au cours de ces quatre années de thèse. Je te suis sincèrement reconnaissant pour cette contribution. Merci.

Aurélie N., je souhaite te remercier du fond du cœur pour la joie de vivre que tu as apporté dans mes journées au laboratoire et pour ta bonne humeur qui a le don de nous remonter le moral. Un moment qui restera à jamais gravé dans ma mémoire est celui que nous avons partagé après le repas de Noël en 2022, sur la terrasse du centre nautique. Merci pour tous ces moments inoubliables et pour ta belle personnalité rayonnante.

Bruno, merci pour ton soutien et ta précieuse aide en histologie et microscopie. Ta compétence et ton dévouement m'ont permis d'obtenir de belles photos de mes larves infectées par OsHV-1. Ta constante

bonne humeur et ton sourire rayonnant ont illuminé mes journées. Merci d'être passé nous voir tous les matins et d'avoir apporté avec toi une énergie positive et motivante. J'espère avoir l'occasion de partager d'autres bières avec Aurélie N. et toi.

Mickaël, merci pour tes précieux conseils en biologie moléculaire. Je garderai en mémoire ta voix grave qui a parfois mis à rude épreuve mes tympans, mais qui a toujours été le signe de discussions passionnantes et enrichissantes. Pour tous les moments de complicité, les soirées inoubliables et les précieux souvenirs que nous avons partagé, merci. Je te souhaite tout le meilleur pour la suite de ton chemin.

Merci à tous les doctorants que j'ai eu l'occasion de croiser durant ma thèse. Nos échanges aussi bien professionnels que personnels ont participé à mon épanouissement durant ces quatre ans.

Sandy, merci d'avoir été mon encadrante de stage en Licence 3. Tu m'as beaucoup appris et surtout tu m'as donné la chance d'intégrer l'unité ASIM. Merci d'avoir cru en moi et de m'avoir fait confiance.

Nicolas, tu as su me comprendre dès notre première discussion et tu es rapidement devenu mon point de repère au sein du laboratoire. Merci de m'avoir écouté et conseillé. Merci pour ta gentillesse, pour les soirées plongées qui se sont souvent poursuivis par des discussions interminables jusqu'à quatre heures du matin. Merci pour ton soutien.

Coralie, merci pour ton dynamisme et ta joie de vivre. Merci pour nos échanges, pour ton écoute et ton soutien. Aurélie M., merci pour ces quelques mois partagés dans le bureau des doctorants. Merci pour ta sagesse et pour tes conseils. J'ai apprécié partager le dernier colloque des doctorants avec toi. Jérémy, notre cohabitation fut courte, mais je garde un bon souvenir de ta présence. Merci pour nos discussions et les conseils partagés sur les présentations orales.

Ajith, I'm glad to have met you! Thank you for the scientific discussions we've had. Take care of yourself and don't leave your things lying around. Hang in there for the rest of your thesis, good luck.

Antoine, attends-moi pour les prochaines sessions surf !! Je pense que tu es conscient de l'importance de ta présence dans ma thèse, que ce soit dans nos moments de loisirs, nos soirées mémorables, ou dans notre collaboration au laboratoire pour partager des connaissances scientifiques sur les huîtres et leurs agents pathogènes. J'ai eu la chance de partager ce superbe bureau sous les toits sans fenêtre avec toi et ta bonne humeur m'a permis de garder le moral. Merci d'avoir été un super collègue et d'être maintenant un ami.

Un merci tout particulier pour toi, Aurélie D.M., pour avoir fait preuve d'une incroyable compréhension face à mes sautes d'humeur et pour avoir su apaiser mon stress en me distrayant. Tu m'as réellement

tiré vers le haut en cette fin de thèse en m'apportant un équilibre essentiel entre concentration, motivation, écoute et détente. Bien évidemment, merci également pour les discussions scientifiques autour du virus OsHV-1, qui m'ont souvent aidé à exprimer mes idées clairement, ou pour les conseils pour améliorer le design de mes figures. Je dirais merci aux Américains pour cette belle rencontre. Bon courage pour ta fin de thèse. Reste comme tu es et aies confiance en toi.

Merci également aux nombreux stagiaires et alternants que j'ai eu l'occasion de côtoyer durant ces quatre ans de thèse : Chloé B., Florent M., Nastasia B., Paul, M., Gabriel B., Juliette H., Marine P., Ewenn C., Jean-Christophe M., Raphael S., Solène P., Fanny T., Grégoire M. La liste est longue mais je ne pouvais pas vous oublier. Merci pour toutes ces aventures, ces rires, ces pleurs, ces embrouilles, ces commérages. Ma vie au laboratoire aurait été beaucoup moins palpitante sans votre présence. Merci beaucoup. Au plaisir de vous recroiser.

Maelle, tu as été ma première stagiaire. Tu m'as beaucoup aidé dans l'ensemble de mes expérimentations, merci. Cette expérience a été extrêmement enrichissante pour moi, et j'espère sincèrement qu'elle l'a été également pour toi.

Tom, merci du fond du cœur. Tu as été un stagiaire un peu particulier. Mon acolyte de la première heure. Pas mal de petites frayeurs, de grands moments de rires et de complicité. Merci d'être encore présent pour moi, souriant et encourageant.

Je tiens à remercier tous mes collègues de l'unité Adaptation Santé des Invertébrés Marins (ASIM), car leur précieuse contribution a largement contribué à façonner cette thèse. Vous êtes nombreux à avoir contribué, que ce soit de près ou de loin, à mon bien-être et à mon développement professionnel. Les remerciements seraient trop longs si je devais exprimer pleinement ma reconnaissance envers chacun d'entre vous. Je vous remercie chaleureusement pour le partage de votre culture scientifique, les moments conviviaux autour des pauses café, votre joie et bonne humeur contagieuses, ainsi que votre inestimable aide. Nos discussions, toujours enrichissantes, qu'elles aient porté sur des sujets professionnels ou plus personnels, ont grandement enrichi mon expérience.

Un grand merci à toutes les personnes qui m'ont permis d'atteindre mes objectifs et qui m'ont soutenu tout au long de mon parcours. Votre contribution, qu'elle ait été sous forme de conseils, d'encouragements, de ressources ou de présence, a été d'une importance cruciale. Votre soutien a été une source d'inspiration et de motivation inestimable. Je suis profondément reconnaissante d'avoir pu compter sur vous.

Aurélie D.M., Marie, Lucie, vous avez été mes petits rayons de soleil en ces derniers mois de thèse. Je tiens à vous exprimer ma plus profonde gratitude pour tous les moments partagés : ces après-midis à

la plage, les couchers de soleil, les batailles de bombes à eau ou les emballages de voiture, nos pique-niques remplis de rires, les moments de détente autour de quelques bières, les toiles de tente qui s'envolent, et bien sûr, nos précieuses habitudes de dégustation de glace. Votre présence et votre amitié ont non seulement rendu ma fin de thèse plus douce, mais aussi lumineuse et pleine de joie. Merci du fond du cœur pour toutes ces merveilleuses expériences partagées.

Lauriane, j'aurais tant de choses à dire… Presque 15 ans que l'on se connait. Je suis consciente que je manque de mots pour véritablement rendre hommage à tout ce que tu as fait pour moi. Au cours de ces quatre années de thèse, mais également bien avant, tu as toujours été là, écoutant mes préoccupations, me soutenant inconditionnellement et apportant une dose constante de positivité dans ma vie. Tu as cette capacité exceptionnelle à me remettre sur le bon chemin lorsque je m'en égare, et pour cela, je te suis infiniment reconnaissante. Merci d'avoir toujours été là, que ce soit dans les moments difficiles où ton soutien était précieux, ou dans les moments de joie durant lesquels tu étais là pour célébrer avec moi. Je sais que je peux toujours compter sur toi, que tu es une épaule solide sur laquelle je peux m'appuyer. Tu sais à quel point je dis rarement ces choses-là, mais je suis tellement reconnaissante de t'avoir dans ma vie.

Merci à la Team des "amoureux de La Rochelle" pour toutes les soirées inoubliables que nous avons partagées ensemble. Vos rires, votre joie de vivre, et les précieux moments que nous avons créé resteront gravés dans ma mémoire. Je suis profondément contente de vous avoir comme amis. Je suis impatiente de créer de nouveaux souvenirs mémorables à vos côtés. Merci pour ces moments de complicité et de rires, et que notre amitié perdure encore longtemps.

C'est avec beaucoup d'émotions que je terminerais ces remerciements en m'adressant à ma famille.

Humoristique pour commencer. Perle, je tiens à t'adresser un merci tout particulier pour ta présence constante dans ma vie, même lorsque tu as décidé que la meilleure place pour faire la sieste était sur mon clavier d'ordinateur. Ta capacité à transformer la maison en un véritable parcours d'obstacles est impressionnante, et j'admire ton art de dérouler le papier toilette comme personne d'autre. Les séances de chasse aux souris invisibles sont toujours divertissantes. Merci pour ces moments de divertissement et de câlins félins. Je te remercie pour ces quatre années de ronronnements et de bêtises.

Mamie Michelle, mamie Monique, papi Roger et papi Jean-Paul, je tiens à vous adresser un message d'affection et de reconnaissance. Votre présence bienveillante dans ma vie a toujours été une source d'inspiration et de réconfort. Vos sages conseils, vos histoires captivantes, et surtout, l'amour que vous avez toujours partagé avec moi ont laissé une empreinte indélébile dans mon cœur. Que ce soit à travers les précieux moments passés ensemble, les repas de famille, ou les valeurs que vous m'avez transmises, votre influence a contribué de manière significative à ma formation en tant qu'individu.

Votre générosité, votre bienveillance et votre soutien constant ont été des piliers essentiels dans ma vie. Je tiens à vous remercier du fond du cœur pour tout ce que vous avez fait pour moi. Vous êtes une source de joie et d'inspiration, et je suis comblée de vous avoir comme grands-parents. Votre amour et vos enseignements continuent de guider ma vie, et je chéris chaque moment que nous partageons ensemble. Merci pour tout.

Amandine, malgré les différences et les désaccords qui peuvent parfois exister entre nous, merci. Merci de m'avoir toujours poussé à me challenger dans les études. C'est peut-être un peu grâce à toi que j'écris ce manuscrit. Le challenge était trop beau. Je suis contente que tu sois ma petite sœur.

Mes chers parents, que serait une thèse sans le soutien inestimable des personnes qui nous ont élevés et guidés tout au long de notre vie ? Vous avez été bien plus que des parents, vous avez été des piliers solides, et je souhaite vous exprimer toute ma reconnaissance. Dans vos yeux et vos paroles, j'ai toujours vu et ressenti la fierté que vous éprouviez à chaque étape de mon parcours. Cette fierté a été ma principale source de motivation. Savoir que je pouvais vous rendre fiers était un moteur puissant qui m'a poussé à donner le meilleur de moi-même, même lorsque les défis semblaient insurmontables. Vous avez été présents à chaque instant où j'avais besoin de vous, que ce soit lorsque les heures de sommeil se faisaient rares, que la mauvaise humeur me gagnait ou lorsque la déprime due à la pandémie de la COVID-19 s'installait. Votre soutien indéfectible a été ma bouée de sauvetage dans les moments difficiles. Je tiens également à vous remercier pour avoir consacré une partie de votre précieux temps libre au cours de ces dernières semaines pour corriger ce manuscrit. Je vous remercie du fond du cœur pour votre amour inconditionnel, votre patience et votre soutien constant. Je suis infiniment reconnaissante d'avoir des parents aussi exceptionnels que vous.

Valorisations

Publications présentées dans le manuscrit

Pelletier C., Chevignon G., Jacquot M., Morga B. Global diversity and dispersal routes of the *Ostreid herpesvirus type 1.* (pour soumission dans *Virus evolution*).

Pelletier C., Faury N., Mege M., Dégremont L., Chevignon G., Jacquot M., Morga B. Simulation of 30 years of *M. gigas* infection with the virus OsHV-1 revealed evolution through positive selection. (pour soumission dans *PCI evolution*).

Pelletier C., Chevignon G., Faury N., Arzul I., Garcia C., Chollet B., Renault T., Morga B., Jacquot M. Genetic differentiation and host specialization among OsHV-1 infecting two oyster species in France. (soumis dans *Journal of virology*).

Dotto-Maurel A., **Pelletier C**., Morga B., Jacquot M., Faury N., Dégremont L., Bereszczynki M., Delmotte J., Escoubas J-M., Chevignon G. (2022) Evaluation of tangential flow filtration coupled to long-read sequencing for *Ostreid herpesvirus* type 1 genome assembly. *Microbial Genomics*, 8(11), 000895 (17p.). Publisher's official version: https://doi.org/10.1099/mgen.0.000895, Open Access version: https://archimer.ifremer.fr/doc/00804/91607/

Delmotte J., **Pelletier C**., Morga B., Galinier R., Petton B., Lamy J-B., Kaltz O., Avarre J-C., Jacquot M., Montagnani C., Escoubas J-M. (2022). Genetic Diversity and Connectivity of the Ostreid Herpesvirus 1 Populations in France: A First Attempt to Phylogeographic Inference for a Marine Mollusc Disease. *Virus Evolution*, 8(1), veac039 (14p.). Publisher's official version : https://doi.org/10.1093/ve/veac039, Open Access version : https://archimer.ifremer.fr/doc/00766/87834/

Morga B., Jacquot M., **Pelletier C.**, Chevignon G., Dégremont L., Biétry A., Pepin J-F., Heurtebise S., Escoubas J-M., Bean T. P., Rosani U., Bai C-M., Renault T., Lamy J-B. (2021). Genomic Diversity of the Ostreid Herpesvirus Type 1 Across Time and Location and Among Host Species. *Frontiers In Microbiology*, 12, 711377 (13p.). Publisher's official version: https://doi.org/10.3389/fmicb.2021.711377, Open Access version: https://archimer.ifremer.fr/doc/00706/81762/

Publications non présentées dans le manuscrit

Battistel C., Mouren J-C., Morga B., **Pelletier C**., Canier L., Garcia C., Arzul I., Auffret P., Leroi L., Chevignon G., Durand P., Jacquot M. MoPSeq-DB: A User-Friendly Web Application for Genomic Data Management and Analysis of Marine Mollusc Pathogens. (pour soumission dans *Scientific data*).

Dotto-Maurel A., **Pelletier C**., Faury N., Heurtebise S., Morga B., Chevignon G. Evaluation of long-read sequencing for Ostreid herpesvirus type 1 genome characterization from host infected tissue. (pour soumission dans *PCI evolutionary Biology*)

Picot S., Faury N., **Pelletier C**., Arzul I., Chollet B., Dégremont L., Renault T., Morga B.(2022). Monitoring Autophagy at Cellular and Molecular Level in Crassostrea gigas During an Experimental Ostreid Herpesvirus 1 (OsHV-1) Infection. *Frontiers In Cellular And Infection Microbiology* , 12, 858311 (14p.) . Publisher's official version : https://doi.org/10.3389/fcimb.2022.858311 , Open Access version : https://archimer.ifremer.fr/doc/00762/87400/

Communications orales et affichées

Battistel C., Mouren J-C., Morga B., **Pelletier C**., Canier L., Garcia C., Arzul I., Leroi L., Durand P., Chevignon G., Jacquot M. (2023). MoPSeq-DB. Molluscs Pathogen Sequences DataBase. JOBIM 2023 - Journées Ouvertes en Biologie, Informatique et Mathématiques. 27 au 30 juin 2023, Nancy.

Pelletier C., Chevignon G., Faury N., Dégremont L., Arzul I., Garcia C., Chollet B., Renault T., Jacquot M., Morga B. (2023). Diversity and evolution of the Ostreid Herpesvirus type 1. 115th Annual meeting of the National Shellfisheries association. 26-30 March 202, Baltimore, United States of America.

Jacquot M., Mouren J-C., Battistel C., Morga B., **Pelletier C**., Canier L., Garcia C., Arzul I., Leroi L., Durand P., Chevignon G. (2022). MoPSeq-DB: a web platform to share, visualise and analyse sequence data of mollusc pathogens. ISEEMPD 2022 - International Symposium on Ecology and Evolution of Marine Parasites and Diseases. 15-18 November 2022, La Rochelle, France. https://archimer.ifremer.fr/doc/00827/93932/

Pelletier C. et Jacquot M. (2022). Les débuts de l'épidémiologie moléculaire des maladies des mollusques marins. Phylomap. 04-06 octobre 2022, Paris, France. https://www.youtube.com/watch?v=MK4NPpDP1bo&t=28s

Pelletier C. (2022). Diversité et évolution de l'*Ostreid Herpesvirus type 1*. Concours Ma Thèse en 180 secondes, finale régionale. 01 avril 2022, Bordeaux, France. https://www.youtube.com/watch?v=bwZaCAlKkzE

Pelletier C. (2021). Diversity and evolution of the OsHV-1 virus. NRLs annual meeting. 16th March 2021, La Tremblade, France.

Picot S., Morga B., Faury N., **Pelletier C**., Degremont L., Arzul I., Renault T. (2017). Monitoring of the autophagy pathway in *C. gigas* haemocytes during an experimental OsHV-1 infection. EAFP - 18th International Conference on Diseases of Fish and Shellfish. 4-8 September 2017, Belfast, UK.

Liste des abréviations

A

ABC *Approximate Bayesian Computation*

ADN Acide DésoxyriboNucléique

AngHV Anguilid herpesvirus

ARN Acide ribonucléique

ARNm Acide ribonucléique messager

B

BEAST *Bayesian Evolutionary Analysis by Sampling Trees*

BTV BlueTongue virus

C

CNC Comité national de conchyliculture

CST *Cross-species transmission* ou transmission inter-espèce

CyHV Cyprinid herpesvirus

D

DGAL Direction Générale de l'ALimentation

dN/dS mutations synonymes / mutations non-synonymes

D_X Duplication du X (isomère)

E

EBV Epstein Barr virus

EE Évolution expérimentale

ELISA E*nzyme-linked immunosorbent assay*

F

FAO *Food and agriculture organization of the United Nations*

G

GaHV-1 Gallid HerpesVirus 1

H

HCMV Cytomégalovirus humain

HSV Herpes Simplex Virus

I

ICTV *International Committee on Taxonomy of Viruses*

IFREMER Institut Français de Recherche pour l'Exploitation de la MER

IHME *Institute for Health Metrics and Evaluation*

I_L/I_S *Inverse long / short* (isomères)

I_{LS} *Inverse long and short* (isomère)

InDels Insertion et délétions

IR_L/IR_S *Inverted repeat long and short*

K

KHV Koi herpesvirus

kb kilobases

L

LNR Laboratoire National de référence

LRUE Laboratoire de Référence de l'Union Européenne

M

MCMC *Markov Chain Monte-Carlo*

MCMV Murine CytoMegaloVirus

MI Maladie Infectieuse

MIE Maladie Infectieuse Émergente

MLST *Multi-Locus Sequence Typing*

MRCA Most Recent Common Ancestor

N

Ne Taille effective d'une population

NGS *Next Generation Sequencing*

N_X No X (isomère)

O

OMS Organisation Mondiale de la Santé

OMSA Organisation Mondiale de la Santé Animale

ORF Open Reading Frame

OsHV-1 Ostreid Herpesvirus type 1

P

PaV-1 Panulirus argus Virus 1

PCR *Polymerase Chain Reaction*

PFGE *Pulse Field Gel Electrophoresis*

POMS *Pacific Oyster Mortality Syndrome*

P Prototype

R

RABV RABies Virus

REPAMO REseau de surveillance des PAthologies des MOllusques marins

RFLP Restriction Fragment Length Polymorphisms

RHV Ranid herpesvirus

RNA-seq RNA sequencing

S

SARS-CoV-2 Severe Acute Respiratory Syndrome CoronaVirus 2

SIDA Syndrome d'ImmunoDéficience Acquise

SNP *Single Nucleotide Polymorphism*

SRAS Syndrome Respiratoire Aigu Sévère

T

TFF Tangential Flow Filtration

TR_L/TR_S *Terminal repeat long / short*

T_X Translocation du X (isomère)

U

U_L/U_S *Unique long / short*

USA *United States of America*

USD *United States dollars*

V

VIH	Virus de l'Immunodéficience Humaine
VZV	Varicella-Zoster Virus

W

WHO	World Health Organization
WOAH	World Organization for Animal Health
WSSV	White Spot Syndrome Virus

Table des matières

Liste des figures ***1***

Liste des encadrés ***3***

Préambule ***5***

Chapitre 1 – État de l'art ***7***

1.1 Maladies infectieuses émergentes **10**

1.1.1 Premières caractérisations d'agents pathogènes 10

1.1.2 Émergence de maladies infectieuses 11

1.1.3 Conséquences des maladies infectieuses 15

1.1.4 Transmission des maladies infectieuses 19

1.2 Les virus : diversité, mécanismes évolutifs et interactions avec leurs hôtes **22**

1.2.1 Diversité du monde viral 23

1.2.2 Mécanismes évolutifs chez les virus 28

1.2.3 Interactions hôte-virus 40

1.3 Étude des maladies infectieuses par des approches d'épidémiologie moléculaire **44**

1.3.1 De la biologie à l'épidémiologie moléculaire 45

1.3.2 Caractérisation génétique et génomique des populations virales 48

1.3.3 Approches phylodynamiques et phylogéographiques pour retracer l'histoire évolutive virale et les chaines de transmission 57

1.4 Cas des *Herpesvirus* en aquaculture et focus sur l'*Ostreid herpesvirus 1* **63**

1.4.1 Aquaculture et émergence d'*herpesvirus* 63

1.4.2 L'ostréiculture Française 67

1.4.3 Cycle de vie de l'huître creuse 69

1.4.4 Mortalités estivales chez le naissain d'huîtres creuses 71

1.4.5 L'Ostreid *Herpesvirus* 1 72

1.5 Problématiques de thèse **86**

Chapitre 2 - Développement méthodologique pour l'étude génomique du virus OsHV-1 ... ***91***

2.1 Résumé de l'article 1 ..... **93**

Article 1 ..... **95**

Chapitre 3 – Etude de l'impact de l'espèce hôte sur la diversité virale 115

3.1 Résumé de l'article 2 117

Article 2.....119

Chapitre 4 – Caractérisation spatio-temporelle de la diversité du virus OsHV-1 chez M. gigas 167

4.1 Résumé de l'article 3 169

4.2 Résumé de l'article 4 170

Article 4.....171

4.3 Résumé de l'article 5 186

Article 5.....189

Chapitre 5 – Identification des mécanismes évolutifs du virus OsHV-1 par une approche expérimentale 225

5.1 Résumé de l'article 6 227

Article 6.....229

Discussion et Perspectives.....267

Levée des verrous technologiques associés à l'étude génomique du virus OsHV-1 271

Le séquençage du génome d'OsHV-1..... 271

L'assemblage du génome d'OsHV-1..... 272

OsHV-1 évolue de façon mesurable.....274

Les variations structurales : moteur de l'évolution du virus OsHV-1..... 274

OsHV-1 évolue à des taux de substitution plus élevés qu'attendu..... 277

Effets de la sélection et de la dérive génétique..... 278

Isolement génétique et spécialisation d'hôte.....279

Historique de diffusion mondiale du virus OsHV-1 chez *M. gigas* 282

Vers une meilleure gestion de la maladie 283

Conclusion 289

Références bibliographiques 291

Annexes.....310

Liste des figures

Figure 1: Facteurs influençant l'émergence d'agents pathogènes impliqués dans les maladies infectieuses (d'après Church, 2004). 13
Figure 3 : Représentation schématique de la chaîne de transmission des agents pathogènes (d'après van Seventer & Hochberg, 2017). 19
Figure 4: Taxonomie et classification des virus en cinq domaines et correspondance avec les classes virales de Baltimore (d'après ICTV taxonomy browser et Koonin *et al.,* 2021). 25
Figure 5 : Phylogénie des *Herpesvirales* basée sur la séquence de l'ADN polymérase (C. Pelletier, Communication personnelle, 2020). 27
Figure 6 : Forces et processus évolutifs qui agissent sur les allèles dans les populations virales (adapté de Adeyemo *et al.*2021 et Sanjuán & Domingo-Calap, 2021). 29
Figure 7 : Les différents types de mutation pouvant agir sur le génome d'un virus 30
Figure 8: Corrélation entre le taux d'évolution et la taille du génome de 223 virus (d'après Sanjuán & Domingo-Calap, 2016). 32
Figure 9 : Représentation schématique des différents types de recombinaison possible chez un virus (d'après Lai *et al.*1992 et Pérez-Losada *et al.,*2015). 34
Figure 10: Différents types de sélection. 35
Figure 11 : Variation aléatoire des fréquences alléliques causées par la dérive génétique au cours de 100 générations. 39
Figure 12 : Représentation schématique de la coévolution (d'après Woolhouse *et al.*2002). 41
Figure 13 : Représentation schématique de l'évolution de la génétique, biostatistique et génétique moléculaire au cours du temps (adapté de Nelson *et al.,* 2013). 45
Figure 14: Représentation schématique des indices de diversité nucléotidiques θ_π et θ_S. 49
Figure 15 : Représentation schématique de l'estimation du D de Tajima. 51
Figure 16 : Comportement de l'indice de fixation F_{ST} dans des scénarii différents de différenciation génétique 52
Figure 17 : Représentation schématique de séquence avec et sans déséquilibre de liaison. (D'après Rafalski, 2002) 53
Figure 18 : Exemple de réseau phylogénétique construit à partir de 42 génomes du virus *Varicella-Zoster* montrant de possibles évènements de recombinaison représentés par les réticulations entre les branches (adapté de Zell *et al.*, 2012). 54
Figure 19 : Exemple schématique de représentation des relations de parenté entre 11 virus sous forme d'arbre phylogénétique. 55

Figure 20 : Illustration d'une reconstruction de coalescent à partir d'échantillons à taille de population constante (d'après Drummond *et al.*2002 et Rosenberg *et al.*2002). 57
Figure 21 : Exemples de modèles d'horloge moléculaire, de coalescent et de diffusion, implémentés dans BEAST pour les analyses phylodynamiques (d'après Drummond & Rambaut, 2007, Volz *et al.*2013, et Nahata *et al.*2021)..... 62
Figure 22 : Impacts du changement climatique sur les relations hôtes-pathogène-environnement marin (d'après Burge *et al.*2014) 66
Figure 23 : Tableau « Déjeuner aux huîtres » peint en 1735 par Jean-François de Troy (1679-1752). 68
Figure 24: Production ostréicole Française et émergence des épizooties entre 1950 et 2019 (d'après FAO, 2020). 69
Figure 25 : Images des principaux stades de développement de l'huître creuse *M. gigas*. 70
Figure 26: Cycle de reproduction biologique de l'huître creuse *M. gigas* (d'après Vogeler *et al.*2016). 70
Figure 27 : Représentation schématique du POMS dans les branchies d'huître (d'après Petton *et al.*2021)..... 72
Figure 28 : Répartition mondiale du virus OsHV-1. 73
Figure 29 : Représentation du virus OsHV-1. 75
Figure 30 : Représentation schématique de l'architecture du génome du virus OsHV-1, des ORFs et du contenu en guanine et cytosine. 76
Figure 31 : Structure du palindrome présent dans le génome d'OsHV-1 entre les positions 65 018 et 65 218 pb (d'après Davison *et al.*2005)..... 77
Figure 32 : Représentation schématique des 16 isomères génomiques du virus OsHV-1 (d'après Davison *et al.*2005). 79
Figure 33 : Nombre d'ORFs du génome d'OsHV-1 codant pour une fonction ou un domaine (d'après Davison *et al.*2005 et Burioli *et al.*, 2017). 80
Figure 34 : Phase lytique du cycle de réplication supposé du virus OsHV-1 (d'après Le Deuff *et al.,* 1994 ; Renault *et al.*, 1994 ; Segarra 2015). 81
Figure 35 : Exemple simplifié d'intégration chromosomique d'un génome d'*herpesvirus* au sein des parties télomériques des chromosomes de l'hôte après recombinaison homologue facilitée par les régions répétées (R) du génome virale (d'après Morissette & Flamand, 2010). 82
Figure 36 : Amorces PCR développées pour l'étude du virus OsHV-1 et du variant µVar. 84
Figure 37 : Représentation graphique des principaux objectifs de ma thèse. 89
Figure 38 : Représentation graphique des principales problématiques abordées dans cette thèse, principaux résultats, impacts pour les professionnels collaborateurs et perspectives. 270
Figure 39 : Pipeline bionformatique utilisé pour l'analyse génomique du virus OsHV-1. 273

Figure 40 : Capture d'écran de la page d'accueil de la plateforme web MoPSeq-DB développée par ASIM pour la gestion et la visualisation des données de séquences des mollusques marins. 274
Figure 41 : Génomique comparative des 9 génomes représentatifs des génotypes du virus OsHV-1 mis en évidence durant ma thèse. 276
Figure 42 : Phylogénie de maximum de vraisemblance montrant les relations évolutives entre les 308 génotypes viraux collectés chez différentes espèces de bivalves dans le monde. 281
Figure 43 : Modèle à compartiments de transmission de *Vibrio aestuarianus* chez les huîtres et dans leur environnement. 284
Figure 44 : Image de la valise développée dans le cadre du projet SEA-MONITOR financé par la Direction de l'Innovation de l'Ifremer pour améliorer la surveillance en temps réel des agents pathogènes viraux des mollusques marins. 285

Liste des encadrés

Encadré 1 : Description des principaux micro-organismes pathogènes impliqués dans les maladies infectieuses 9
Encadré 2 : Endémie, épidémie, épizootie ou pandémie ? 11
Encadré 3 : Maladies humaines d'origine zoonotique. 14
Encadré 4 : Définition d'un phénotype, sérotype et génotype viral 15
Encadré 5 : Définition d'un écosystème 16
Encadré 6 : Interactions biologiques entre deux individus 22
Encadré 7 : Qu'est-ce qu'un allèle ? 28
Encadré 8 : Balayage sélectif 36
Encadré 9 : Goulot d'étranglement 39
Encadré 10 : Transmission inter-espèce ou *cross species transmission* 43
Encadré 11 : Coefficient de sélection 58
Encadré 12 : Méthode de Monte-Carlo par chaînes de Markov (MCMC) 59
Encadré 13 : Inférence des routes de dispersion virales en phylogéographie 59
Encadré 14 : Isomères génomiques 78

Préambule

La compréhension des agents pathogènes chez les mollusques bivalves revêt une importance cruciale dans le contexte de la santé des populations et de la sécurité alimentaire. Parmi ces agents pathogènes, l'Ostreid Herpesvirus type 1 (OsHV-1) se démarque en raison de son impact dévastateur sur les huîtres creuses *Magallana gigas* et sa détection occasionnelle chez d'autres espèces de bivalves. Cependant, l'étude du virus OsHV-1 a été entravée par des défis techniques majeurs, notamment l'absence de lignée cellulaire pour sa culture *in vitro* et la complexité de son génome.

Dans ce contexte, la thèse présentée ici s'est efforcée de relever ces défis en utilisant des nouvelles générations de séquençage (NGS) et une approche bioinformatique spécifique. L'objectif principal était d'acquérir, de séquencer et d'assembler un grand nombre de génomes du virus OsHV-1 pour mieux comprendre *i*) la distribution de ses génotypes chez différentes espèces hôtes, en particulier les huîtres plates et creuses, *Ostrea edulis et Magallana gigas*, et *ii*) la structuration de sa diversité génétique spatio-temporelle *iii*) ses routes de dispersion, et *iv*) les mécanismes évolutifs sous-jacents aux patrons de diversité observés.

Ce manuscrit introduit les principales notions et concepts abordés pendant mes travaux de thèse de doctorat, établit les problématiques traitées et relate les résultats clés de mes recherches sous forme de chapitres-articles, mettant en lumière l'importance de l'utilisation de technologies de pointe pour améliorer la compréhension des agents pathogènes marins et leur évolution au sein des populations de bivalves.

Les découvertes présentées dans cette thèse sont discutées et mises en perspectives car elles contribuent non seulement à notre connaissance des agents pathogènes chez les mollusques bivalves, mais elles ont également des implications significatives pour la mise en place de stratégies de prévention et de contrôle de maladie d’élevage aquacole.

Chapitre 1 – État de l’art

C. Pelletier

Dans notre monde contemporain, connecté, où les barrières géographiques disparaissent et où les écosystèmes sont perturbés, les maladies infectieuses émergent à un rythme accéléré, rappelant l'inévitable pouvoir d'adaptation et de propagation des micro-organismes pathogènes (Bonneaud & Longdon, 2020). Les virus, les bactéries, les protozoaires, les champignons et les vers parasites sont les principaux micro-organismes pathogènes impliqués dans les maladies infectieuses (Encadré 1).

Encadré 1 : Description des principaux micro-organismes pathogènes impliqués dans les maladies infectieuses.

	Virus	Bactérie	Champignon	Protozoaire	Ver parasite
Taille	0,05 - 0,3 µm	0,1 - 10 µm	2 - 10 µm	10 - 800 µm	200 µm - 1 m
Organisme	Non	Unicellulaire	Multicellulaire	Unicellulaire	Multicellulaire
Noyau	Non	Non	Oui	Oui	Oui
Génome	ARN/ADN 10^2-10^5 pb	ADN 10^5-10^7 pb	ADN 10^7-10^8 pb	ADN 10^7-10^9 pb	ADN 10^7-10^9 pb
Temps d'incubation	Rapide + (jours)	Rapide (jours/mois)	Lent (mois/années)	Rapide (jours/mois)	Lent + (mois/années)
Durée de vie	Court (heures/jours)	Long (jours/mois)	Long + (années)	Long (mois/années)	Long + (années)
Exemples	*Bluetongue virus* *Rabies virus* *Avian infleunza A* *SARS-CoV-2* *Human herpesvirus 1* *Ostreid herpesvirus 1*	*Mycobacterium tuberculosis* *Borrelia burgdorferi* *Staphylococcus aureus* *Bacillus anthracis* *Brucella abortus* *Vibrio spp.*	*Candida spp.* *Aspergillus fumigatus* *Cryptococcus neoformans* *Scedosporium spp.* *Lomentospora prolificans*	*Plasmodium spp.* *Trypanosoma spp.* *Haplosporidium spp.* *Marteilia refringens* *Bonamia ostreae*	*Taenia spp.* *Ancylostoma duodenale* *Necator americanus* *Fasciola hepatica* *Dactylogyrus spp.*

Dans cet état de l'art, je commencerai par définir les maladies infectieuses émergentes, en abordant ensuite leurs principales conséquences, tout en mettant en lumière les facteurs essentiels associés.

Ensuite, je décrirai les interactions entre les espèces hôtes et leurs agents pathogènes, en mettant en lumière les mécanismes évolutifs qui facilitent leur adaptation, tout en explorant les méthodes disponibles pour l'étude de ceux-ci. Enfin, je présenterai le modèle d'étude que j'ai utilisé dans mes travaux de thèse l'*Ostreid Herpesvirus 1*, puis je terminerai par les problématiques et les questions scientifiques abordées pendant mon doctorat.

1.1 Maladies infectieuses émergentes

Une maladie est dite « émergente » ou « ré-émergente » dès lors que son incidence augmente de façon significative dans une population et sur un territoire donné. La première notion de maladie infectieuse émergente (MIE) a été proposée par l'épidémiologiste Stephen S. Morse (Morse, 1995) qui décrit les maladies infectieuses (MI) comme un processus en deux étapes : *i)* introduction de l'agent pathogène nouveau ou d'un variant dans une nouvelle population d'hôtes, puis *ii)* implantation et dispersion de celui-ci dans la nouvelle population. Le contrôle des MIE est crucial pour la santé publique, mais également pour la préservation des populations animales sauvages ou d'élevage, et la santé des plantes (Brooks *et al.*, 2022; Hayek, 2022). Celui-ci passe par la caractérisation des micro-organismes pathogènes impliqués dans ces maladies, le plus précocement possible, pour limiter une propagation rapide à l'échelle mondiale (Hayek, 2022).

1.1.1 Premières caractérisations d'agents pathogènes

La découverte des premiers agents pathogènes a joué un rôle central dans l'histoire de la médecine et de la microbiologie. Au cours du XIXe siècle, plusieurs chercheurs pionniers ont été essentiels dans l'identification des agents pathogènes impliqués dans les maladies et la compréhension des mécanismes infectieux. Le terme "pathogène" a été utilisé pour la première fois dans les années 1880 (Casadevall & Pirofski, 2014; Cunliffe, 2008), pour désigner un organisme capable de causer une maladie. En 1882, le médecin allemand Robert Koch a identifié le bacille responsable de la tuberculose, *Mycobacterium tuberculosis*. C'est la première découverte d'un agent pathogène spécifique à l'Homme. Il a utilisé une combinaison de techniques de coloration et de cultures bactériennes pour isoler et caractériser l'agent pathogène, établissant ainsi le concept de l'agent causal spécifique d'une maladie.

En 1886, Mayer fut le premier à décrire les propriétés infectieuses de la mosaïque du tabac. Il démontra que cette maladie pouvait être transmise en mettant en contact des feuilles de tabac saines avec un broyat de feuilles de tabac malades (Lecoq, 2001; Mayer, 1886). Ces travaux furent ensuite poursuivis par Dmitri Ivanovski en 1892, qui identifia l'agent pathogène responsable de la mosaïque du tabac comme étant une "très petite bactérie capable de passer à travers les filtres de Chamberland". Martinus Beijerinck confirma en 1898, le pouvoir infectieux des extraits de plantes malades après filtration uniquement à l'intérieur de cellules vivantes, et nomma "virus" l'entité responsable de la maladie (Lecoq, 2001). Cette découverte fondamentale a établi la nature des virus

en tant qu'entités infectieuses, distinctes des bactéries et des autres agents pathogènes connus à ce moment-là.

Louis Pasteur est également un précurseur en matière d'études des agents pathogènes. Entre 1865 et 1870, il observe le champignon responsable de la pébrine du ver à soie. Cette maladie est caractérisée par des tâches marrons sur le corps de la larve ou de la chenille, ainsi que sur l'adulte. Elle entrave la production des fibres de soie nécessaires à la construction du cocon (Pasteur, 1922).

Les découvertes de Koch, Mayer, Ivanovski, Beijerinck et Pasteur ont posé les bases de la microbiologie, bactériologie et virologie moderne et, ont ouvert la voie à de nombreuses avancées majeures dans la compréhension, le diagnostic et le traitement des MI.

1.1.2 Émergence de maladies infectieuses

Des MI émergent ou ré-émergent régulièrement, provoquant des épidémies, c'est à dire une augmentation soudaine de l'incidence d'une maladie par rapport à ce qui est attendu dans une région. Parfois ces épidémies peuvent s'amplifier jusqu'à devenir une pandémie, c'est à dire une épidémie qui se répand dans le monde entier (Encadré 2).

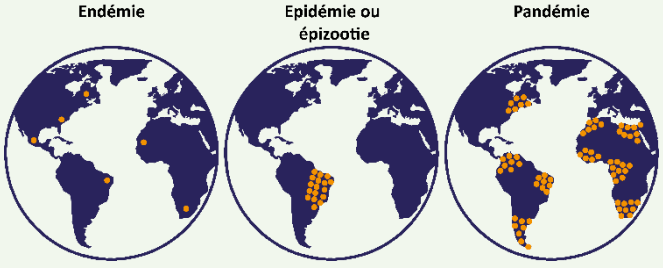
Encadré 2 : Endémie, épidémie, épizootie ou pandémie ?

Endémie : apparition et propagation d'une maladie dont le nombre de cas est constant dans une région précise.

Épidémie : apparition et propagation d'une maladie qui affecte un grand nombre d'individus plus ou moins en même temps au sein d'une communauté, d'une population ou d'une région (*e.g.* Dengue).

Épizootie : épidémie qui touche les animaux (*e.g.* grippe aviaire).

Pandémie : épidémie qui se répand dans plusieurs pays ou continents (*e.g.* Covid-19).



Différences entre une endémie, une épidémie et une pandémie (d'après Center for Disease Control and Prevention,2012 et Brabaw, 2020).

(Center for Disease Control and Prevention, 2012 et Brabaw, 2020)

Régulièrement, l'Organisation mondiale de la santé (OMS), l'Organisation mondiale de la santé animale (OMSA) et l'organisation des nations unies pour l'alimentation et l'agriculture (FAO) lancent des alertes concernant la rapidité sans précédent de l'émergence de nouvelles maladies infectieuses humaines et animales. Une nouvelle base de données mise en place par l'OMSA, fournit notamment des informations en temps réel, sur différentes épidémies en cours dans le monde (https://wahis.woah.org/#/home).

L'émergence et la propagation des MI est un phénomène complexe et multifactoriel (van Seventer & Hochberg, 2017)(Figure 1). La modification de l'environnement et le changement climatique peuvent favoriser l'apparition de nouvelles maladies en modifiant ou créant de nouveaux habitats pour les hôtes, vecteurs et agents pathogènes (Nova *et al.*, 2022). La mondialisation, *via* l'augmentation des trajets et les échanges commerciaux, facilite l'introduction d'agents pathogènes dans de nouvelles aires géographiques, et donc, la propagation des MI (Knobler *et al.*, 2006; Nova *et al.*, 2022). Il est évoqué une « mondialisation épidémique ». Les facteurs socio-économiques peuvent également jouer un rôle important dans la propagation des MI, notamment par le fait que la transmission dans les populations les plus vulnérables est souvent facilitée par des conditions d'hygiène moins bonnes ou des densités de population plus élevées (Nova *et al.*, 2022).

De plus, la plupart des agents pathogènes sont dotés de capacités d'adaptation à des espèces hôtes ou à des environnements différents. Ces adaptations représentent un enjeu majeur pour la santé publique mondiale, étant donné que 61% des MI et 75% des MIE chez l'homme sont d'origine zoonotique (Encadré 3), c'est-à-dire qu'elles se transmettent de l'animal à l'homme (Magouras *et al.*, 2020; Polack *et al.*, 2015; Said *et al.*, 2023; van Doorn, 2014). Au contraire, une maladie est dite épizootique si elle ne se transmet pas à l'humain. Les MIE récentes sont fréquemment caractérisées par une transmission entre plusieurs espèces hôtes et sont causées le plus souvent par des virus, comme par exemple pour les virus Sars-Cov-2, MERS-Cov et Nipah virus, et des bactéries, comme pour la borréliose de Lyme (Steere *et al.,*2004).

D'autres causes de réémergence de maladies infectieuses incluent les incidents en laboratoire, comme les problèmes de traitement des aérosols (*i.e.* infection de laborantins puis de 10 000 personnes par la brucellose (*Brucella melitensis*) en Chine en 2019, Pappas, 2022), les mutations (voir section 1.2.2.1) ou la recombinaison (voir section 1.2.2.2) chez les vaccins vivants atténués, pouvant entraîner la réversion du virus vers une souche sauvage ou l'apparition d'une nouvelle souche virulente (Hanley, 2011). Parfois des cas extrêmement rares de réémergence sont liés au bioterrorisme, comme l'utilisation de l'anthrax aux États-Unis en 2001 (Morens *et al.*, 2004).

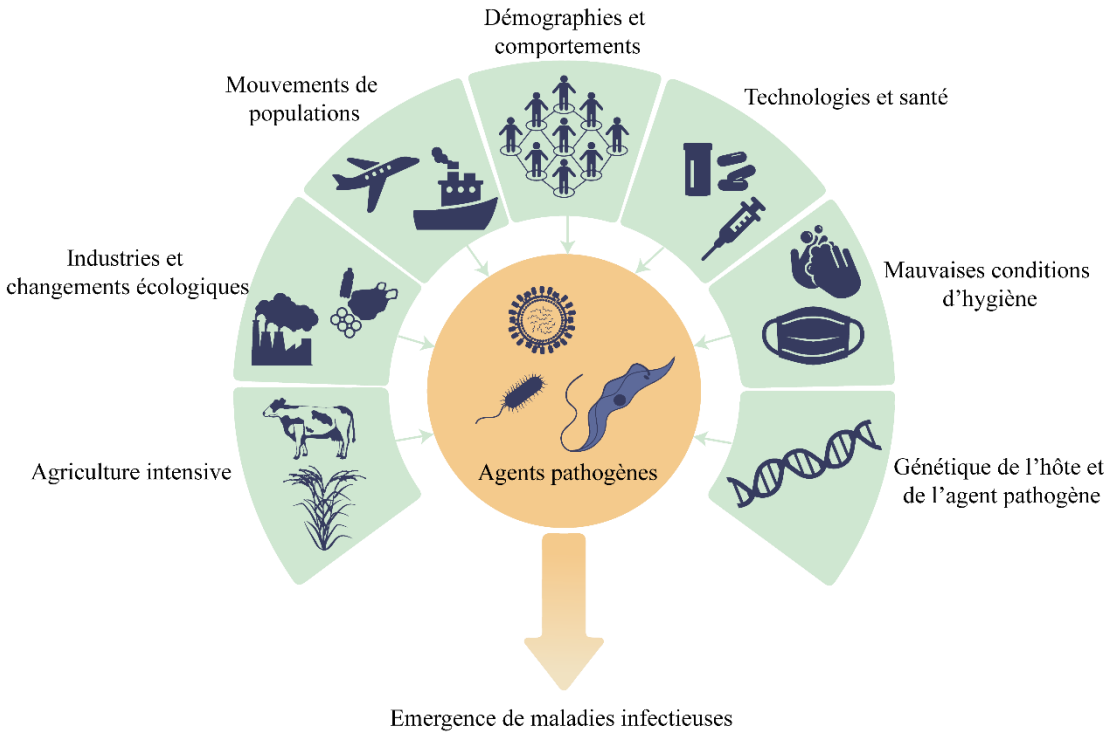


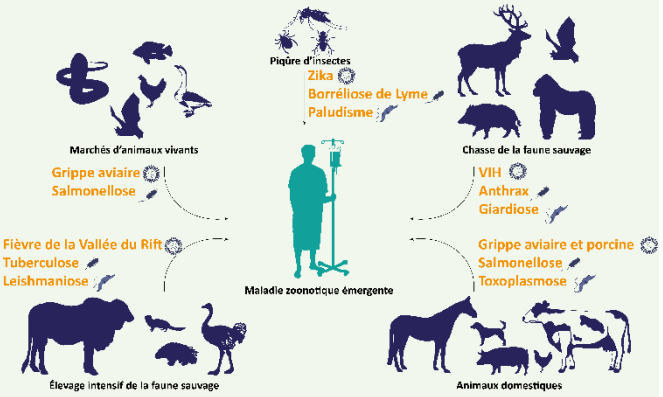
Figure 1: Facteurs influençant l'émergence d'agents pathogènes impliqués dans les maladies infectieuses (d'après Church, 2004).

Les facteurs favorisant l'émergence d'agents pathogènes sont multiples. L'agriculture intensive, le développement des industries et les changements écologiques, peuvent être à l'origine de l'émergence d'un agent pathogène. L'augmentation des mouvements de populations en lien avec la mondialisation ainsi que la taille et le comportement des populations, peuvent favoriser la transmission et la dispersion de celui-ci. Certaines technologies ou matériels utilisés dans le cadre des systèmes de santé peuvent être la cause de l'émergence ou de l'accélération de la dispersion d'un agent pathogène. Les mauvaises conditions d'hygiène, que ce soit dans les systèmes de santé publiques ou dans les exploitations animales, sont également une source de développement d'agents pathogènes, et peuvent entrainer la propagation de celui-ci. Enfin, un facteur important dans l'émergence de maladies infectieuses est la génétique de l'hôte et de l'agent pathogène lui-même. Celle-ci permet l'adaptation plus ou moins rapide de l'agent pathogène à son hôte et détermine donc le développement de la maladie infectieuse.

Parmi les émergences, nombreuses sont celles qui impliquent des virus. En effet, entre 1900 et 2023 plus de 300 virus ont émergé et ont été responsables de crises épidémiques ou pandémiques majeures. L'une des plus importantes pandémies, la grippe Espagnole, causée par le virus de la grippe Influenza A (H1N1), a tué entre 30 et 50 millions de personnes entre 1918 et 1920. On peut également citer le virus Ebola apparu en 1976 au Soudan et au nord de la république du Congo (Weyer *et al.,*2015), qui ré-émerge régulièrement en Afrique, où il cause des pertes humaines considérables (*i.e.* 11 000 morts entre 2013 et 2016 en Afrique de l'Ouest, 2 300 morts entre 2018 et 2020 en République du Congo) (Weyer *et al.,*2015, Jacob *et al.,*2020). Très récemment, le SARS-Cov-2 responsable de la pandémie de Covid-19 aurait coûté la vie à près de 7 millions de personnes dans le monde entre 2019 et 2023 (https://covid19.who.int/, consultée le 25/06/2023).

Encadré 3 : Maladies humaines d'origine zoonotique.

Les maladies zoonotiques sont des maladies causées par des organismes pathogènes (*i.e.* virus, bactéries, protozoaires, etc) qui sont naturellement transmises de l'animal à l'homme par un événement de transmission inter-espèce. Parmi les exemples récents de maladies zoonotiques figurent la Toxoplasmose (causée par le protozoaire *Toxoplasma gondii*), la Salmonellose (causée par des bactéries du genre *Salmonella*) ou encore la grippe a*via*ire (causée par le virus *Avian Influenza A*).



Exemples de maladies zoonotiques qui ont émergé ou réémergé (d'après Magouras *et al.*, 2020 et Polack *et al.*2015).

Ces types de maladies représentent un enjeu majeur pour la santé publique car leur émergence n'est pas prévisible, elles peuvent être responsables de pandémies et leur gestion extrêmement compliquée en lien avec la multiplicité d'hôtes possibles et les mouvements de populations associés. Ces maladies émergent très souvent dans des régions pauvres où les conditions d'élevages sont mauvaises et les mesures d'hygiène sont moindres.

Magouras *et al.*, 2020 et Polack *et al.*2015

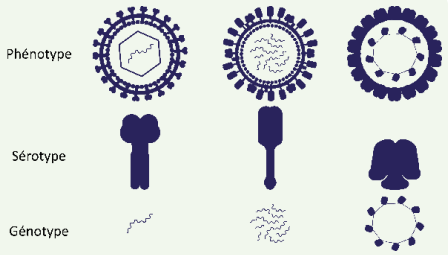
Les émergences virales existentent également chez d'autres espèces animales vertébrés et invertébrés. Un grand nombre de maladies infectieuses d'origine virale se développent au sein des populations d'animaux sauvages (*e.g.* mangoustes, chameaux, ratons laveurs, daurades, huîtres), ou dans les élevages d'animaux domestiques (*e.g.* vaches, cochons, chèvres), en raison de la proximité entre les individus (Bayry, 2013). Parmi les maladies d'animaux d'élevage, figure la fièvre catarrhale causée par le virus Bluetongue (BTV) depuis les années 1850 (Maclachlan *et al.*, 2015). Cette maladie transmise uniquement par des insectes, et notamment par des moucherons piqueurs, affecte principalement les ruminants sauvages (*e.g.* antilopes, buffles) et d'élevage (*e.g.* chèvres, bovins, moutons) (WOAH, 2023b). La fièvre catarrhale est originaire d'Afrique, mais actuellement détectée en Asie, en Europe, en Australie, et en Amérique du Nord (WOAH, 2023b). Elle cause principalement des diminutions de rendement et des avortements voire, parfois, des mortalités, et a donc des conséquences économiques importantes. Malgré la mise au point de stratégies de luttes basées sur le dépistage à grande échelle ou la restriction de mouvements des animaux ainsi que sur le développement de vaccins, la fièvre catarrhale reste difficile à contrôler. Actuellement, 28 sérotypes (Encadré 4) ayant des virulences variables circulent à travers le monde (Maclachlan *et al.*, 2015). L'aquaculture est un milieu propice à l'émergence des virus. Plusieurs virus sont problématiques et répertoriés sur les listes de l'OIE (Kibenge, 2019). C'est le cas de l'infection des turbots et des truites arc-en-ciel par le virus de l'hémorragie septicemie virale (VHS) ou, de l'infection des Salmonidés par le

virus Infectieux hematopoiétique nécrosant (IHNV) qui causent d'importantes mortalités dans les élevages de poissons (Kibenge, 2019).

Encadré 4 : Définition d'un phénotype, sérotype et génotype viral

Phénotype : caractéristiques observables d'un virus, telles que la structure morphologique ou les processus de réplication. Le phénotype peut, par exemple, être étudié *via* des approches de microscopie.

Sérotype : ensemble des caractéristiques antigéniques (*e.g.* glycoprotéines) d'un virus permettant de différencier des souches virales appartenant à une même espèce.



Différences entre phénotype, sérotype et génotype.

Les sérotypes sont étudiés par le dosage des anticorps circulants dans l'individu hôte infecté ou par des méthodes immuno-enzymatiques (ELISA) qui permettent la conjugaison d'antigènes spécifiques avec les protéines virales.

Génotype : patrimoine génétique (ADN ou ARN) du virus, dépendant de l'ensemble des gènes composant le génome. Les génotypes sont étudiés par des approches de génotypage qui permet l'identification de variation génétique pour une position génétique précise, ou de séquençage de génome qui permet l'identification de l'ensemble des variations génétiques présentes dans le génome.

(Taylor et Lewontin, 2017, Maclachlan *et al.*, 2015)

1.1.3 Conséquences des maladies infectieuses

L'histoire des MI est intimement liée à celle de l'humanité, de l'agriculture et de la domestication des animaux. Leur impact sur les populations a été profond et souvent dévastateur (Debré, 2020). Au cours de l'histoire, de nombreuses maladies infectieuses ont émergé et se sont propagées à travers les populations avec de nombreuses conséquences variables, et étroitement liées les unes aux autres (**Erreur ! Source du renvoi introuvable.**).

Tout d'abord, les MI peuvent avoir des conséquences démographiques significatives. Elles peuvent entrainer de fortes diminutions des populations hôtes touchées. En effet, les MI sévères peuvent entraîner des taux de mortalité élevés, réduisant ainsi considérablement les effectifs de la population infectée. Chez l'homme, les maladies infectieuses sont responsables de 14 millions de décès dans le monde chaque année (Gray & Sharara, 2022). Dans les élevages, il peut être nécessaire de sacrifier volontairement des individus pour limiter voire éradiquer la maladie infectieuse (*e.g.* abattage de poules et de canards en cas d'épidémie de grippe aviaire, Gerritzen *et al.*, 2006). Certaines MI peuvent affecter de manière préférentielle certains groupes d'âge ou de sexes (Gay *et al.*, 2021; Glynn & Moss,

2020), créant ainsi un déséquilibre de la structure démographique habituelle de la population. Par exemple, la Covid-19 a particulièrement impacté les personnes âgées ainsi que les personnes fragiles atteintes de comorbidités (*e.g.* diabètes, surpoids, problèmes cardiaques ou respiratoires), pour lesquelles des mortalités plus importantes ont été observées (Djaharuddin *et al.*, 2021), créant ainsi un déséquilibre en faveur des individus les plus résistants. Ce type de perturbation peut avoir des conséquences sur la dynamique reproductive et le taux de renouvellement de la population. Les MI peuvent aussi avoir un impact sur la distribution géographique des populations animales. Afin d'éviter la maladie, les animaux peuvent se déplacer vers d'autres régions ou habitats, ce qui modifie leur répartition spatiale (Loehle, 1995). Ces changements peuvent avoir des répercussions sur les écosystèmes (Encadré 5) ou les interactions entre les espèces, et sont en lien avec des conséquences écologiques.

En effet, les MI peuvent également entraîner des modifications écologiques, telles que des modifications des habitats, de la biodiversité, des chaînes alimentaires ou des interactions avec les écosystèmes (Buck & Weinstein, 2020).

Encadré 5 : Définition d'un écosystème

En écologie, un écosystème désigne une communauté biologique d'organismes (biocénose) interagissant entre eux et avec leur environnement (biotope). L'échelle d'étude d'un écosystème peut varier d'une échelle globale comme la Terre, à une échelle réduite comme une goutte d'eau.

Illustration d'un écosystème marin.

Les écosystèmes sont des réseaux complexes et dynamiques d'interactions durables qui permettent un équilibre de vie pour toutes les espèces présentes.
Ainsi, ils permettent la mise en place et la stabilité des réseaux trophiques, de la prédation ou de la reproduction des différentes espèces présentes dans l'écosystème.

Cependant, les écosystèmes sont fortement impactés et déséquilibrés par les activités anthropiques, ce qui entraînent des dérèglements, aussi bien au niveau du biotope (*e.g.* pollution, réchauffement climatique) que de la biocénose (*e.g.* extinction d'espèce, diminution de la valeur sélective).

(Rao & Larsen, 2010)

Les écosystèmes sains et équilibrés dépendent de la présence et de l'interaction de différentes espèces (Rao & Larsen, 2010). La diminution de la population d'une espèce animale, voire son extinction, en raison d'une épizootie par exemple, peut avoir des répercussions sur la diversité des autres espèces et la santé des écosystèmes (Rao & Larsen, 2010). Un bon exemple de conséquences écologiques des MI est l'extinction de certaines espèces d'oiseaux présentes sur l'île d'Hawaii à la fin du 19ème siècle. Suite à l'introduction de moustiques *Culex quinquefasciatus* infectés par le virus birdpox et par le paludisme a*via*ire, les populations d'oiseaux ont été décimées. Ces extinctions ont eu pour conséquences une diminution de la pollinisation de certains arbres et une limitation de leur reproduction (Cox, 1983; Rao & Larsen, 2010; Samuel *et al.*, 2015).

Les épizooties peuvent de fait perturber les chaînes alimentaires et les interactions écologiques. Par exemple, la diminution de la taille de population d'une espèce animale peut avoir des répercussions sur les populations d'espèces prédatrices qui dépendent de cette espèce comme source de nourriture (Rao & Larsen, 2010). Les populations de lapin (*Oryctolagus cuniculus*) dans les îles Anglaises ont été décimées par la myxomatose, maladie causée par le virus myxomateux. Cela a entrainé une diminution drastique de la taille des populations prédatrices du lapin dans ces zones géographiques, à savoir les renards, les fouines, les chouettes ou les buses (Sumption & Flowerdew, 1985). Certaines espèces de prédateurs ont alors changé de proies, permettant une réaugmentation de la taille de la population prédatrice (Sumption & Flowerdew, 1985).

Les conséquences comportementales et sociologiques d'une MI peuvent être directes ou indirectes. Parmi les conséquences directes figurent l'isolement des individus, la peur collective, les tensions politiques ou encore la réorganisation des systèmes de santé. Chez l'Homme, la pandémie de SARS-CoV-2 réunit, à elle seule, toutes ces conséquences directes. Chez les animaux, les langoustes saines refusent de s'abriter avec les individus infectés par le virus mortel *Panulirus argus virus 1* (PaV1). Elles pratiquent ainsi un isolement du cas infecté pour limiter la transmission du virus (Behringer *et al.*, 2006). D'autres exemples similaires existent chez les Mandrills, pour lesquels il est question de « système immunitaire comportemental » (Poirotte *et al.*, 2017).

Les conséquences indirectes d'une épidémie peuvent être de l'ordre de la stigmatisation. Par exemple le rejet d'un groupe d'individus primo-infectés par l'agent pathogène ou celui des personnels de santé et humanitaires aidant les populations locales (*e.g.* isolement social, abus verbaux et peur de la contagion dans le cas de maladies comme Ebola ou le SIDA ; Chollier *et al.*, 2016; James *et al.*, 2020). Dans le cas des épizooties dans les élevages, les conséquences peuvent être la perte du travail et des revenus de l'éleveur, mais aussi un impact sur la sécurité alimentaire. Lorsque des maladies infectieuses touchent les animaux d'élevage ou les espèces sauvages comestibles,

l'approvisionnement alimentaire peut être compromis (Brooks *et al.*, 2022). Cela peut entraîner une augmentation des prix des denrées alimentaires, une réduction de la disponibilité des produits d'origine animale, et des conséquences nutritionnelles pour les communautés qui dépendent de ces ressources pour leur alimentation (Brooks *et al.*, 2022). Les conséquences comportementales des MI peuvent entraîner des conséquences économiques.

Les conséquences économiques des MI sont considérables pour un pays mais aussi à l'échelle mondiale, aussi bien pour les épizooties que pour les épidémies. Lorsqu'une épizootie se déclare, elle peut entraîner des pertes importantes dans le secteur agricole produisant des animaux d'élevage. Les maladies infectieuses peuvent se propager rapidement au sein des troupeaux, entraînant la mort de nombreux animaux et causant des pertes économiques substantielles pour les éleveurs (Saatkamp *et al.*, 2016; WOAH, 2007). Les coûts liés aux soins vétérinaires, au traitement des animaux malades et à la mise en place de mesures de biosécurité peuvent également être élevés (Saatkamp *et al.*, 2016; WOAH, 2007).

De plus, les épizooties peuvent avoir un impact significatif sur les chaînes d'approvisionnement et le commerce international. Lorsqu'une MI est détectée dans une région ou un pays, des restrictions commerciales peuvent être imposées pour prévenir la propagation de la maladie. Cela peut entraîner des diminutions d'exportation, des baisses de production et des échanges internationaux (Saatkamp *et al.*, 2016; WOAH, 2007). Les répercussions économiques peuvent être ressenties par les travailleurs de l'industrie, notamment les éleveurs, les ouvriers des abattoirs, les vendeurs et les transporteurs (WOAH, 2007). En plus des pertes économiques liées à l'arrêt de la production, il est également nécessaire d'investir dans des consommables (*e.g.* masques, gants, blouses, *etc.*) pour la gestion de l'épidémie, voire de pratiquer la vaccination des individus sains pour protéger les populations (Saatkamp *et al.*, 2016; WOAH, 2007). Ces pratiques ont également un coût économique élevé. En Allemagne par exemple, le coût économique de l'épidémie de BTV-8, infectant le bétail, entre 2006 et 2009 a été estimé entre 157 et 203 millions d'euros de pertes pour le pays (Gethmann *et al.*, 2020). A Madagascar, les exploitations de crevettes sont touchées par le virus *White spot syndrome virus* (WSSV) qui cause chaque année près d'un million de dollars de perte, pour un total de plus de 15 millions de dollars depuis son émergence en 1991 (Onihary *et al.*, 2021).

Pour limiter les conséquences des MI, les scientifiques tentent d'identifier l'ensemble des hôtes des agents pathogènes, ainsi que les chaines de transmission qui permettent la dispersion de ces derniers. L'objectif final est d'aider à la prise de décision par les politiques de santé publiques et, de développer des mesures de contrôle qui visent notamment à casser les chaines de transmission.

1.1.4 Transmission des maladies infectieuses

La transmission d'une MI est un processus complexe, impliquant la propagation de l'agent pathogène d'individus infectés à des individus sains. Elle peut être représentée sous forme de chaîne d'infection (Figure 2, van Seventer & Hochberg, 2017). Elle correspond à un ordonnancement précis de maillons qui permettent la transmission de l'agent pathogène (van Seventer & Hochberg, 2017). Cette chaîne de transmission contient six maillons : *i)* les agents infectieux, *ii)* les hôtes réservoirs, *iii)* les voies de sortie utilisées par l'agent pathogène, *iv)* les modes de transmission, *v)* les voies d'entrée dans l'hôte sensible et *vi)* les hôtes sensibles.

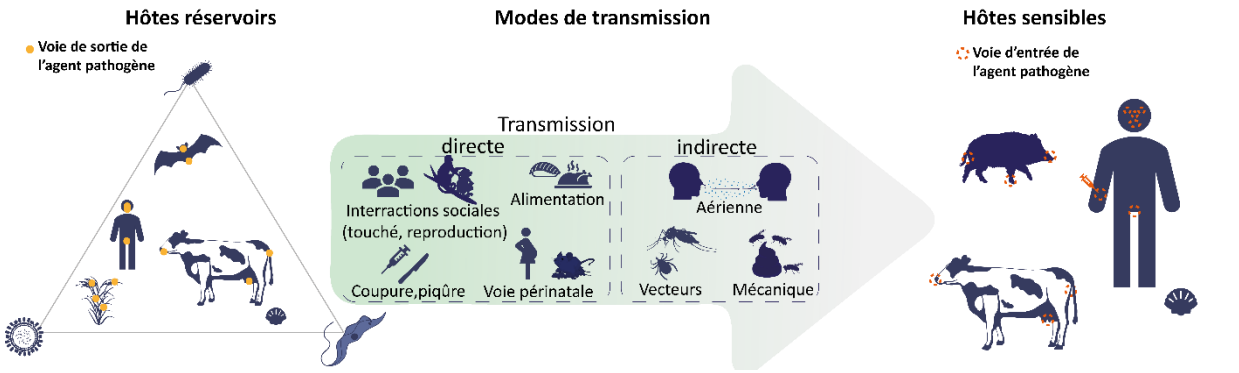


Figure 2 : Représentation schématique de la chaîne de transmission des agents pathogènes (d'après van Seventer & Hochberg, 2017).

A) Hôtes réservoirs représentés par des organismes vivants. Les hôtes réservoirs servent pour la multiplication et la dissémination des agents pathogènes (portes de sorties). B) Modes de transmission des agents pathogènes classés en deux groupes : directe et indirecte. Le mode de transmission sert de transport à l'agent pathogène pour aller infecter un nouvel hôte. C) Hôtes sensibles des agents pathogènes. Les hôtes sensibles sont des individus sains, non immunisés, qui vont développer la maladie suite à l'infection, et qui présentent des voies d'entrées pour l'agent pathogène.

Le premier maillon de la chaîne de transmission est l'agent infectieux caractérisé par ses niveaux d'infectiosité (*i.e.* sa capacité à infecter un hôte), de pathogénicité (*i.e.* sa capacité à causer une maladie) et de virulence (*i.e.* l'intensité de sa capacité à provoquer une maladie). Certains agents pathogènes, tels que les bactéries, champignons ou certains virus, peuvent persister longtemps en dehors de leur hôte (*e.g. Bacillus anthracis* ou anthrax capable de survivre des décennies sous forme de spores ; WHO, 2008, le virus de l'hépatite A dans un milieu hydrique peut survivre plus de 200 jours ; Rzeżutka & Cook, 2004, ou le champignon *Aspergillus fumigatus* capable de survivre plus de 30 jours sur du polyuréthane ; Neely & Orloff, 2001). Au contraire, d'autres agents pathogènes ont des temps de survie très court dans l'environnement (*e.g.* survie du virus Influenza hors de son hôte jusqu'à 24h ; Bean *et al.,*1982). Cette faculté de survie longue en dehors de l'hôte augmente les opportunités d'exposition et d'infection sur de nouveaux individus, en particulier si les conditions environnementales sont favorables. La variabilité génétique joue également un rôle crucial dans la

transmission des agents pathogènes car elle permet, notamment, l'évolution et l'adaptation (voir section 1.2.2) au système immunitaire de l'hôte. Elle détermine également leur pathogénicité, leur capacité d'infection, et leur aptitude à contaminer de nouvelles espèces hôtes. La capacité de l'agent pathogène à infecter un hôte sensible est également déterminée par le réservoir et le mode de transmission utilisé pour sa dissémination.

Le second maillon de la chaîne de transmission est l'hôte réservoir. Il est essentiel pour la survie, la multiplication et la transmission des agents pathogènes (van Seventer & Hochberg, 2017). L'hôte réservoir correspond à l'espèce animale qui héberge un agent pathogène de manière durable et à partir duquel débute le cycle de transmission de l'agent pathogène vers une espèce cible (Belding, 1965; van Seventer & Hochberg, 2017) (Figure 2). Dans certains cas, l'agent pathogène peut être stocké dans un environnement abiotique tel que l'eau ou le sol, on parle alors simplement de réservoir (van Seventer & Hochberg, 2017). Les agents pathogènes peuvent avoir plusieurs hôtes réservoirs domestiques ou sauvages (van Seventer & Hochberg, 2017). C'est le cas pour l'un des virus les plus problématiques à l'échelle mondiale, affectant les élevages de crevettes, le *White Spot Syndrome virus* (WSSV). Pas moins de 98 hôtes réservoirs potentiels ont été identifiés, parmi lesquels les crabes, les homards ou les écrevisses (Small & Pagenkopp, 2011). Deux problèmes sont associés à cette multiplicité d'hôtes réservoirs : *i)* c'est un facteur déterminant dans la rapidité de transmission de la maladie et *ii)* elle rend difficile l'identification des sources d'épidémie potentielle et donc le contrôle de la maladie. De plus, les hôtes réservoirs peuvent être asymptomatiques ou porteurs sains, tout en participant activement à la transmission de la maladie. A titre d'exemple, plusieurs études sur le paludisme ont montré l'importance du nombre de porteurs asymptomatiques dans la propagation de la maladie. Plus les porteurs sains sont nombreux, plus les moustiques propagent la maladie (Agaba *et al.,*2022; Lin *et al.,*2014; Tadesse *et al.,*2018).

Le troisième maillon de la chaîne d'infection, la « voie de sortie », est étroitement lié à l'hôte réservoir. En effet, l'agent pathogène doit quitter son hôte réservoir pour infecter l'hôte sensible. L'espèce hôte réservoir peut être la même que l'espèce hôte sensible. La porte de sortie correspond au site où l'agent pathogène est localisé et va quitter l'hôte réservoir (van Seventer & Hochberg, 2017). Suivant l'agent pathogène, les portes de sorties peuvent être multiples : voies respiratoires, tractus gastro-intestinal, voies génitales, peau ou mucus. Dans le cas des infections de moutons par le BTV, la porte de sortie est le sang. En effet, celui-ci va être absorbé par les moucherons qui servent de vecteurs lors de piqûre pour se nourrir. Les virus de la grippe a*via*ire (Influenza A) sont excrétés par les fèces et par les sécrétions respiratoires des oiseaux infectés (WOAH, 2023a). Chez l'humain, le SARS-CoV-2 est un virus causant principalement des dommages respiratoires, il va donc être principalement excrété par les

voies respiratoires, mais il est également excrété dans l'urine, les fèces ou le plasma et reste infectieux (WHO, 2020).

En effet, différentes stratégies de transmission des maladies infectieuses existent dans l'environnement. Elles dépendent des agents pathogènes (Figure 2, van Seventer & Hochberg, 2017). La voie de transmission la plus fréquemment utilisée est probablement la transmission directe qui peut être assimilée à une transmission *via* l'environnement (*i.e.* ingestion d'aliments ou d'eau contaminés, coupure, morsure), par contact physique (*i.e.* touché, baisé, gouttelettes d'éternuement ou de toux) ou par la voie périnatale (*i.e.* grossesse, accouchement ou allaitement). Il existe d'autres voies de transmission qui sont dites indirectes (van Seventer & Hochberg, 2017). La transmission indirecte se produit lorsque l'agent pathogène est transmis par l'intermédiaire d'objets, de surfaces, de particules en suspension dans l'air, ou d'un hôte intermédiaire. Les transmissions indirectes d'agents pathogènes sont classées en trois grandes catégories en fonction de l'intermédiaire qui permet la transmission. Les transmissions indirectes biologiques ou vectorisées font intervenir un hôte intermédiaire qui sert de vecteur dans la transmission de la maladie (*e.g.* Dengue, fièvre de la vallée du Nil et virus de West Nile transmis par l'intermédiaire de moustiques, BTV transmis par l'intermédiaire de moucherons, maladie de Lyme transmise par les tiques ; Steere *et al.,*2004; Trovato *et al.,*2020). Les transmissions indirectes mécaniques impliquent un déplacement de l'agent pathogène par une partie du corps d'un arthropode (*e.g.* transmission des bactéries du genre *Shigella spp*. par le transport de fèces sur les pattes de mouches, transmission de *Poxviridae* impliqués dans la dermatose nodulaire contagieuse par des mouches ; (Issimov *et al.*, 2020; Levine & Levine, 1991). Les transmissions indirectes peuvent également avoir lieu par voie aérienne (*e.g.* bactéries responsables de la tuberculose transportées sur de très longues distances et pendant des temps longs tout en restant infectieuses ; Churchyard *et al.,*2017).

Lorsqu'un agent pathogène cherche à infecter un hôte sensible et à accéder aux cellules cibles, il doit trouver une "voie d'entrée" appropriée dans cet hôte. Ces voies d'entrée partagent souvent des similitudes avec les voies de sortie des hôtes réservoirs, telles que les voies respiratoires, les voies génitales ou le tractus digestif (van Seventer & Hochberg, 2017).

L'hôte sensible est un individu qui est susceptible d'être infecté par l'agent pathogène (Figure 2). Certains hôtes sont plus sensibles que d'autres du fait de leur mode de vie (*i.e.* alimentation, vie en cohortes, migrations), leur génétique (*i.e.* immunité innée et adaptative) ou encore, d'une immunisation (*i.e.* vaccination, immunisation suite à une précédente infection). Il existe une compatibilité entre l'agent pathogène et son hôte qui favorise son entrée et sa survie dans l'organisme cible, en exploitant des mécanismes spécifiques à ces voies d'accès. Une fois qu'ils ont réussi à

s'installer à l'intérieur de l'hôte, certains agents pathogènes développent des mécanismes d'évasion sophistiqués vis-à-vis de la réponse immunitaire de l'hôte, ce qui leur confère la capacité de persister de manière prolongée et de déclencher des infections chroniques. Cette capacité à éviter la reconnaissance du système immunitaire de l'hôte est une adaptation essentielle pour les agents pathogènes afin de maintenir leur présence dans l'organisme hôte et de favoriser leur reproduction et leur propagation.

Dans le domaine de la virologie, la compréhension de la chaine de transmission des virus est essentielle pour étudier leur diversité, leurs mécanismes évolutifs, ainsi que leurs interactions avec leurs hôtes.

1.2 Les virus : diversité, mécanismes évolutifs et interactions avec leurs hôtes

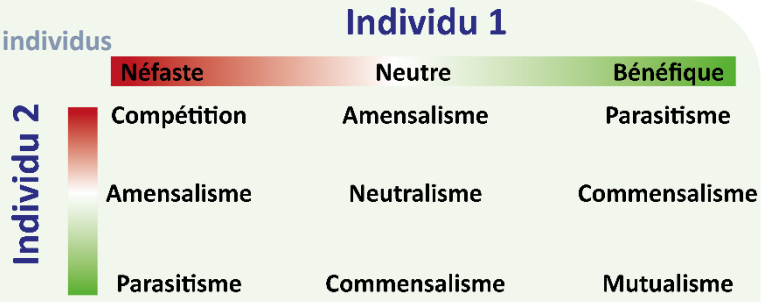
Les interactions hôtes-pathogènes représentent un réseau complexe et varié qui influence la dynamique des maladies infectieuses dans les populations. Pour mieux cerner et approfondir notre compréhension des mécanismes impliqués dans ces interactions, nous allons désormais restreindre notre étude spécifiquement aux virus.

Les virus présentent une grande diversité et ont la capacité d'infecter un large éventail de cellules hôtes, incluant les animaux, les plantes ou encore les amibes. Leurs interactions avec leurs hôtes peuvent varier du mutualisme symbiotique au parasitisme agressif (Encadré 6), induisant la mise en place de processus évolutifs spécifiques qui leur permettent de s'adapter rapidement à leur hôte (Ottmann, 2022).

Encadré 6 : Interactions biologiques entre deux individus

Mutualisme : interaction durable obligatoire ou non obligatoire dans laquelle les deux organismes trouvent un intérêt.

Commensalisme : interaction entre deux individus dans laquelle un seul tire profit de l'association sans être néfaste pour le second.



Matrice des interactions biologiques classées dans les catégories néfaste, neutre et bénéfique pour deux individus existants dans un même écosystème (d'après Lewin, 1982 et Overstreet & Lotz, 2016).

Neutralisme : interaction entre deux individus dans laquelle il n'y a aucun bénéfice ou effet néfaste pour les deux organismes.

Amensalisme : interaction n'ayant aucun impact pour l'un des individus mais néfaste pour l'autre.

Parasitisme : interaction durable nuisible pour un individu hôte et bénéfique pour le parasite.

Compétition : interaction concurrentielle nuisible pour les deux individus de la combinaison.

(Lewin, 1982, Overstreet & Lotz, 2016)

1.2.1 Diversité du monde viral

1.2.1.1 Virus à ADN vs virus à ARN

Le monde viral est caractérisé par une vaste diversité de structures génomiques, de cycles de vie, d'hôtes, d'environnements de vie, ainsi que de mécanismes évolutifs et adaptatifs (Sanjuán & Domingo-Calap, 2021). Selon la classification de Baltimore (1971), les virus sont regroupés en sept grandes classes en fonction de leurs caractéristiques génomiques et de leurs capacités de réplication (Figure 3, Baltimore, 1971). Les classes I et II incluent les virus ayant un génome d'ADN double brin et simple brin, comme les *Herpesviridae* ou les *Parvoviridae*, tandis que les classes III à VII sont composées de différents types de virus à ARN. Les classes III à V englobent respectivement les virus à ARN double brin, et ARN simple brin de polarité positive et négative. Les classes VI et VII regroupent les rétrovirus à ARN simple brin de polarité positive et les rétrovirus à ADN double brin circulaire, comme les *Retroviridae* et les *Hepadnaviridae* (Koonin *et al.*, 2021).

Actuellement, cette classification n'est pas la plus utilisée. Elle est basée uniquement sur des caractéristiques structurelles de l'acide nucléique contenu dans la capside et l'information génétique transmise (Koonin, 2016). Basée sur les analyses phylogénétiques et dans l'objectif d'obtenir des classes monophylétiques, la classification de Baltimore a été revue par le comité international sur la taxonomie des virus (International Committee on Taxonomy of Viruses, ICTV). L'ICTV a effectué des recherches approfondies utilisant des méthodes de séquençage génétique et des données phylogénétiques pour revoir la classification des virus. Comme pour les classifications du monde vivant, cette classification est organisée en domaines, règnes, phylums, classes, ordres, familles, genres et espèces. En plus des caractéristiques génomiques, morphologiques et biologiques, cette nouvelle classification prend en compte les distances entre les différents virus et est plus en phase avec les avancées scientifiques récentes. Elle permet d'identifier plus précisément des groupes de virus et de mieux comprendre leur diversité (Koonin *et al.,*2021). Bien que la taxonomie soit cruciale, elle reste sujette à des modifications constantes en réponse aux nouvelles informations génomiques, génétiques ou moléculaires sur les virus connus et nouvellement découverts. Dans la suite de ce manuscrit, j'utiliserai uniquement la classification décrite par l'ICTV.

Une caractéristique qui différencie les virus ADN des virus ARN, étroitement corrélée avec la structure de leur acide nucléique, est leur capacité d'évolution plus ou moins rapide (voir section 1.2.2). Les virus à ARN sont connus pour leur capacité à évoluer rapidement, notamment, par des mutations ponctuelles induites durant leur cycle de réplication (Duffy, 2018). En comparaison, les capacités d'évolution des virus à ADN sont moins importantes et l'évolution est plus lente que chez les virus à

ARN, en raison des polymérases utilisées pour la réplication de leur génome qui sont plus fidèles et capables de corriger les erreurs lors de la synthèse de l'ADN. C'est particulièrement le cas chez les *herpesvirus* qui utilisent des ADN polymérases multifonctionnelles, capables par exemple d'allonger la séquence nucléotidique par une activité polymérase, ou de supprimer les erreurs dans la séquence ADN grâce à une activité exonucléase (Zarrouk *et al.*, 2017).

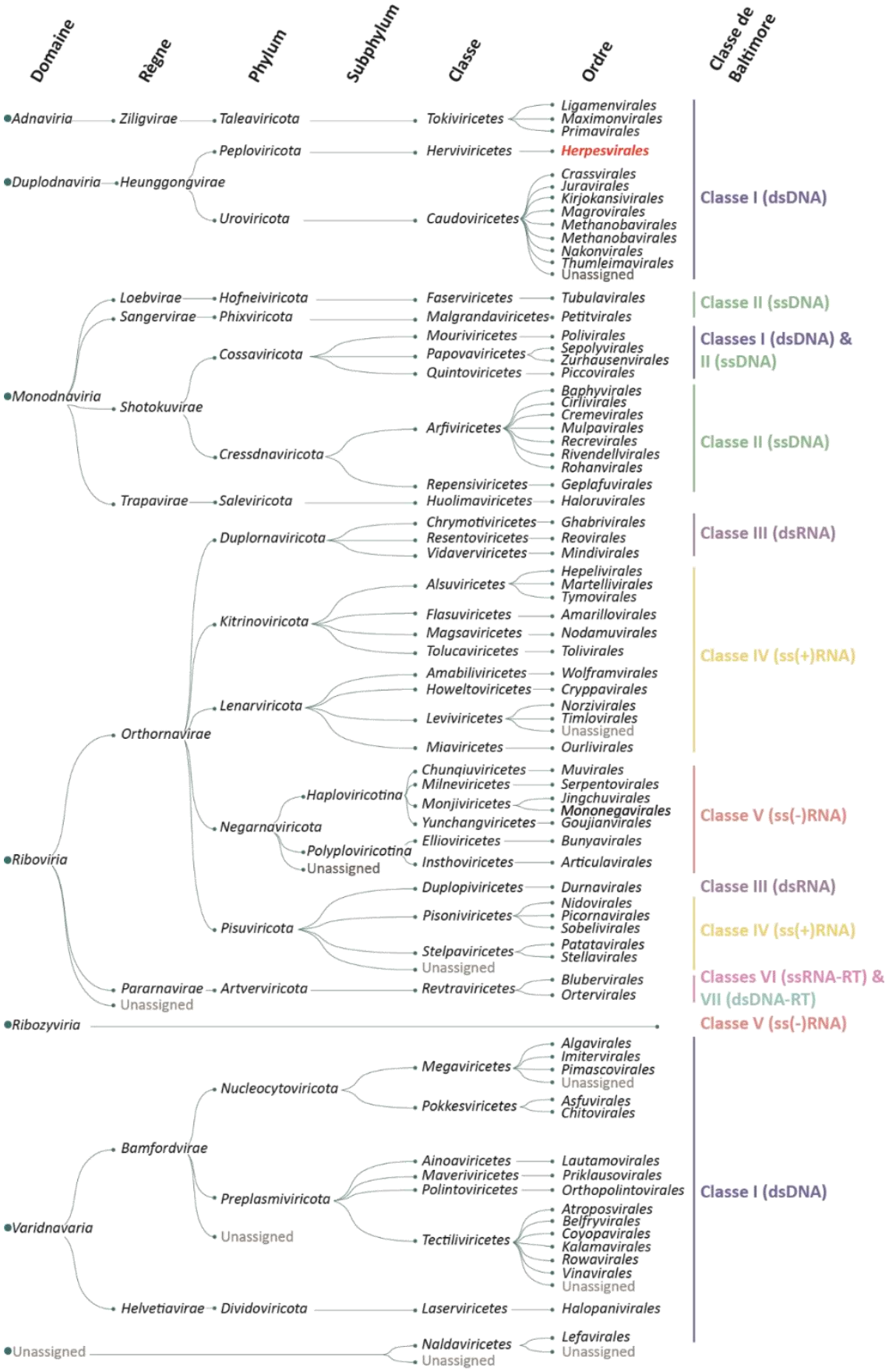


Figure 3: Taxonomie et classification des virus en cinq domaines et correspondance avec les classes virales de Baltimore (d'après ICTV taxonomy browser et Koonin *et al.*, 2021).

1.2.1.2 Les *Herpesvirus*

L'ordre des *Herpesvirales* représente actuellement 133 espèces d'*herpesvirus*, réparties en trois familles (Figure 3) : les *Alloherpesviridae*, les *Orthoherpesviridae* et les *Mallacoherpesviridae* (Davison, 2010; Davison *et al.*, 2009; Lefkowitz *et al.*, 2018). Les *Alloherpesviridae* infectent les chauves-souris et certains poissons comme la carpe, le poisson-chat et le saumon. Parmi les plus notoires, on trouve le virus de l'herpès cyprin (*Cyprinus herpesvirus*, CyHV), le virus de l'herpès de l'anguille (*Anguilid Herpesvirus*, AngHV) et des batraciens (*Ranid herpesvirus*, RHV). Les hôtes des *Orthoherpesviridae* sont plus diversifiés que ceux des *Alloherpesviridae*, comprenant des oiseaux (*Gallid alphaherpesvirus*, *Columbid alphaherpesvirus*), des reptiles (*Varanid herpesvirus* et *Lacerta viridis herpesvirus*) et des mammifères (*Bovine gammaherpesvirus* et *Equid alphaherpesvirus*). Huit virus de type herpès de la famille des *Orthoherpesviridae* infectent uniquement l'Homme (*e.g. Herpes simplex virus* (HSV), *herpesvirus varicellæ zoster* (VZV), *Epstein-Barr virus* (EBV) ou *Cytomégalovirus Humain* (HCMV)). La famille des *Malacoherpesvirus* contient uniquement deux genres : les *Aurivirus* et les *Ostreavirus*, infectant des espèces de mollusques marins.

Les symptômes causés par les *Herpesvirus* sont très variables d'une espèce sensible à une autre. Chez l'humain, HSV-1 est responsable de l'éruption de lésions vésiculaires sur les voies orales ou génitales, communément appelés boutons de fièvres, accompagnés par de la fièvre et, en cas d'atteinte génitale, de difficultés à uriner. Dans les cas les plus extrêmes, le virus peut déclencher des encéphalites léthales (Szpara *et al.*, 2014).

Chez les anguilles, AngHV-1 est responsable d'hémorragies au niveau de la peau, des nageoires, des branchies et du foie (Hanson *et al.*, 2011). Chez les poules, l'*herpesvirus* a*via*ire *Gallid alphaherpesvirus 1* (GaHV-1) est responsable de toux, d'éternuements, de tremblements de tête, de léthargies, d'écoulement des yeux et des narines (parfois sanguinolent) et de difficultés respiratoires entraînant parfois la mort de l'animal (Gowthaman *et al.*, 2020). Chez les chevaux, l'*herpesvirus* Equine *herpesvirus* 2 se manifeste par une pharyngite, une lymphadénopathie, une rhinite avec des écoulements nasals, de la toux et de la fièvre accompagnés par une inappétence et des performances médiocres (Fortier *et al.*, 2009).

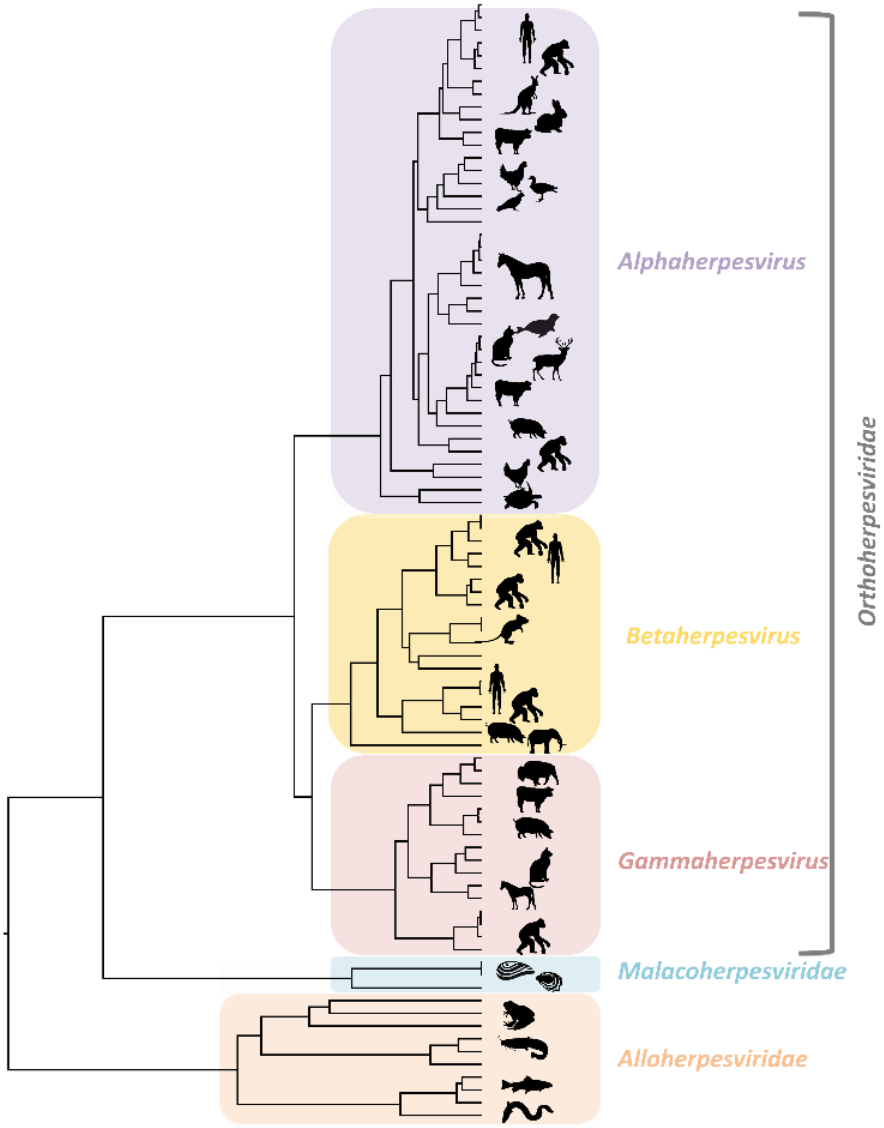


Figure 4 : Phylogénie des *Herpesvirales* basée sur la séquence de l'ADN polymérase (C. Pelletier, Communication personnelle, 2020).

1.2.1.3 Cycle de réplication viral

Tous les virus, à l'exception des virus géants, sont des parasites intracellulaires obligatoires, ils ont besoin des cellules de leur hôte pour se maintenir et se répliquer. Les cycles viraux sont étroitement liés à leurs caractéristiques génomiques, et sont donc variables d'une famille à l'autre. Ils peuvent cependant être résumés en quatre grandes étapes communes : l'entrée du virus dans une cellule hôte, la réplication de son génome, la formation des capsides et la sortie des particules virales (Ryu, 2017). Les stratégies de réplication virale diffèrent d'une famille virale à l'autre, et notamment en fonction de la classe à laquelle ils appartiennent (Louten, 2016). A titre d'exemple, HSV-1, pour lequel le cycle de réplication est très bien décrit, est capable de réaliser un cycle lytique, responsable de l'infection

primaire de l'hôte, mais également un cycle latent, responsable du maintien silencieux du virus dans les cellules de son hôte (Moin et al., 2021).

Au cours de ce cycle, les virus peuvent subir des évolutions génétiques, pouvant conduire à des changements dans les caractéristiques phénotypiques du virus, telles que la virulence, la transmissibilité ou l'aptitude à infecter de nouveaux hôtes.

1.2.2 Mécanismes évolutifs chez les virus

Les virus sont dotés d'une capacité à évoluer rapidement qui leur permet de s'adapter aux changements de l'environnement, de contourner les défenses immunitaires des hôtes et de développer de nouvelles stratégies d'infection (Sanjuán & Domingo-Calap, 2021). Les forces évolutives qui façonnent les patrons de diversité des virus sont multiples et comprennent la mutation génétique, les flux de gènes ou de matériel génétique (*via* la recombinaison, ou le réassortiment), la sélection naturelle et la dérive génétique (Figure 5). Ces forces interagissent de manière complexe, permettant aux virus de générer de la diversité génétique et de modifier la fréquence des allèles (Encadré 7) dans une population au cours du temps (Sanjuán & Domingo-Calap, 2021). La compréhension des processus évolutifs est essentielle pour anticiper et lutter contre les épidémies, ainsi que pour développer de nouvelles stratégies thérapeutiques et vaccinales efficaces (Sanjuán & Domingo-Calap, 2021).

Encadré 7 : Qu'est-ce qu'un allèle ?

En génétique, un allèle est défini comme chacune des versions possibles d'un même gène. Ainsi les allèles sont générés par les processus évolutifs, qui modifie la séquence ADN des virus. Les allèles déterminent l'ensemble des traits du virus.

Elston *et al.*, 2012

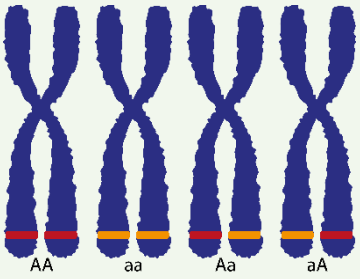


Illustration de quatre chromosomes portant différentes combinaisons de l'allèle A.

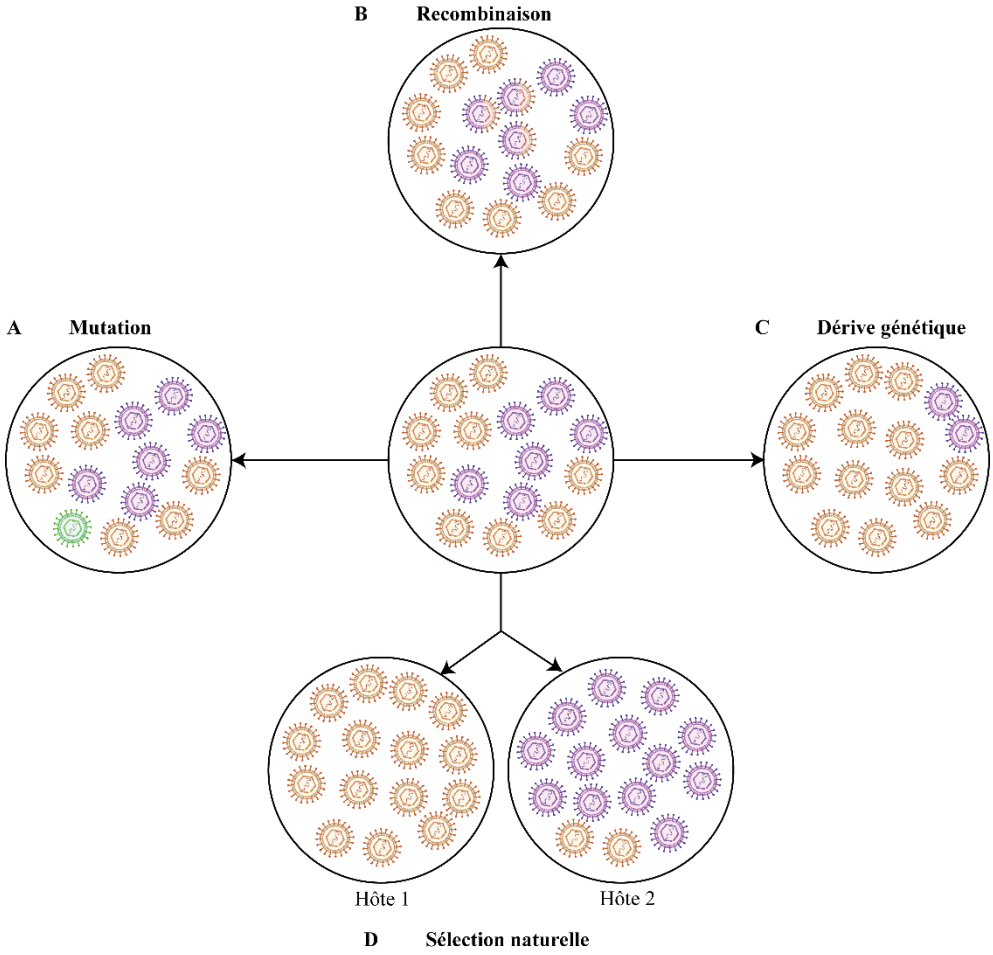


Figure 5 : Forces et processus évolutifs qui agissent sur les allèles dans les populations virales (adapté de Adeyemo *et al.*2021 et Sanjuán & Domingo-Calap, 2021).

A) Les mutations impliquent un changement de nucléotide et l'introduction potentielle d'un nouvel allèle dans la population. B) La recombinaison est un phénomène qui implique la combinaison de deux portions de génome provenant de deux particules virales différentes sur un seul brin ADN ou ARN. C) La dérive génétique implique le changement aléatoire de fréquences alléliques d'une population. D) La sélection naturelle implique un changement de fréquences alléliques dans une population, étroitement lié à la valeur sélective du virus en fonction de son environnement ou de son hôte.

1.2.2.1 La mutation

La mutation génétique est l'un des mécanismes évolutifs les plus fondamentaux. Une mutation est définie comme une modification ponctuelle de la séquence d'acides nucléiques d'un génome. En raison de leur structure génomique et de leurs mécanismes de réplication, les virus sont sujets à des taux de mutation élevés par rapport aux organismes plus complexes (Biek *et al.*, 2015). Les mutations surviennent lorsqu'il y a des erreurs dans la réplication de l'ARN ou de l'ADN viral, ou sous l'action d'une source exogène, ce qui entraîne des modifications dans la séquence génétique (Figure 6). Ces mutations peuvent impliquées le remplacement d'une seule base nucléotidique, elles sont alors

appelées substitution. Elles peuvent être plus complexes, impliquant des insertions, des délétions ou des réarrangements de séquences génétiques (Alveshere, 2019).

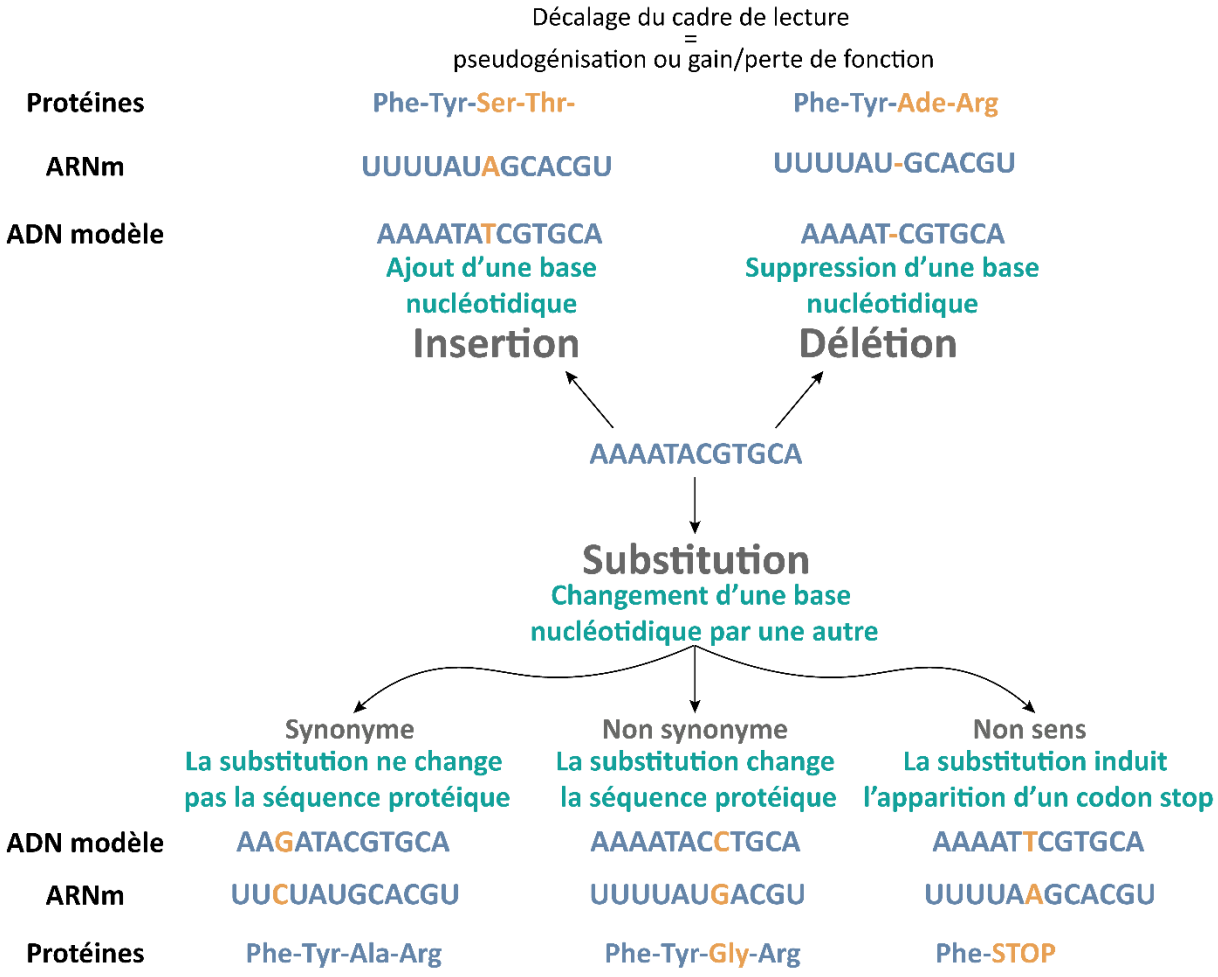


Figure 6 : Les différents types de mutation pouvant agir sur le génome d'un virus.

Les mutations observées peuvent être des insertions, c'est-à-dire un ajout d'une base nucléotidique ou plusieurs bases nucléotidiques dans le génome, des délétions qui correspondent à la suppression d'une base nucléotidique ou plusieurs bases nucléotidiques dans le génome. Ces Indels peuvent être responsable d'un décalage du cadre de lecture provoquant potentiellement une pseudogénisation ou un gain/perte de fonction. Il existe également, les substitutions qui sont un changement d'une base nucléotidique par une autre. Trois types de substitutions peuvent être observées : les substitutions synonymes qui ne changent pas la séquence protéique, les substitutions non synonymes qui changent la séquence protéique et enfin les substitutions non-sens qui induisent l'apparition d'un codon STOP dans la séquence nucléotidique.

Les virus à ARN, tels que les *Coronaviridae*, le virus de la grippe ou le virus Ebola, ont tendance à présenter des taux de substitutions plus élevés que les virus à ADN (Sanjuán, 2012; Sanjuán *et al.*,2010; Sanjuán & Domingo-Calap, 2016, 2021). En effet, les taux d'évolution des virus ARN varient entre 10^{-6} et 10^{-4} substitutions/nucléotide/an (s/n/a) tandis que les taux des virus à ADN sont généralement de l'ordre de 10^{-8} à 10^{-6} s/n/a (*e.g.* 9x10^{-5} s/n/a pour le *Poliovirus 1* et 5,9x10^{-8} s/n/a pour l'*Herpes simplex virus* ; Sanjuán & Domingo-Calap, 2016). Cela est dû à l'absence de mécanismes de correction des

erreurs dans la réplication de l'ARN (absence d'activité 3'-exonuclease proofreading), ce qui entraîne une accumulation plus rapide de mutations (Roberts *et al.*, 1988; Sanjuán & Domingo-Calap, 2021; Steinhauer *et al.*, 1992).

De plus, certaines études ont mis en évidence une corrélation entre la taille du génome et le taux de mutation. Plus le génome est petit, comme c'est le cas chez les virus à ARN, plus le taux de mutation est élevé (Figure 7, Lynch, 2010; Sanjuán & Domingo-Calap, 2016). De même, les virus ayant un génome simple brin ont tendance à avoir des taux de mutation plus élevés que les virus ayant un génome double brins (Sanjuán *et al.,*2010). Une explication possible est la prédisposition des acides nucléiques simple brin aux dommages oxydatifs et chimiques (Sanjuán & Domingo-Calap, 2016). Du fait de leur structure simple brin et de l'absence de brin complémentaire, ils n'ont pas non plus de possibilité de corrections post-réplicatives.

Bien que les mutations puissent survenir de manière spontanée lors de la réplication virale, les sources abiotiques peuvent également induire des mutations chez les virus. Les sources abiotiques comprennent des facteurs physiques tels que les radiations ionisantes, les produits chimiques mutagènes et les rayonnements ultraviolets. Plus spécifiquement, les rayonnements ultraviolets (UV : 100-200 nm, UV-A : 315-400 nm, UV-B : 280-315 nm et UV-C : 200-280 nm) provenant du soleil sont une source importante de mutations chez les virus. Les rayons UV peuvent endommager directement l'ADN ou ARN viral. Ils engendrent des cassures, et forment des ponts chimiques entre les acides nucléiques, provoquant des dimers nucléotidiques et modifiant la stabilité des liaisons entre acides nucléiques (Gómez-López *et al.*, 2021; Kong *et al.*, 2021; Liu & Shan, 2023). Cela peut entraîner des erreurs lors de la réplication virale, conduisant à de nouvelles mutations, voire empêcher le bon déroulement des processus de réplication, transcription ou traduction (Kong *et al.*, 2021). Certaines radiations sont utilisées pour inactiver des particules virales et, notamment, dans les processus de désinfection (Gómez-López *et al.*, 2021; Hessling *et al.*, 2022; Kong *et al.*, 2021).

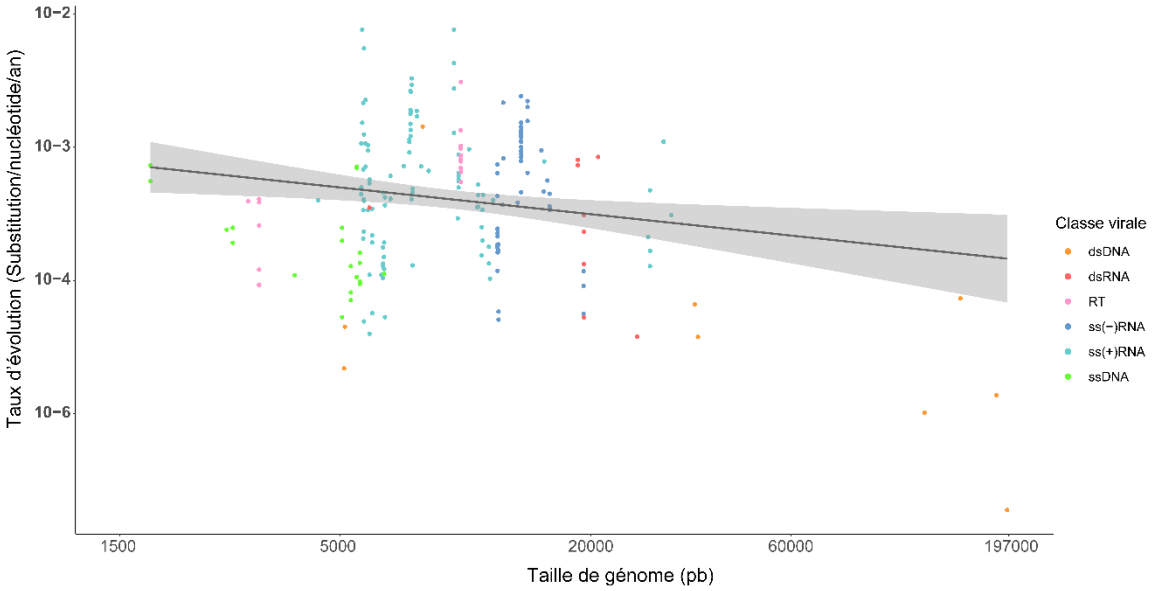


Figure 7: Corrélation entre le taux d'évolution et la taille du génome de 223 virus (d'après Sanjuán & Domingo-Calap, 2016).

Des taux d'évolution élevés peuvent conduire à l'émergence rapide de nouvelles lignées virales avec des propriétés différentes. Celles-ci peuvent, par exemples, permettre l'adaptation du virus à de nouvelles espèces hôtes, engendrer des modifications de virulence, améliorer les capacités du virus à reconnaitre les récepteurs cellulaires ou à échapper à la réponse immunitaire de l'hôte (Sanjuán *et al.*, 2010; Sanjuán & Domingo-Calap, 2016, 2021). En complément des mutations génétiques, les virus sont capables de réorganiser leur matériel génétique, voire d'échanger du matériel génétique avec d'autres virus, pour générer de la diversité génétique au travers d'un processus appelé recombinaison génétique.

1.2.2.2 La recombinaison

La recombinaison est un phénomène essentiel dans l'évolution des virus. Alors que les mutations permettent de générer de nouveaux allèles, la recombinaison permet de générer de nouvelles combinaisons d'allèles (Pérez-Losada *et al.,*2015). Il s'agit d'un processus moléculaire au cours duquel au moins deux virus distincts échangent des segments de leur matériel génétique lorsqu'ils infectent la même cellule hôte par croisement de deux brins d'acides nucléiques (Pérez-Losada *et al.,*2015). Ce processus de recombinaison génétique peut conduire à la création de nouvelles souches virales dotées d'un génome hybride, combinant les caractéristiques des virus parentaux. Plusieurs facteurs influencent la recombinaison et notamment une homologie génétique entre les souches virales, la

taille des populations au sein de l'hôte, la co-circulation de différentes souches virales au sein d'un hôte ou encore les niveaux de co-infection (Pérez-Losada *et al.,*2015).

La recombinaison peut se produire de différentes manières. La plus courante est la recombinaison homologue (Figure 8A), qui implique la rupture et la réunion de brins d'ARN ou d'ADN dans des sites similaires (*i.e.* région génomique homologue) entre deux virus (Lai, 1992; Pérez-Losada *et al.,*2015). Cette recombinaison peut être facilitée par des enzymes virales, notamment les enzymes actives lors des étapes de réplication de l'ADN. Ces dernières sont capables d'inverser des régions de plus de 50 kpb dans les génomes chez les *Herpesvirus* (Rennekamp & Lieberman, 2010). Les transcriptases inverses, peuvent, quant à elles, copier et intervertir les brins d'ARN (Lai, 1992). Une autre forme de recombinaison est la recombinaison non homologue (Figure 8B), au cours de laquelle les brins d'ARN ou d'ADN provenant de virus différents se combinent sans correspondance de séquence, ce qui produit souvent des structures aberrantes (Lai, 1992; Pérez-Losada *et al.,*2015). Ces processus ont, par exemple, été observés entre les coronavirus bovins, murins et Humains et le virus Influenza C, conduisant à l'apparition, chez ces coronavirus, d'un gène non essentiel codant pour une hémagglutinine-estérase (Lai, 1992).

Il existe également le réassortiment génomique (Figure 8C). Il concerne les virus ayant des génomes à ARN segmentés (*e.g.* Influenza) qui leur permettent d'échanger des segments complets de génome, et donne ainsi lieu à des nouvelles combinaisons de segments génomiques (Marshall *et al.,*2013; Vijaykrishna *et al.,*2015). De plus, la recombinaison peut avoir lieu au sein d'un même génome et sera dénommé recombinaison intra-génomique (Figure 8D), ou, entre des brins ADN/ARN originaires de différents génomes ce qui correspond à de la recombinaison inter-génomique. Dans les quatre types de recombinaison, l'échange de matériel génétique n'est cependant pas réciproque, ce qui implique que le génome qui reçoit une séquence génomique n'en donne pas obligatoirement en retour pour remplacer la perte génétique du génome donneur (Pérez-Losada *et al.,*2015).

La recombinaison est fréquente et particulièrement bien décrite chez les virus ARN. Cependant, les mécanismes de recombinaison restent très peu caractérisés chez les virus ADN, et seuls quelques exemples sont disponibles dans la littérature. L'un des virus de type herpès les plus étudié pour la recombinaison, est l'*Herpesvirus simplex 1* (HSV-1). Il a été démontré que la recombinaison joue un rôle important dans la diversité génétique de HSV-1 et sa capacité à échapper aux réponses immunitaires de l'hôte (Pérez-Losada *et al.*, 2015). Les mécanismes de recombinaison chez HSV-1 impliquent des événements de recombinaison à la fois homologue et non homologue (Pérez-Losada *et al.*, 2015). La recombinaison homologue chez HSV-1 se produit généralement lors de la réplication de l'ADN viral. Pendant le processus de réplication par « rolling-circle », des « branches » de

concatémères (*i.e.* longue molécule d'ADN continue qui contient plusieurs copies de la même séquence d'ADN liées en série) se fixent sur des régions hautement similaires et génèrent des extrémités cohésives appelées bras de recombinaison (Pérez-Losada *et al.,*2015). En plus de la recombinaison homologue, la recombinaison non homologue a également été observée chez HSV-1 (Muylaert *et al.,*2011; Pérez-Losada *et al.,*2015).

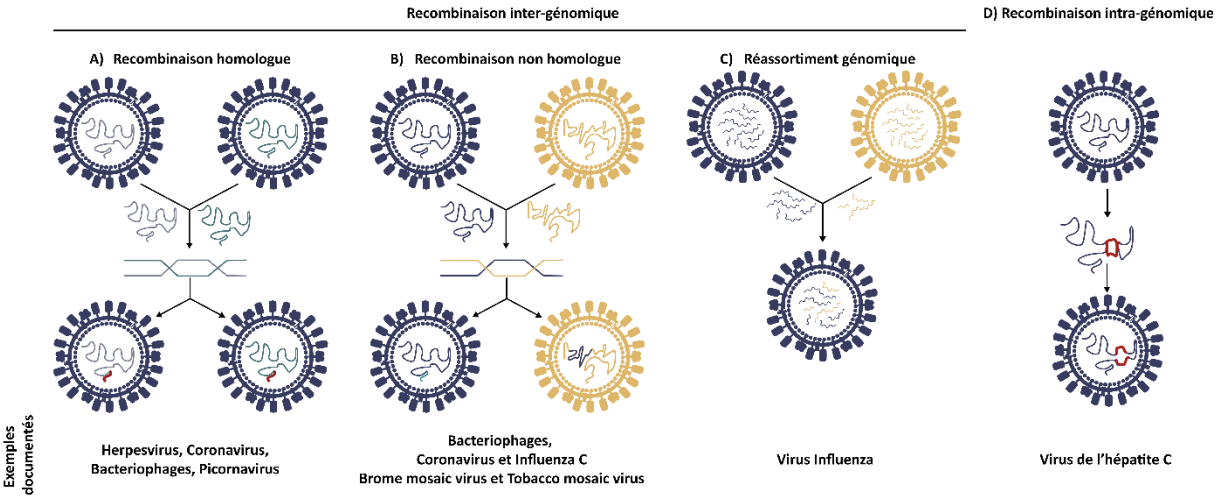


Figure 8 : Représentation schématique des différents types de recombinaison possible chez un virus (d'après Lai *et al.*1992 et Pérez-Losada *et al.,*2015).

A) Recombinaison homologue, B) recombinaison non homologue, C) réassortiment génomique, D) recombinaison intra-génomique.

Comme la mutation, l'échange de matériel génétique peut avoir des conséquences majeures sur l'évolution des virus. Elle peut entraîner l'apparition de nouveaux variants viraux avec des propriétés différentes de celles des virus parentaux. La destinée de ces lignées virales reste tout de même extrêmement liée aux processus de sélection naturelle et de dérive génétique (Sanjuán & Domingo-Calap, 2021).

1.2.2.3 La sélection naturelle

La sélection naturelle est un processus évolutif clé, façonnant les caractéristiques et les adaptations des organismes vivants. Ce concept fondamental, développé en 1889 par Charles Darwin dans son « *Journal of researches* » au travers de l'exemple des Pinsons des Galapagos, repose sur l'idée que certains traits favorables confèrent un avantage sélectif aux individus dans leur environnement, augmentant ainsi leur capacité de survie et de reproduction (Darwin, 1889). Ces traits avantageux sont transmis aux générations suivantes, tandis que les traits désavantageux ont une faible probabilité de

se propager (Perdoncini Carvalho *et al.,*2022). La sélection naturelle agit sur les virus en fonction de leur capacité à infecter, se répliquer et se propager efficacement dans les populations hôtes. Les virus présentant des adaptations génétiques favorables, telles que des allèles qui améliorent leur capacité de fixation aux récepteurs cellulaires, leur résistance aux défenses immunitaires de l'hôte ou leur transmission entre individus, ont un avantage sélectif dans leur survie et leur propagation (valeur sélective plus forte). La plupart des mutations qui surviennent dans les génomes viraux étant non avantageuses, elles ne sont pas fixées dans les populations à cause des processus de sélection (Otto & Whitlock, 1997). La sélection naturelle peut également entraîner la spécialisation d'hôte d'un virus, tandis que d'autres pourront infecter et se propager efficacement dans un groupe d'espèces hôtes étroitement liées (Bandín & Dopazo, 2011; Woolhouse *et al.*, 2001). Dans la nature, on observe trois types de sélection naturelle qui peuvent se succéder : la sélection directionnelle, la sélection stabilisante et la sélection diversifiante (Figure 9) (González-Candelas *et al.,*2018).

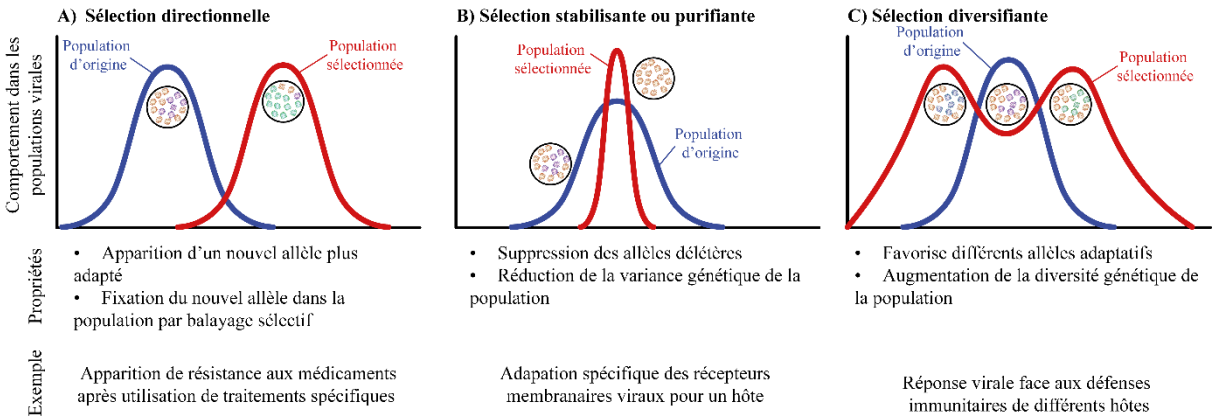


Figure 9: Différents types de sélection.

A) Sélection directionnelle : apparition d'un génotype adapté et fixation dans la population. B) sélection stabilisante : suppression des allèles délétères et réduction de la variabilité génétique de la population. C) sélection diversifiante : plusieurs génotypes favorisés, augmentation de la diversité génétique de la population.

1.2.2.3.1 La sélection directionnelle

La sélection directionnelle favorise un seul phénotype ou génotype, entraînant un déplacement continu des fréquences alléliques dans une direction spécifique (Figure 9A). Ce type de sélection est souvent associé à des événements de balayage sélectif (Encadré 8) au cours desquelles, un allèle bénéfique se répand rapidement dans la population au détriment des autres allèles. Les mutations neutres autour de cette mutation bénéfique voient également leur fréquence allélique augmenter et finissent par être fixées dans la population par auto-stop génétique lors d'un événement de balayage

sélectif (González-Candelas *et al.*, 2018; Maynard & Haigh, 2007; Stephan, 2019). Un exemple intéressant de sélection directionnelle est observé chez les patients immunodéprimés, pour lesquels les réponses immunitaires sont absentes. En cas d'infection par un virus, les processus de réplication virale sont beaucoup plus importants en termes de quantité de particules virales générées, et durent plus longtemps que chez un patient non-immunodéprimé. Ces différences offrent aux virus des opportunités accrues de mutations qui peuvent parfois être bénéfiques, et se fixer dans la population pour conférer des propriétés de résistance aux traitements antiviraux. Il a été observé, chez HSV-1, l'émergence de variants résistants du virus suite à l'utilisation de traitements antiviraux (Field, 2001). L'hypothèse la plus probable pour expliquer cette émergence, est que les variants résistants, initialement présents à de faibles fréquences alléliques, se sont rapidement multipliés, devenant ainsi la population virale dominante et provoquant une infection.

Encadré 8 : Balayage sélectif

En génétique des populations, un balayage sélectif (ou *selective sweep*) est un mécanisme par lequel un allèle ayant un effet positif sur la valeur sélective des individus porteurs se propage et se fixe au sein de la population, entraînant ainsi une augmentation de sa fréquence. Sous l'effet du balayage sélectif, la variabilité au sein d'une population peut être diminuée, voire supprimée pour la portion d'ADN entourant la mutation sélectionnée.

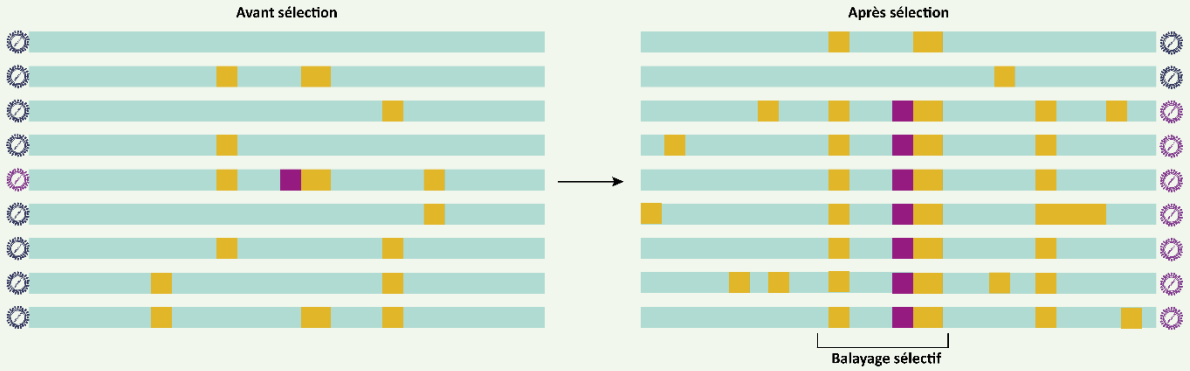


Illustration des conséquences d'un balayage sélectif fort sur une population virale. Un nouveau variant bénéfique subit une sélection positive immédiate, se propage dans la population et entraîne la disparition des autres variants présents, dont la valeur sélective est plus faible.

(Stephan, 2019)

1.2.2.3.2 La sélection stabilisante ou purifiante

La sélection stabilisante entraîne une diminution de la variance génétique de la population lorsqu'un phénotype ou un allèle est favorisé et sélectionné (Figure 9B). La plupart des nouveaux allèles sont délétères et sont éliminés de la population, préservant ainsi les séquences d'ADN. Par exemple, ces

pressions de sélection peuvent être observées dans le cas d'une spécialisation des récepteurs membranaires vis-à-vis d'une espèce hôte, notamment après un événement de transmission inter-espèces. La diversité intra-hôte d'un marqueur est fortement restreinte en raison de cette spécificité pour l'hôte. Chez le *Cytomegalovirus murin* (MCMV), un *herpesvirus* infectant les souris, une large proportion codante du génome évolue sous sélection purifiante, probablement en lien avec l'exposition du génome viral face au système immunitaire de l'hôte pendant de très longues périodes (Smith *et al.*, 2013). De même chez les Cyprinivirus (*Alloherpesvirus* infectant les cyprins), certaines portions du génome, telle que l'ADN polymérase, l'hélicase et les terminases sont soumises à des pressions de sélection purifiante importantes ce qui permet de limiter les mutations dans ces régions (Donohoe *et al.*, 2021).

1.2.2.3.3 La sélection diversifiante ou positive

La sélection diversifiante favorise le maintien de plusieurs allèles dans une population (Figure 9C), ce qui se traduit par une plus forte diversité génétique (González-Candelas *et al.,*2018). Des pressions de sélection diversifiantes sont généralement exercées sur les gènes impliqués dans les mécanismes d'échappement des virus aux défenses immunitaires des hôtes, ou liés à l'infection, et notamment à l'entrée du virus dans son hôte. Chez HSV-1 et HSV-2 les glycoprotéines, et particulièrement la glycoprotéine C impliquée dans l'attachement du virus sur la surface cellulaire, subissent de fortes pressions de sélection positive. Ces régions accumulent de la diversité, probablement pour diversifier les récepteurs d'attache du virus à la cellule hôte (Lamers *et al.*, 2015).

Les pressions de sélection peuvent changer au cours du temps, et varier selon les environnements et les interactions entre les virus et leurs hôtes. Ces processus ne sont pas isolés et peuvent agir simultanément ou se succéder. Il est donc parfois difficile de déterminer précisément quel type de sélection est en action à un instant donné. Par exemple, ces successions de pressions de sélection ont notamment été observées dans le cas des virus de la grippe Influenza A. Pour les virus de la grippe, on observe des alternances d'épisodes de pression de sélection diversifiante et purifiante sur de courtes périodes (McHardy & Adams, 2009; Pybus & Rambaut, 2009), ce qui est responsable de l'apparition de différents variants. D'autre part, les pressions de sélection qui s'exercent, varient le long des génomes (González-Candelas *et al.*, 2018). En effet, elles sont étroitement liées aux fonctions des gènes codés. Les régions du génome essentielles pour la réplication virale et la production de protéines structurales subissent généralement une sélection purifiante intense, où les variants porteurs de mutations délétères sont éliminés, assurant ainsi la stabilité de ces fonctions vitales. Les régions du génome qui sont impliquées dans les interactions hôte-pathogène, telles que les sites de liaison aux récepteurs cellulaires ou les protéines reconnues par le système immunitaire de l'hôte, sont souvent soumises à

des pressions de sélection positive ou diversifiante, augmentant ainsi la capacité du virus à infecter de nouveaux hôtes ou à échapper à la réponse immunitaire (González-Candelas *et al.*, 2018).

Bien que la sélection joue un rôle prépondérant dans les variations des fréquences alléliques, ces fréquences peuvent également être modifiées par la dérive génétique, un processus aléatoire qui exerce une influence au sein des populations au cours du temps.

1.2.2.4 La dérive génétique : évolution aléatoire des populations virales

La dérive génétique est une force évolutive importante. Il s'agit d'un phénomène au cours duquel la fréquence des allèles change dans les populations de manière aléatoire (*i.e.* stochasticité due à l'échantillonnage aléatoire des parents dans la population), et son impact s'intensifie avec la diminution de la taille de la population (Figure 10, M. Lynch *et al.*, 2016; Sanjuán & Domingo-Calap, 2021). La dérive génétique est également influencée par le taux de mutation (M. Lynch *et al.*, 2016). Dans les populations virales, qui sont souvent caractérisées par leur taille importante et leur reproduction rapide, des mutations spontanées surviennent fréquemment lors des cycles de réplication virale. Ces mutations engendrent des variations génétiques au sein de la population virale. Dans certaines situations, comme lors de transmission virale inter-espèces, le taux de transmission limité d'un virus d'un hôte à un autre, ou l'isolement de populations virales dans des environnements spécifiques, créent un contexte propice à la dérive génétique (Sanjuán & Domingo-Calap, 2021). En effet, durant une transmission inter-espèces, très peu de virus sont transmis (*i.e.* population fondatrice dans un nouvel hôte estimée à moins de 10 virus) et beaucoup des allèles de la population virale sont perdus (Sanjuán & Domingo-Calap, 2016). Ces évènements sont appelés goulot d'étranglement (Encadré 9), et résultent en une diminution drastique de la taille effective de la population, ainsi qu'en une perte considérable de diversité, qui peut mener à la fixation de mutations aléatoires neutres ou délétères (Sanjuán & Domingo-Calap, 2021).

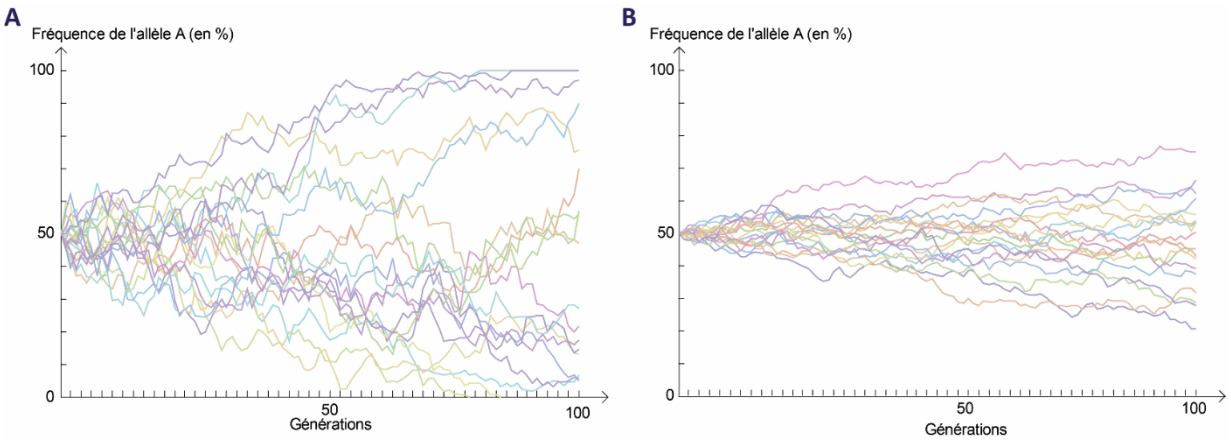


Figure 10 : Variation aléatoire des fréquences alléliques causées par la dérive génétique au cours de 100 générations.

Simulation des variations des fréquences de 50 allèles induites par la dérive génétique dans une population de taille effective (Ne*)* : A) Ne=100 et B) Ne=1000 et pour des allèles ayant une fréquence initiale de 50%. Chaque ligne de couleur représente une simulation de dérive génétique pour l'allèle A dans la population étudiée (Simulations réalisées avec https://www.pedagogie.ac-nice.fr/svt/productions/derive-genetique/).

Encadré 9 : Goulot d'étranglement

Les goulots d'étranglement, dans les populations virales, représentent des phases de réduction drastique de la diversité génétique de la population virale. Cette réduction peut conduire à la perte d'allèles spécifiques et à l'homogénéisation génétique.

Ces événements surviennent généralement suite à des pressions environnementales sévères, telles que des changements environnementaux rapides ou des contraintes immunologiques. Les goulots d'étranglement peuvent être causés par des facteurs divers tels que, les fluctuations démographiques, la migration restreinte, la sélection naturelle intense ou la dérive génétique.
Les goulots d'étranglement ont des implications importantes pour l'évolution virale, car ils peuvent favoriser l'accumulation rapide de mutations neutres ou délétères, ainsi que la fixation rapide d'allèles avantageux.
Les études des goulots d'étranglement dans les populations virales utilisent des approches phylogénétiques, des simulations informatiques et des analyses génétiques pour identifier les signes de réductions passées de la taille de la population et, pour estimer leur impact sur la diversité génétique actuelle.

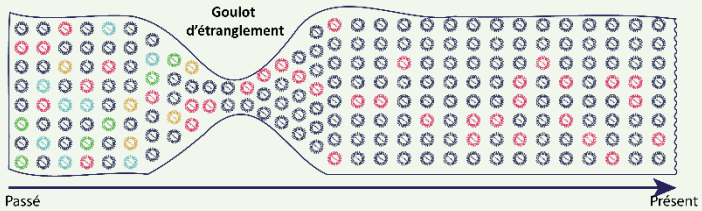


Illustration d'un goulot d'étranglement dans une population virale qui évolue au cours du temps. Cette population est initialement diversifiée. Après l'évènement de goulot d'étranglement, la diversité de la population se retrouve réduite

(McCrone et Lauring, 2018)

1.2.3 Interactions hôte-virus

1.2.3.1 Coévolution

Selon la théorie de la reine rouge, l'interaction entre un agent pathogène et son hôte est caractérisée par une évolution perpétuelle visant à maintenir une valeur reproductive optimale (McLaughlin & Malik, 2017). Le concept de coévolution a été introduit par Ehrlich et Raven en 1964, et désigne l'adaptation réciproque entre un hôte et son parasite au cours du temps, pour maintenir une valeur sélective optimale pour chacun d'entre eux. La coévolution entre un hôte et un virus est un processus évolutif dynamique résultant de l'interaction étroite entre les deux parties (Kaján *et al.,*2020). Les virus, sont soumis à des pressions de sélections exercées par les mécanismes de défense de leurs hôtes (Kaján *et al.,*2020; Simmonds *et al.,*2019). De même, les hôtes subissent des pressions sélectives pour développer des mécanismes de défense contre les virus (Kaján *et al.,*2020; Obbard & Dudas, 2014). Cette interaction constante conduit à des adaptations génétiques réciproques, qui peuvent se produire à des échelles de temps variables entrainant des coûts et des bénéfices aussi bien pour l'agent pathogène que pour l'hôte (Figure 11A, Woolhouse *et al.*, 2002). Ces adaptations successives et réciproques se traduisent au niveau génétique par le changement de fréquence d'un allèle avantageux, (*e.g.* augmentation de la férquence d'un allèle de résistance chez l'hôte), suivit plus ou moins rapidement par le changement de fréquence d'un allèle avantageux (*e.g.* augmentation de la virulence) chez le virus, qui permet de contrebalancer l'adaptation de l'hôte (Figure 11B, Woolhouse *et al.*, 2002).

Chez les virus, la coévolution peut se manifester par l'émergence de variants viraux capables d'échapper aux mécanismes de défense de l'hôte (Adler *et al.,*2017; Kaján *et al.,*2020). Les virus peuvent subir des mutations génétiques aléatoires ou acquérir de nouveaux gènes par recombinaison génétique, leur permettant de contourner les récepteurs de reconnaissance de l'hôte ou de modifier leurs antigènes pour échapper à la réponse immunitaire. Ces changements peuvent conférer aux virus une meilleure capacité de réplication et de transmission, leur permettant de persister dans les populations d'hôtes (Obbard & Dudas, 2014).

De leur côté, les hôtes développent également des adaptations pour résister aux infections virales (Obbard & Dudas, 2014). Les systèmes immunitaires des hôtes peuvent évoluer pour reconnaître et éliminer les virus plus efficacement. Par exemple, les hôtes peuvent adapter leurs récepteurs cellulaires qui empêchent l'entrée des virus, produire des anticorps spécifiques qui neutralisent les particules virales, ou activer des voies de signalisation immunitaire pour éliminer les cellules infectées

(Obbard & Dudas, 2014). Ces mécanismes de défense peuvent également varier selon les populations hôtes en fonction de leur histoire évolutive et de l'exposition antérieure aux virus (Netea *et al.*, 2019).

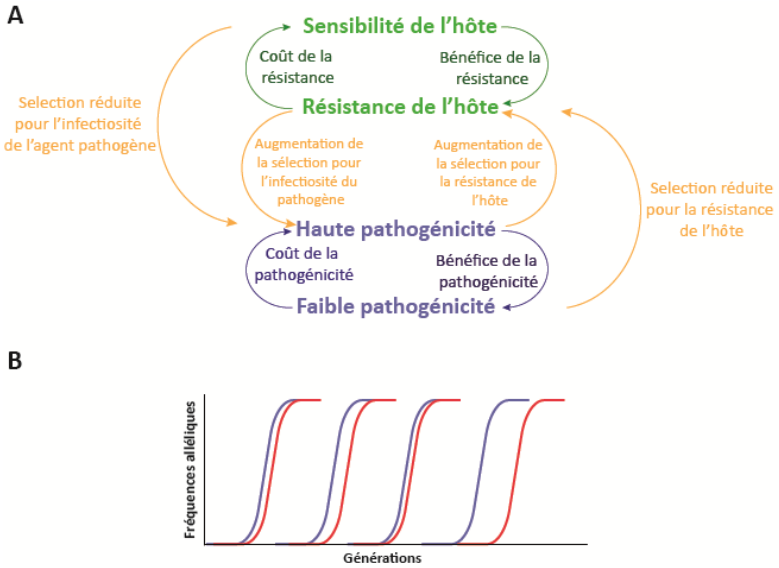


Figure 11 : Représentation schématique de la coévolution (d'après Woolhouse *et al.*2002).

A) Réciprocité des changements de fréquences alléliques due à la sélection exercée par l'hôte sur le pathogène et inversement. Les flèches vertes représentent la sélection intra-spécifique chez l'hôte, c'est-à-dire les processus de sélection qui se produisent à l'intérieur de la même espèce d'hôte. Les flèches bleues indiquent la sélection intra-spécifique chez l'agent pathogène, c'est-à-dire la sélection qui se produit à l'intérieur de la même espèce d'agent pathogène. Enfin, les flèches jaunes indiquent la sélection inter-spécifique, c'est-à-dire les processus de sélection qui se produisent entre l'hôte et l'agent pathogène, qui appartiennent à des espèces différentes. B) Changement de fréquences alléliques imposées par la co-évolution. Le graphe de gauche représente une série de balayages sélectifs imposés aux allèles de l'hôte (ligne bleue) et du pathogène (rouge) dans le cadre d'une sélection directionnelle. La co-évolution induit une adaptation réciproque qui se traduit par un changement de fréquences alléliques dans les deux populations en un temps plus ou moins court.

La coévolution entre les *herpesvirus* et leurs hôtes est un exemple fascinant de l'interaction complexe et dynamique entre un virus et son hôte. Ces virus ont coévolué avec leurs hôtes pendant des millions d'années, entraînant des adaptations réciproques (Adler *et al.,*2017; McGeoch *et al.,*2000, 2006). Par exemple, ces adaptations visent à supprimer ou inhiber la réponse immunitaire innée de l'hôte, altérer l'expression de leurs antigènes viraux pour éviter la reconnaissance immunitaire, ou établir une latence persistante dans les cellules hôtes, échappant ainsi à la détection immunitaire (De Pelsmaeker *et al.,*2018; Griffin *et al.,*2010). Les *herpesvirus* ont également évolué pour optimiser leur transmission et leur persistance dans les populations hôtes (Weidner-Glunde *et al.,* 2020). C'est le cas de certains *herpesvirus* maintenus de manière latente dans le génome de leur hôte en se circularisant sous forme d'épisomes attachés aux chromosomes par des protéines virales et cellulaires (Weidner-Glunde *et al.,*2020), ou par intégration de leur génome au sein des chromosomes de leur hôte (Morissette &

Flamand, 2010). Ils peuvent ensuite se réactiver périodiquement, conduisant à des épisodes récurrents d'infection. Cette stratégie permet au virus de persister dans l'hôte pendant de longues périodes. Ces phénomènes de réactivation du virus après une phase de latence sont régulés par les cellules de l'hôte (Adler *et al.*, 2017).

D'autre part, les hôtes ont également développé des mécanismes de défense spécifiques pour contrôler les infections par les *herpesvirus* (Griffin *et al.,*2010). L'hôte peut produire des anticorps qui neutralisent les particules virales, mobiliser des cellules tueuses naturelles (NK) pour détruire les cellules infectées, ou déclencher des réponses immunitaires cellulaires pour éliminer les virus (De Pelsmaeker *et al.,*2018).

La coévolution entre les hôtes et les virus peut également être influencée par d'autres facteurs, tels que la diversité génétique des hôtes et des virus, les taux de mutation et de recombinaison, ainsi que les changements environnementaux. Par exemple, des changements dans l'environnement peuvent favoriser l'émergence de nouveaux virus ou altérer la distribution et l'abondance des hôtes, ce qui peut entraîner des changements dans les pressions sélectives exercées sur les deux parties (Northfield & Ives, 2013).

1.2.3.2 Spécificité d'hôte

De nombreux virus, tels que les virus de la grippe *Influenza* ou les *Ranavirus*, ont un spectre d'hôte très large (Bandín & Dopazo, 2011). Mais, la course à l'évolution, notamment à travers la coévolution, conduit certains virus (*e.g. Herpesvirus*, *Betanodaviruses*) à développer une adaptation extrême à leur hôte. Cela se traduit par une spécificité d'hôte et signifie qu'ils ont évolué pour infecter et se reproduire principalement dans une ou quelques espèces hôtes spécifiques (Bandín & Dopazo, 2011).

La spécificité d'hôte des virus est souvent déterminée par des interactions moléculaires spécifiques entre les récepteurs cellulaires de l'hôte et les protéines de surface virales, mais également par d'autres étapes du cycle de réplication viral qui nécessitent la machinerie de traduction de la cellule hôte pour produire les protéines essentielles à sa réplication (Bandín & Dopazo, 2011). Les virus doivent se lier à des récepteurs spécifiques présents sur les cellules hôtes pour pouvoir pénétrer et infecter ces cellules. Ces récepteurs peuvent être des molécules membranaires spécifiques ou des récepteurs moléculaires présents à la surface des cellules hôtes. Les virus subissent une pression de sélection pour infecter efficacement les cellules de leur hôte, et ont évolué pour développer des protéines de surface qui reconnaissent spécifiquement ces récepteurs, permettant ainsi leur

attachement et leur entrée dans les cellules hôtes appropriées. C'est le cas chez les *Herpesvirus*, qui sont connus pour leur spécificité d'hôte étroite (Azab *et al.*, 2018; Bandín & Dopazo, 2011).

Il est important de noter que la spécificité d'hôte n'est pas une caractéristique fixe et immuable des virus. Au cours de leur évolution, les virus peuvent développer la capacité à infecter de nouvelles espèces hôtes. Cette transition, connue sous le terme de transmission inter-espèces ou *cross species transmission* (CST), peut s'avérer fructueuse dans de rares cas du fait des mutations génétiques ou de la recombinaison génétique. Ces altérations engendrent des modifications au niveau des interactions entre le virus et les récepteurs cellulaires de l'hôte (Bandín & Dopazo, 2011; Brito & Pinney 2020; Parrish *et al.*, 2008). Il existe deux types de CST suivant le succès de transmission du virus d'une espèce A à une espèce B (Encadré 10). Le débordement d'espèce correspond à la transmission du virus d'une espèce hôte A vers une espèce hôte B, mais le virus ne peut pas être transmis au sein de l'espèce B, tandis que dans le saut d'espèce, le virus peut être transmis au sein de l'espèce B (Longdon *et al.*, 2014).

Encadré 10 : Transmission inter-espèce ou *cross species transmission*

La transmission inter-espèces ou *cross-species transmission* (CST) d'un virus correspond au passage du virus d'une espèce hôte A pour laquelle il est adapté à une espèce hôte B pour laquelle il n'est pas adapté. Pour réussir une CST, il est nécessaire que le virus passe outre les barrières (*e.g.* environnementales, adaptations de l'hôte) qui limitent les CST. Elle regroupe deux types d'événements selon le succès de la transmission :

Le débordement d'espèce, ou *spillover,* décrit la transmission d'un agent pathogène depuis une espèce A vers une espèce B, mais qui ne donnera pas lieu à des évènements de transmission entre individus de l'espèce B.

Le saut d'espèce ou *host shift* décrit un évènement de transmission du virus depuis une espèce A vers une espèce B, à la suite duquel, des évènements de transmission au sein de l'espèce B auront lieu.

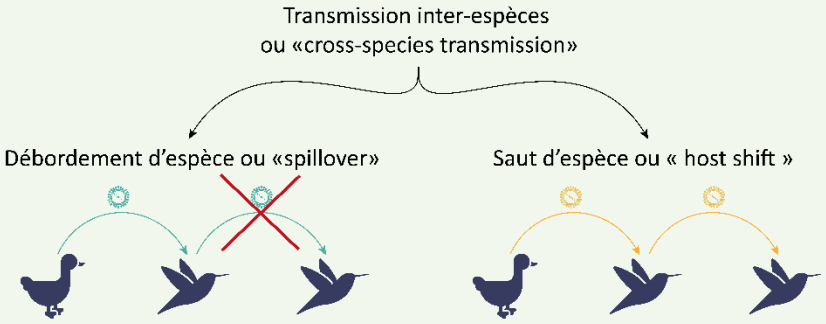


Illustration du débordement et du saut d'espèce, deux types de transmission inter-espèce.

(Longdon *et al.*, 2014)

Certains facteurs environnementaux et comportementaux peuvent favoriser les CST chez les *Herpesvirus*. La déforestation, l'urbanisation, l'exploitation forestière et l'expansion de l'agriculture intensive conduisent souvent à des contacts plus étroits entre les animaux sauvages, les animaux domestiques et les humains, créant ainsi des opportunités pour la transmission de virus entre différentes espèces (Ellwanger & Chies, 2021). Cela peut conduire à des épidémies émergentes et à la transmission de virus d'animaux sauvages aux populations humaines (Parrish *et al.,*2008). Certains exemples bien connus comprennent le virus de l'herpès B chez les macaques (Woźniakowski & Samorek-Salamonowicz, 2015) et le virus EBV chez les grands singes (Woźniakowski & Samorek-Salamonowicz, 2015). Ils ont été transmis aux humains résultant en un saut d'espèce. Le virus de l'herpès équin de type 9 (EHV-9) qui a été transmis des chevaux vers différents animaux sauvages tels les zèbres (*i.e.* saut d'espèce), les ours polaires, les girafes ou les gazelles (*i.e.* pas de de saut d'espèce démontrée mais débordement d'espèce) dans certains parcs animaliers, est également un bon exemple de saut d'espèce (Schrenzel *et al.*, 2008).

L'identification et la caractérisation des cycles épidémiologiques fournissent une vue d'ensemble essentielle de la propagation des MI et des interactions entre un virus et ses potentielles espèces hôtes. L'épidémiologie moléculaire, en combinant la biologie moléculaire et l'épidémiologie permet la caractérisation précise de ces chaines de transmission. Elle permet d'identifier des marqueurs moléculaires spécifiques et de suivre l'évolution des souches infectieuses au fil du temps.

1.3 Étude des maladies infectieuses par des approches d'épidémiologie moléculaire

En 1974, Mausner et Bahn ont défini l'épidémiologie comme « l'étude de la distribution, des déterminants des maladies et blessures au sein des populations humaines ». En 2018, Riley et Blanton ont actualisé cette définition pour inclure tous les animaux et pour intégrer la dimension moléculaire des nouvelles approches complémentaires à l'épidémiologie. L'épidémiologie moléculaire vise à décrire de manière qualitative et quantitative *i*) les occurrences et la distribution des maladies dans le temps, *ii)* les hôtes primaires ou secondaires des agents pathogènes, *iii)* les modes et profils de transmission, *iv)* les facteurs environnementaux, pathogène-spécifiques ou hôte-spécifiques favorisant la dissémination, et *v)* lier les causes d'une maladie à l'agent pathogène associé (Riley & Blanton, 2018).

Face à la répidité d'émergence des MI observée ces dernières années, il est essentiel de comprendre la transmission et l'évolution des agents pathogènes impliqués dans ces maladies. Cette compréhension constitue l'un des défis majeurs pour le déploiement de stratégies de prévention et de

mesures de gestion efficaces. Il est ainsi nécessaire de mieux caractériser la diversité des agents pathogènes, leur interaction avec leurs hôtes et avec les autres agents pathogènes, mais également les potentiels vecteurs impliqués dans leur transmission et l'impact de l'environnement.

Les approches d'épidémiologie moléculaires sont particulièrement utilisées pour l'étude des virus, *via* l'étude de leur diversité génétique et de leur évolution à des échelles de temps compatibles avec les processus épidémiologiques (Honardoost *et al.,*2018).

1.3.1 De la biologie à l'épidémiologie moléculaire

La caractérisation moléculaire et génomique des virus, des populations virales et de leur diversité joue un rôle essentiel dans la compréhension de la biologie virale, de l'évolution des virus et de leurs interactions avec les hôtes. L'histoire de l'application de la biologie moléculaire à l'épidémiologie virale est étroitement liée aux progrès technologiques dans le domaine de la génétique, de la biologie moléculaire, du séquençage et de la bio-informatique (Figure 12). Initialement, les techniques de laboratoire, dont la biologie moléculaire, ont été utilisées pour analyser un large panel d'échantillons

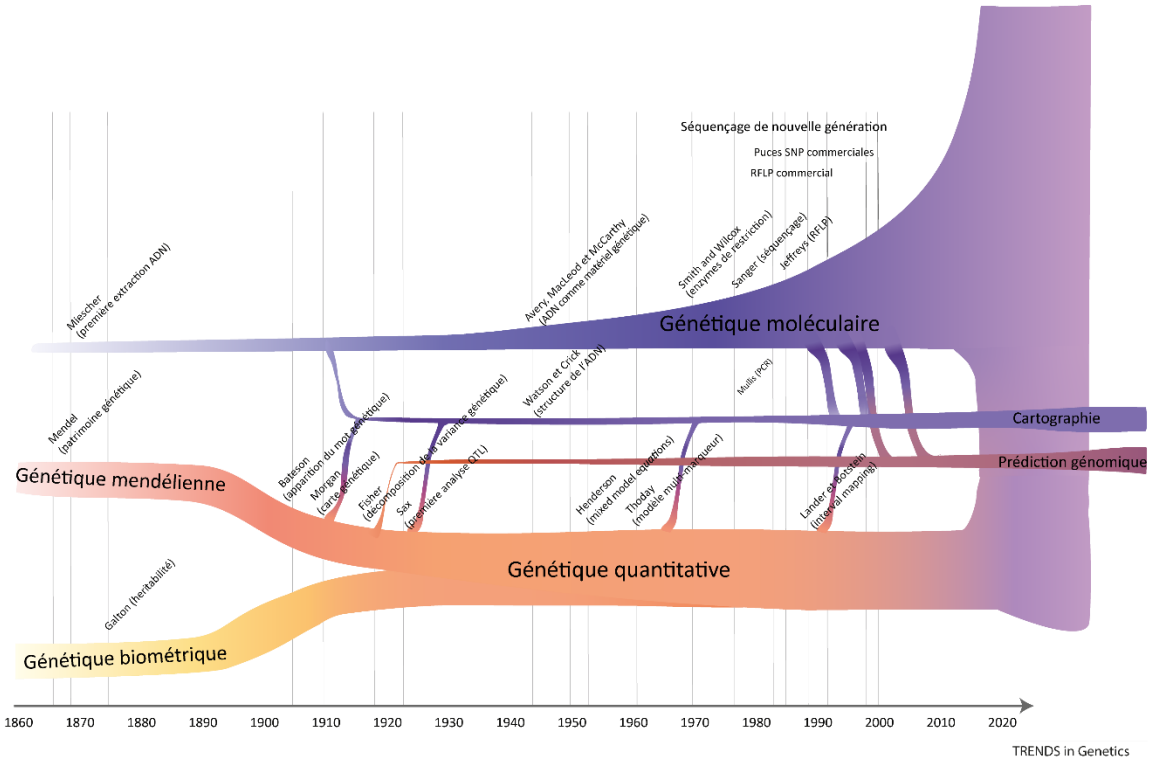


Figure 12 : Représentation schématique de l'évolution de la génétique, biostatistique et génétique moléculaire au cours du temps (adapté de Nelson *et al.,* 2013).

dans l'objectif de comparer les populations pour établir un lien de causalité entre une maladie et un agent pathogène (Eybpoosh *et al.,*2017; Grad & Lipsitch, 2014).

Les premières approches utilisées pour caractériser les populations virales étaient basées sur des méthodes telles que l'histologie et la microscopie. Elles permettaient une description visuelle des agents pathogènes (Laketa, 2018). Par la suite, les études se sont concentrées sur la caractérisation du polymorphisme nucléotidique, en utilisant des approches relativement limitées et spécifiques. Parmi celles-ci, les profils de polymorphisme de longueur de fragments de restriction (RFLP) associés à une électrophorèse en champ pulsé (PFGE), permettent, en se basant sur une petite portion spécifique du génome, de déterminer le génotype présent dans un échantillon. Des méthodes utilisant l'amplification de marqueurs spécifiques, telles que la PCR ou la PCR-RFLP, sont également utilisées (Artika *et al.,*2020; Grad & Lipsitch, 2014).

La nécessité d'améliorer la résolution de la caractérisation des agents pathogènes, notamment celle des virus, a entraîné le développement de différentes méthodes de typage de lignées, comme la Multilocus Sequence Typage (MLST), qui est devenue l'une des méthodes les plus utilisées pour la caractérisation des organismes pathogènes (Maiden *et al.*, 2013; Pérez-Losada *et al.*, 2013). Le principe de la MLST repose sur le séquençage simultané de plusieurs régions génomiques ou *loci.* Initialement cette méthode a été développée pour étudier le transfert vertical et horizontal de gènes chez les bactéries. C'est une méthode précise et reproductible, qui peut être facilement partagée par l'intermédiaire des bases de données. En revanche, la résolution du typage effectué dépend du ou des marqueurs étudiés qui peuvent être soumis à des pressions de sélection différentes et ainsi compliquer leur analyse conjointe (Maiden *et al.*, 2013).

Le développement des méthodes de séquençage a révolutionné l'épidémiologie moléculaire en permettant l'obtention de génomes complets des agents pathogènes. Une étape majeure a été franchie en 1977 avec le séquençage du premier génome complet du phage phiX174 grâce à la méthode de Sanger (Sanger *et al.,*1977). Cette avancée considérable a ouvert la voie à l'analyse approfondie de la structure génétique des agents pathogènes impliqués dans les maladies. Elle a permis de caractériser les gènes impliqués dans la virulence ou dans la résistance, d'étudier et de comparer les mutations présentes dans les génomes, amorçant ainsi les études de diversité génétique des populations et phylogénétiques (Young & Gillung, 2020).

Aujourd'hui, grâce au séquençage à haut débit, il est désormais possible de caractériser finement les agents pathogènes (*i.e.* augmentation de la résolution des informations obtenues), à moindre coût. Cela permet d'analyser en détail leur structure génétique, d'identifier les variants viraux émergents et de surveiller leur propagation (Van Borm *et al.*, 2014). Les approches basées sur les séquençages de

nouvelle génération (NGS) ont également révolutionné la surveillance épidémiologique et ont contribué à une meilleure compréhension des mécanismes de transmission des maladies infectieuses (Van Borm *et al.*, 2014). Dans le cadre de la pandémie de Covid-19 causée par le SARS-Cov-2 par exemple, la surveillance génomique a permis de constituer une base de données mondiale des séquences génomiques du SARS-CoV-2, ce qui a facilité la surveillance épidémiologique et la détection de variants émergents préoccupants. De plus, l'identification de ces variants a permis le développement de vaccins utilisés dans le cadre de la pandémie (Dhama *et al.,*2023). Ces techniques sont largement utilisées également dans le cadre des épizooties, comme la grippe a*via*ire ou les *herpesvirus* épizootiques, pour établir un diagnostic, suivre la maladie et proposer ou mettre en place des mesures de contrôle (Van Borm *et al.*, 2014).

Les séquences génomiques sont notamment utilisées pour détecter les variations génétiques qui se trouvent réparties à travers le génome de ces génotypes et suivre leur propagation géographique permettant ainsi de surveiller la répartition spatiale du virus et d'identifier les clusters d'infections liés à des variants spécifiques (Grad & Lipsitch, 2014). De même, la reconstruction d'arbres phylogénétiques permet de mettre en évidence les relations entre les différentes souches virales. Ces informations sont essentielles pour comprendre la dynamique de transmission du virus, identifier les sources d'infection et mettre en place des mesures de contrôle et de prévention ciblées *via*, par exemple, des mesures d'isolement ou le développement de vaccins et de thérapies antivirales (Grad & Lipsitch, 2014).

Les progrès fulgurants en biologie moléculaire ont ouvert de nouvelles perspectives pour la compréhension des mécanismes évolutifs des virus. Ces avancées ont permis de développer des approches expérimentales sophistiquées, telles que les expérimentations d'évolution en laboratoire qui, combinées aux NGS, ont amélioré nos connaissances sur l'adaptation virale (Kawecki *et al.*, 2012). Les évolutions expérimentales (EE) reposent sur une conception minutieuse des conditions de laboratoire pour simuler et accélérer les processus évolutifs. En exposant les populations virales à des environnements contrôlés, il est désormais possible de suivre l'évolution en temps réel des génomes viraux et d'étudier les différentes forces évolutives auxquelles le virus est soumis (Kawecki *et al.*, 2012). Lorsqu'elles sont menées sur plusieurs générations virales successives, les EE permettent d'observer la fréquence et l'évolution des mutations cruciales pour l'adaptation virale dans des conditions standardisées (Kawecki *et al.*, 2012). Ces études expérimentales offrent un aperçu précieux des trajectoires évolutives des virus, de la manière dont ils s'adaptent à de nouveaux hôtes ou environnements, et de la dynamique sous-jacente aux épidémies virales. Cependant, les EE se déroulent dans des environnements artificiels, souvent simplifiés par rapport à la complexité des écosystèmes naturels (Kawecki *et al.*, 2012). Cela pourrait limiter la généralisation des résultats

obtenus en laboratoire aux contextes réels. De plus, les expériences peuvent être sujettes à des biais expérimentaux, tels que la sélection expérimentale involontaire de certaines mutations, la sous-représentation de la diversité génétique initiale ou des tailles de populations insuffisantes peuvent limiter l'apparition de mutations (Kawecki *et al.*, 2012).

Finalement, l'intégration des avancées en biologie moléculaire avec les approches d'évolution expérimentale et de séquençage à haut débit a considérablement élargi la vision de l'évolution virale, ouvrant la voie à des stratégies innovantes de prédiction et de contrôle de l'émergence de nouvelles souches virales. Ces prédictions peuvent notamment être effectuées par des approches de génétique des populations et de génomique.

1.3.2 Caractérisation génétique et génomique des populations virales

L'utilisation des indices de génétique des populations dans la description des populations virales est une approche essentielle en virologie et en épidémiologie (Fuhrmann *et al.,*2021). Ces indices permettent d'analyser les caractéristiques génétiques des virus collectés dans différentes populations, afin de mieux comprendre leur évolution, leur propagation et leur adaptation à l'hôte. Les indices traditionnellement utilisés en génétique des populations permettent *i)* d'estimer la diversité nucléotidique des séquences génétiques, *ii)* de caractériser la différentiation génétique entre individus et entre populations, *iii)* de caractériser les mécanismes impliqués dans l'évolution des agents pathogènes ou encore *iv)* de comprendre la dynamique des populations (Fuhrmann *et al.,*2021).

1.3.2.1 Diversité nucléotidique

L'analyse de la diversité nucléotidique au sein et le long des séquences génétiques repose sur différents indices qui fournissent des informations précieuses sur la variabilité génétique d'une population ou d'une espèce. Ces indices sont basés sur la fréquence relative des différents nucléotides présents dans une séquence donnée (Fuhrmann *et al.,*2021; Lauring, 2020; Nelson & Hughes, 2015).

L'un des indices les plus couramment utilisés est l'indice de diversité nucléotidique, également connu sous le nom de theta pi (θ_π). Cet indice mesure le nombre moyen de différences par site entre paires de séquences (Figure 13). La diversité nucléotidique est estimée comme la moyenne du nombre de différences nucléotidiques entre les séquences d'un échantillon prises deux à deux, divisée par la longueur du fragment d'ADN en nombre de nucléotides :

$$\theta_\pi = \frac{n}{n-1}\sum_{ij} x_i x_j \pi_{ij}$$

avec n le : nombre de séquences dans l'échantillon ; x_i et x_j : les fréquences respectives des séquences i et j et π_{ij} : nombre de différences nucléotidiques par base entre les séquences i et j.

Plus la valeur de θ_π est élevée, plus le nombre d'allèles est grand (Fuhrmann *et al.,*2021; Lauring, 2020; Nelson & Hughes, 2015).

L'indice de ségrégation des sites polymorphes (θ_S) est également important. Cet indice quantifie la proportion de sites dans une séquence qui présentent une variabilité génétique, c'est-à-dire des positions où différents nucléotides sont observés chez les individus de la population (Figure 13). Plus précisément, il s'agit de la fréquence relative de ces sites polymorphes par rapport à l'ensemble des sites étudiés.

$$\theta_S = \frac{S_{ij}}{n}$$

avec S_{ij} le nombre de sites ségrégeant et n le nombre de bases dans la séquence.

Un indice θ_S élevé indique un grand nombre de sites polymorphes (Nelson & Hughes, 2015).

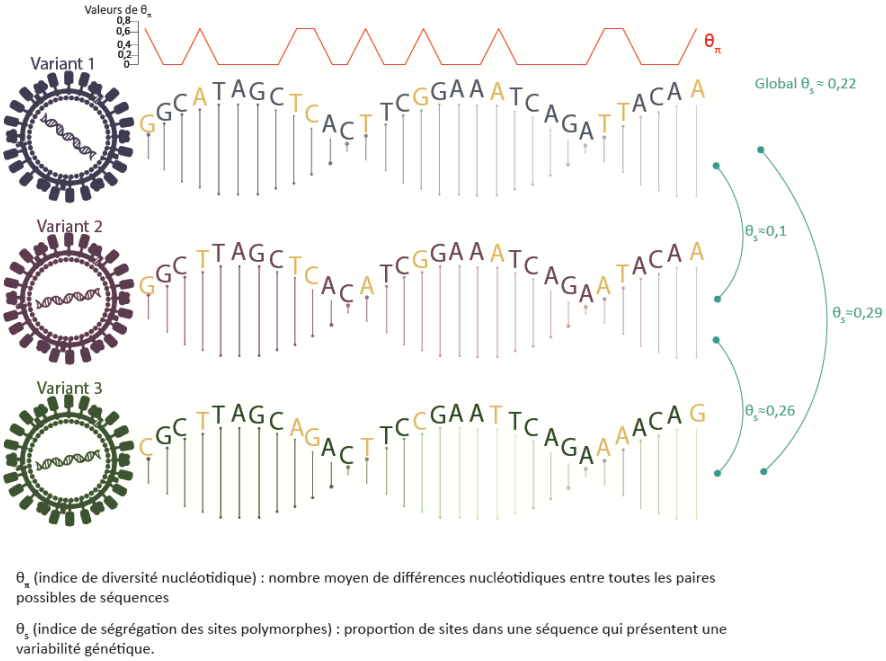


Figure 13: Représentation schématique des indices de diversité nucléotidiques θ_π et θ_S.

Exemple de trois séquences ADN de 31 pb correspondant à trois variants viraux différents : le variant 1 en bleu, le variant 2 en violet et le variant 3 en vert. Chaque nucléotide en jaune correspond à un site polymorphe. La courbe rouge au-dessus des séquences représente les valeurs de θ_π le long des séquences. A droite des séquences sont indiquées les valeurs de θ_S calculées par paires de séquences et pour la totalité des séquences.

1.3.2.2 Détection de la sélection

En 1989, le généticien japonais Fumio Tajima a proposé un test statistique utilisé en génétique des populations pour distinguer une séquence évoluant au hasard d'une séquence évoluant sous sélection (Tajima, 1989). Le D de Tajima est un test d'écart à la neutralité basé sur l'hypothèse nulle de neutralité génétique qui postule que les mutations génétiques se produisent au hasard et qu'aucune pression de sélection n'est exercée sur la population (Figure 14). Le principe de ce test est de comparer les deux estimateurs du paramètre de mutation θ décrit précédemment : le nombre de sites polymorphes (S) et le nombre moyen de différences nucléotidiques entre toutes les paires possibles de séquences (π). A taille de population constante, si la diversité génétique est conforme à l'hypothèse d'évolution neutre, θ_π et θ_S devraient être équivalents, ce qui conduirait à une valeur de D égale à zéro. Cependant, des valeurs positives ou négatives de D indiquent des scénarios évolutifs différents de l'hypothèse de neutralité ou des processus démographiques influençant la diversité génétique de la population.

Une valeur D de Tajima positive suggère que la valeur de θ_π est supérieure à la valeur de θ_S, ce qui indique une diversité plus élevée que celle attendue si le locus évoluait sous sélection neutre. Ceci peut notamment être expliqué par une diminution de la taille de population, durant laquelle les allèles de fréquences intermédiaires sont favorisés. Cet effet sera d'autant plus important si la taille de la population est petite. De plus, un D de Tajima positif peut également traduire un ou des événements de sélection balancée en faveur d'allèles de fréquences intermédiaires.

Au contraire, une valeur de D de Tajima négative indique que la valeur de θ_π est inférieure à la valeur de θ_S, ce qui suggère une diversité plus faible que celle attendue si le locus évoluait sous sélection neutre. Une expansion de la population survenue à la suite d'un goulot d'étranglement (Encadré 9) durant laquelle les allèles rares sont maintenus peut expliquer une valeur de D négative. De plus, ces valeurs peuvent également être expliquées par un balayage sélectif (Encadré 8) récent sur les allèles rares.

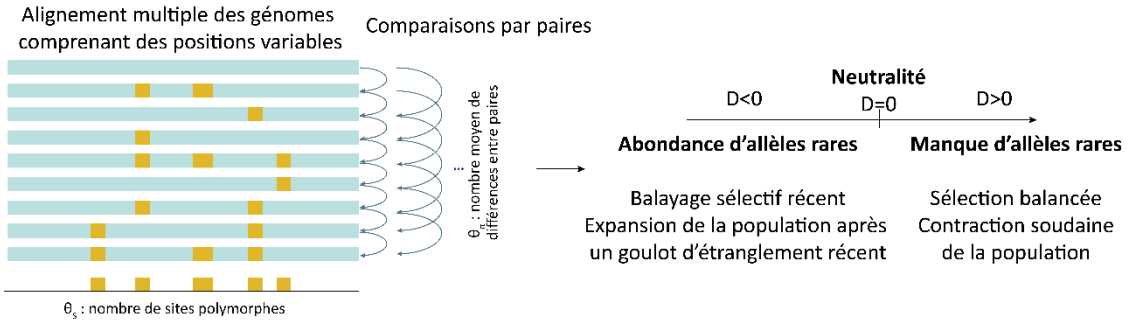


Figure 14 : Représentation schématique de l'estimation du D de Tajima.

A partir d'un alignement multiple de génomes contenant des comparaisons deux à deux des séquences sont effectuées, au sein de chaque population, et entre les deux populations. A partir de ces comparaisons, les valeurs de θ_π et θ_S pour l'alignement global ou pour des fenêtres glissantes le long de l'alignement sont estimées puis comparées par le test du D de Tajima. A partir des valeurs de D, il de testé l'hypothèse d'évolution aléatoire des séquences. Si D=0, les génomes évoluent aléatoirement. Si D<0, un excès d'allèles rares est observé, indiquant un potentiel évènement de balayage sélectif ou une expansion de la population après un goulot d'étranglement. Au contraire, si D>0, un manque d'allèles rares est observé, ce qui indique une sélection balancée ou une contraction soudaine de la population.

Similairement au test du D de Tajima, il est également possible de détecter des évènements de sélection par la comparaison du nombre de substitutions synonymes dS par rapport au nombre de substitutions non synonymes dN. Le rapport dN/dS devient supérieur à 1 lorsque la sélection naturelle favorise les substitutions dans la séquence protéique (sélection positive), tandis qu'un rapport inférieur à 1 indique que la sélection supprime la diversité de la séquence protéique (sélection purifiante).

1.3.2.3 Différenciation génétique

L'isolement des populations par la distance ou les pressions de sélection exercées par l'hôte sont responsables de différenciation des populations au niveau génétique. L'indice de fixation de Wright (F_{ST}) est un indice largement utilisé pour mesurer la différenciation génétique entre populations (Wright, 1931). Il évalue la variation génétique à l'intérieur de chaque population en la comparant à la variation génétique totale présente dans l'ensemble des populations, en se basant sur les fréquences alléliques.

Les valeurs de F_{ST} varient entre 0 et 1 (Figure 15). Une valeur de 0 indique une absence de différenciation génétique entre les sous-populations, ce qui signifie que les gènes se mélangent librement et que les populations sont génétiquement homogènes. À l'opposé, une valeur de 1 indique une différenciation génétique totale entre les sous-populations, impliquant une absence de flux génétique et des populations génétiquement distinctes.

Les valeurs de F_{ST} peuvent varier d'une espèce à une autre (*e.g.* F_{ST} = 0,34 entre les Shiba Inu et les Golden retriever), mais également au sein d'une même espèce, d'une sous-population à l'autre (*e.g.* F_{ST} = 0,11 entre les Golden retriever des USA et d'Allemagne, Karlsson *et al.*, 2007). Chez les virus, la différentiation entre deux lignées peut être particulièrement importante à cause de leur taux d'évolution rapide. Par exemple, la comparaison génétique des lignées virales de Iris yellow spot orthotospovirus (IYSV) par l'indice de fixation F_{ST} a mis en évidence une différenciation génétique importante entre les virus présents en Hollande et ceux présents au Brésil (F_{ST} = 0,72). Cette différenciation génétique est moins marquée entre ces populations et les autres populations échantillonnées dans le monde (Tabassum *et al.*, 2021).

Les variations de différenciations génétiques peuvent être la conséquence par exemple de pression de sélection, de dérive génétique, des taux de migrations, de la taille de la population ou du taux d'isolement géographique (Tills, 2008; Wright, 1931).

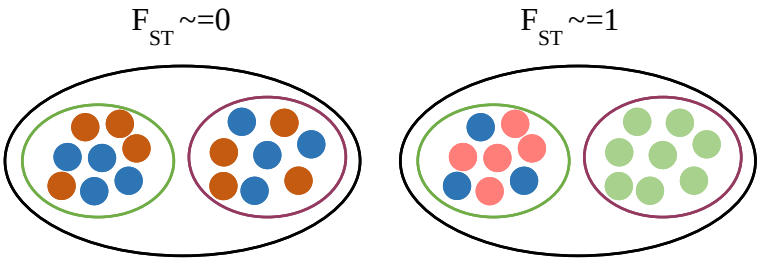


Figure 15 : Comportement de l'indice de fixation F_{ST} dans des scénarii différents de différenciation génétique.

Lorsque les fréquences alléliques sont proches dans deux sous-populations (Sous-population 1 et Sous-population 2), alors les deux sous-populations ne sont pas différenciées et l'indice de fixation calculé sera proche de 0 (à gauche). Si les fréquences des allèles dans deux sous-populations, sont totalement différentes alors les deux sous-populations sont différenciées et l'indice de fixation calculé sera proche de 1 (à droite).

1.3.2.4 Détection de la recombinaison

La détection de la recombinaison génétique est essentielle en génétique des populations, parce qu'elle permet d'étudier l'évolution des populations virales et représente un biais majeur dans l'analyse phylogénétique des séquences d'ADN (Pérez-Losada *et al.,*2015; Slatkin, 2008). En effet, le résultat des évènements de recombinaison implique la présence de fragments génétiques ayant des ancêtres communs différents dans une seule séquence nucléotidique résultant en des incongruences phylogénétiques (voir section 1.2.2.2) (Pérez-Losada *et al.*, 2015). Grâce aux NGS et à l'acquisition de données génétiques en grande quantité, l'analyse de la recombinaison génétique a connu une

révolution. Le séquençage à haut débit permet d'obtenir d'importantes quantités de données génomiques, ce qui facilite l'identification des régions recombinées (Pérez-Losada *et al.*, 2015).

Plusieurs approches ont été développées pour détecter la recombinaison génétique en fonction du type de marqueurs utilisés. L'une des approches les plus couramment utilisées consiste à détecter le déséquilibre de liaison entre deux marqueurs ou sites polymorphes. Le déséquilibre de liaison est observé lorsque les fréquences des allèles entre paires de SNP diffèrent significativement de celles attendues dans un régime aléatoire. En l'absence de recombinaison, les allèles de deux locus différents s'associent de manière non aléatoire au sein des populations (*i.e.* qu'ils ségregent ensemble), créant ainsi un déséquilibre de liaison (Figure 16, Lewontin, 1964).

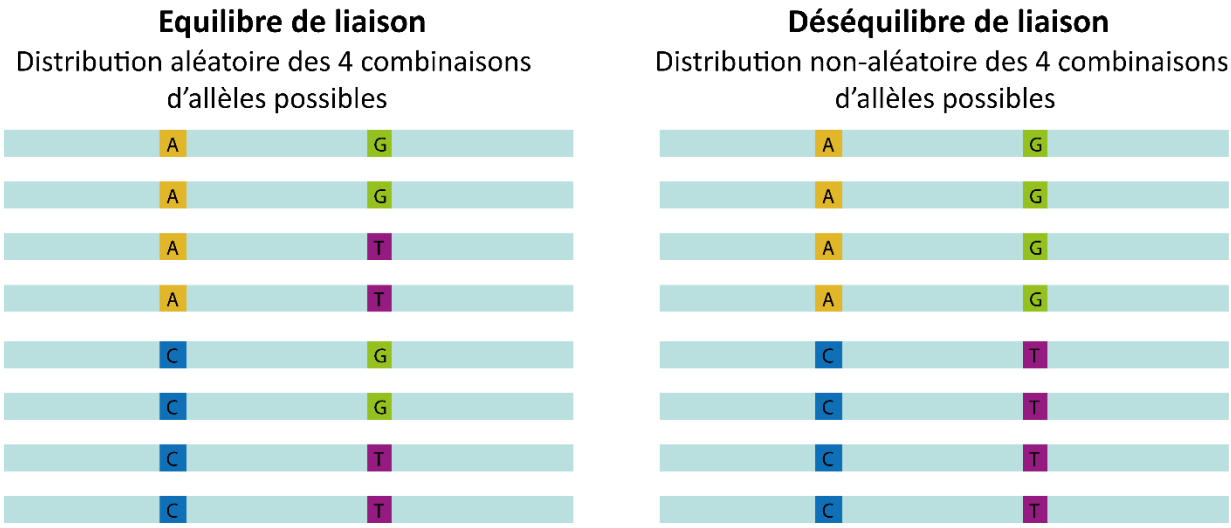


Figure 16 : Représentation schématique de séquence avec et sans déséquilibre de liaison. (D'après Rafalski, 2002)

A) Représentation d'allèles en équilibre de liaison. Dans cette configuration, chaque allèle est présent en quantité équimolaire ou presque et chaque haplotype (4 différents) a ici une fréquence de 0,25 dans la population, ce qui a pu être provoqué par un évènement de recombinaison au deuxième locus. B) Au contraire dans le cas d'un déséquilibre de liaison (D=1), une distribution non aléatoire des haplotypes est observée ce qui correspond à une situation où il n'y a pas eu de recombinaison. Certains allèles s'associent préférentiellement avec un autre. On trouve alors deux haplotypes à une fréquence de 0.5 chacun.

Les réseaux phylogénétiques (Figure 17) offrent également la possibilité de visualiser les ambiguïtés à travers des réticulations entre les branches (Huson & Bryant, 2006). Ces ambiguïtés peuvent être liées à la recombinaison ou à des événements de transfert de gènes.

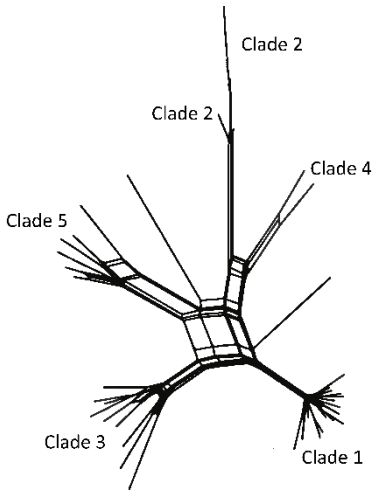


Figure 17 : Exemple de réseau phylogénétique construit à partir de 42 génomes du virus *Varicella-Zoster* montrant de possibles évènements de recombinaison représentés par les réticulations entre les branches (adapté de Zell *et al.*, 2012).

Enfin, le ratio r/m mesure la probabilité qu'un site soit altéré par la recombinaison (r) ou la mutation (m) est couramment utilisé pour décrire les évènements de recombinaison. Il permet notamment de mesurer l'importance relative des évènements de recombinaison par rapport aux évènements de mutation dans la génération de la diversité des lignées virales (González-Torres *et al.*, 2019).

1.3.2.5 Approches phylogénétiques

L'analyse phylogénétique est actuellement l'un des outils les plus puissants pour étudier l'évolution des organismes vivants, et notamment pour l'étude des lignées virales. Les arbres phylogénétiques permettent de reconstruire les liens de parenté entre les différentes souches virales (Figure 18) et de classer les organismes en fonction de leurs similarités et différences (Gorbalenya & Lauber, 2017; Pace *et al.*, 2012). Les premières études de classification hiérarchiques prenaient en compte uniquement les données morphologiques mais très rapidement les données moléculaires ont été intégrées pour gagner en précision et décrire de multiples caractères simultanément (Pace *et al.*, 2012).

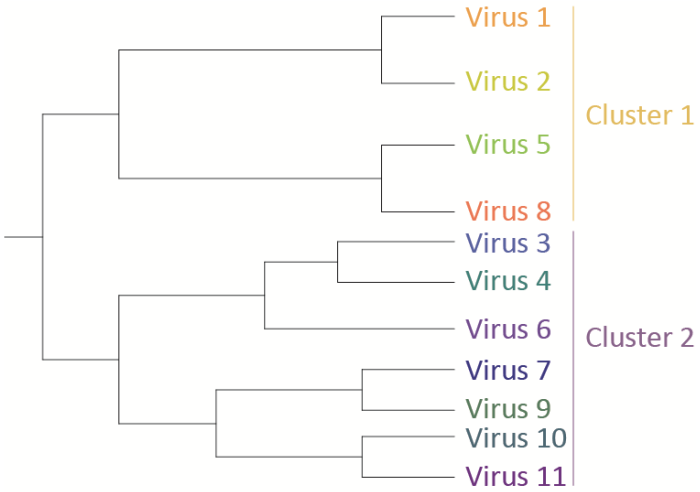


Figure 18 : Exemple schématique de représentation des relations de parenté entre 11 virus sous forme d'arbre phylogénétique.

La longueur des branches corresponds aux nombres de changements observés entre les différents nœuds et les feuilles de l'arbre (*i.e.* distance génétique). Cette distance génétique est souvent exprimée en nombre de substituitions par site. Un cluster phylogénétique est un groupe de séquences apparentées.

L'une des principales applications de la phylogénétique dans l'étude des populations virales est de représenter les relations évolutives entre les différentes souches virales. Ces arbres permettent de visualiser la diversité génétique des virus et d'identifier les liens de parenté entre les différentes souches, et d'établir des groupes ou « cluster » de séquences homologues ayant un ancêtre commun (Gorbalenya & Lauber, 2017).

La phylogénétique permet également d'identifier les facteurs qui contribuent à l'évolution des populations virales, tels que les taux de mutation ou de recombinaison (Gorbalenya & Lauber, 2017). Les taux d'évolution sont estimés en comparant les séquences nucléotidiques apparentées et en quantifiant les changements nucléotidiques apparus au fil du temps (Gorbalenya & Lauber, 2017). En ce qui concerne la recombinaison, la phylogénétique met en évidence des signaux de mélange génétique lorsque des séquences provenant de différentes lignées virales s'intègrent dans un même cluster d'arbre phylogénétique ou par l'utilisation de réseaux phylogénétiques (voir section 1.2.2.2). Les événements de recombinaison peuvent également être mis en évidence par comparaison d'arbres construits à partir de séquences de gènes avec des arbres construits à partir de génomes complets (Gorbalenya & Lauber, 2017; Pérez-Losada *et al.*, 2015). Il s'agit alors de détecter les incongruences statistiques significatives entre ces différentes topologies d'arbres (Pérez-Losada *et al.*, 2015). L'identification de ces événements permet de mettre en évidence les régions du génome viral qui sont sujettes à des échanges génétiques.

L'étude simultanée des processus épidémiologiques et évolutifs d'un pathogène infectieux à travers les phylogénies est décrite comme la phylodynamique, dont la plupart des méthodes d'inférence sont basées sur des modèles de dynamique de populations tels que les modèles de coalescent.

1.3.2.6 Modèles de coalescent

L'utilisation de modèles de coalescent dans l'analyse de la diversité et de l'évolution des populations virales est d'un intérêt majeur pour comprendre les mécanismes fondamentaux qui façonnent la dynamique génétique de ces populations. Ces modèles sont des outils mathématiques puissants qui permettent de retracer l'histoire évolutive d'une population en remontant le temps à partir des séquences génétiques actuelles (Drummond *et al.*, 2002; Rosenberg & Nordborg, 2002; Sigwart, 2009). Ils reposent sur la théorie de la coalescence, processus qui correspond à la fusion des lignées génétiques jusqu'à un ancêtre commun le plus récent (Figure 19, Drummond *et al.*, 2002; Rosenberg & Nordborg, 2002; Sigwart, 2009).

Grâce à la généalogie des échantillons permise par les modèles de coalescent, il est possible d'estimer des paramètres clés de l'évolution des populations virales, tels que la taille effective de la population ou le taux de mutation (Sigwart, 2009). Il est également possible d'identifier des évènements de goulot d'étranglement (Encadré 9) ou des émergences dans les populations virales. Ces estimations permettent d'obtenir des informations précieuses sur les dynamiques passées et présentes des populations virales, ainsi que sur les facteurs qui influencent leur diversité génétique. L'un des avantages majeurs des modèles de coalescent est leur capacité à prendre en compte des schémas complexes de diversité génétique, y compris structurés, c'est-à-dire prenant en compte plusieurs populations, au sein et entres lesquelles des évènements de recombinaison peuvent avoir lieu, ou dont les tailles de populations sont variables au cours du temps (Sigwart, 2009).

Alors que les modèles de coalescent se concentrent sur la généalogie des échantillons pour estimer des paramètres clés de l'évolution virale, les approches de phylodynamiques prennent en compte un niveau supérieur de complexité, en intégrant la généalogie, mais aussi les informations sur le temps et l'espace, ainsi que les interactions entre les populations virales.

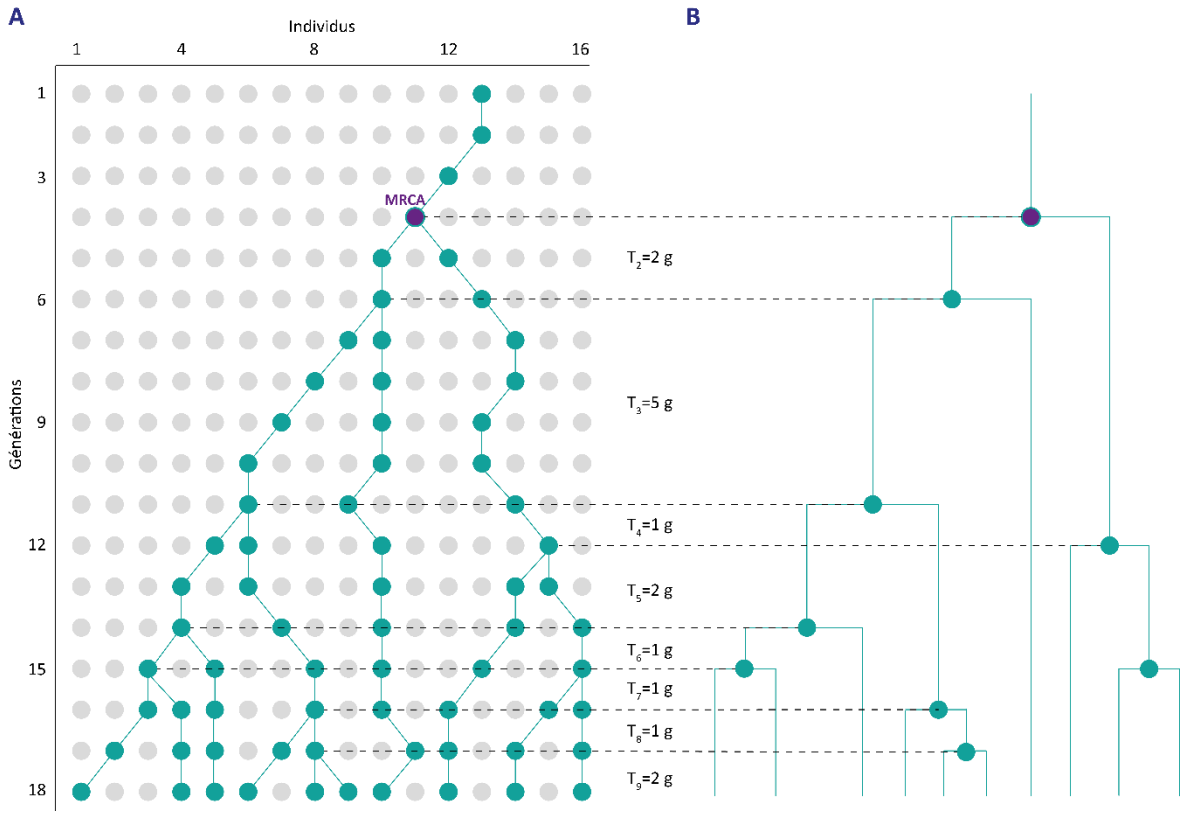


Figure 19 : Illustration d'une reconstruction de coalescent à partir d'échantillons à taille de population constante (d'après Drummond *et al.*2002 et Rosenberg *et al.*2002).

A) Généalogie complète d'une population de 16 virus au cours de 18 générations. A la génération 18, 10 virus sont échantillonnés (rond bleu en génération 18). Le but du coalescent est de retracer la généalogie, c'est-à-dire les liens de parenté de ces 10 virus avec les générations ancestrales (lignes bleues), jusqu'à l'ancêtre commun le plus récent (MRCA : Most Recent Common Ancestor), représenté en violet sur cette figure. B) Généalogie des 10 lignées virales parentes sélectionnées à la génération 18.

1.3.3 Approches phylodynamiques et phylogéographiques pour retracer l'histoire évolutive virale et les chaines de transmission

La phylodynamique combine la phylogénétique et la dynamique des populations pour analyser les relations évolutives entre les souches virales, ainsi que les facteurs démographiques et écologiques qui influencent leur émergence, leur diffusion et leur diversité génétique (Attwood *et al.*, 2022). En combinant les informations phylogénétiques avec des données épidémiologiques telles que la date d'infection, la localisation géographique (phylogéographie) et les caractéristiques cliniques des individus infectés, les modèles implémentés pour la phylodynamique permettent l'estimation de paramètres démographiques et épidémiques au cours du temps comme les changements de taille de population ou des coefficients de sélection (Encadré 11, Attwood *et al.*, 2022). Ces estimations fournissent des informations cruciales pour la compréhension de la dynamique des épidémies virales et l'évaluation de l'efficacité des mesures de contrôle (Attwood *et al.*, 2022). Elles permettent d'étudier

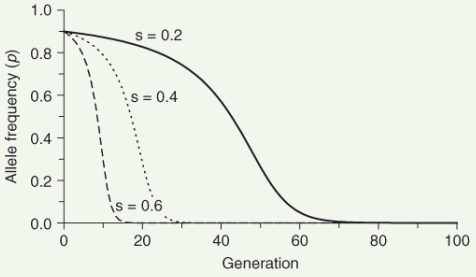
précisément l'importance des interactions entre le virus et son hôte au cours du temps et notamment de mettre en évidence des événements de transmissions inter-espèce (Gorbalenya & Lauber, 2017). En utilisant des méthodes d'inférence phylogénétique avancées, il est possible d'estimer les temps de divergence entre les souches virales, c'est-à-dire quand elles ont divergé d'un ancêtre commun, et de retracer les chemins évolutifs empruntés par les virus au cours du temps (Gorbalenya & Lauber, 2017).

De plus, la phylodynamique permet d'analyser les schémas de transmission des virus au sein des populations. En combinant des données épidémiologiques et des informations phylogénétiques, il est possible de détecter une émergence, de reconstruire les chaînes de transmission des infections, d'identifier les clusters de cas et les sources et de quantifier les taux de transmission des virus. Cette approche, joue un rôle clé dans la surveillance et le contrôle des maladies virales, en permettant d'identifier les sources d'infection et de mettre en place des stratégies de prévention ciblées (Gorbalenya & Lauber, 2017).

Encadré 11 : Coefficient de sélection

En génétique des populations, le coefficient de sélection (s) correspond à une mesure quantitative de la sélection subie par un génotype. Il exprime l'avantage ou le désavantage relatif d'un trait spécifique ou d'un génotype par rapport à une référence, en prenant en compte les aspects de survie et de reproduction.

Plus le coefficient de sélection est proche de 1, plus la sélection sera forte et donc les allèles avantageux seront sélectionnés et les allèles délétères éliminés.



Exemple de variation de la fréquence d'allèles dominants délétères au cours des générations sous différents coefficients de sélection.

(Lande & Arnold, 1983)

La phylodynamique peut également être utilisée pour étudier les facteurs écologiques et environnementaux qui influencent la diversité et l'évolution des populations virales. Par exemple, en intégrant des données sur les hôtes, telles que les espèces animales ou les populations humaines, la phylodynamique permet de mieux comprendre les processus de saut d'espèces et les interactions entre les virus et leurs hôtes (Attwood *et al.*, 2022). La plupart des méthodes d'inférence phylodynamique sont basées sur des modèles de dynamique de populations tels que des modèles de coalescent (voir section 1.3.2.6) ou de naissance et de mort, et utilisent une approche d'inférence bayésienne reposant sur les chaînes de Markov (*Markov chain Monte-Carlo,* MCMC, Encadré 12).

Encadré 12 : Méthode de Monte-Carlo par chaînes de Markov (MCMC)

Les méthodes de MCMC sont des algorithmes qui permettent l'échantillonnage de valeurs de paramètres à partir de distributions de probabilité pour effectuer des simulations. Cet algorithme permet d'explorer efficacement un espace important de paramètres.
Ces méthodes sont appliquées dans le cadre de l'inférence Bayésienne car elles sont très performantes.

(Larget et Simon, 1999)

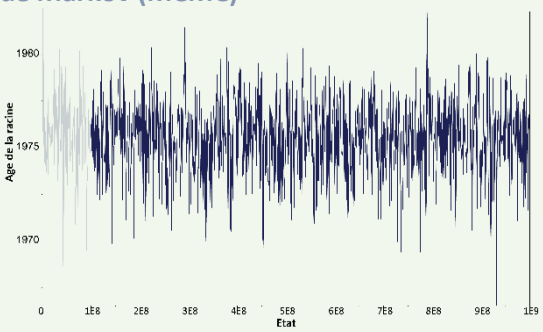


Illustration d'une chaine MCMC utilisée pour échantillonner les valeurs du paramètre « âge de la racine ».

De plus, en tenant compte des lieux d'échantillonnage, des mouvements géographiques des individus et des flux migratoires, la phylogéographie peut évaluer les schémas de dispersion (Encadré 13), les mouvements migratoires, et les processus évolutifs ayant façonné la distribution géographique des populations virales (Attwood *et al.*, 2022). Les données génétiques obtenues à partir d'échantillons de populations à différentes localités géographiques fournissent des informations cruciales sur les voies de colonisation, les événements de fragmentation, les migrations passées, ainsi que les réponses des populations aux changements climatiques et environnementaux (Attwood *et al.*, 2022). Actuellement, l'outil le plus utilisé pour ses implémentations d'approches phylodynamiques est le logiciel BEAST (Bayesian Evolutionary Analysis by Sampling Trees, Drummond & Rambaut, 2007). Il permet la reconstruction précise d'arbres phylogénétiques en utilisant la modélisation statistique, l'estimation rigoureuse des taux d'évolution moléculaire, des temps de divergence ainsi que des dynamiques démographiques au moyen de données génomiques.

Encadré 13 : Inférence des routes de dispersion virales en phylogéographie

En phylogéographie virale, la reconstruction des routes de dispersion d'un virus est une approche cruciale pour comprendre les schémas de transmission et de propagation de l'agent pathogène à travers le temps et l'espace. Cette méthode repose sur l'analyse des données génétiques du virus provenant de différentes localités géographiques et permet de retracer les liens de parenté entre les souches virales présentes dans ces différentes régions. En utilisant des modèles de diffusion et méthodes statistiques, il est possible d'estimer les taux de migration ou de diffusion du virus entre les populations hôtes dans différentes régions géographiques.

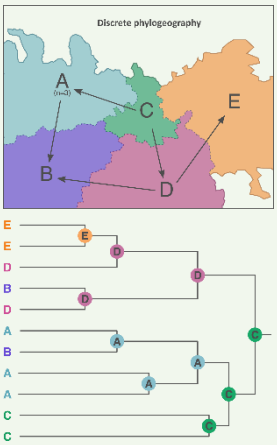


Illustration d'une inférence phylogéographique en espace discret pour la reconstruction des routes de dispersion du virus de la rage *Rabies virus* (RABV) (Nahata *et al.*2021).

(Nahata *et al.*2021)

Les modèles de diffusion occupent une place centrale et cruciale en phylogéographie. Ces modèles fournissent un cadre mathématique permettant de simuler et d'analyser les processus de dispersion des populations ou les localisations ancestrales des populations à partir de séquences génétiques. Il existe deux types de modèles de diffusion (Figure 20) : *i)* les modèles de diffusion dans un espace continu, et *ii)* les modèles de diffusion dans un espace discret.

L'utilisation des modèles de diffusion en espace continu sous-entend que les origines géographiques des génotypes viraux peuvent ainsi être représentées comme une séquence de points dans un espace continu (distribution spatiale « aléatoire »). Par ailleurs, ce type de modèle suppose une diffusion virale constante entre les différentes zones géographiques (Dellicour *et al.*, 2022). Ces chaînes permettent notamment de déterminer les directions des mouvements, les taux de diffusion et les origines ancestrales des génotypes. Ces modèles se révèlent très puissants, car ils capturent avec précision les schémas de diversité spatiale et fournissent des informations plus réelles sur les processus de dispersion (Bloomquist *et al.*, 2010; Lemey *et al.*, 2010). Ils conviennent davantage aux virus se propageant selon des processus aléatoires, tel que le transport par des animaux (*e.g.* la dissémination de la grippe a*via*ire par différentes espèces d'oiseaux ; Lemey *et al.*, 2012).

Dans le cadre des modèles de diffusion en espace discret, l'espace est divisé en unités spatiales, appelées états discrets, dans chacune desquelles les conditions de dispersion et d'évolution de l'agent pathogène sont considérées homogènes. Ces modèles trouvent particulièrement leur intérêt pour l'étude d'échantillons collectés sur des zones géographiques étendues qui ne sont pas en communication les unes avec les autres (*e.g.* un prélèvement à l'échelle d'un pays ou d'une région, Dellicour *et al.*, 2022). Par rapport aux modèles en espace continu, les modèles en espace ont l'avantage de permettre la limitation des biais liés à l'incertitude du lieu d'échantillonnage et de réduire le nombre d'inférences à réaliser puisque les estimations portent seulement sur les transitions d'une unité spatiale à une autre sans tenir compte des coordonnées GPS de chaque nœud (Kalkauskas *et al.*, 2021).

Les modèles de diffusion peuvent être combinés aux modèles de coalescent (Figure 20), utilisés, en phylodynamique, pour inférer les changements de taille de population, de taux de migration, ou d'autres paramètres démographiques qui ont influencé l'histoire évolutive des virus étudiés (Nahata *et al.*, 2021; Volz *et al.*, 2013). Ces modèles peuvent être utilisés pour tester des hypothèses sur les facteurs qui ont influencés la diffusion et la diversification des populations virales (Nahata *et al.*, 2021; Pybus *et al.*, 2000; Volz *et al.*, 2013). Il existe des modèles paramétriques qui permettent l'estimation d'histoires démographiques simples (*i.e.* taille de population constante dans le temps) et non

paramétriques qui permettent de modéliser des histoires démographiques plus complexes (*i.e.* taille de population fluctuante dans le temps).

Enfin, dans un dernier temps, les modèles de diffusion et démographiques peuvent être associés à des horloges moléculaires plus ou moins complexes (Figure 20) qui permettent d'estimer les taux et la rapidité d'évolution moléculaire (Drummond *et al.*, 2006; Gorbalenya & Lauber, 2017). Deux modèles d'horloge moléculaires existent : *i)* l'horloge moléculaire stricte qui assume un taux d'évolution constant pour toutes les lignées de l'arbre phylogénétique (Gorbalenya & Lauber, 2017), et *ii)* l'horloge moléculaire « relâchée » qui permet au taux d'évolution de varier au sein des lignées de l'arbre phylogénétique. L'horloge moléculaire « relâchée » est particulièrement utile lorsque des lignées subissent des pressions de sélection différentes par exemple (Drummond *et al.*, 2006).

Bien que la phylodynamique soit devenue un outil incontournable pour décrypter l'évolution des maladies infectieuses humaines, son application reste limitée dans les domaines vétérinaires et, plus spécifiquement, aquacole. Il est évident que ces domaines connaissent des défis similaires en termes de propagation de maladies et d'adaptation génétique des agents pathogènes, pour lesquels les approches phylodynamiques seraient extrêmement pertinentes. Dans le contexte de l'aquaculture, où les maladies peuvent causer des pertes économiques massives et menacer la sécurité alimentaire, l'application de la phylodynamique pourrait apporter des perspectives précieuses pour la gestion durable de l'aquaculture.

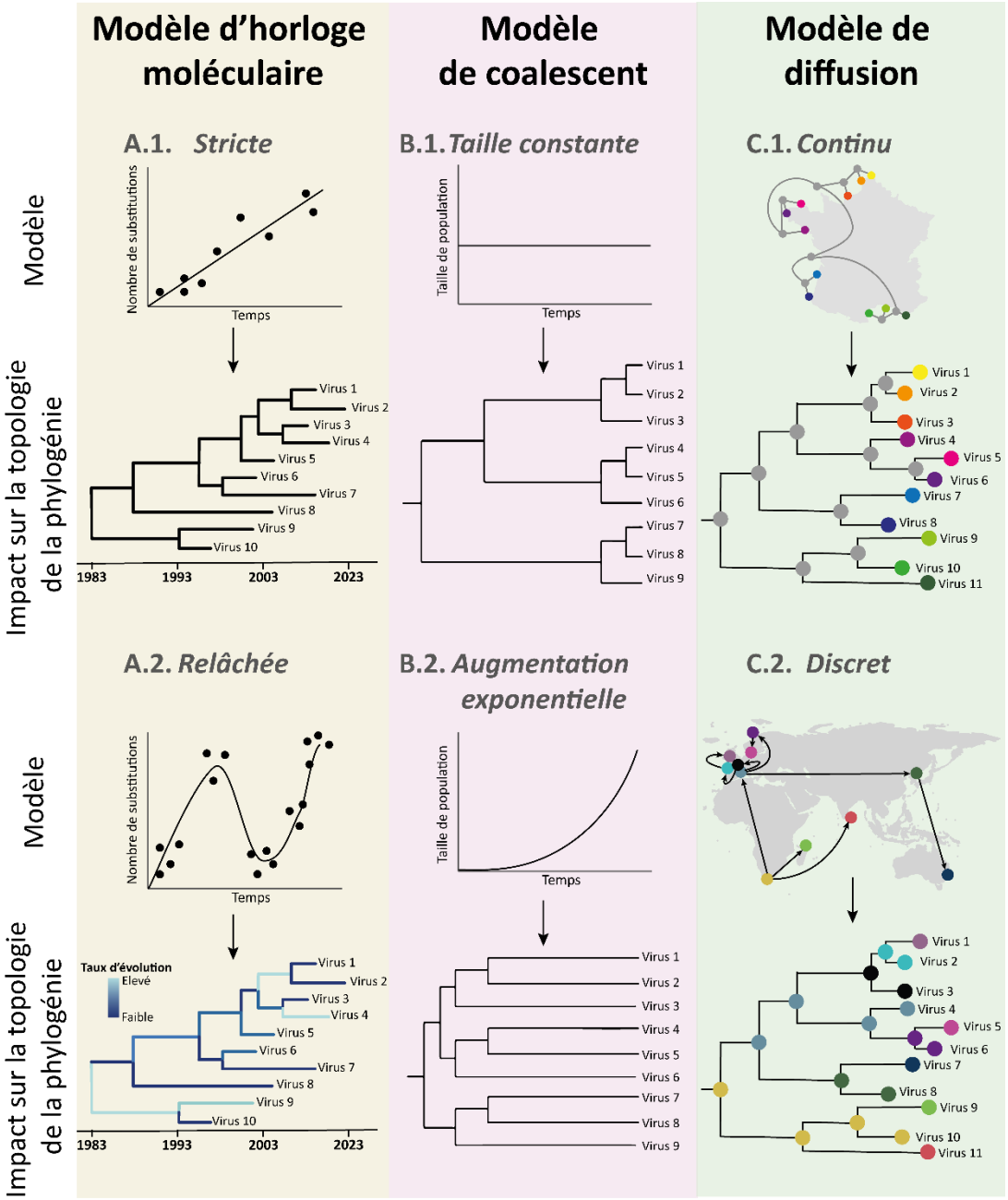


Figure 20 : Exemples de modèles d'horloge moléculaire, de coalescent et de diffusion, implémentés dans BEAST pour les analyses phylodynamiques (d'après Drummond & Rambaut, 2007, Volz *et al.*2013, et Nahata *et al.*2021).

A) Exemples de deux modèles d'horloges moléculaires pouvant être utilisés pour inférer des phylogénies calibrées dans le temps. L'horloge moléculaire stricte (A.1.) est utilisée dans le cas où l'évolution des lignées virales est assumée similaire pour l'ensemble de la phylogénie, ce qui se traduit par un seul taux d'évolution pour l'ensemble de la phylogénie. Au contraire, le modèle d'horloge moléculaire relâchée (A.2.) est utilisé dans le cas où le taux d'évolution varie au cours du temps et en fonction des lignées virales, ce qui se traduit par des taux d'évolution plus ou moins élevé en fonction des branches de l'arbre phylogénétique. B) Exemples de modèles démographiques pouvant être utilisés dans BEAST. Dans le modèle démographique de taille constante (B.1.), la taille de population est assumée constante au cours du temps, ce qui se traduit par des longueurs de branches internes plus importantes que celles des branches externes sur la phylogénie. Au contraire, dans le modèle d'augmentation exponentielle de la taille de population (B.2.), il est assumé que la taille de population augmente de façon exponentielle, ce qui se traduit par des longueurs de branches externes supérieures à celles des branches internes. C) Exemples de deux modèles de diffusion implémentés dans BEAST. Le premier modèle correspond à un modèle en espace continue (C.1.) et le second a un modèle de diffusion en espace discret (C.2.).

1.4 Cas des *Herpesvirus* en aquaculture et focus sur l'*Ostreid herpesvirus 1*

1.4.1 Aquaculture et émergence d'*herpesvirus*

L'aquaculture est une activité économique importante répondant à la demande croissante de produits marins dans le monde. Le terme « aquaculture » est très large et englobe toute la production animale et végétale en milieu aquatique. En 1990, l'aquaculture marine mondiale représentait à peine 15 millions de tonnes (15 % de la production aquatique globale, pêche et aquaculture ; FAO, 2020). Depuis les années 90, on observe une croissance rapide des cultures aquacoles. En 2021, l'aquaculture a atteint une production de 123 millions de tonnes (46 % de la production marine globale, pêche et aquaculture), représentant ainsi une denrée économique importante. Actuellement, l'aquaculture marine est dominée par la production de mollusques (17 millions de tonnes, soit 56 % de l'aquaculture marine mondiale), et, notamment par la production d'huîtres creuses (6 millions de tonnes, soit 35,2 % de la production de mollusques) (FAO, 2023).

Tout comme en agriculture, les pratiques culturales, vétérinaires, et l'intensification de la production en aquaculture impliquent l'émergence et le développement rapide de MI qui peuvent menacer la durabilité des élevages (Krkosek, 2010). Les maladies virales touchent diverses espèces, parmi lesquelles le saumon, la truite, les carpes, les crevettes ou les huîtres. Les virus qui causent ces maladies peuvent appartenir à différentes familles virales (Figure 3), notamment les *Reoviridae* (Kibenge, 2019; Zhao *et al.*, 2021), les *Rhabdoviridae* (Kibenge, 2019; Zhang et al., 2022), les *Orthomyxoviridae* (Kibenge, 2019) et les *Allo- et Malacoherpesviridae* (Figure 4, Davison, 1998; Davison *et al.*, 2005, 2009; Donohoe *et al.*, 2021).

1.4.1.1 Les *herpesvirus* en aquaculture

Les *Herpesvirus* sont des agents pathogènes ayant un impact significatif pour l'industrie aquacole mondiale. Plusieurs *herpesvirus* sont impliqués dans des maladies affectant les poissons d'élevage. Parmi les plus notables, le virus de l'herpès du saumon (*Salmonid Herpesvirus*, SalHV-1), touche principalement les espèces de salmonidés, telles que le saumon, la truite et l'omble. Le virus de l'*herpesvirus* de la carpe (*Cyprinid Herpesvirus*, CyHV-1, CyHV-2 et CyHV-3) affecte les carpes communes et les carpes koï (Hanson *et al.*, 2011). Les élevages d'anguilles Européennes, Japonaises et Américaines sont touchés par l'herpesvirus de l'anguille (*Anguilid herpesvirus 1*, AngHV-1, Sano *et al.*, 1990). De même, les élevages d'ormeaux sont fortement impactés par l'*Haliotid herpesvirus* (Chang *et al.*, 2005). Enfin, l'*Ostreid herpesvirus 1* (OsHV-1) est responsable d'infections provoquant de fortes

mortalités chez les huîtres et cause des problèmes importants pour l'ostréiculture (Davison *et al.*, 2005; Nicolas *et al.*, 1992).

Les infections à *herpesvirus* peuvent avoir des effets dévastateurs sur l'industrie aquacole. Les espèces infectées peuvent présenter une morbidité élevée, entraînant des taux de mortalité significatifs dans les populations d'élevage (Hanson *et al.*, 2011). Ces épidémies peuvent provoquer des pertes économiques substantielles pour les producteurs, en raison de la diminution de la production et de la nécessité de mettre en œuvre des mesures de contrôle coûteuses. Ces coûts supplémentaires concernent notamment la quarantaine des poissons nouvellement introduits, la vaccination, l'amélioration génétique pour la résistance virale, ainsi que le suivi épidémiologique régulier pour détecter rapidement les éventuelles épidémies (Combe *et al.*, 2023).

Les modes de transmission varient selon les *herpesvirus*, mais les voies les plus courantes comprennent le contact direct entre individus infectés, les sécrétions corporelles, la contamination *via* de l'eau de mer, les débris organiques, les vecteurs tels que les parasites et les prédateurs, ainsi que l'introduction d'individus infectés dans des installations aquacoles (Krkosek, 2010). Dans certains cas, les *herpesvirus* peuvent être introduits ou maintenus par des mauvaises mesures d'hygiène, comme un manque de désinfection des équipements utilisés lors de l'entretien des installations aquacoles.

La susceptibilité à l'infection par les *herpesvirus* varie d'une espèce à l'autre et peut être influencée par différents facteurs, tels que l'âge, la santé de l'individu, les conditions environnementales et le niveau de stress des animaux (Hanson *et al.*, 2011). La densité élevée des populations dans les bassins aquacoles peut également favoriser la propagation du virus.

L'émergence et la propagation rapide des infections par des *Herpesvirus* sont étroitement liées aux facteurs environnementaux offerts par le milieu aquatique.

1.4.1.2 Facteurs environnementaux facilitant l'émergence d'*herpesvirus* en milieu aquatique

L'émergence des *herpesvirus* en milieu aquatique est étroitement liée à divers facteurs environnementaux qui influencent la survie, la transmission et la virulence de ces virus (Figure 21, Kibenge, 2019). Tout d'abord, le milieu aquatique offre un milieu propice à la propagation des virus de manière générale du fait des courants qui facilitent la transmission des particules virales entre les individus. La température de l'eau joue un rôle crucial dans la dynamique des *herpesvirus*. Les variations saisonnières de la température peuvent influencer leur taux de réplication et leur persistance dans l'environnement (Burge et al., 2014; Combe et al., 2023; Petton et al., 2013; Zhang et

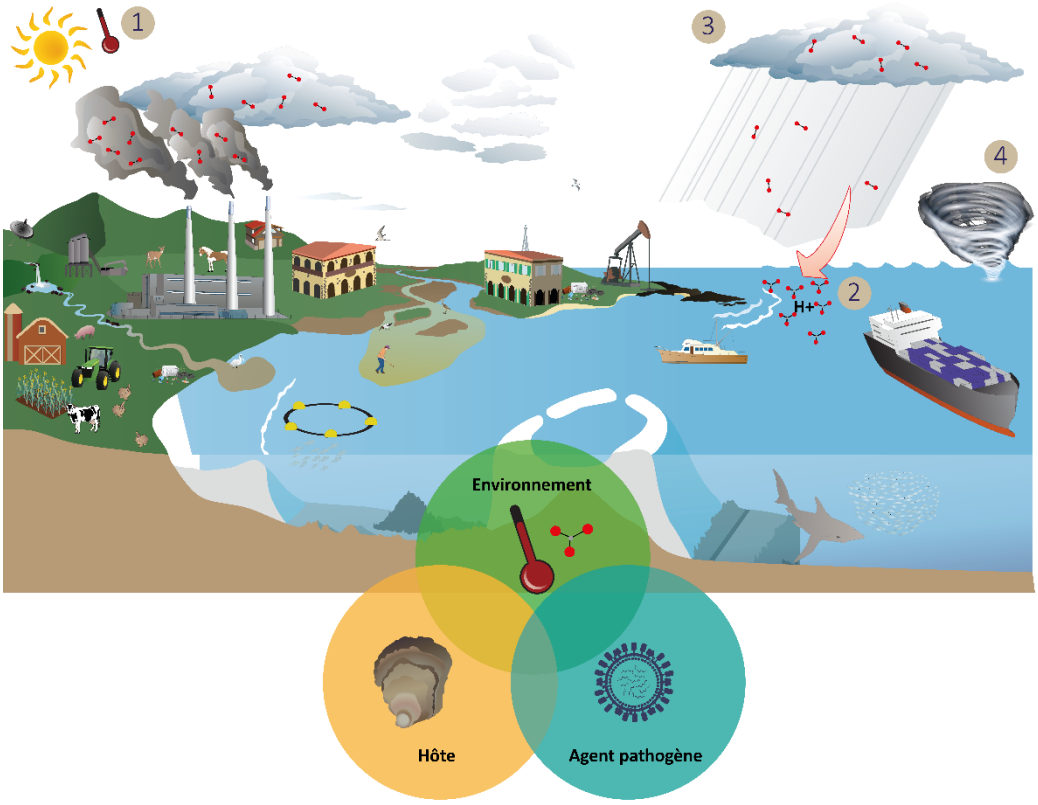
al., 2022). Des températures plus élevées peuvent accélérer la multiplication virale et favoriser ainsi l'infection des populations aquatiques (Combe *et al.*, 2023). De plus, des périodes prolongées de températures élevées peuvent affaiblir le système immunitaire des poissons et des autres organismes aquatiques, les rendant plus sensibles aux infections virales (Burge *et al.*, 2014).

La qualité de l'eau est un autre facteur clé. La présence de polluants, de produits chimiques et de contaminants peut affaiblir la résistance des animaux aquatiques aux *herpesvirus* en perturbant leur système immunitaire (Moreau *et al.*, 2015). De même, une teneur trop élevée en dioxyde de carbone peut entraîner une augmentation du pH de l'eau, créant ainsi une acidification et produisant des réactions chimiques qui affaiblissent les organismes (*e.g.* problème de calcification des coquilles chez les huîtres ; Miller *et al.*, 2009). Une faible teneur en oxygène dissout dans l'eau, émanant d'une eutrophisation ou d'une dégradation de l'environnement, souvent observé en aquaculture, peut créer des conditions de stress pour les poissons, les rendant plus susceptibles à l'infection virale (Burge *et al.*, 2014).

Les variations de salinité peuvent également influencer la propagation des *herpesvirus* (Zhang et al., 2022). Certains virus peuvent être plus stables dans des eaux salées ou saumâtres, tandis que d'autres prospèrent mieux dans des eaux douces. Par conséquent, les changements dans la salinité peuvent moduler la distribution des *herpesvirus* dans différents habitats aquatiques (Zhang et al., 2022).

Les facteurs liés à la migration et au mouvement des espèces aquatiques jouent également un rôle important dans l'émergence des *herpesvirus*. Les mouvements saisonniers des poissons migrateurs peuvent faciliter la dispersion des virus sur de grandes distances, tandis que la concentration d'animaux aquatiques dans des zones spécifiques, comme les frayères, peut favoriser la transmission inter-espèces des virus (Hanson *et al.*, 2011).

Enfin, les changements climatiques peuvent avoir un impact significatif sur l'émergence des *herpesvirus* en milieu aquatique. Les modifications des températures de l'eau, des régimes de précipitations, des courants océaniques et des conditions météorologiques peuvent influencer la distribution géographique des virus et favoriser leur propagation vers de nouvelles régions (Combe *et al.*, 2023).



Changements chez l'hôte	Changements environnementaux	Changements chez l'agent pathogène
• Augmentation de la sensibilité • Changement de comportements • Invasions • Changement de distribution géographique	1	• Augmentation de la réplication • Changement de virulence • Nouvelles émergences • Invasions • Changement de distribution géographique
• Augmentation de la sensibilité • Changement dans la survie et le recrutement larvaire • Changement de comportements • Réduction de la diversité	2 H+	• Changement de virulence • Changement d'abondance
• Augmentation de la sensibilité • Stress nutritionnel • Stress lié à la salinité • Stress lié à la pollution	3	• Changements de distrbution géographique • Changement de virulence • Nouvelles émergences • Invasions • Plus de nutriments
• Augmentation de la sensibilité • Blessures physiques	4	• Changements de distrbution géographique • Nouvelles émergences • Invasions • Résurgence de nutriments

Figure 21 : Impacts du changement climatique sur les relations hôtes-pathogène-environnement marin (d'après Burge *et al.*2014)

Le réchauffement climatique entraîne des modifications physiques de l'océan : des changements de température ; des augmentations des concentrations de CO_2 et acidification de l'eau ; des variations de précipitations entrainant des modifications de la salinité.

1.4.2 L'ostréiculture Française

L'élevage des huîtres ou « ostréiculture », est une filière économique importante dans le monde avec une valeur marchande de 7,6 milliard USD en 2021 (FAO, 2023). En Europe, la production ostréicole a dépassé 114 000 tonnes en 2021, représentant 568 millions USD, avec une part considérable pour la France métropolitaine de 86 000 tonnes (455 millions USD), soit 76 % de la production Européenne.

Son histoire en France débute au milieu du XVIIIème siècle, avec la pêche d'huîtres plates *Ostrea edulis* (Linnaeus, 1758), espèce endémique des côtes Européennes à cette période. Cette espèce fut décimée par la surexploitation alimentaire des gisements naturels entre 1700 et 1750 (Figure 22). A partir de 1854, le développement et l'utilisation de nouvelles techniques de captage des larves permet d'assurer une production d'huîtres plates afin de renouveler les populations (Buestel *et al.,*2009). *O. edulis* devient alors l'huître majoritairement exploitée jusque dans les années 1920, date à laquelle les premières mortalités sont observées, entraînant un deuxième déclin de cette espèce (Dollfus, 1921). En 1927, l'huître plate est presque décimée. L'ostréiculture Française s'oriente alors vers l'huître portugaise *Magallana angulata* (Lamarck, 1819) introduite accidentellement lors de l'échouage d'un bateau remplit de ces huîtres dans les années 1860 (Boudry *et al.*, 1998). Durant la période 1950-1970, l'huître portugaise est l'huître la plus cultivée en France (Figure 23). En 1966, un irridovirus-like est suspecté de provoquer l'extinction de *M. angulata* sur les côtes Françaises (Comps, 1988). En parallèle, au début des années 1970 une première diminution de la production d'huîtres plates est observée en lien avec la détection d'un premier agent pathogène, le parasite protozoaire *Marteilia refringens*, responsable de la Marteiliose. Puis, un second agent pathogène, le parasite protozoaire *Bonamia ostreae,* responsable de la Bonamiose, est détecté suite à de fortes mortalités à la fin des années 70 . Ces deux parasites protozoaires ont considérablement affecté la production voire même provoqué une forte diminution des stocks naturels d'huîtres plates. Afin de palier à la diminution de production de l'huître plate Européenne et à l'extinction de l'huître Portugaise, l'huître creuse du pacifique *Magallana gigas* (Thunberg, 1793) est alors importée dans les années 1970 du Japon et du Canada pour maintenir l'activité ostréicole en France (Boudry *et al.*, 1998; Grizel & Heral, 1991).

Figure 22 : Tableau « Déjeuner aux huîtres » peint en 1735 par Jean-François de Troy (1679-1752).

Ce tableau illustre les repas dans les appartements du roi Louis XV où les huîtres étaient consommées en grande quantité.

Actuellement, l'huître plate *O. edulis* est listée comme « espèce menacée et en déclin » par la convention Oslo-Paris (MNHN, 2014). L'ostréiculture est aujourd'hui largement dominée par la production d'huîtres creuses (*i.e.* 98% de production de *M. gigas* en France, FAO 2023), mais sa production fluctue du fait de nouvelles épizooties détectées depuis une trentaine d'années. Très rapidement, au début des années 1990, un virus fut suspecté d'être responsable de ces mortalités (Nicolas *et al.*, 1992; Renault *et al.*, 1994a, 1994b). Cette hypothèse sera d'abord confirmée par observation microscopique, l'assimilant ainsi à un *Herpesvirus*, ce qui a ensuite pu être confirmé par le séquençage de son génome (Davison *et al.*, 2005; Le Deuff & Renault, 1999; Renault, 1998; Renault, Cochennec, *et al.*, 1994). Ce virus, nommé *Ostreid Herpevirus type 1* ou OsHV-1, a causé des mortalités estivales importantes, et, continue actuellement, de causer de fortes mortalités sur les huîtres juvéniles (Arzul *et al.*, 2017). De plus, en 2012, la bactérie *Vibrio aestuarianus* a aussi été associée à de fortes mortalités chez les huîtres adultes (François *et al.*, 2013; Goudenège *et al.*, 2015). Ces deux agents pathogènes majeurs de l'huître creuses ont fortement impacté la production ostréicole ces dernières années (Petton *et al.*, 2021).

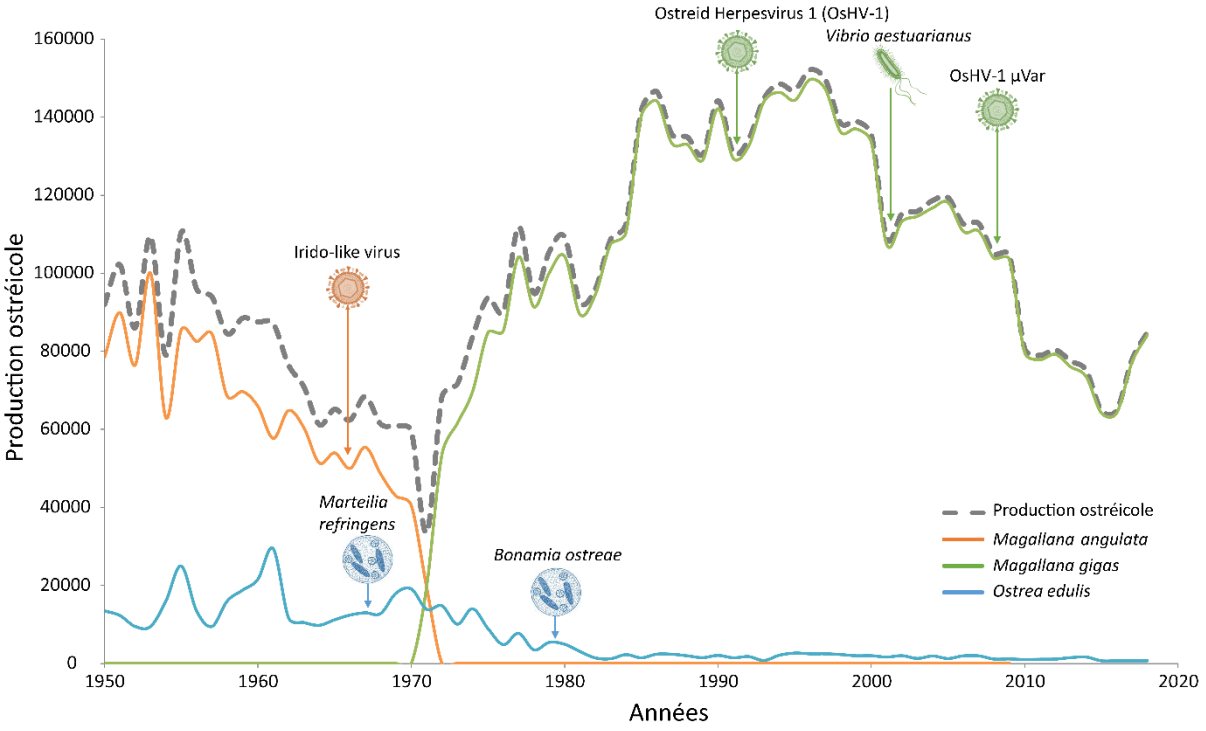


Figure 23: Production ostréicole Française et émergence des épizooties entre 1950 et 2019 (d'après FAO, 2020).

Les courbes représentent la production d'huîtres globale (en gris), Portugaise *M. angulata* (en orange), du Pacifique *M. gigas* (en vert) et Européenne *O. edulis* (en bleu). Les épizooties sont indiquées par des flèches colorées en fonction de l'espèce d'huître infectée.

1.4.3 Cycle de vie de l'huître creuse

L'huître creuse est un mollusque marin bivalve hermaphrodite successif, c'est-à-dire qu'un individu est successivement mâle puis femelle, et seulement dans de rares cas, les individus présentent les deux sexes simultanément (Amemiya, 1929). Le déterminisme du sexe chez l'huître creuse semble influencé par de nombreux paramètres extérieurs et génétiques (Broquard, 2019). Sa reproduction est externe. L'ensemble des gamètes mâles et femelles sont expulsés dans l'eau pour réaliser une fécondation, qui est à l'origine d'une cellule œuf. Le cycle de vie de l'huître creuse (Figure 25) est ensuite constitué de deux stades : un stade de vie larvaire et un stade de vie benthique (*i.e.* vie fixée sur le substrat de fond). Le début de la vie larvaire commence avec le développement de la cellule œuf vers une larve trocophore (*i.e.* larve ciliée spécifique des mollusques, annélides et plathelminthes) en moins de 6 heures, disposant de nombreux cils qui lui permettent de se déplacer (Vogeler *et al.,*2016). Vingt-quatre heures après la fécondation, une coquille commence à se dessiner et donne une forme de « D » à la larve, qui lui donne son nom. Elle mesure 60 µm, possède des cils et commence à développer un

velum (*i.e.* voile cilié) qui lui permet de se déplacer et de capturer des petites proies. A partir du deuxième jour et jusqu'au 20ème jour, la larve est appelée larve véligère. Elle est pélagique et continue de développer son velum et sa coquille, ainsi que la charnière de la coquille, qui permettra ensuite l'ouverture de la coquille. A partir du vingtième jour, la larve pédivéligère (300 µm, Figure 24A) commence sa métamorphose, la larve tombe sur le fond et cherche à se fixer sur un substrat adapté, avant de s'y fixer définitivement (Figure 25, Figure 24B). Durant cette période, un pied ainsi qu'une tache noire sur la coquille (larve oeillée) apparaissent. Après ces étapes de métamorphose et de fixation, la larve entame sa vie benthique et devient du micro-naissain, (Figure 25, Figure 24B et C) puis du naissain (Figure 25, Figure 24D). Pour obtenir sa taille adulte, le naissain devra encore se développer pendant deux ans et demi.

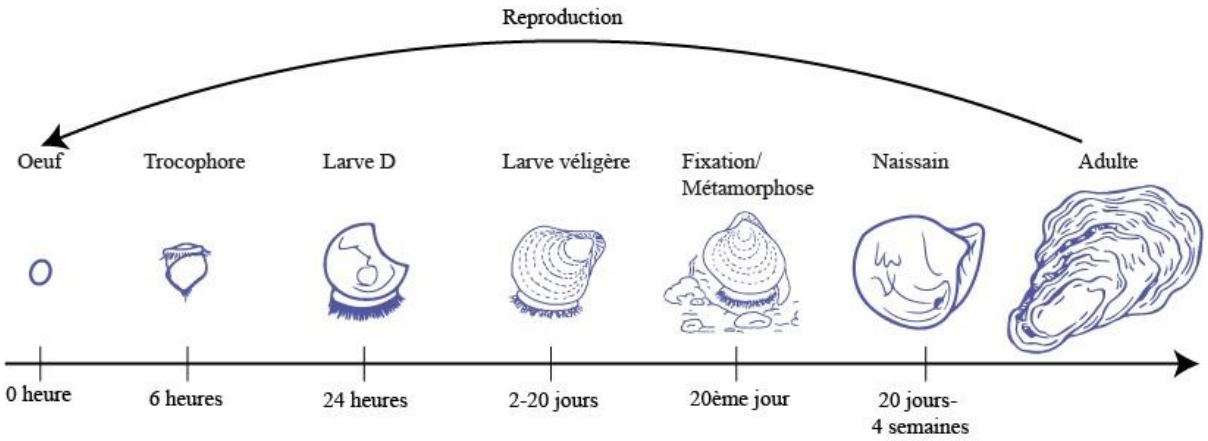


Figure 25: Cycle de reproduction biologique de l'huître creuse *M. gigas* (d'après Vogeler *et al.*2016).

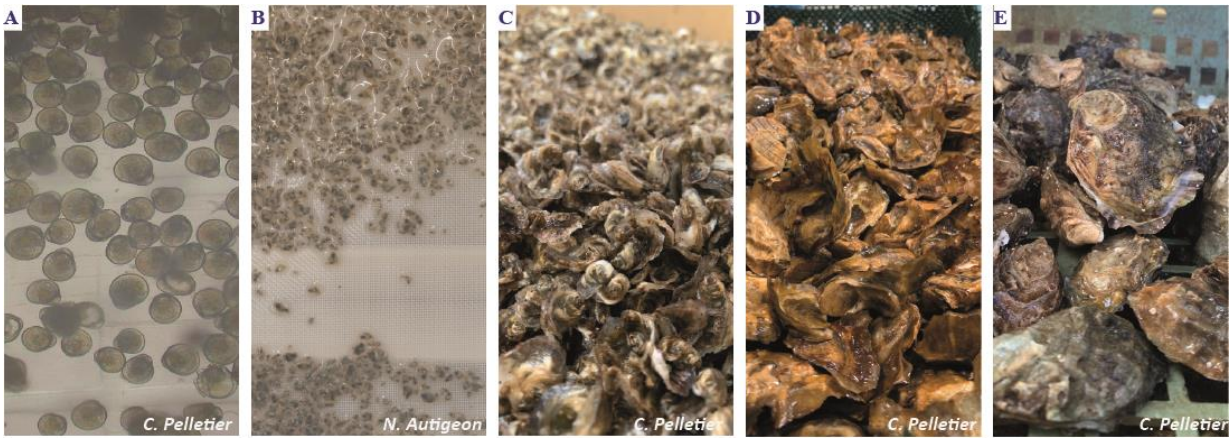


Figure 24 : Images des principaux stades de développement de l'huître creuse *M. gigas*.

A) Stade larve véligère, B-C) stade micro-naissain, D) stade naissain et E) stade adulte.

1.4.4 Mortalités estivales chez le naissain d'huîtres creuses

Tous les ans, durant la saison estivale, les naissains d'huîtres creuses sont touchés par d'importants épisodes de mortalités. En 2008, jusqu'à 100% de mortalité a été rapportées pour les stades larvaires et juvéniles chez *M. gigas* sur le littoral Français. Ces mortalités ont été associées à la détection d'un variant du virus OsHV-1 nommé OsHV-1 µVar (Segarra *et al.*, 2010).

Récemment, une étude a rapporté que cette maladie pourrait être polymicrobienne panzootique (appelé POMS pour Pacific Oyster Mortality Syndrome), affectant principalement les larves et le naissain d'huîtres creuses (de Lorgeril *et al.,*2018; Petton *et al.,*2021). Plusieurs agents pathogènes interviennent dans ce syndrome. Le naissain est exposé à l'OsHV-1 qui va infecter les hémocytes des cellules de l'hôte (Figure 26). La physiologie des hémocytes est alors altérée, induisant une immunodépression chez l'hôte, ce qui favorise la prolifération de bactéries opportunistes (*i.e.* bactéries du genre *Vibrio*). Le développement bactérien entraine une lyse des hémocytes permettant une colonisation des tissus profonds par les bactéries, ce qui, à terme, provoque la mort de l'animal (Petton *et al.,*2021).

Quatre paramètres essentiels entrent en jeu dans cette maladie multifactorielle :

i) La présence d'OsHV-1 et de bactéries opportunistes simultanément dans le microbiote de l'huître (de Lorgeril *et al.,*2018).

ii) le fond génétique des huîtres et leur niveau de sensibilité à l'infection virale par OsHV-1 (Dégremont, 2013; Segarra *et al.,*2014).

iii) leur nutrition, même si le mécanisme par lequel la disponibilité en nourriture affecte la résistance des huîtres est encore très mal connu, il semblerait que des huîtres à jeun aient un meilleur taux de survie que des huîtres bien nourries (Pernet *et al.,*2019).

iv) la température, puisque le développement de cette maladie est permis dans une fenêtre restreinte de température de l'eau, allant de 16°C à 24°C (Fleury *et al.,*2020; Pernet *et al.,*2014).

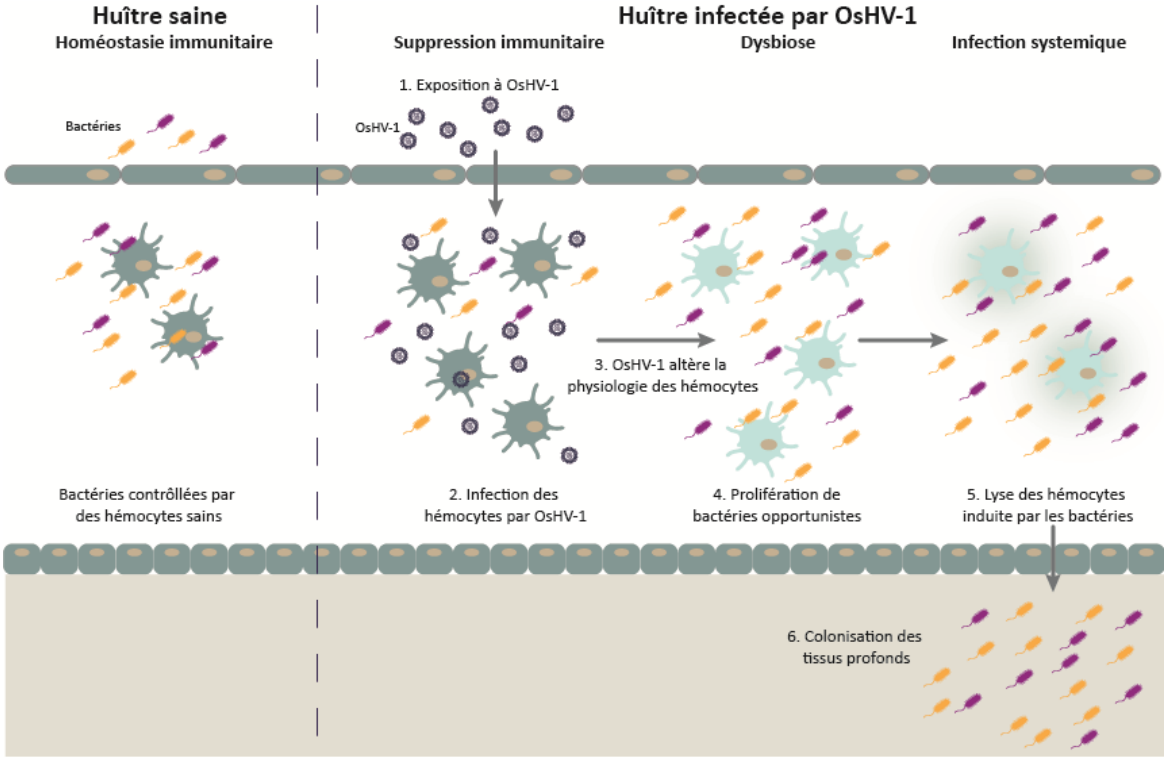


Figure 26 : Représentation schématique du POMS dans les branchies d'huître (d'après Petton *et al.*2021).

POMS est une maladie polymicrobienne. Les hôtes sont d'abord exposés au virus OsHV-1 (1.), qui infecte les hémocytes de l'hôte (2.) et altère leur physiologie (3.). Cette détérioration permet le développement de bactéries opportunistes (4.) qui provoquent la lyse des hémocytes (5.) et finissent par coloniser les tissus profonds de l'hôte (6.) induisant sa mort.

1.4.5 L'Ostreid *Herpesvirus* 1

1.4.5.1 Distribution géographique et espèces hôtes de l'OsHV-1

Actuellement, OsHV-1 est majoritairement détecté aux stades larvaires et juvéniles d'huîtres creuses, mais il a également été détecté dans une moindre mesure chez des adultes de *M. gigas* (Arzul *et al.,*2002, 2017; Dundon *et al.,*2011; Lipart & Renault, 2002) ou chez d'autres espèces hôtes telles que :

- les pétoncles *Azumapecten farreri* (Bai *et al.*,2019).
- les palourdes *Anadara broughtonii* (Xia *et al.*,2015), *Ruditapes philippinarum* (Arzul, Nicolas, *et al.*,2001; Arzul, Renault, & Lipart, 2001; Arzul, Renault, Lipart, *et al.*,2001; Renault & Arzul, 2001; Renault, 1998) et *Ruditapes decussatus* (Renault & Arzul, 2001; Renault, 1998).
- les coquilles Saint-Jacques *Pecten maximus* (Arzul, Nicolas, *et al.,*2001).

- les huîtres plates *Ostrea edulis* (Arzul, Renault, & Lipart, 2001; Arzul, Renault, Lipart, *et al.*,2001; Comps & Cochennec, 1993; Renault *et al.*,2000).
- les huîtres Portugaises *M. angulata* (Arzul, Renault, Lipart, *et al.*,2001).

Il a été détecté sur l'ensemble des zones de culture d'huîtres creuses mondiales (Figure 27, Tableau 1).

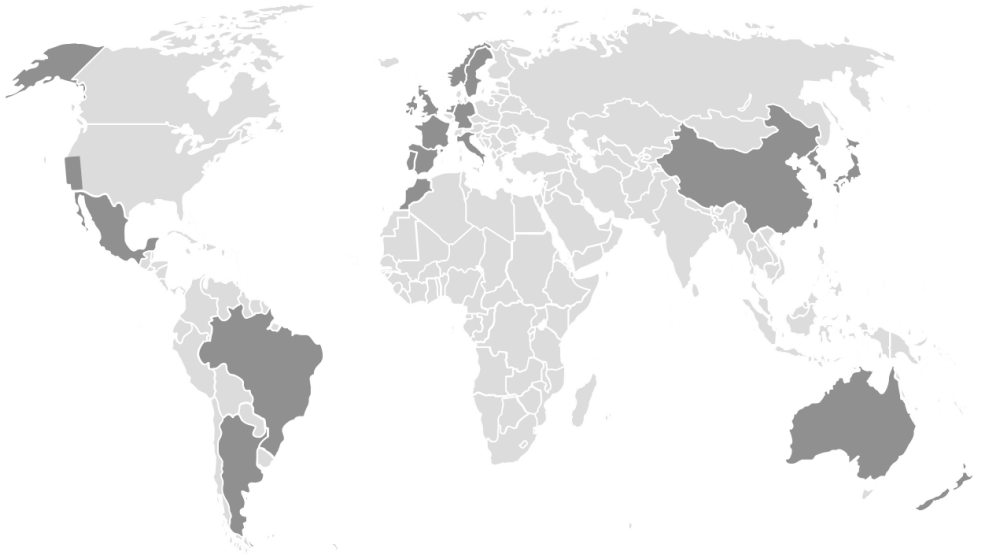


Figure 27 : Répartition mondiale du virus OsHV-1.

L'Ostreid *Herpesvirus* 1 a été détecté en Amérique, notamment en Californie, au Mexique, au Brésil, en Argentine et en Alaska. Il a pu être détecté dans différents bassins de productions Européens, comme en France, en Italie, au Portugal, en Espagne, en Suède et en Norvège, en Allemagne et aux Pays-Bas, et en Irlande, ainsi qu'en Angleterre. Il a également été détecté en Chine, en Corée, au Japon, à Taïwan, en Australie, en Nouvelle-Zélande et au Maroc (voir le tableau 1 pour les précisisions sur les origines géographiques et les espèces hôtes du virus OsHV-1).

Tableau 1 : Détection du virus OsHV-1 dans le monde et espèce hôte associée.

Continent	Lieu de détection	Espèce hôte	Publications
USA	Californie	*M. virginica*, *M. gigas*	Burge *et al.*, 2006, 2021 ; Farley *et al.*, 1972 ; Friedman *et al.*, 2005
	Mexique	*M. gigas, M. corteziensis*	Grijalva-Chon *et al.*, 2013 ; Martinez *et al.*, 2017 ; Vasquez Yeomans *et al.*, 2004
	Brésil	*M. gigas*, *M. brasiliana*	Mello *et al.*, 2018
	Argentine	*M. gigas*	Barbieri *et al.*, 2019
	Alaska	*M. gigas*	Meyers *et al.*, 2009
Europe	France	*M. gigas*, *R. philippinarum*, *R. decussatus*, *O. edulis*, *P. maximus*	Arzul, Nicolas, *et al.*, 2001 ; Arzul, Renault, & Lipart, 2001 ; Arzul, Renault, Lipart, *et al.*, 2001 ; Boudry *et al.*, 1998 ; Burioli *et al.*, 2017 ; Comps & Cochennec, 1993 ;

			Martenot *et al.*, 2012 ; Nicolas *et al.*, 1992 ; Renault, Cochennec, *et al.*, 1994 ; Renault, Deuff, *et al.*, 1994 ; Renault *et al.*, 2000 ; Segarra *et al.*, 2010
	Italie	*M. gigas*	Abbadi *et al.*, 2018 ; Dundon *et al.*, 2011
	Portugal	*M. angulata*	Batista *et al.*, 2015
	Espagne	*M. gigas, M. angulata*	Roque *et al.*, 2012
	Suède	*M. gigas*	Mortensen *et al.*, 2016
	Norvège	*M. gigas*	Mortensen *et al.*, 2016
	Allemagne	*M. gigas*	Gittenberger *et al.*, 2016
	Pays-Bas	*M. gigas*	Gittenberger *et al.*, 2016
	Irlande	*M. gigas*	Burioli *et al.*, 2017 ; Clegg *et al.*, 2014 ; S. A. Lynch *et al.*, 2012 ; Morrissey *et al.*, 2015
	Angleterre	*M. gigas*	Peeler *et al.*, 2012
Asie	Chine	*M. honkongensis*, *M. ariakensis*, *M. sikamea*, *M. gigas, A. irradians, R. philippinarum*	Bai *et al.*, 2015 ; Bai *et al.*, 2019 ; Moss *et al.*, 2007 ; Ren *et al.*, 2013 ; Xia *et al.*, 2015
	Corée	*M. honkongensis, M. ariakensis, M. sikamea, M. gigas*	Hwang *et al.*, 2013 ; Jee, Bo Young *et al.*, 2013 ; Moss *et al.*, 2007
	Japon	*M. honkongensis, M. ariakensis, M. sikamea, M. gigas*	Hwang *et al.*, 2013 ; Jee, Bo Young *et al.*, 2013 ; Moss *et al.*, 2007 ; Ren *et al.*, 2013 ; Shimahara *et al.*, 2012
	Taïwan	*M. angulata, M. gigas*	Boudry *et al.*, 1998
Océanie	Australie	*O. angasi, S. glomerata, M. gigas*	Cochennec *et al.*, 2011; Hine & Thorne, 1997; Jenkins *et al.*, 2013; S. Keeling *et al.*, 2014; Paul-Pont *et al.*, 2014
	Nouvelle-Zélande	*T. chilensis, M. gigas*	S. Keeling *et al.*, 2014 ; Webb *et al.*, 2007
Afrique	Maroc	*M. angulata*	Fabioux *et al.*, 2002

En 1995, à partir de larves infectées d'huître creuse *M. gigas* provenant d'une écloserie Française, des particules virales d'OsHV-1 ont été isolées et étudiées en cryo-miscroscopie, ce qui a permis de décrire la structure du virus OsHV-1 (Davison et al., 2005).

1.4.5.2 Structure de l'OsHV-1

La structure des particules virales d'OsHV-1 est similaire à celles des autres *Herpesvirus*. Des approches moléculaires, morphologiques et de microscopie électronique à transmission ont permis de mettre en évidence la présence d'une capside, d'un tégument et d'une enveloppe (Figure 28 ; Davison *et al.,*2005). L'enveloppe est constituée d'une double couche lipidique associée à la surface externe du tégument (Davison *et al.,*2009) et, est composée de différentes protéines de structure dont certaines sont impliquées dans la reconnaissance et la pénétration du virus dans la cellule hôte (Segarra, 2015; Segarra *et al.,*2014). Son ADN double brin est situé dans la capside et son génome linéaire de 207 kb a été séquencé par Davison *et al.* en 2005 (numéro d'accession : NC_005881.2).

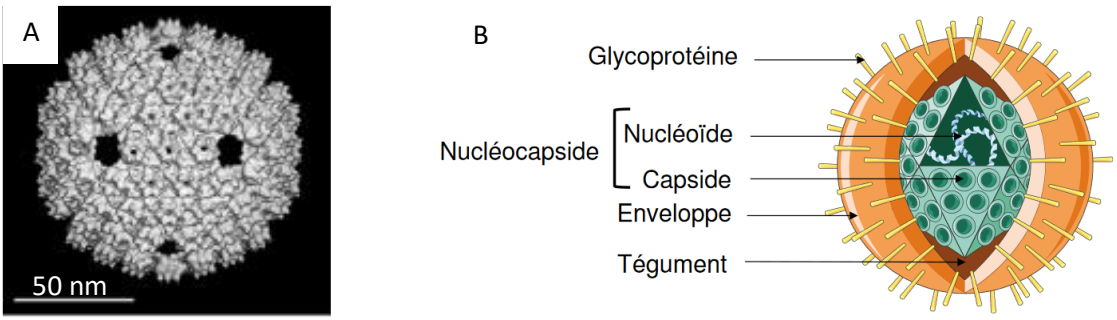


Figure 28 : Représentation du virus OsHV-1.

A) Reconstruction de la capside du virus OsHV-1 par cryo-microscopie (Davison *et al.,*2005) ; B) représentation schématique et hypothétique du virus OsHV-1, basée sur les connaissances des autres *herpesvirus* (Segarra, 2014).

1.4.5.3 Génome d'OsHV-1

Le génome d'OsHV-1 (Davison *et al.*, 2005 ; numéro d'accession : NC_005881.2) a une taille d'environ 207 kpb et est constitué de régions uniques (U_L : Unique Long, environ 167,8 kpb ; U_S : Unique Short, ~3,4 kpb et X, ~1,5 kpb) encadrées par des régions répétées terminales (TR_L : Terminal Repeat Long, ~7,6 kpb ; TR_S : Terminal Repeat Short, ~9,8 kpb) et des régions répétées inversées (IR_L : Inverted Repeat Long, ~7,6 kpb ; IR_S : Inverted Repeat Short ~9,8 kpb). Le pourcentage d'homologies entre TR_L/IR_L d'une part et TR_S/IR_S d'autre part est de 100 %. Ainsi, le génome d'OsHV1 peut être représenté schématiquement sous la forme TR_L-U_L-IR_L-X-IR_S-U_S-TR_S (Figure 29, Davison *et al.,*2005). La composition nucléotidique du génome d'OsHV-1 est de 38,7 % G+C (Davison *et al.*, 2005), et plusieurs régions, notamment autour des jonctions entre les régions répétées et les régions uniques, sont constituées par des homopolymères $(C)_n$ ou $(A)_n$, ce qui induit des fluctuations de G+C % (Figure 29). Au sein de la région U_L, un palindrome (*i.e.* séquence ADN identique qu'elle soit lue dans le sens 5' vers 3' ou

inversement) de 200 pb correspond à une structure en épingle à cheveux ou « stem-loop » qui peut se replier en deux ou trois parties (Figure 30, Davison *et al.*, 2005). Par homologie avec les *herpesvirus* de vertébrés, ce palindrome serait candidat pour être une origine de réplication ADN (Davison *et al.*, 2005).

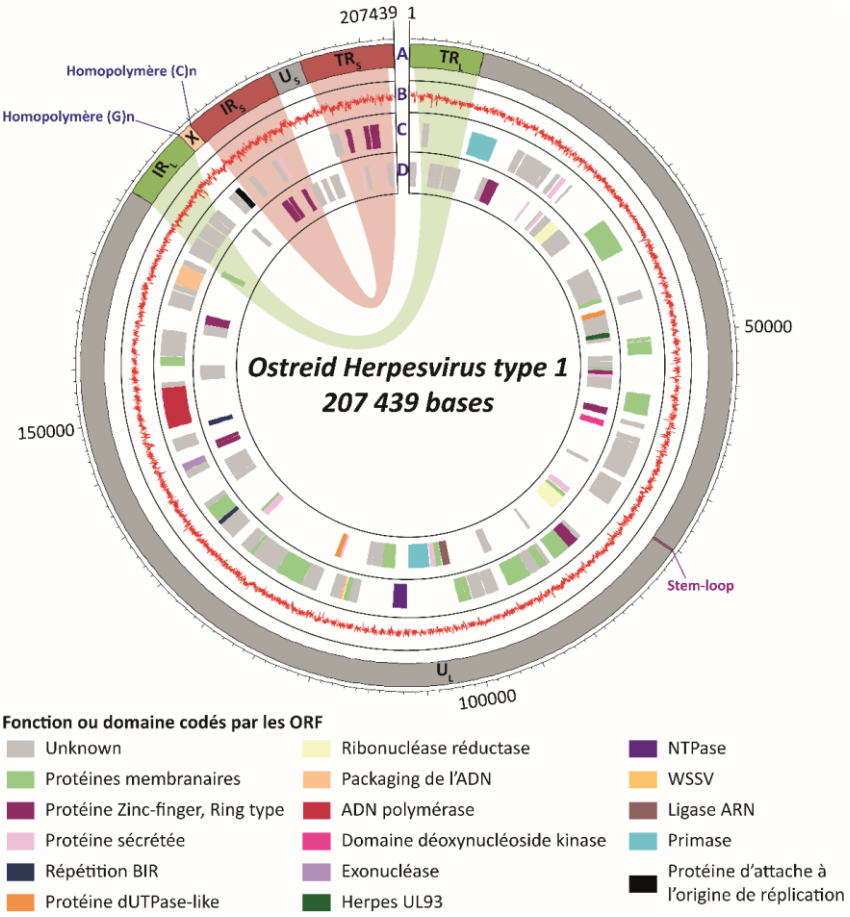


Figure 29 : Représentation schématique de l'architecture du génome du virus OsHV-1, des ORFs et du contenu en guanine et cytosine.

A) Architecture génomique du virus OsHV-1 découpé en régions uniques et en régions répétées. TR_L/IR_L : Terminal/Inverted Repeat Long ; U_L : Unique Long ; TR_S/IR_S : Terminal/Inverted Repeat Short ; U_S : Unique Short. B) Variations du contenu G+C le long du génome d'OsHV-1. C) Position des ORF sens et D) position des ORFs antisens présents dans le génome viral. La couleur des ORFs correspond à la fonction ou au domaine associé aux protéines qu'ils codent. Le lien vert, entre les régions TR_L et IR_L, et rouge, entre les régions TR_S et IR_S, représente les régions homologues.

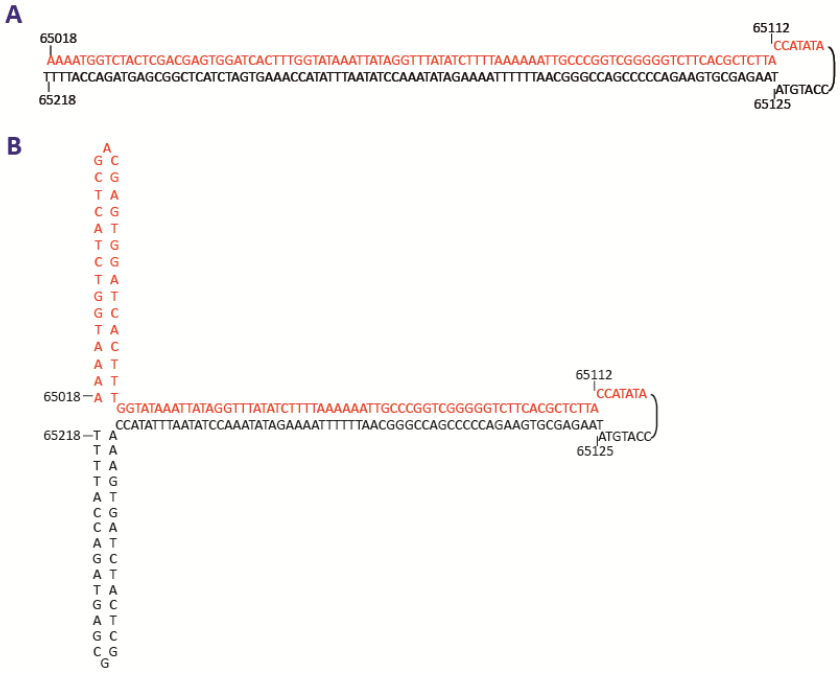


Figure 30 : Structure du palindrome présent dans le génome d'OsHV-1 entre les positions 65 018 et 65 218 pb (d'après Davison *et al.*2005).

A) Structure bipartite du palindrome. B) Structure tripartite du palindrome.

Quatre isomères génomiques (Encadré 14) majeurs (P : prototype, I_L : inversion long, I_S : inversion short, I_{LS} : inversion long and short) ont été mis en évidence et sont caractérisés par des inversions des régions U_L et U_S (Figure 31A), présents en proportion équimolaire dans les populations infectant l'huître creuse *M. gigas* (Davison *et al.*, 2005). Plus précisément, l'isomère de référence est la forme P dont l'arrangement des régions est celle usuellement utilisées comme génome de référence et décrite par Davison *et al.* (2005). Dans cet isomère, les régions U_L et U_S sont orientées dans le sens 5' vers 3' (Figure 31A, isomère P). En comparaison à cet isomère de référence, les isomères I_L et I_S contiennent respectivement une inversion de la région U_L et une inversion de la région U_S (Figure 31A, isomères I_L et I_S). Enfin, le dernier isomère, I_{LS}, est caractérisé par une inversion des deux régions uniques U_L et U_S (Figure 31A, isomère I_{LS}). D'autre part, pour chaque isomère majeur caractérisé, il existe trois isomères détectés en proportion mineure caractérisés par des variations de la région X (Figure 31B). Certaines particules virales contiennent *i)* une région X additionnelle à l'extrémité 5' du génome en plus de celle présente entre les région IR_L et IR_S (Figure 31B, isomère D_X), *ii)* une translocation du X à l'extrémité 5' du génome (Figure 31B, isomère T_X), *iii)* une absence de région X (Figure 31B, isomère N_X ; Davison *et al.*, 2005). Au total, 16 isomères du génome d'OsHV-1 ont pu être caractérisés (Davison *et al.*, 2005).

Encadré 14 : Isomères génomiques

Les isomères génomiques sont des structures moléculaires de l'ADN présentant une séquence nucléotidique identique, mais dont l'arrangement spatial diffère.
Les isomères génomiques sont particulièrement détectés dans les génomes contenant des répétitions inversées larges, comme les génomes d'*herpesvirus* par exemple. Les régions répétées favorisent les événements de recombinaison ce qui permet des réarrangements (inversions, duplications, translocations) au sein des génomes viraux qui co-infectent une cellule hôte.
A l'heure actuelle, la fonction de ces isomères est inconnue mais ces altérations dans la structure de l'ADN pourraient influencer de manière significative l'expression génique, la réplication de l'ADN et la reconnaissance des protéines, jouant ainsi un rôle crucial dans la régulation de la stabilité du génome et l'évolution des organismes.

Exemple d'isomère pouvant être observé dans un génome viral. En bleue est représentée la région unique longue (UL) qui a subi une isomérisation et résulte en une inversion de la région. En gris sont représentée les autres régions génomiques qui ne sont pas impactées par l'isomérisation (dans cet exemple).

(Chaitanya, 2019 ; Mahiet *et al.,* 2012 ; Roizman, 1979)

Le génome de référence d'OsHV-1 possède un total de 136 Open Reading Frame (ORF), dont 24 sont dupliqués dans les régions répétées, et 78% de son génome est codant (Davison *et al.,*2005). Actuellement, seuls 60 ORFs (44 %) sont fonctionnellement annotés et ont une fonction ou un domaine associé (Figure 32 et 30). Parmi ces ORF, on retrouve l'ORF109, codant pour une ATPase, caractéristique des *Herpesvirus* et l'ORF110 qui code pour l'ADN polymérase très conservée chez tous les *Herpesvirus*. Cet ORF est le seul qui présente une forte homologie avec ceux des autres *Herpesvirus* de vertébrés (Davison *et al.,*2005). La plupart des ORFs qui possèdent une annotation, codent pour des fonctions essentielles pour le virus, telles que, la réplication et la transcription ADN, la formation de la capside, ou encore la facilitation de l'entrée des particules virales dans les cellules d'huître. Vingt-deux ORFs codent pour des glycoprotéines structurales transmembranaires ou des protéines membranaires. Chez les virus ces protéines sont principalement impliquées dans la reconnaissance des récepteurs de l'hôte, puis dans la fusion avec la membrane des cellules hôtes, permettant ainsi l'entrée de la particule virale dans la cellule hôte afin d'initier son cycle de réplication (Banerjee & Mukhopadhyay, 2016). Ces protéines sont donc cruciales pour l'infection (Banerjee & Mukhopadhyay, 2016). Une autre protéine importante retrouvée dans le génome du virus OsHV-1 est la protéine Zinc-finger, Ring type qui est un facteur de transcription impliqué dans de nombreux processus cellulaires et notamment dans la régulation des gènes (X. Li et al., 2022). Dix ORF sont annotés comme codant pour des protéines sécrétées, et quatre codant pour des protéines d'inhibition de l'apoptose

(Baculovirus inhibitor of apoptosis protein repeat), c'est-à-dire pour inhiber le système immunitaire de l'huître (Burioli *et al.*, 2017; Davison *et al.*, 2005). Enfin, d'autres protéines codant pour des primases, des ligases, des exonucléases ou des déoxynucléoside kinase sont également retrouvées dans le génome d'OsHV-1.

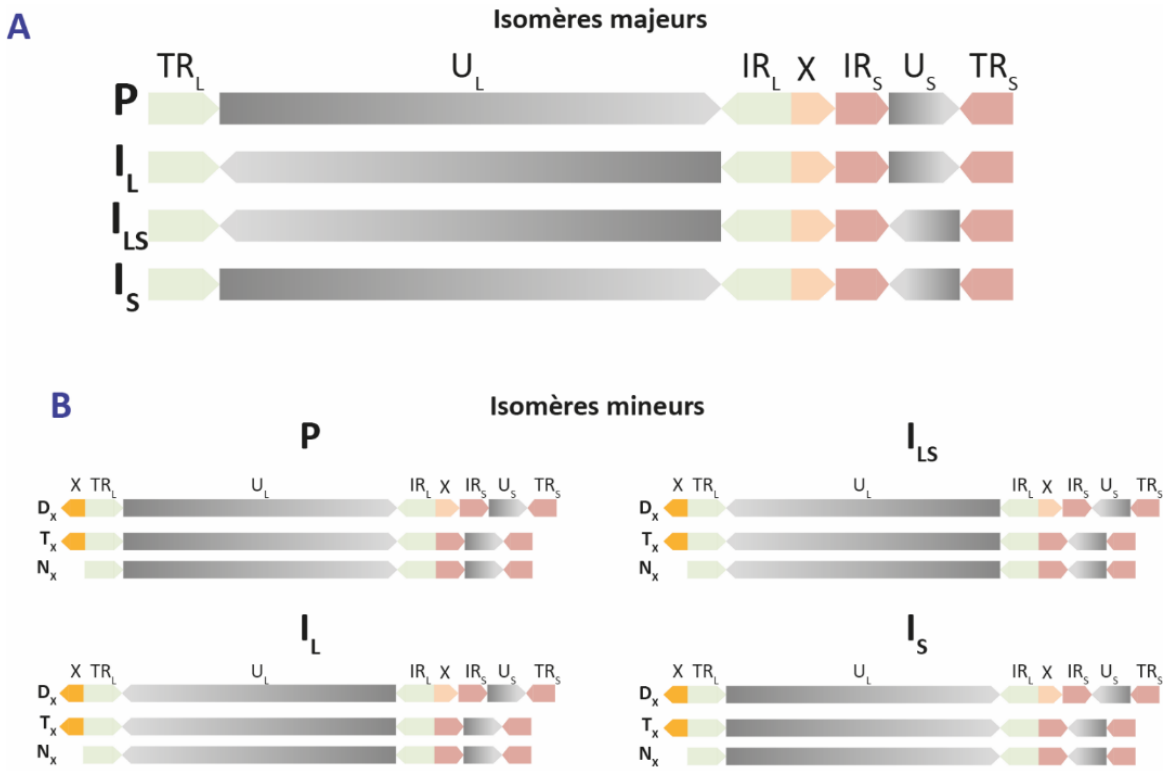


Figure 31 : Représentation schématique des 16 isomères génomiques du virus OsHV-1 (d'après Davison *et al.*2005).

A) Le génome viral est organisé en région unique longue et courte (U_L et U_S en dégradé de gris), en région répétées terminales ou inversées longues (TR_L et IR_L en vert) et courtes (TR_S et IR_S en rouge). Une dernière région unique nommée X est située au centre du génome (en orange clair).
Les particules virales peuvent contenir quatre isomères génomiques majeurs en fonction de l'orientation des régions U_L et U_S :

P (Prototype) : structure de référence du génome OsHV-1 organisé sous la forme TR_L-U_L-IR_L-X-IR_S-U_S-TR_S.
I_L (Inverse Long) : même structure que P mais contient une inversion de la région U_L.
I_{LS} (Inverse Long and Short) : même structure que P mais contient une inversion des régions U_L et U_S.
I_S (Inverse Short) : même structure que P mais contient une inversion de la région U_S.

B) Pour chaque isomère génomique majeur (P, I_L, I_{LS}, I_S), il existe également trois isomères mineurs en fonction de la position de la région X (représentée en orange) :

D_X (Duplication du X) : isomère qui contient une duplication inversée du X en début de génome.
T_X (Translocation du X) : isomère qui contient une translocation inversée du X en début de génome.
N_X (No X) : isomère qui porte une délétion du X.

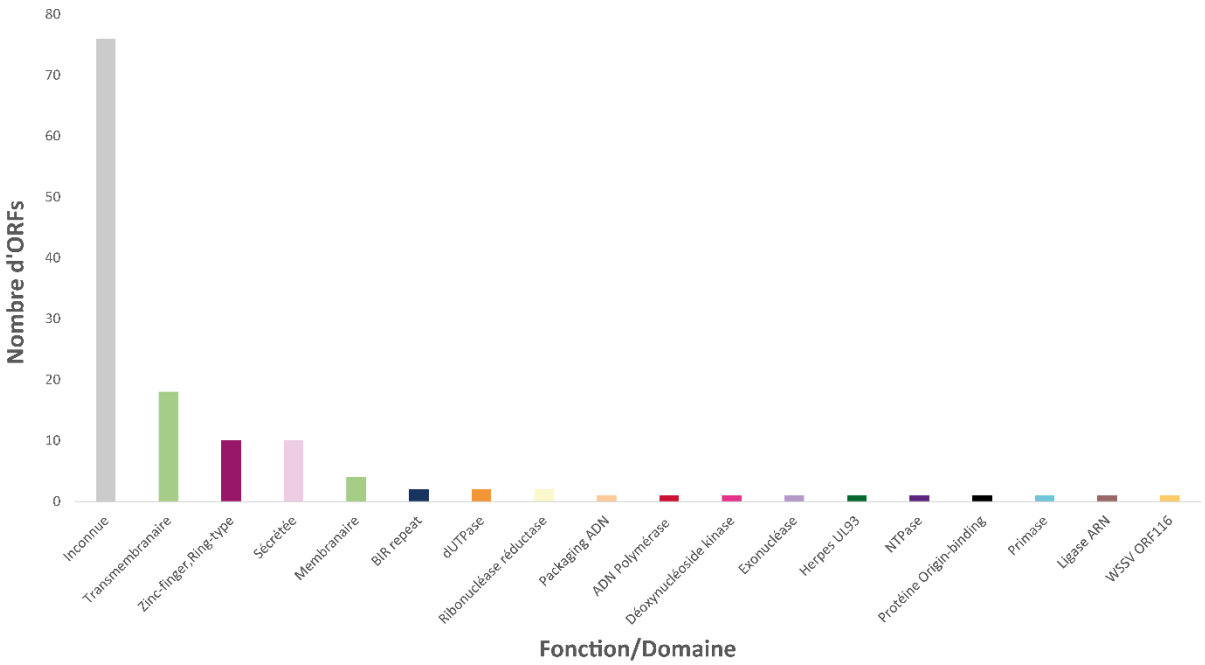


Figure 32 : Nombre d'ORFs du génome d'OsHV-1 codant pour une fonction ou un domaine (d'après Davison *et al.*2005 et Burioli *et al.*, 2017).

1.4.5.4 Cycle de réplication du virus OsHV-1

Aucune lignée cellulaire permettant de multiplier OsHV-1 n'a été produite à ce jour, ce qui complique l'étude de son cycle de réplication. Cependant, la microscopie à transmission a permis d'établir que son cycle de réplication dans la cellule hôte est très proche de celui des *Herpesvirus* (Segarra, 2015). Le cycle du virus OsHV-1 est constitué de deux phases : une phase productive ou lytique et une phase de persistance ou latente (Segarra, 2015).

La phase productive correspond à la réactivation du virus à la suite d'une stimulation extérieure (phase lytique), conduisant à la production puis la libération de nouveaux virus (Figure 33). Tout d'abord, le virus se fixe sur les récepteurs membranaires de la cellule hôte (Figure 33, étape 1). Des études suggèrent l'implication de protéines de type glypican (Renault *et al.*, 2011; Segarra, *et al.*, 2014) et intégrine (Jouaux *et al.*, 2013) dans la reconnaissance d'OsHV-1 par les cellules de son hôte. Suite à l'adsorption du virus sur la membrane cellulaire, les membranes virales et cellulaires vont fusionner, pour que le virus pénètre dans le cytoplasme de la cellule (Figure 33, étape 2), et perde son enveloppe (Figure 33, étape 3) (Segarra, 2015). Les protéines structurales de la capside virale vont alors commencer à être dégradées pour libérer le nucléoïde dans le cytoplasme avant qu'il soit acheminé vers les pores de la membrane nucléaire grâce aux microtubules (Figure 33, étape 4) (Segarra, 2015).

La réplication de l'ADN viral a ensuite lieu dans le noyau des cellules hôtes (Davison *et al.,*2009). L'ADN viral linéaire se circularise dans le noyau de la cellule hôte et la réplication est initiée à partir d'une ou plusieurs origines suivant un mécanisme de « rolling-circle » (Figure 33, étape 5) (Segarra, 2015). En parallèle, une nouvelle capside est assemblée (Figure 33, étape 6) et, l'ADN viral est conditionné à l'intérieur de celle-ci pour former une nucléocapside (Figure 33, étape 7) (Segarra, 2015). La nucléocapside s'attache alors à la membrane interne du noyau, ce qui lui permet d'acquérir le tégument (Figure 33, étape 8) (Segarra, 2015). La nucléocapside est alors libérée dans le cytoplasme (Figure 33, étape 9) et la nouvelle enveloppe se forme autour de la capside (Figure 33, étape 10). La particule virale mature (Figure 33, étape 11) est ensuite libérée par exocytose (Figure 33, étape 12) (Segarra, 2015).

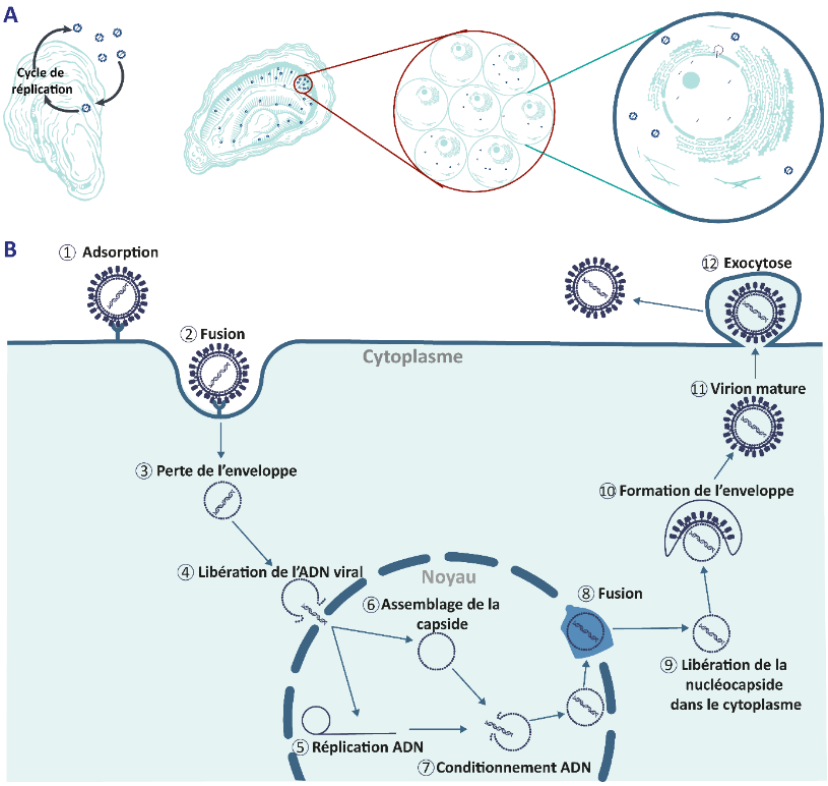


Figure 33 : Phase lytique du cycle de réplication supposé du virus OsHV-1 (d'après Le Deuff *et al.*, 1994 ; Renault *et al.*, 1994 ; Segarra 2015).

A) Cycle du virus OsHV-1 dans les cellules de l'huître. B) Phase lytique du cycle de réplication du virus OsHV-1 découpé en 12 étapes : 1)Adsorption de la particule virale sur les récepteurs de la cellule hôte grâce aux glycoprotéines transmembranaires, 2) fusion des membranes virale et cellulaire et pénétration de la particule virale dans le cytoplasme de la cellule hôte, 3) perte de l'enveloppe, 4) dégradation des protéines de la capside virale et libération de l'ADN viral dans le noyau de la cellule hôte par le pore nucléaire, 5) circularisation de l'ADN viral et réplication par « rolling-circle », 6) assemblage de la capside, 7) conditionnement de l'ADN dans la nouvelle capside, 8) attachement de la nucléocapside à la membrane nucléaire interne et acquisition du tégument, 9) libération de la nucléocapside dans le cytoplasme, 10) formation de l'enveloppe, 11) particule virale mature et 12) libération des particules virales par exocytose.

La phase de persistance correspond au maintien d'un virus dans certaines cellules hôte, parfois caractérisée par des phases d'infections lytiques silencieuses qui ne tuent pas et ne produisent pas de dommages conséquents dans les cellules de l'hôte (Boldogh *et al.*, 1996). La persistance peut prendre la forme d'infections latentes, chroniques ou ralenties (Boldogh *et al.*, 1996). Chez OsHV-1 la phase de persistance n'a jamais été caractérisée mais est suspectée car le virus OsHV-1 est détecté et associé tous les ans au printemps et en été à des mortalités dans les bassins de production ostréicole, sans être détecté l'hiver. Chez les autres *Herpesvirus*, notamment l'*Herpesvirus humain* 6 (HHV-6) et le virus Epstein-Barr (EBV), la phase de persistance est caractérisée par une infection latente, c'est-à-dire qu'ils échappent au système immunitaire de l'hôte après leur inactivation et leur entrée dans une phase de dormance. Chez les *herpesvirus*, la forme génomique principale observée lors de la latence est la forme d'épisome (*i.e.* ADN viral circulaire observée par exemple chez EBV), mais parfois, le génome viral peut également être intégré au sein des parties télomériques des chromosomes de leur hôte (*e.g.* HHV-6 intégré dans les chromosomes humains, Figure 34 ; Cohen, 2020; Morissette & Flamand, 2010; Serquiña & Ziegelbauer, 2017).

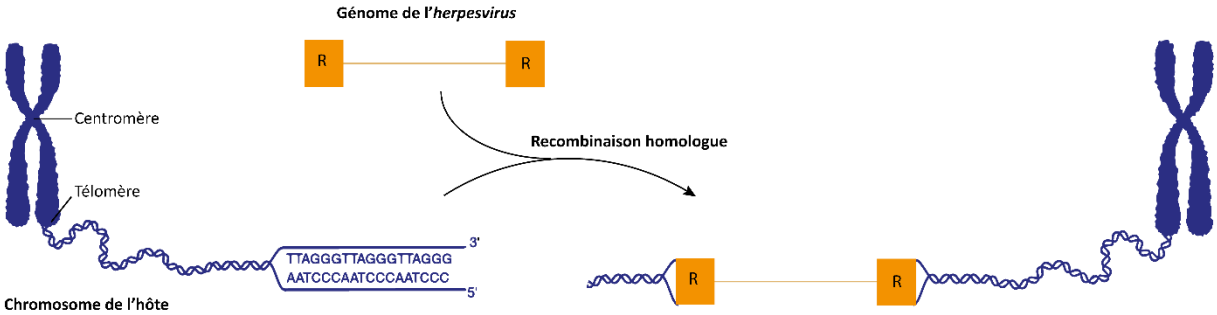


Figure 34 : Exemple simplifié d'intégration chromosomique d'un génome d'*herpesvirus* au sein des parties télomériques des chromosomes de l'hôte après recombinaison homologue facilitée par les régions répétées (R) du génome virale (d'après Morissette & Flamand, 2010).

1.4.5.5 Études épidémiologiques et analyse de diversité appliquées au virus OsHV-1

Depuis 1992, dans le cadre de la Directive Européenne 2006/88/CE, l'IFREMER sous la tutelle du ministère de l'Agriculture et de l'Alimentation au travers de la Direction Générale de l'Alimentation (DGAL), a mis en place un réseau de surveillance pour la santé des mollusques marins à l'échelle Française : REPAMO (REseau de Pathologie des Mollusques). Ce réseau a initialement été mis en place pour :

i) surveiller l'état de santé des mollusques et dresser une image de référence,

ii) prévenir l'introduction et la propagation d'agents infectieux notamment pour les maladies réglementées et

iii) surveiller l'évolution des agents pathogènes déjà présents dans le milieu (François *et al.,*2013).

Dans le cadre de ce réseau, les mortalités anormales de mollusques ont été régulièrement répertoriées et notamment dans le cas de l'huître du Pacifique infectée par le virus OsHV-1. Depuis décembre 2022, ce réseau a été repris par le comité national de conchyliculture (CNC) et n'est plus sous la tutelle de l'IFREMER. Les échantillons collectés continueront tout de même d'être stockés dans les collections du Laboratoire National de Référence (LNR, IFREMER). Seules les analyses de détection des agents pathogènes réglementés au niveau Européen seront financées par l'état (Canier *et al.*, 2019)

Les premières études épidémiologiques reposaient sur l'étude des cas de mortalité déclarées, au sein du réseau REPAMO, associées à la détection du virus OsHV-1, rapidement, suppléaient par l'apport de la biologie moléculaire et plus récemment du séquençage. Si très peu d'études de diversité ont été réalisées sur le virus OsHV-1 à l'échelle du génome, plusieurs variant du virus OsHV-1 ont néanmoins déjà été identifiés, grâce à ces méthodes, chez différentes espèces hôtes et dans différentes régions du monde (Abbadi *et al.,*2018; Arzul, Nicolas, *et al.,*2001; Arzul, Renault, & Lipart, 2001; Arzul, Renault, Lipart, *et al.,*2001; Bai *et al.,*2019; Burioli *et al.,*2017; Friedman *et al.,*2005; Moss *et al.,*2007; Ren *et al.,*2013; Renault & Arzul, 2001; Renault *et al.,*2012; Xia *et al.,*2015).

Les premières études pionnières dans l'analyse de la diversité chez OsHV-1 sont basées uniquement sur une ou plusieurs régions génomiques et utilisent le séquençage Sanger (Arzul, Nicolas, *et al.,* 2001; Renault *et al.,*2012). Le premier variant mis en évidence a été détecté chez des larves de palourdes et d'huîtres creuses en 1997 (Arzul, Renault, Lipart, *et al.,* 2001) puis chez des larves de coquilles Saint Jacques en 2001 (Arzul, Nicolas, *et al.,* 2001). Ce variant a été nommé OsHV-1 Var et il est caractérisé par une délétion d'au moins 2,8 kb dans la région TRL en comparaison du génome de référence du virus OsHV-1 (Arzul, Nicolas, *et al.,*2001). Ces variations génétiques ont été mises en évidence par une approche de PCR dans les régions A (région U_L, en amont de l'ORF 83), B (région U_L, ORF 99) et C (région TR_L/IR_L), couplée à une digestion par des enzymes de restriction et enfin par le séquençage des amplicons PCR (Figure 35). Ce variant n'a été que très rarement détecté en France.

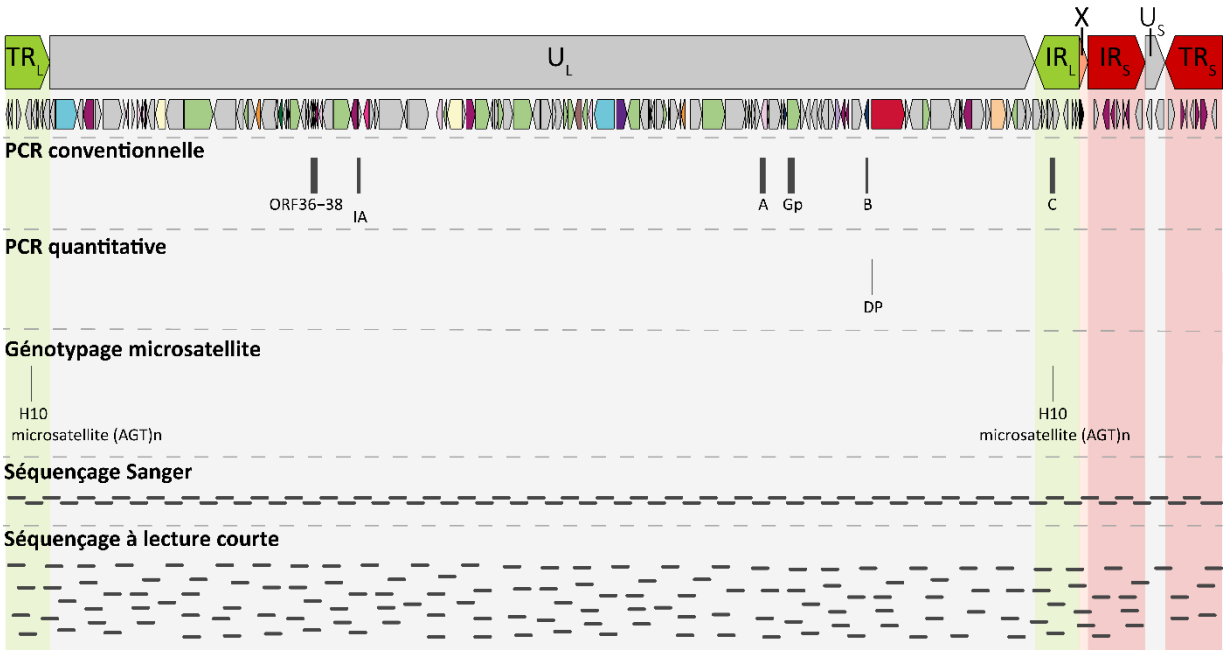


Figure 35 : Amorces PCR développées pour l'étude du virus OsHV-1 et du variant µVar.

Plus récemment, en 2008, des mortalités massives de naissains d'huîtres creuses ont été observées en Europe. Ces mortalités ont été associées à l'émergence d'un nouveau variant détecté pour la première fois en France (Martenot *et al.,*2011; Renault *et al.,*2012; Segarra *et al.,*2010), puis en Irlande (Peeler *et al.,*2012) et en Grande-Bretagne (Lynch *et al.,*2012). Ce variant a été nommé µVar, en raison des variations génétiques détectées au niveau d'un microsatellite en amont de l'ORF4, dans l'ORF4 et dans les ORF42/43 (OIE, 2019). Ces variants ont également été caractérisés par un séquençage Sanger d'amplicons. La variabilité génétique mise en évidence pour ces variant a été détecté dans les zones C et IA (ORF 43 codant pour un Inhibiteur d'apoptose) (Segarra *et al.,*2010). Depuis 2008, le génotype µVar est le principal génotype détecté dans les bassins ostréicoles.

Jusqu'en 2017, la détection et l'étude de la diversité des OsHV-1 a donc été réalisée par des approches ciblant des régions génomiques de moins de 1000 pb (Figure 35, Renault *et al.,*2012, 2014). En 2017, Burioli *et al.* ont publiés deux génomes micro-variants du virus OsHV-1 : OsHV-1 µVarA (numéro d'accession : KY242785.1) et µVarB (numéro d'accession : KY271630.1). La publication de ces génomes a permis les premières comparaisons de génome complet, entre le génome de référence du virus OsHV-1 publié en 2005 par Davison *et al.* (numéro d'accession : NC_005881.2) et ces deux micro-variants. Par comparaison, il a été mis en évidence, que le génome des micro-variants a accumulé de larges délétions de matériel génétique (6546 pb au total), notamment dans l'ORF11 d'une taille de 1385 pb et codant pour une protéine transmembranaire et dans les ORF 62 et 63 d'une taille cumulée de 3549 pb et qui codent pour une protéine sécrétée et protéine membranaire de classe I

respectivement. Burioli *et al.* (2017) posent l'hypothèse que l'augmentation de virulence des micro-variants pourrait être expliquée par ces nombreuses mutations apparaissant dans les régions codantes et impactant parfois de larges régions génomiques.

Récemment, de nouveaux génomes de variants du virus OsHV-1 ont été publiés dans les bases de données publiques (Tableau S1). C'est le cas d'OsHV-1 PT (numéro d'accession : MG561751.2 ; Abbadi *et al.,*2018), séquencé à partir d'échantillons d'ADN de naissains d'huîtres creuses *M. gigas* infectées en Italie, d'OsHV-1 ZK118 (numéro d'accession : MF509813.1 ; Bai *et al.,*2018) séquencé à partir d'échantillons ADN de pétoncles *C. farreri* moribonds et enfin du variant OsHV-1 SB (numéro d'accession : KP412538.1 ; Xia *et al.,*2015) séquencé à partir d'échantillons ADN de palourdes *S. broughtonii*.

Avant le début de ce projet de thèse, une seule étude menée par Delmotte *et al.* (2020) s'intéressait à la diversité du virus OsHV-1. Les auteurs ont utilisé une approche d'appel de variants dans des jeux de données RNA-seq produits à partir d'ARN d'huîtres, provenant de quatre fonds génétiques différents, infectés par le virus OsHV-1 en milieu naturel dans la baie de Logonna-Daoulas et dans l'étang de Thau. L'appel de variant a été réalisé en alignant les séquences des librairies contre le génome d'OsHV-1 µVarA. Plus de 400 positions variables ont été détectées, dont 268 spécifiques aux échantillons provenant de l'Atlantique et, 59 spécifiques aux échantillons provenant de la Méditerranée, mettant ainsi en évidence deux populations virales différentes. Les auteurs de cette étude mettent également en évidence qu'au sein d'un même environnement, les différents fonds génétiques d'huitres ne sont pas infectés par les mêmes populations virales. Cependant, différents points sont problématiques dans cet article. Tout d'abord, ils utilisent des génomes de référence publiés dans les bases de données, mais pas adaptées aux librairies de séquençage utilisées dans leur étude, créant un biais de détection de SNP qui n'existent peut-être pas réellement dans les librairies. Les librairies utilisées par Delmotte *et al.* (2020) ont été produites à partir d'ARNm et non d'ADN. Les parties non codantes ne sont donc pas représentées, ce qui entraine un biais de représentativité de la diversité virale observée. De plus, elles manquent de profondeur de séquençage pour réaliser les analyses de diversité. Enfin, cette étude a été menée sur un pool de séquences issues des librairies produites pour un même site d'échantillonnage (*i.e.* un pool Atlantique et un pool Méditerranée), toute l'information de diversité intra- et inter-individuelle au sein d'un site ostréicole est donc perdue.

Ces premières analyses ont montré l'émergence de nouveaux génotypes du virus OsHV-1 dans différentes régions du monde et au sein de différents hôtes. Elles ont servies de point de départ pour mon projet de thèse. En 2019, seulement six génomes d'OsHV-1 ou de ses variants étaient disponibles. Il était donc important de générer de nouvelles séquences pour mieux comprendre la structuration

des populations d'OsHV-1, son cycle de vie, les interactions avec son hôte ou encore son adaptation et son évolution.

1.5 Problématiques de thèse

Depuis la première détection du virus OsHV-1 en 1970 chez *M. virginica* en Californie, plusieurs variants ont été détectés. Certains se sont propagés dans le monde et au sein de différentes espèces hôtes. Ces variants ont été caractérisés par des approches ciblées en biologie moléculaire et beaucoup plus rarement par le séquençage de génomes complets. En 2019, seulement six génomes du virus OsHV-1 étaient disponibles dans les bases de données publiques. Les premières études épidémiologiques prenaient en compte principalement des données de cas et des données moléculaires ciblées telles que les microsatellites ou des séquences d'amplicons sur certains ORF (*i.e.* ORF 4 à 5, ORF 36 à 38, ORF 42 à 43, ORF 82 à 83, ORF 88, ORF 99). Plus récemment, des études de diversité et de génomique comparative menées à l'échelle du génome complet ont montrées des différences significatives entre les génotypes échantillonnés dans le monde et chez différentes espèces hôtes. Du fait de l'insuffisance données génomiques à l'échelle du génome du virus OsHV-1, aucune étude de diversité à l'échelle du génome portant sur un nombre conséquent d'échantillons n'avait été menée. L'insuffisance des données génomiques disponibles pour le virus OsHV-1 explique la subsistance de nombreuses incertitudes quant aux mécanismes impliqués dans son évolution.

Dans ce contexte, les travaux présentés dans ce manuscrit ont pour objectifs de répondre aux questions suivantes :

i) Comment séquencer le virus OsHV-1 et analyser sa diversité ?

ii) Les lignées virales présentes chez *M. gigas* et *O. edulis* sont-elles les mêmes et les événements de transmission inter-espèce sont-ils fréquents ?

iii) Quelle est la diversité spatio-temporelle du virus OsHV-1 chez *M. gigas* ?

iv) Comment le virus OsHV-1 évolue en fonction du fond génétique de son hôte, *M.gigas* ?

Ainsi, le second chapitre de ce manuscrit s'intéresse au développement méthodologique d'approches permettant d'une part la purification et le séquençage du virus, et d'autre part le développement d'un pipeline bio-informatique. Ces développements sont nécessaires pour une caractérisation plus précise du génome viral et de la diversité génétique observée au sein et entre différentes populations virales. L'impossibilité de produire des cultures cellulaires pour propager le virus rend difficile l'étude de son génome du fait de la contamination des échantillons par l'ADN de l'hôte. Par conséquent, jusqu'en 2019, la diversité du virus OsHV-1 n'avait été étudiée qu'au travers d'approches ciblées (*i.e.* PCR, microsatellites, séquençage d'amplicons) se focalisant sur des régions génomiques spécifiques du

génome d'OsHV-1. Les informations résultant de ces approches étaient donc incomplètes. L'objectif de cette première partie était de développer un pipeline bioinformatique qui permette d'assembler des génomes du virus OsHV-1 de bonne qualité (*i.e.* sans contaminant et représentatifs de la lignée virale de chaque échantillon) à partir de séquences courtes d'ADN (*i.e.* 150 pb) générées *via* la technologie Illumina afin d'étudier la diversité génomique du virus à différentes échelles (*i.e.* variations structurales, diversité intra- et inter-individuelle).

Le troisième chapitre de ce manuscrit traitera de la différenciation génétique d'OsHV-1 chez deux espèces d'huîtres, l'huître creuse du Pacifique *M. gigas* et l'huître plate Européenne *O. edulis*. Le virus OsHV-1 a déjà été détecté au sein de plusieurs espèces hôtes secondaires comme le pétoncle *A. farreri,* les palourdes *A. broughtonii*, *R. philippinarum* et *R. decussatus,* la coquille Saint-Jacques *P. maximus*, l'huître plate *O. edulis*, ou encore l'huître Portugaise *C. angulata*. Ces différentes détections laissent penser que des événements de transmission inter-espèce du virus OsHV-1 peuvent survenir naturellement dans les populations d'hôtes. D'autre part, les *herpesvirus* sont connus pour co-évoluer avec leurs hôtes, conduisant à des lignées très différenciées et une diversité spécifique à l'hôte infecté (Kaján *et al.*, 2020; McGeoch *et al.*, 2006), l'implication des hôtes « alternatifs » dans le cycle viral d'OsHV-1 ainsi que la caractérisation de la diversité du virus au sein de ses différents hôtes n'ont jamais été étudiées. Dans ce contexte, il était important de caractériser les lignées virales infectant différentes espèces hôtes. Au vu des capacités de co-évolution des *herpesvirus*, l'hypothèse de départ était que les populations virales échantillonnées chez les deux espèces hôtes avaient probablement évoluées indépendamment et seraient aujourd'hui totalement différenciées. A partir d'échantillons prélevés entre 2005 et 2010 chez des huîtres creuses et des huîtres plates infectées par OsHV-1 en France et en Suède puis séquencés par une approche short-read, des génomes complets du virus OsHV-1 ont pu être reconstruits. La diversité des populations virales dans chacune des deux espèces a ensuite pu être analysée à l'échelle du consensus, notamment par l'étude des variations structurales et l'utilisation d'indices de diversité, mais également à l'échelle intra-individuelle par l'analyse des variations ponctuelles (*i.e.* SNP, InDels). Ces résultats ont été complétés par des approches de phylodynamique, et particulièrement une reconstruction des états ancestraux en trait discret pour l'espèce hôte, qui ont permis de retracer l'histoire évolutive du virus OsHV-1 au sein et entre ces deux espèces.

Le quatrième chapitre de ma thèse s'intéresse à la caractérisation de la diversité spatio-temporelle du virus OsHV-1. Comme évoqué précédemment, la diversité de ce virus n'a été étudié qu'au travers d'approches ciblées, et on compte deux études portant sur sa diversité à l'échelle du génome complet. La dynamique des populations virales mondiales n'a jamais été étudiée et aucune étude de génomique comparative prenant en compte des échantillons anciens et de diverses origines n'a été conduite. L'objectif ici est de caractériser ces différentes populations virales au niveau des quatre plus

importants bassins de production ostréicoles Français, mais également de comparer ces populations virales avec celles d'autres bassins de production mondiaux au cours du temps, et de retracer les mouvements des populations virales entre les différents bassins ostréicoles. L'hypothèse initiale pour cette partie était que la structuration des populations virales est influencée par la localisation ainsi que l'année de collecte de l'échantillon. Dans ce contexte, des collaborations avec des laboratoires étrangers ont été mises en place pour collecter des échantillons viraux originaires de plusieurs bassins de production dans le monde. En parallèle, tous les ans depuis 2019, des échantillons d'huîtres infectées ont été collectés, en France, en période estivale lors des épisodes de mortalités massives des huîtres. En combinant des approches de génomique comparative et de phylodynamique, la structuration des populations du virus OsHV-1 dans le monde et au cours du temps a pu être établie et mise en lien avec les pratiques culturales.

Enfin, le dernier chapitre de ma thèse portera sur la caractérisation des mécanismes évolutifs du virus OsHV-1 au travers d'une approche expérimentale. Les interactions entre le virus et son hôte sont encore très mal connues. A l'heure actuelle, une seule étude réalisée par Delmotte *et al.* (2020) décrit les prémices des mécanismes évolutifs du virus OsHV-1. Celle-ci porte uniquement sur les protéines membranaires du virus qui semblent évoluer sous l'action de la sélection naturelle. Les études évolutives sont pourtant essentielles pour une meilleure compréhension de l'émergence, de la dispersion et de la pathogénicité de nouveaux génotypes viraux. Au vu des travaux précédemment menés, l'hypothèse sur laquelle reposent les expérimentations était que le génome du virus OsHV-1 évolue sous l'effet de la sélection naturelle opérée par l'hôte infecté. Dans cette partie, la première évolution expérimentale sur le modèle huître-OsHV-1 a été mise en place, *in vivo,* puisqu'aucune culture cellulaire n'est disponible actuellement. Des infections expérimentales successives par cohabitation ont été réalisées et les lignées virales ont été séquencées toutes les cinq générations. Ces séquençages ont permis de suivre l'évolution des fréquences alléliques au sein de chaque lignée d'huîtres et au cours des générations. Les résultats ont montré que le fond génétique des huîtres infectées joue un rôle important dans l'évolution du génome viral.

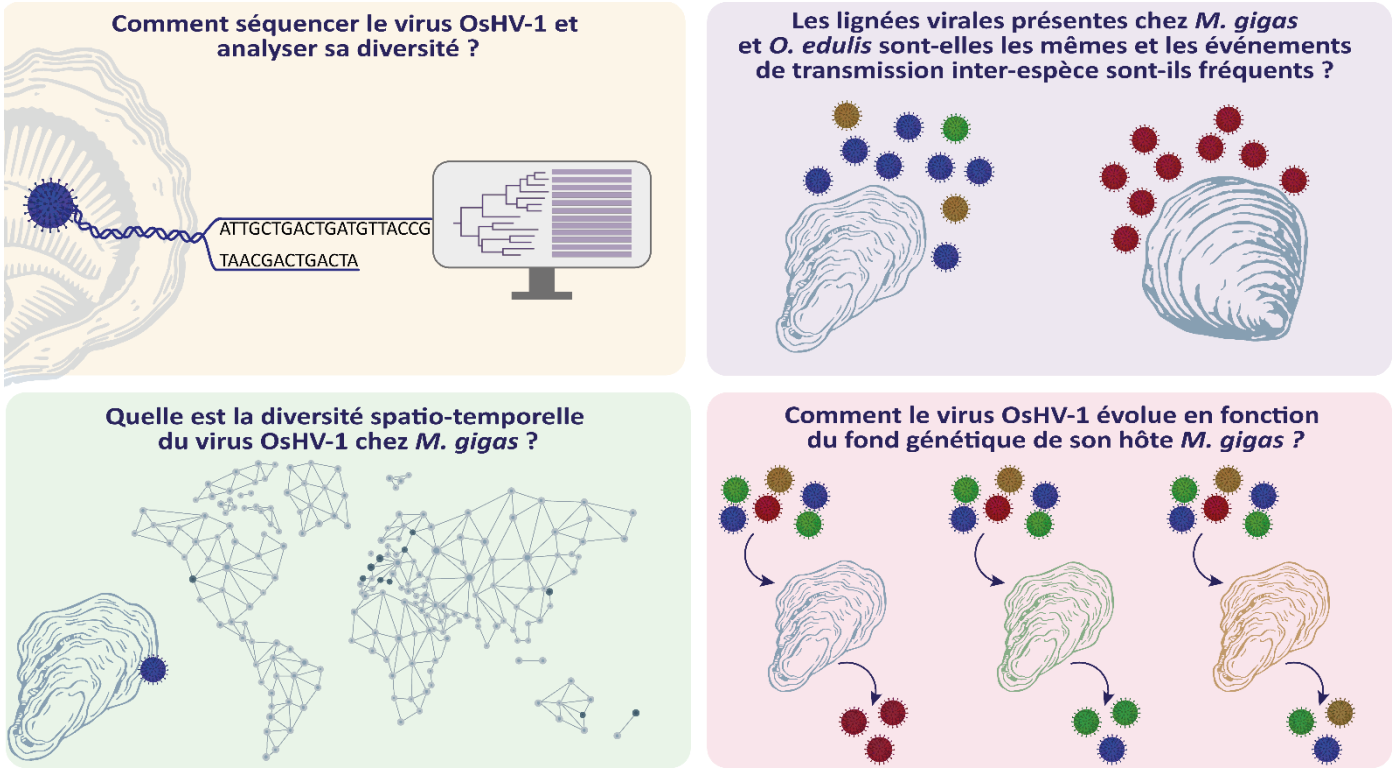


Figure 36 : Représentation graphique des principaux objectifs de ma thèse.

Chapitre 2 - Développement méthodologique pour l'étude génomique du virus OsHV-1

2.1 Résumé de l'article 1

Comme mentionné précédemment, le virus OsHV-1 n'est pas cultivable, ce qui signifie que sa caractérisation repose sur notre capacité à séquencer directement le virus à partir de tissus hôtes ou d'échantillons d'eau. Étant donné que le matériel génétique viral dans de tels échantillons est rare, les approches de génomique ne permettent parfois pas la reconstruction de génomes de ces virus par manque de profondeur de séquençage. De plus, le génome du virus OsHV-1 est complexe, structuré en régions unique et en régions répétées qui posent problème lors de l'étape d'assemblage des génomes. Certaines régions, telles que les homopolymères et la stem-loop posent également des problèmes importants dans la reconstruction des génomes, et d'autant plus lorsque la quantité de séquences virales n'est pas suffisante. Dans ce contexte, il était nécessaire de développer des outils permettant de purifier les particules virales puis de séquencer ces virus plus facilement pour améliorer les étapes de reconstruction génomique en développant un pipeline bioinformatique adapté.

Dans le cadre de ma thèse, je me suis uniquement intéressée à l'utilisation de séquençage de court fragments ADN *via* la technologie Illumina ainsi qu'au développement du pipeline bioinformatique qui a ensuite pu être utilisé pour l'analyse de la diversité du virus OsHV-1 dans le reste de ma thèse.

Actuellement, les technologies de séquençage de court fragments ADN de type Illumina sont majoritairement utilisées pour le séquençage massif d'échantillons. Le coût de séquençage *via* ces technologies a fortement diminué et les temps de séquençage ont été largement réduit, ce qui permet une utilisation à plus grande échelle. De plus, la quantité d'ADN nécessaire pour effectuer ce type de séquençage est dérisoire, ce qui permet de séquencer des échantillons de bonnes qualités mais également des échantillons de collection, stockés depuis des années, qui seraient mal conservés ou dont les concentrations ADN seraient négligeables. Cette technologie est idéale pour identifier des variations à petite échelle telles que les SNP puisqu'elle permet un séquençage précis, c'est-à-dire avec des faibles taux d'erreurs de séquençage, et avec de très grande profondeur. Ces caractéristiques ont été décisives dans le choix de la technologie utilisées pour le séquençage des échantillons collectés durant ma thèse.

Bien que cette technologie soit idéale pour identifier des variations à petite échelle telles que les SNP, les séquences courtes ont du mal à résoudre des variations à plus grande échelle telles que les variations de structure (SV). Ces limitations peuvent maintenant être surmontées par des approches de séquençage à longues lectures. Le séquençage par nanopores a été testé avec succès sur des *herpesvirus* tels que le virus herpès humain de type 1 (HHV-1), le cytomégalovirus humain (HCMV),

l'*herpesvirus* équin 1 (EHV-1) et le virus herpès koi (KHV). Le séquençage de longues séquences ADN *via* la technologie ONT est une perspective majeure en cours de développement au laboratoire pour le séquençage des prochains échantillons du virus OsHV-1.

De plus, il était nécessaire que je développe un pipeline bioinformatique me permettant d'effectuer toutes les étapes pour assembler les génomes du virus et étudier la diversité virale à partir de librairies de séquences brutes. Ce pipeline reprend les étapes classiques des pipelines utilisés pour l'analyse de la diversité virale chez d'autres espèces, notamment pour l'analyse du SARS-Cov-2. Ce pipeline permet de trier les séquences brutes en fonction de leur qualité, puis de récupérer uniquement les séquences virales. En effet, le séquençage des échantillons est effectué sur des ADN extraits à partir de manteau d'huître, ce qui induit que plus de 90% des séquences obtenues dans les librairies correspondent à des séquences de l'hôte. A partir des séquences non-huîtres, l'assemblage de contigs est effectué à l'aide de l'outil metaSpades. L'architecture génomique du virus OsHV-1, organisée en régions répétées et en régions uniques, ainsi que la présence de nombreux homopolymères et d'une stemp-loop, ne permettent pas la reconstruction automatique du génome complet, ce qui implique des étapes d'assemblage manuel des contigs pour obtenir un assemblage complet et de bonne qualité pour chaque génome viral. Les analyses de diversité ont ensuite été effectuées à l'échelle du génome complet avec l'analyse des variations structurales, mais également à l'échelle locale intra-individuelle par appel de variant et inter-individuelle par des analyses de génétique des populations. Enfin, des modules d'inférences phylogénétiques et phylogéographiques ont également pu être intégrés.

Article 1

Evaluation of tangential flow filtration coupled to long-read sequencing for *Ostreid herpesvirus* type 1 genome assembly

Aurélie Dotto-Maurel, Camille Pelletier, Benjamin Morga, Maude Jacquot, Nicole Faury, Lionel Dégremont, Maëlis Bereszczynki, Jean Delmotte, Jean-Michel Escoubas, Germain Chevignon.

MICROBIAL GENOMICS

RESEARCH ARTICLE

Dotto-Maurel *et al.*, *Microbial Genomics* 2022;8:000895

DOI 10.1099/mgen.0.000895

MICROBIOLOGY SOCIETY

OPEN DATA OPEN ACCESS

Evaluation of tangential flow filtration coupled to long-read sequencing for ostreid herpesvirus type 1 genome assembly

Aurélie Dotto-Maurel[1]†, Camille Pelletier[1]†, Benjamin Morga[1], Maude Jacquot[1], Nicole Faury[1], Lionel Dégremont[1], Maëlis Bereszczynki[2], Jean Delmotte[2], Jean-Michel Escoubas[2,*] and Germain Chevignon[1,*]

Abstract

Whole-genome sequencing is widely used to better understand the transmission dynamics, the evolution and the emergence of new variants of viral pathogens. This can bring crucial information to stakeholders for disease management. Unfortunately, aquatic virus genomes are usually difficult to characterize because most of these viruses cannot be easily propagated *in vitro*. Developing methodologies for routine genome sequencing of aquatic viruses is timely given the ongoing threat of disease emergence. This is particularly true for pathogenic viruses infecting species of commercial interest that are widely exchanged between production basins or countries. For example, the ostreid herpesvirus type 1 (OsHV-1) is a Herpesvirus widely associated with mass mortality events of juvenile Pacific oyster *Crassostrea gigas*. Genomes of Herpesviruses are large and complex with long direct and inverted terminal repeats. In addition, OsHV-1 is unculturable. It therefore accumulates several features that make its genome sequencing and assembly challenging. To overcome these difficulties, we developed a tangential flow filtration (TFF) method to enrich OsHV-1 infective particles from infected host tissues. This virus purification allowed us to extract high molecular weight and high-quality viral DNA that was subjected to Illumina short-read and Nanopore long-read sequencing. Dedicated bioinformatic pipelines were developed to assemble complete OsHV-1 genomes with reads from both sequencing technologies. Nanopore sequencing allowed characterization of new structural variations and major viral isomers while having 99,98% of nucleotide identity with the Illumina assembled genome. Our study shows that TFF-based purification method, coupled with Nanopore sequencing, is a promising approach to enable in field sequencing of unculturable aquatic DNA virus.

DATA SUMMARY

All sequence data have been deposited in NCBI under Bioproject ID PRJNA823871. Nanopore data have been deposited to SRA under accession number SAMN27363714 with base-called fastq reads under accession number SRR19025776 and raw fast5 and slow5 file under accession number SRR19039499; Illumina fastq reads have been deposited to SRA under accession number SAMN27363675. OsHV-1 assembled genome from Nanopore data have been deposited to GenBank under accession number ON411413; OsHV-1 assembled genome from Illumina data have been deposited to GenBank under accession number ON411412.

Received 13 May 2022; Accepted 06 September 2022; Published 10 November 2022

Author affiliations: [1]Ifremer, ASIM, F-17390 La Tremblade, France; [2]IHPE, Univ. Montpellier, CNRS, Ifremer, UPVD, F-34095 Montpellier, France.
***Correspondence:** Jean-Michel Escoubas, jean.michel.escoubas@ifremer.fr; Germain Chevignon, germain.chevignon@ifremer.fr

Keywords: *Crassostrea gigas*; Illumina; ostreid herpesvirus type 1; oxford nanopore technologies; tangential flow filtration; virus.

Abbreviations: ASW, artificial seawater; Cp, crossing point; cp, copy number; DP, DNA polymerase; FSW, filtered seawater; GU, genomic units; IRL, inverted repeats long; IRS, inverted repeats small; mNGS, metagenomic next-generation sequencing; ONT, Oxford Nanopore Technologies; ORF, open reading frames; OsHV-1, ostreid herpesvirus type; P, permeate; POMS, pacific oyster mortality syndrome; R, retentate; SNP, single nucleotide polymorphism; SV, structure variations; TFF, tangential flow filtration; TRL, tandem repeats long; TRS, tandem repeats small; UL, unique long; US, unique small; VLP, virus-like particles.

All sequence data have been deposited in NCBI under Bioproject ID PRJNA823871; biosample IDs SAMN27363714 and SAMN27363675; and to GenBank under accession number ON411412 and ON411413.

†These authors contributed equally to this work

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Three supplementary figures and two supplementary tables are available with the online version of this article.

000895 © 2022 The Authors

This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Impact Statement

Several aquatic viruses are pathogenic to species of commercial interest, leading to massive value losses during disease outbreaks. To better understand the origin, transmission and propagation of such viruses, more genomic data is required and needs to be analysed. However genomic characterization of viruses that are mainly unculturable is a major challenge. In the present study, we have successfully tested the combination of tangential flow filtration (TFF) and Nanopore sequencing for virus purification and genome sequencing. Our results demonstrate that it is now possible to easily sequence the complete genome of large DNA viruses directly from host infected tissues without prior *in vitro* propagation.

INTRODUCTION

Viruses found in aquatic environments are often unculturable and their genomic characterization often relies on our ability to enrich samples for viral nucleic acid. To date, several methods have been implemented to this end and different workflows are applied depending on the type of virus studied: small or large virus, RNA or DNA virus, virus with linear or circular genome, single or double strand virus. Enrichment methods range from rapid and simple steps involving low-speed centrifugation, filtration and nuclease treatment, to more complex methods involving ultracentrifugation, tilling PCR, rolling circle amplification and hybridization-based capture. While the former shows higher efficiency at purifying viral nucleic acid, these methods habitually rely on *a priori* knowledge of the targeted virus and are limited at detecting a new variant or new pathogen [1–4]. Recently, metagenomic next-generation sequencing (mNGS) has been successfully tested for broad-spectrum pathogen identification for clinical diagnosis of infectious diseases, outbreak surveillance, and pathogen discovery [5, 6]. However, the cost, sequencing depth and background contamination are true limitations to the accuracy of mNGS-based diagnostics relative to specific targeting testing.

Given that aquaculture is continuously faced with the challenge of identifying novel pathogenic viruses infecting different fish, crustaceans, and shellfish species, there is an urgent need to develop tools enabling those viruses to be sequenced easily and with limited bias to develop timely disease control strategies.

Ostreid herpesvirus type 1 (OsHV-1) is a double strand DNA virus belonging to the Malacoherpesviridae family and is the only species of the genus Ostreavirus [7]. OsHV-1 has been associated with mass mortality events of naive Pacific oyster *Crassostrea gigas* [8] from larvae to adults [9, 10]. It has plagued the oyster production worldwide, from Europe to north and south America, Australia and New Zealand, and in a lesser extent in Asia, for more than a decade [11–17]. OsHV-1 mostly affects spat and juvenile (less than 18 months) in oyster farms, and leads to mass mortalities that can reach 100% within days [18–20]. Research efforts have revealed a series of factors contributing to the disease, including seawater temperature, oyster age and genetics [8, 9, 16, 18, 21–24]. Recently, holistic molecular approaches uncovered the pathogens associated to Pacific oyster mortality syndrome (POMS) [25]. These studies showed that an infection by OsHV-1 is the critical step of the infectious process leading to an immune-compromised state followed by a microbiota destabilization that 'opens the door' to opportunistic bacterial pathogens (e.g. *Vibrio* spp.) that target haemocytes to induce their lysis [26]. The infectious process is completed with subsequent bacteraemia ultimately leading to oyster death (reviewed in [27]).

Since the characterization of the emergent genotype OsHV-1 µVar, associated with the 2008 high mortality events [18], several variants of the µVar genotype have been described (reviewed in [28]). However, most studies on OsHV-1 genetic diversity have been based on PCR molecular markers focusing on a few variable regions of the viral genome, which do not reflect the whole genomic diversity [29, 30]. Indeed, OsHV-1 µVar whole-genome sequencing [31] and its comparison with the reference genome published in 2005 [7] showed that both genomes differed at the nucleotide level (single nucleotide polymorphism, SNP) as well as at the structural level by the loss or gain of several genomic regions carrying many open reading frames (ORFs). This observation indicates that whole-genome sequencing is needed to fully reveal viral diversity and better understand the structure and evolution of the virus populations [31].

OsHV-1 particles harbour a linear genome of approximately 207 kilobases (kb) with the majority of the 128 ORFs detected (54.7%), having no known homologies with other organisms [31]. OsHV-1 genomic organization is characteristic of other Herpesviruses consisting of two invertible unique regions (U_L for 'Unique long' 170 kbp; U_S for 'Unique small' 3.5 kbp) each flanked by tandem and inverted repeats (TR_L/IR_L: 7.5 kbp; TR_S/IR_S: 9.7 kbp) separated by a third unique region (X: 1.5 kbp) with a global genomic organization TR_L–U_L–IR_L–X–IR_S–U_S–TR_S. Furthermore, several isomers of the virus cohabit at equimolar levels in the same samples [7].

To date, 50 fully assembled OsHV-1 genomes are available in public databases, those genomes have been sequenced and assembled using various protocol and sequencing technologies (Table S1, available in the online version of this article). Although recent genomes have been obtained from short-read mNGS [31–34], OsHV-1 complete genome assembly remains a complex process. Indeed, a short-read approach is ideal for the identification of small-scale variation like SNPs but this technology struggles to identify variations at a larger scale like structure variations (SV) or large inverted and repeated regions. Due to its intrinsic

Herpesvirus genome architecture with high G+C content regions, long repeated regions (TR_L/IR_L and TR_S/IR_S), and the presence of four major isomers differing in the relative orientations of the two unique regions, it is extremely difficult, if not impossible, to solve the complete structure of OsHV-1 genome using short-read sequencing [35].

These limitations can now be addressed by long-read sequencing approaches of which Oxford Nanopore Technologies (ONT) is one example. Indeed, ONT has already demonstrated its ability to sequence and assemble several types of viruses from RNA [36–42] as well as DNA [43–46] matrices. Nanopore sequencing has also been successfully tested on Herpesviruses like the Human Herpes Virus Type 1 (HHV-1) [47], the Human cytomegalovirus (HCMV) [48], the Equine herpesvirus 1 (EHV-1) [49] and the koi herpesvirus (KHV) [50]. In addition, ONT already allowed researchers to sequence viruses infecting aquaculture species like the White spot syndrome virus (WSSV) infecting shrimp [51] or the Salmonid alphavirus (SAV) infecting Atlantic salmon [52]. Finally, ONT has also demonstrated efficiency to sequence viruses from sea-water samples [53–55]. However, it has never been tested for the characterization of the OsHV-1 genome.

In order to implement ONT long-read sequencing, high molecular weighted DNA is required in a large quantity. However, due to the lack of a cell line, that could be used for the propagation and isolation of viral particles, OsHV-1 must be extracted from infected animal tissues. Therefore, it is extremely challenging to obtain viral DNA in sufficient quantity to perform long-read mNGS.

To overcome these difficulties, we need a new approach for viral particle purification coupled with long-read sequencing to assemble complex OsHV-1 genomes efficiently and accurately.

One approach to purify viral particles is tangential flow filtration (TFF). This method has been widely used for purification of a large variety of virus particles, such as particles of human influenza A virus [56], insect baculovirus [57, 58], HIV-like virus [59], oncolytic measles virus [60] and of several waterborne viruses like norovirus and adenovirus [61]. While TFF has been used to reduce oyster pathogen concentration in seawater to bio-secure shellfish farms [62], it has never been used with the goal of purifying those pathogens. Recently, TFF coupled with Nanopore sequencing, has been used to concentrate a large amount of seawater for the characterization of bacteriophage population [63].

In the present study, we have developed a TFF-based method allowing purification and concentration of OsHV-1 infectious particles from infected oyster tissues and we have evaluated the relative benefits of long-read (Nanopore) and short-read (Illumina) sequencing for genome assembly. Our results show that TFF purification of OsHV-1 viral particles coupled with ONT sequencing is a promising strategy for OsHV-1 whole-genome reconstruction and to decipher genetic diversity and evolution of virus populations.

METHODS

Oyster production

Experimental *C. gigas* spat was produced at the Ifremer hatchery in La Tremblade, (Marennes-Oléron Bay, Charente-Maritime, France). Oysters belonged to a line that had been characterized for its higher susceptibility to OsHV-1 [64]. Thirty adult oysters were conditioned in seawater at 20 °C for 2 months to favour gametogenesis. Then, ripe oysters were triggered for a mass spawning event in March 2015. Larvae and spat were cultured as described by [64, 65] and they were always maintained in our controlled facilities using filtered and UV-treated seawater to protect from pathogens naturally present in the Marennes-Oléron Bay. Prior to the experiment, spat (2 g and 4 cm) were acclimatized at 19 °C for 2 weeks in flow-through 120 l tanks using filtered and UV-treated seawater enriched in phytoplankton (*Skeletonema costatum, Isochrisis galbana* and *Tetraselmis suecica*).

Viral inoculum preparation and OsHV-1 quantification

OsHV-1 inoculum were prepared from tissue of infected oysters following protocols of Schikorsky *et al.* [66]. Briefly, gills and mantle were dissected from moribund oysters naturally or experimentally infected with OsHV-1 in previous trials and frozen at −80 °C. Then, ten volumes of 0.2 µm filtered and artificial seawater (ASW) were added and tissues were crushed on ice using an Ultraturax mixer (3×5 s). The tissue mash was clarified by centrifugation (1000 ***g***, 5 min, 4 °C), and then filtered consecutively in sterile conditions using syringe filters at 8, 0.45 and 0.22 µm pore sizes (Millipore; Billerica, USA).

Quantification of OsHV-1 was carried out using quantitative PCR (qPCR) using primers amplifying a 197 bp fragment of the OsHV-1 DNA polymerase (DP) gene. DNA was extracted from viral inoculum using the QiaAmp DNA Mini Kit (Qiagen) according to the manufacturer's protocol. DNA purity ($1.8 \leq$ A260 nm/280 nm ≤ 2.0; $2.0 \leq$ A260 nm/230 nm ≤ 2.20) and concentration were checked using a Nano-Drop ND-1000 spectrometer (Thermo Scientific) and Qubit® dsDNA HS assay kits (Molecular Probes Life Technologies), respectively.

qPCRs were performed in duplicate using a Mx3005 P Thermocycler (Agilent). Amplification reactions were each performed in a total volume of 20 µl. Each reaction contained 5 µl of DNA extracted from viral inoculum or from TFF fractions (readjusted to initial inoculum volume) or 5 µl of Milli-Q water for the negative control, 10 µl of Brilliant III Ultra-Fast SYBR® Green PCR

Master Mix (Agilent), 2 µl of primer OsHVDP For (forward) 5′-ATTGATGATGTGGATAATCTGTG-3′ and 2 µl of primer OsHVDP Rev (reverse) 5′-GGTAAATACCATTGGTCTTGTTCC-3′ [67] at the final concentration of 550 nM each, and 1 µl of Milli-Q water. qPCR cycling conditions were as follows: 3 min at 95 °C followed by 40 cycles of amplification at 95 °C for 5 s and 60 °C for 20 s. For absolute quantification, the amplification product of the DP gene was cloned into the pCR4-Topo vector and replicated in *Escherichia coli* DH5α (Invitrogen). Plasmid was extracted using the Wizard Plus SV miniprep DNA purification system (Promega) and standard curves of known concentration of plasmid were generated according to the Applied Biosystems manual for absolute real-time PCR quantification. Absolute quantification of OsHV-1 genome copies in oyster samples was then estimated by comparing the observed crossing point (Cp) values to known plasmid standards from 10^3 to 10^9 copies of the DP product. The results were expressed as the log-transformed copy number of viral DNA per micro litter of DNA extracts (cp μl^{-1}).

Virus-like particle quantification by epifluorescence microscopy

Virus-like particles (VLP) were quantified using epifluorescence microscopy after staining with nucleic acid dye [68, 69]. Briefly, viral inoculum were first fixed with glutaraldehyde (Sigma-Aldrich G6257, final concentration 0.5%) and diluted in 1% potassium citrate solution (pH 7) to prevent aggregation of viral particles. Inoculum were vacuum filtrated to capture viruses on a 0.02 µm aluminium oxide filter (Whatman 6809–6002 Anodisc 25 inorganic membrane filter), before subsequent staining (SYBR Gold nucleic acid gel stain Invitrogen S11494, diluted at 0.2% in Tris-EDTA) and washing by 4 ml of 1% potassium citrate solution. Filters were then mounted on glass slides and VLP were observed and enumerated under epifluorescence microscopy (Olympus BH2-RFCA).

Experimental infections

For experimental infections, oysters were previously myorelaxed in hexahydrate MgCl2 (ACROS, 50 g l^{-1}) [70]. Oysters were injected using 26-gauge needle attached to a multi-dispensing hand pipette, with 100 µl of inoculum (diluted in ASW to reach the desired viral concentration) in the abductor muscle. For each experiment, a group of oysters was myorelaxed and injected with ASW as the control. Oysters were replaced in seawater in plastic tanks at 20 °C and mortality was monitored daily.

Virus purification and concentration by tangential flow filtration (TFF)

To develop a TFF-based process for the concentration of OsHV-1 particles, we used MicroKros® hollow fibre filter modules (Spectrum®) with three different pore sizes: 750 kDa (CO2-E7500-05-N), 500 kDa (CO2-E500-05-N) and 300 kDa (CO2-E300-05-N).

The first step of the process was to determine the pore size of the appropriate membrane to retain the OsHV-1 particles. To this end, we have filtered 10 ml of OsHV-1 inoculum containing 5.9×10^4 genomic units (GU)/µl with the 750 kDa TFF-device until the volume reached 1 ml. This 1 ml retentate (R), consisting of particles larger than the pore size of the hollow fibre, was diluted to the initial volume (10 ml) with filtered (0.02 µm) seawater and stored at 4 °C. The permeate (P), composed of particles smaller than the pore size of the hollow fibre, was collected, the volume adjusted to 10 ml with filtered (0.02 µm) seawater and was then filtered on the 500 kDa TFF device. The 500 kDa retentate was then processed similarly to the 750 kDa retentate and the 500 kDa permeate was then filtered on the 300 kDa TFF-device. At the end of this filtering process, we obtained three retentates (R-750; R-500 and R-300) and one permeate (P-300) (Fig. 1a).

The second step consisted of purifying and concentrating OsHV-1 particles to extract enough DNA for both short- and long-read sequencing. For this purpose, we prepared a viral inoculum (see Viral inoculum preparation and OsHV-1 quantification for details) from 31.12 g of tissue obtained from 28 susceptible oysters (family F15) naturally infected by OsHV-1 in the Thau Lagoon in 2016 and stored at −80 °C (for more details on oyster origin and infection process see [71]). We obtained 462 ml of 0.22 µm filtrated inoculum, which were processed on a 500 kDa hollow fibre filter. Concentrated viral particles contained in 82 ml of the 500 kDa retentate were recovered by centrifugation (2 h at 18 000 r.p.m. at 4 °C). Pellets were then re-suspended in 1.2 ml of PBS and DNA was extracted and quantified as described above. We finally obtained 17.9 µg of viral DNA.

Illumina short-read sequencing

Illumina DNA-Seq library preparation and sequencing were performed by the Genome Quebec Company (Genome Quebec Innovation Center, McGill University, Montreal, Canada) using 120 ng of DNA extracted from the OsHV-1 inoculum concentrated on 500 kDa TFF-device [as described at the end of Virus purification and concentration by tangential flow filtration (TFF)], the NxSeq AmpFREE Low DNA Library Kit (Lucigen) and the NovaSeq 6000 Sequencing system (Illumina®) (paired ends, 150 bp).

Nanopore long-read sequencing

For long-read sequencing, we used the Oxford Nanopore Technologies (ONT) MinION platform (Mk1b). Briefly, 10 µg of DNA extracted from the OsHV-1 inoculum concentrated on 500 kDa TFF-device [as described at the end of Virus purification and concentration by tangential flow filtration (TFF)] were fractionated on Agencourt AMPure XP beads (Beckmann Coulter, 0.5× final concentration) to eliminate small DNA molecules (< 300 bp). DNA size range analysis on Agilent Fragment Analyzer

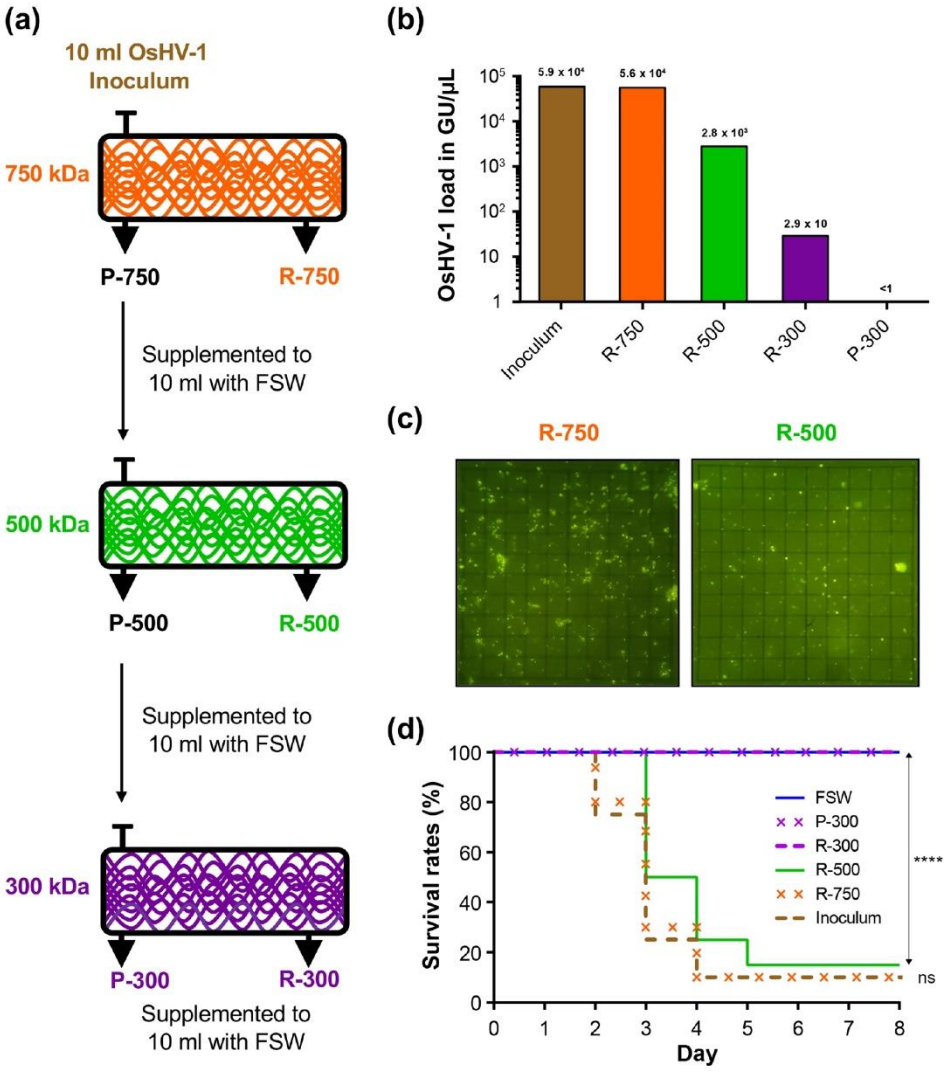


Fig. 1. (a) Workflow for the processing of OsHV-1 inoculum on tangential flow filtration (TFF) devices with three membrane pore sizes (750, 500 and 300 kDa). After filtration on each TFF-device, the volumes of retentates (R) and permeates (P) were supplemented to 10 ml with filtered seawater (FSW, 0,02 µm). (b) qPCR monitoring of OsHV-1 loads (in genomic units per µl) at the different steps of TFF. (c) Visualization and quantification of virus-like particles (VLP) by epifluorescence microscopy in retentates of TFF devices of 750 kDa (R-750) and 500 kDa (R-500). An eyepiece with a 10×10 grid reticle was used for counting VLP (note in our setup the grid reticle is 100×100 µm). (d) Kaplan-Meier survival curves of oysters injected with 100 µl of OsHV-1 inoculum, or retentates of TFF-devices of 750, 500, 300 kDa (R-750, R-500 and R-300, respectively), or permeate of TFF device 300 kDa (P-300) or FSW as negative control. ****P <0.0001; ns not significant (Mantel–Cox Log-rank test).

(Genomic DNA analysis kit DNF-467–33) revealed that the average size was over 60 kb with a genomic quality number (GQN) of 8.9 (threshold 10 kb). Then, 400 ng of AMPure purified DNA were used to construct the sequencing library using SQK-RAD0004 tagmentation kit and sequencing was performed as per manufacturer's guidelines using R9.4 flowcells (FLO-MIN106) and run for 20 h.

Illumina genome assembly

Illumina raw reads files were quality filtered (PHRED score >31) and adapters trimmed with Fastp version 0.20.1 [72]. To evaluate the origin of reads, fastq reads were aligned to the oyster *C. gigas* genome (accession: GCA_902806645.1) and to OsHV-1 µVar

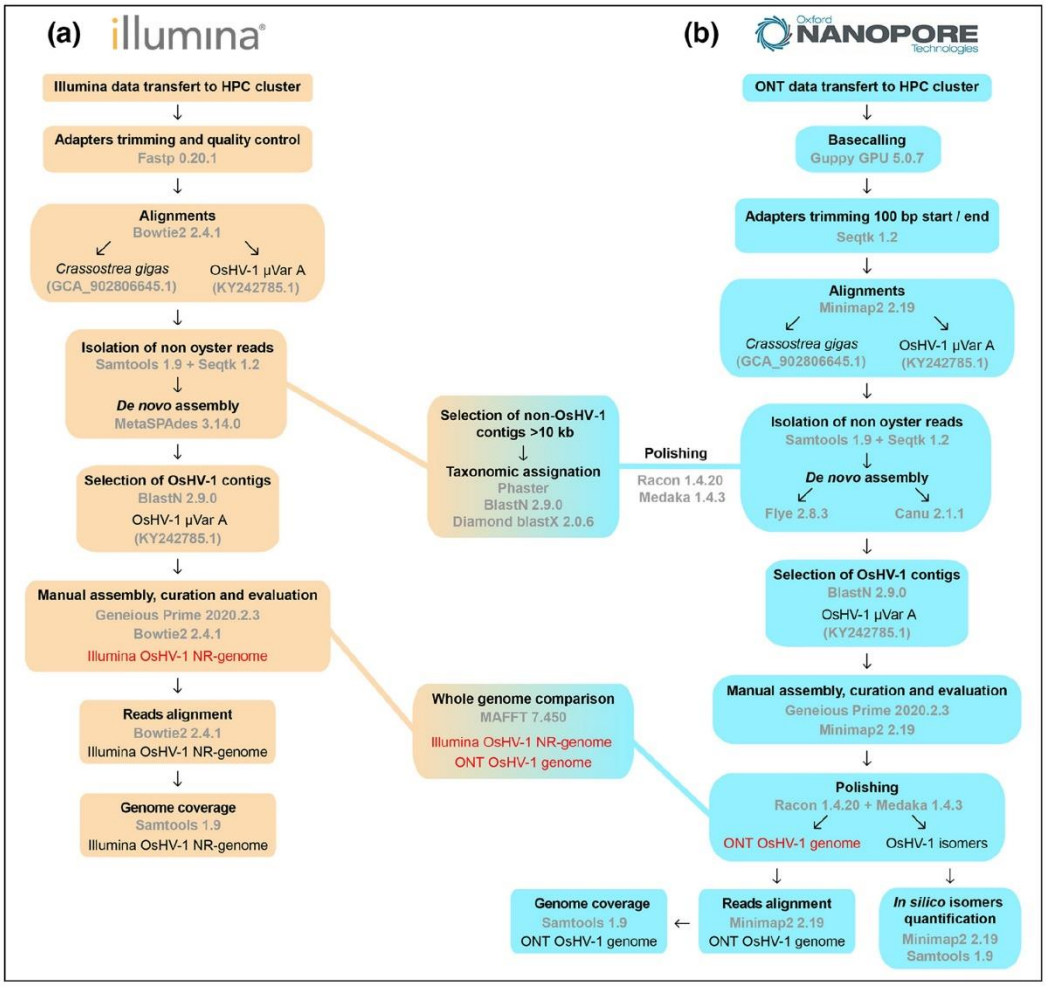


Fig. 2. Flow diagram to illustrate the bioinformatic pipelines used to assemble OsHV-1 genome from (a) Illumina short-read or (b) Nanopore long-read.

A genome (accession: KY242785.1) with Bowtie2 version 2.4.1 [73]. All the non-oyster reads were then selected with Samtools version 1.9 [74] and Seqtk version 1.2 (https://github.com/lh3/seqtk). Those reads were then used for *de novo* assembly with SPAdes version 3.14.0 with parameter –meta [75]. The *de novo* OsHV-1 contigs generated by SPAdes were then identified with BLASTN version 2.6.0 [76] using OsHV-1 µVar A sequence as a query (accession number: KY242785.1). The identified OsHV-1 contigs were then manually scaffolded in Geneious Prime (version 2020.2.3) to reconstruct the common virus architecture (U_L-IR_L-X-IR_S-U_S) [7]. To overcome the problem of inverted repeated regions that cannot be confidently positioned with short-read data, we constructed non-redundant genomes (NR genomes) that contain only one copy of each repeated region (U_L-IR_L-X-IR_S-U_S) an approach commonly used for Herpesvirus genome analyses [34, 77, 78]. Quality filtered reads were then aligned to the new constructed OsHV-1 NR genome to evaluate coverage uniformity. This pipeline is summarized in Fig. 2a.

Nanopore genome assembly

Nanopore fast5 file were GPU base called with Guppy version 5.0.7 with standard parameters except -config dna_r9.4.1_450bps_hac.cfg (https://community.nanoporetech.com/downloads). Fastq reads were then 100 bp trimmed at their beginning and at their end to ensure proper adapter removal with SeqTk version 1.3 with parameters -b 100 and -e 100 (https://github.com/lh3/seqtk). To assign reads to either host or viral origin, fastq reads were aligned to the oyster *C. gigas* genome (accession: GCA_902806645.1) and to OsHV-1 µVar A genome (accession: KY242785.1) with Minimap2 version 2.22 with standard parameters [79]. Then two

types of *de novo* assemblies were performed with non-oyster reads with two different *de novo* assemblers. All fastq reads or only the top 10% of the longest reads were used as input for *de novo* assembly with Flye version 2.8.3 and parameters -meta [80] or with Canu version 2.1.1 and parameters -nanopore [81]. OsHV-1 *de novo* assembled contigs were then identified with BLASTN [76] against the OsHV-1 µVar A genome and were manually trimmed in Geneious version 2020.2.4 (https://www.geneious.com). All *de novo* contigs plus the new assembled OsHV-1 genome were then polished using Nanopore reads with three run of Racon version 1.4.20 with parameters --no-trimming -m 8x–6g –8w 500 [82] and one run of Medaka version 1.4.3 with parameters medaka_consensus (https://github.com/nanoporetech/medaka).

For OsHV-1 isoform coverage evaluation, the four contigs produced by Canu with the top 10% of the longest reads were trimmed on the U_L side to keep a 10 kb U_L fragment with Geneious version 2020.2.4 (https://www.geneious.com). The four major isomers were then used as a reference for reads mapping with Minimap2 version 2.22 with standard parameters -axe map-ont (https://github.com/lh3/minimap2). Samtools version 1.10 [83] were then used to filter out all the shortest aligned read and all the alignment with MapQ <40, keeping only the reads spanning at least a fragment of the U_L and a fragment of the U_S (i.e. 28 kb long). A pipeline summary can be found in Fig. 2b.

Non-OsHV-1 contigs' classification

Polished non-OsHV-1 contigs with a size larger than 10 kb were then subjected to taxonomic identification and classification with PHASTER [84] accessed in September 2021, with BLASTN version 2.0.9 with parameters -max_hsps 1 -max_target_seqs 100 -evalue 0.0001 [76] or with DIAMOND BLASTX [85] version 2.0.6 with parameters --sensitive --max-hsps 1 --max-target-seqs 0 --range-culling -F 15 --evalue 0.0001 against the refseq database accessed in September 2021.

Comparison of Illumina and Nanopore OsHV-1 genomes

Before whole-genome alignment evaluation, both genomes were modified to fit the same isomers' version of the genome (i.e. TR_L/U_L->/$IR_L/X/IR_S/U_S$->/TR_S). To evaluate the capacity of Nanopore sequencing to produce an accurate OsHV-1 genome, the manually constructed Illumina OsHV-1 non-redundant (NR) genome was aligned with the manually curated Nanopore OsHV-1 complete genome with MAFFT version 1.4.0 using default parameters [86] implemented in Geneious version 2020.2.4 (https://www.geneious.com).

OsHV-1 phylogeny

In order to illustrate the global phylogenetic relationships among our and other available genomes, 13 OsHV-1 NR-genomes (11 available in GenBank: AY509253, GQ153938, KP412538, KY242785, KY271630, MG56175, MF509813, OM811579, OM811584, OM811586, OM811594 and the two from this study) were aligned with progressiveMauve algorithm [87] implemented in Geneious version 2020.2.4 to local collinear blocks. Reordered genomes block were then aligned with MAFFT version 7.4 using default parameters [86] implemented in NGPhylogeny.fr [88]. Alignment was cleaned with BMGE with default parameters [89] implemented in NGPhylogeny.fr. The best fitting model of nucleotide substitution was chosen according to the Akaike Information Criterion with jModelTest version 2.1.10 with default parameter [90]. Tree was constructed with phyML version 3.3 [91] with default parameters except '-b 100 m HKY85 -f m -c 8 s BEST -a e -v e -o tlr' implemented in NGPhylogeny.fr. Phylogenetic tree was midpoint rooted and arranged in FigTree version 1.4.4 (https://github.com/rambaut/figtree).

RESULTS

Calibration of tangential flow filtration (TFF) device to purify and concentrate OsHV-1 particles

To develop a tangential flow filtration-based process for the purification and concentration of OsHV-1, we tested TFF-devices with three membrane pore sizes [molecular weight cut-off (MWCO) of 750, 500 and 300 kDa]. Then, we used an OsHV-1 inoculum as starting material and successively filtered it through the three TFF-devices (Fig. 1a; see Methods for further detail). Finally, we have obtained three retentates (R-750, R-500 and R-300) and one permeate (P-300), which were then used to analyse the OsHV-1 load and infectivity.

To evaluate the ability of the three TFF-devices to retain OsHV-1, we have quantified OsHV-1 load by qPCR in the initial inoculum, in the three retentates and in the permeate (Fig. 1b). The number of OsHV-1 genomic units (GU) in the R-750 fraction (5.6×10^4 GU) was of the same order of magnitude as it was in the initial inoculum (5.9×10^4 GU), which suggests that most of the viruses (95%), were retained by 750 kDa MWCO. However, a small fraction (4.8%) of GUs present in the initial inoculum were also found in the R-500 retentate (2.8×10^3 GU) and traces of amplification (below standard curve) were also observed in the R-300 (2.9×10 GU). Epifluorescence microscopy did not reveal virus-like particles (VLP) in the R-300 and P-300 fractions (data not shown) while in the R-750 and R-500 fractions we were able to enumerate 2.71×10^4 ($7.05 \times 10^3\pm$SD) and 4.88×10^3 ($2.79\times10^3\pm$SD) VLP μl^{-1}, respectively (Fig. 1c). Epifluorescence microscopy gave similar results to qPCR and confirms that most of the viral particles present in the starting inoculum were preserved in the R-750 fraction.

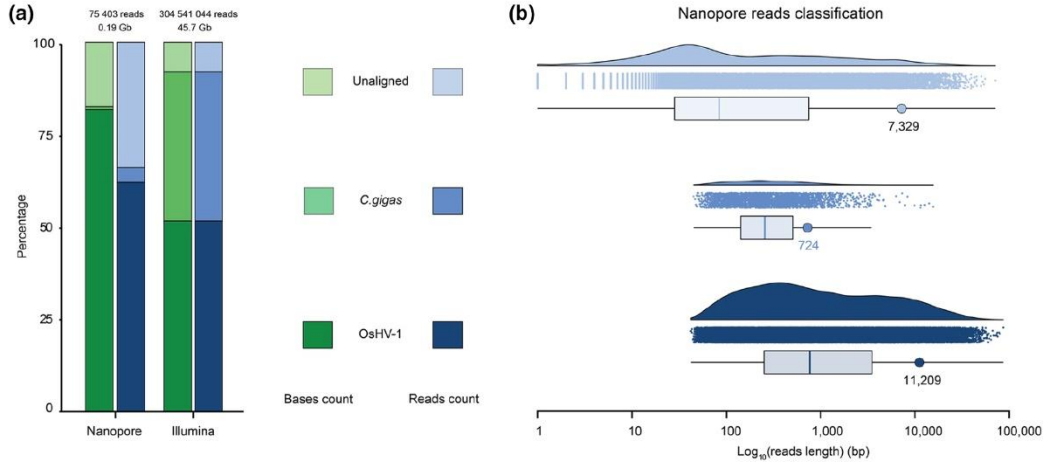


Fig. 3. Global sequencing statistics. (a) Histogram representing percentage of sequenced bases (green) and reads (blue) that aligned to OsHV-1 (dark) or to *C. gigas* (regular) genomes or that did not align to these genomes (light). (b) Raincloud plot representing the size distribution of the Nanopore reads that aligned to OsHV-1 (dark blue) or to *C. gigas* (blue) genomes or that did not align to these genomes (light blue). Each panel is made of a density plot, a dot plot where each dot corresponds to one Nanopore read and a box plot showing average read size (vertical coloured bar) and N50 (coloured circle).

Finally, we compared the infectious capacity of the different fractions with the initial inoculum by injecting them into oysters. No mortality was observed for oysters injected with fractions R-300 and P-300 as well as for oysters injected with filtered sea water (Fig. 1d). However, oysters that had been injected with R-750 and R-500 fractions exhibit survival curves, which are not significantly different from those of oysters injected with the starting inoculum. Interestingly, survival curves of R-750 and R-500 fractions are not significantly different although R-500 fraction contained around 20 times less OsHV-1 GU than the R-750 fraction. These results showed that filtration on 500 kDa TFF-device allowed the purification and concentration of fully infective OsHV-1 viral particles. Consequently, we used the 500 kDa TFF-device to purify and concentrate viral particles from an inoculum prepared from 28 oysters associated from mortality event, which have occurred in 2016. We have been able to extract ~18 µg of viral DNA, an amount sufficient to perform both short- and long-read sequencing (see Methods for further detail).

General statistics of the two sequencing technologies

Viral DNA extracted from TFF purified viral particles was sequenced using two different sequencing technologies: (i) a short-read approach based on the Illumina platform and (ii) a long-read approach based on the Nanopore platform.

The emergence of mNGS, has allowed researchers to put aside complex techniques such as ultracentrifugation coupled to sucrose gradient for OsHV-1 genome sequencing. The direct consequence of the lack of viral purification is the production of a large amount of non-viral data, resulting in the waste of a significant portion of the sequencing run. In recent studies, sequencing of the OsHV-1 genome from DNA extracted directly from infected tissues has led to the acquisition of only 0.01–13.21% OsHV-1 reads (Table S1).

As expected TFF purification of OsHV-1 particles allowed both sequencing technologies to obtain a large enrichment in OsHV-1 reads compared to host reads, leading to a mean coverage of 740 reads/bp and 213 790 reads/bp for Nanopore and Illumina sequencing, respectively. In more details, the Illumina sequencing run yielded more than 304 million paired-end reads of 150 base pairs (bp) corresponding to 45.68 Gb and distributed as follows: 52% of OsHV-1 reads, 40% of *C. gigas* reads and 8% of unaligned reads with a similar proportion of sequenced base number (Fig. 3a). The Nanopore sequencing run yielded 75 403 reads corresponding to 0,19 Gb and distributed as follows: 62 and 82% of OsHV-1 reads and bases, respectively, 4 and 0.7% of *C. gigas* reads and bases, respectively, and 33 and 17% of unaligned reads and bases, respectively (Fig. 3a). Since the Nanopore technology sequence native DNA molecules present in the DNA library, a large range of reads' size has been sequenced. In detail, OsHV-1 reads ranged from 42 to 86 234 bp with a N50 of 11 209 bp, *C. gigas* reads ranged from 45 to 15 635 bp with a N50 of 724 bp and unaligned reads ranged from 1 to 70 966 bp with a N50 7 329 bp (Fig. 3b). The difference in N50 between host and viral DNA (11 209 vs 724 bp) was probably due to oyster DNA fragmentation during purification steps while viral DNA was protected by the viral envelope and capsid. This observation has probably led to the observed difference in host reads' percentage between both sequencing approaches. Indeed, before and during the Nanopore libraries' construction several Ampure XP bead-based DNA purification were performed, and this type of DNA purification is known to induce a loss of small size DNA molecules.

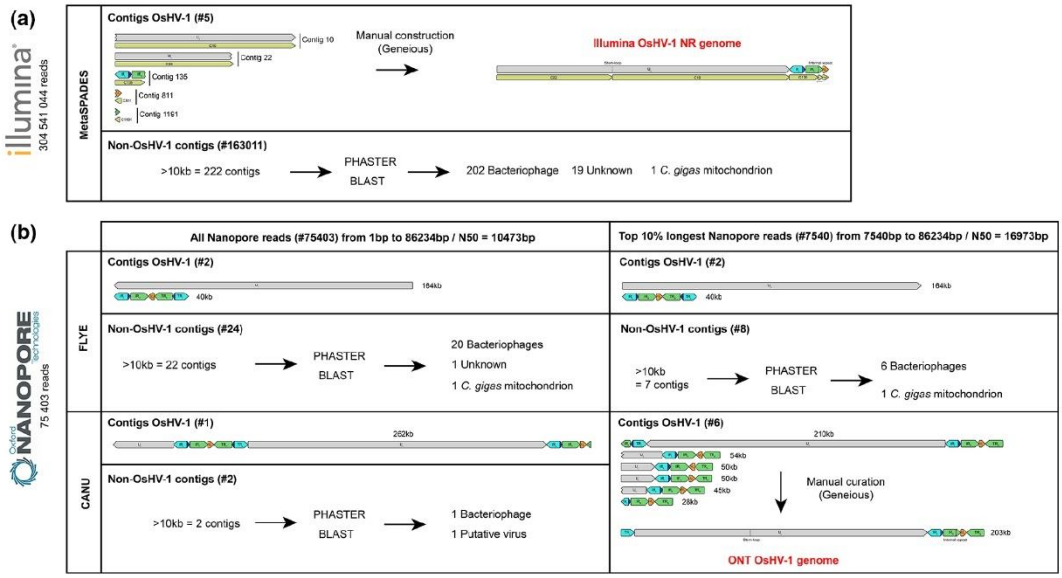


Fig. 4. Schematic representation of OsHV-1 whole-genome assembly outputs for (a) Illumina short-read sequencing and (b) Nanopore long-read sequencing.

OsHV-1 genome assembly using Illumina reads

Our short-read pipeline based on the *de novo* assembler metaSPAdes produced five OsHV-1 contigs ranging from 2 429 to 99 197 bp (Fig. 4a). Those five contigs were then manually assembled to construct an OsHV-1 NR genome in which only one copy of each repeat was conserved (U_L-IR_L-X-IR_S-U_S, see Methods for further details). The re-alignment of the Illumina reads on the new OsHV-1 NR genome allowed us to correct the junction between contigs C1191 and C811 (addition of one A/T bp) and to confirm that the other metaSPAdes contigs were correctly assembled.

As expected, we observed that the reads' coverage within IR_L and the IR_S is twice the reads' coverage for unique regions U_L, U_S and X given that the NR genome contains only one copy of theses duplicated regions (Fig. 5 upper panel). Interestingly, two decreases of reads' coverage were observed (Fig. 5 upper panel). The first decrease of coverage was at the junction of the C22 and C10 contigs and takes place in a stem-loop suspected to be the viral origin of replication [7]. The presence of this stem-loop could have induced a difficulty to access the DNA in this region by the sequencing polymerase leading to a decrease of sequencing yield around this palindromic region and the production of two distinct contigs by the *de novo* assembler (Fig. S1). The second decrease of reads' coverage takes place within the C135 contig, in a long C22 homopolymer of the IR_L region. This type of bias in read coverage has already been observed for long homopolymer sequenced by Illumina [92] (Fig. 5 upper panel and Fig. S2).

OsHV-1 genome assembly using Nanopore reads

Since this is the first time that an OsHV-1 genome has been assembled from ONT long-read, we decided to evaluate the two most used long-read *de novo* assembler Flye and Canu with two datasets: (i) all the Nanopore reads and (ii) only the top 10% longest Nanopore reads (see Methods for further detail) (Fig. 4b).

Whatever the dataset used, Flye yielded the same two contigs; a long one (164 kb) corresponding to the complete U_L and a short one (40 kb) made of the U_S surrounded by one copy of IR_L-X-IR_L on its left side and one copy of TR_S-X-TR_L on its right side (Fig. 4b). With the dataset including all Nanopore reads, Canu yielded one large OsHV-1 contig (262 kb). This contig is made of one fragment of the U_L (32 kb) followed by a fragment made of the U_S surrounded by one copy of IR_L-X-IR_S on its left side and one copy of TR_S-X-TR_L on its right side (40 kb), followed by the complete U_L (164 kb) and followed by another fragment made of U_S surrounded by one copy of IR_L-X-IR_S on the left side and one copy of the TR_L (24 kb) on its right side (Fig. 4b).

With the dataset composed of only the top 10% longest Nanopore reads, Canu yielded six OsHV-1 contigs ranging from 28 to 210 kb (Fig. 4b). The 28 kb contig consisted of the U_S surrounded on its left side by one copy of TR_L-X-IR_S and on its right side by one copy of TR_S. The four contigs ranging from 45 to 54 kb share the same general structure; that is to say, a fragment of the U_L region followed by a copy of IR_L-X-IR_S-U_S-TR_S. These four contigs are differentiated by the relative

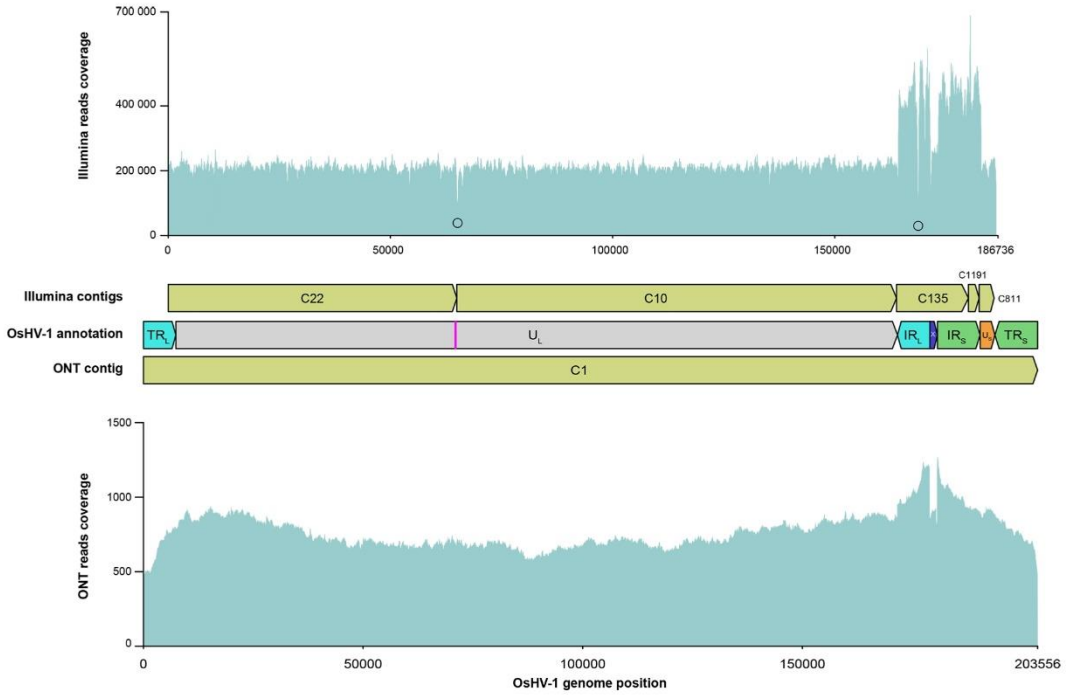


Fig. 5. OsHV-1 whole-genome assemblies' evaluation by read mapping. Upper panel represent Illumina short-read coverage on the Illumina OsHV-1 NR genome. Black circles indicate bottom line of the two losses of coverage identified in the Illumina OsHV-1 NR genome. Middle panel represent Illumina *de novo* assembled contigs (green), OsHV-1 genomic region annotation with the pink bar representing the stem-loop and the ONT *de novo* assembled contig (green). Bottom panel represent ONT long-read coverage on the ONT OsHV-1 genome. C: Contig.

orientation of the U_L and U_S fragments (same or inverted orientations) and their orientation with respect to the IR_L-X-IR_S fragment (Fig. 4b). It was likely that these four contigs correspond to the four major OsHV-1 isomers described by Davison *et al.* [7]. In order to verify this observation and to quantify the relative abundance of the four isomers, we have screened all the reads long enough to cover at least a part of both U_L and U_S regions. We identified 92 long-read divided into 18 <-U_L/<-U_S, 24 U_L->/U_S->, 26 <-U_L/U_S-> and 24 U_L->/<-U_S (Fig. 6). This result confirms the OsHV-1 genome isomerization and shows that the viral population is composed of approximately equimolar amounts of the four isomers. Finally, Canu assembled one large contig (210 kb) made of the complete U_L surrounded by one copy of IR_S-X-TR_L on its left side and one copy of IR_L-X-IR_S-U_S-TR_S on its right side (Fig. 4b). This former contig have been manually curated to create the complete ONT OsHV-1 genome by removing the IR_S-X fragment on the left side of the U_L and by reverse complement the U_L to fit with the most used representation of OsHV-1 genomes isomer [7].

To assess the robustness of the manual assembly, we aligned all ONT reads back to the ONT OsHV-1 genome and observed a continuity of coverage throughout the genome, which validates the assembly of the ONT OsHV-1 genome with our pipeline (Fig. 5 lower panel). Interestingly, we can see a small decrease of coverage at both extremities of the genome and an increase of coverage on IR_L and IR_S at the junction surrounding the X region (Fig. 5 lower panel). This observation reflects the use of one of the 16 minor isomers described by Davison *et al.* to represent our ONT OsHV-1 genome [7]. Indeed, in addition to the four isomers described previously, Davison *et al.* showed that some of these isomers contained an additional copy of the X region at the left or right terminus whereas some might have lacked the X region internally. Since only one copy of the X region is presented in the ONT OsHV-1 genome, the reads corresponding to the minor variant of the X region might tend to align around the X region between IR_L and IR_S instead of aligning at the extremities of the genome (TR_L and TR_S) leading to a higher coverage of the region surrounding X compared to the rest of the genome (Fig. 5 lower panel).

Taken together these results validates the assembly of the ONT OsHV-1 genome, confirms for the first time the OsHV-1 genome isomerization and shows that the viral population is composed of approximately equimolar amounts of the four major isomers.

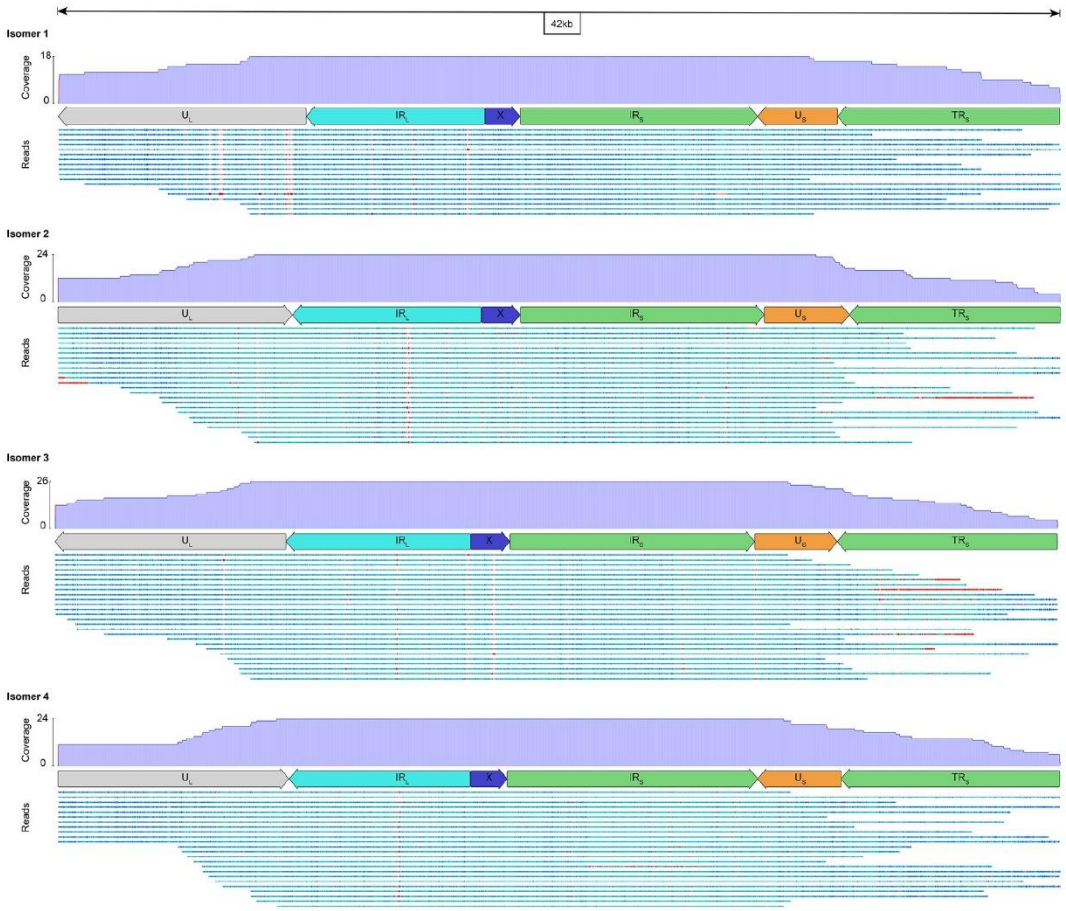


Fig. 6. OsHV-1 major isomer validation. Each panel corresponds to a 42 kb long OsHV-1 genomic region made of U_L fragment plus IR_L, X, IR_S, U_S and TR_S. The different panels correspond to the different OsHV-1 isomers made by permutation of U_L and U_S regions. Each panel is made of a diagram of reads coverage, a diagram of OsHV-1 region annotation and a diagram representing long-read alignment that span at least one fragment of the U_L and one fragment of the U_S. Read bases are color-coded from conserved (dark blue), to less conserved (light blue) with gaps in sequence colored in red.

Comparison of Illumina OsHV-1 NR genome and ONT OsHV-1 genome

Since it is almost impossible to assemble a complete OsHV-1 genome from short-read technology due to the presence of large inverted repeats, and because long-read technology tends to produce longer OsHV-1 genome than expected, both technologies still requires manual editing to join the different fragments or to remove extra repeats to obtain a genome structure as accurate as possible. Furthermore, ONT is known to produce a high error rate on raw sequences compared to the Illumina technology. Here, rather than looking at the raw sequence error rate, we attempted to quantify the residual error by comparing to each other, the polished ONT OsHV-1 genome, and the Illumina OsHV-1 NR genome. To do so, we used a pairwise alignment and were able to identify a total of 187 bp difference between the two genomes (Fig. 7).

Those differences could be classified in three categories: (i) SNPs, (ii) copy number variation of a 115 bp repeat and (iii) large homopolymer size variation. We observed a total of 31 SNPs between the two genomes, 24 SNPs correspond to InDel, which take place at the extremity of small homopolymers (from 5 to 11 bp), and 7 SNPs were distributed elsewhere in the genome. The 115 bp copy number variation takes place into the IR_S region. While the Illumina OsHV-1 genome contains two copies of the 115 bp sequence, the ONT OsHV-1 genome harbour three copies of this repetition (Fig. S3). Finally, we observed an homopolymer size variation in the IR_L region corresponding to 41 bp insertion in the ONT OsHV-1 genome compared to Illumina OsHV-1 genome. Since sequencing error in homopolymer is a well-known caveat of Nanopore sequencing [93], and because long homopolymer are known to be error prone also in Illumina sequencing [92, 94], we have performed a careful analysis of the read alignment for

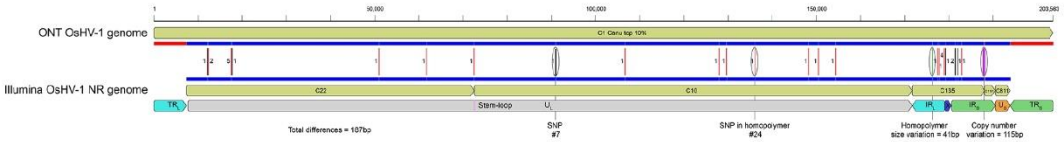


Fig. 7. ONT and Illumina OsHV-1 whole-genome comparison. A total of 187 bp differences are identified between both genomes. Overall, 7 bp correspond to SNPs (black bar), 24 bp correspond to SNPs in small homopolymer (red bar), 41 bp correspond to a homopolymer size variation (green bar), and 115 bp correspond to a copy number variation (purple bar). Numbers next to the bars correspond to the number of variable bases at that position.

both technologies in this specific region. Interestingly, alignment observation in this large homopolymer identify heterogeneity in homopolymer size at the level of individual reads for both sequencing technologies. While size variation for Illumina reads ranged from 7 to 22 bp, size variation for Nanopore reads ranged from 5 to 98 bp (Fig. S2A and B respectively).

Because of the impossibility to accurately distinguish error in this large homopolymer for both Illumina and Nanopore sequencing, and because we cannot accurately assign the number of Cs inside this large homopolymer, we have decided to keep the most represented number of Cs for each of the sequencing techniques. We therefore did not consider this large homopolymer in our final pairwise homology comparison of both sequencing technologies.

Finally, if we disregard the additional 115 bp repeat highlighted by ONT sequencing, the two *de novo* assembled genomes only differ by a total of 31 bp and show 99.98% of nucleotide identity.

OsHV-1 phylogenetic analysis

In order to determine the relationships of the OsHV-1 genomes from this study with those known so far, we constructed a phylogenetic tree from the alignment of the two genomes we generated in the present study (viruses collected in 2016), seven OsHV-1 genomes available in public databases (viruses collected before 2016) and four representatives of the 21 genomes acquired by Delmotte-Pelletier *et al.* [34] (viruses collected in 2018). This analysis confirms that the OsHV-1 virus sampled in Thau lagoon

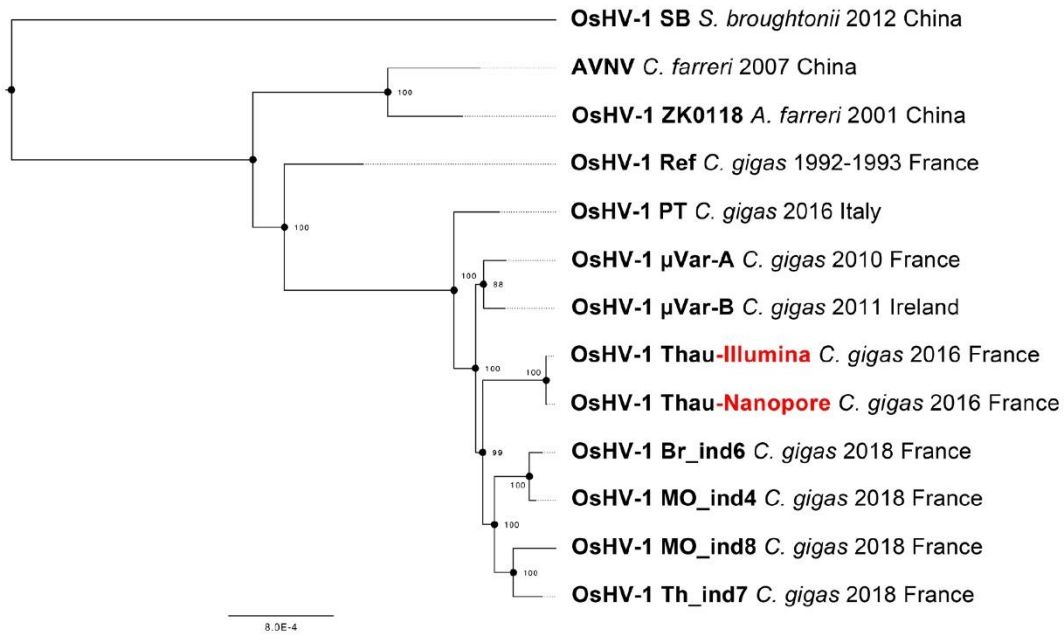


Fig. 8. Phylogenetic tree of the 11 OsHV-1 NR version of the genomes available in public database, plus the Nanopore and Illumina NR genomes generated in the present study (red). The tree was mid-point rooted. Bootstrap values are shown at nodes (100 replicates).

(France) in 2016 belongs to µVar clade and that genomes reconstructed from the same sample with either the Nanopore or the Illumina sequencing data are really close to each other (Fig. 8).

Non-OsHV-1 contigs' classification

Interestingly, our TFF-based purification process led to the characterization of other viral populations present in oyster tissues. To obtain a global overview of this viral metagenome, we made an extensive taxonomic classification of the non-OsHV-1 contigs produced by both ONT and Illumina sequencing. While the Nanopore sequencing produced 22 non-OsHV-1 contigs (from 14 374 to 163 592 bp) (Fig. 4b, Table S2A), Illumina sequencing produced 222 non-OsHV-1 contigs with size greater than 10 kb (from 10 017 bp to 172 478 bp) (Fig. 4a, Table S2B). Among the 22 Nanopore contigs 14 were also detected in the Illumina non-OsHV-1 contigs (Table S2A). Nanopore identified 20 contigs coding for putative bacteriophages, 1 contig coding for a CRESS DNA virus and 1 contig coding for the *C. gigas* mitochondrial DNA. Illumina identified 202 contigs' coding for putative bacteriophages, 19 contigs with no hit and 1 contig coding for the *C. gigas* mitochondrial DNA (Table S2B).

DISCUSSION

Viral whole-genome sequencing has recently become a standard in molecular epidemiology and evolutionary genomics, however this approach can be challenging when implemented on DNA viruses with large and complex genomes such as Herpesviruses [95, 96]. Indeed, it is often difficult to purify these types of viruses in sufficient quantity to allow whole-genome sequencing and reconstruction. Moreover, whole-genome assembly is a real burden considering the viral complexity due to the presence of long inverted repeated regions and structural variants. This is even more challenging to achieve for uncultured viruses derived from clinical samples or from livestock.

We implemented TFF to purify and to concentrate OsHV-1. To this end, we tested TFF devices with different pore sizes and quantified the virus load and infectious capacity of each fraction. Our results showed that TFF device with a MWCO of 500 kDa was able to retain and to concentrate most the OsHV-1 infectious particles contained in the inoculum. Interestingly, viral load quantification by real-time PCR and epifluorescence microscopy were of the same order of magnitude suggesting that TFF not only concentrates the virus but also eliminates most of the free DNA in the starting inoculum as previously suggested by Loewe *et al.* [60]. This is supported by the tremendous increase of the signal (OsHV-1 reads) to noise (oyster reads) ratio obtained in the sequencing data. Indeed, more than 60% of the sequenced reads belonged to OsHV-1 whereas, in previous work, the ratio obtained from infected tissues ranged from 0,01% to 13,21% [31, 34]. While other virus enrichment methods like PCR or capture-based approaches can reach up to 100% of viral data, the TFF approach does not suffer from amplification or hybridization bias caused by sequence variation over time due to genomic divergence [97]. We believe this method will be very helpful to improve the quantification of infectious viral particles in samples from organisms for which there is no cell culture available. Indeed, for tissue samples, the quantification of viruses is mainly based on qPCR approaches, which quantify the genome copy number but not the number of infectious viral particles. Since infected tissues could contain viral genomes undergoing replication or non-canonical genomes [48], the quantification of genome packaged in viral particle reflect a more robust estimate of the true quantification of viable virion in the sample. In addition, we have shown that TFF is an alternative particle enrichment method to centrifugation-based purification. TFF has many advantages including low cost, transportability, rapid purification, and enrichment in infectious particles (i.e. 500 kDa for OsHV-1). We therefore demonstrate, that TFF is a powerful tool to purify unculturable viruses from clinical and environmental isolates.

Then, we performed a high-quality DNA extraction that allowed us to compare the capabilities of long- and short-read sequencing technologies (Illumina and Oxford Nanopore Technologies) to *de novo* assemble OsHV-1 whole genome.

To date and to our knowledge, there is no maintained bioinformatic pipeline for reference-free and *de novo* assembly of long DNA viruses with large inverted-repeat like OsHV-1 for either short- or long-read sequence. Usually, long- and short-read sequencing technologies are combined to *de novo* assemble Herpesvirus whole genomes [48, 98]. Conversely to what is usually done, we have successfully tested two bioinformatic pipelines to independently analyse the long- and short-read datasets, and to reference-free and *de novo* assemble two OsHV-1 whole genomes. The *de novo* assembly of the short-read pipeline produced five OsHV-1 contigs that have been manually scaffolded to produce a complete OsHV-1 genome. The *de novo* assembly of the long-read pipeline produced one large contig that has been manually curated to remove one extra duplicated region. Because of uncertainty with the short-read assembly to position duplicated regions relative to unique region, we have depicted the Illumina assembly as a Non-Redundant (NR) genome, with only one copy of each R_L and R_S, as frequently done in the literature [34, 77, 78, 99]. In contrast, the long-read assembly resulted in a genome with both inverted repeats with the help of a minimal manual edit.

Pairwise comparison of the two assembled OsHV-1 genomes showed that the Nanopore genome was 99,98% similar to its Illumina counterpart. Our results demonstrate that Nanopore sequencing is now a mature sequencing technology for the small-scale genomic characterization of difficult long DNA virus with large inverted repeat. This breakthrough sequencing technology coupled with TFF virus purification is now a promising solution to drastically increase the number of sequenced OsHV-1 genomes allowing in-depth molecular evolution and epidemiology characterization.

In addition to being accurate at the base level, ONT sequencing also allowed to resolve large-scale genomic variation. For the first time here, the Nanopore sequencing allowed the identification of one extra copy of a 115 bp multi-copy repetition inside the IR_S region that was not previously resolved by short-read approach. This long-read sequencing approach also confirmed the presence of the four major OsHV-1 isomers in viral particles for the first time since the acquisition of the first OsHV-1 genome [7]. Isomers are characterized by permutation of the unique regions U_L and U_S with respect to the IR_L-X-IR_S fragment. Some of the long-read obtained with Nanopore sequencing were large enough to cover one part of the two unique regions allowing us to quantify and confirm that the different OsHV-1 isomers were present in a stoichiometric proportion among viral particles. Interestingly, this observation is different from the one made in herpes simplex virus 1 (HSV-1), where four isomers have also been characterized, but with variable tropisms specific to cell lines and tissues [100]. Since the presence of isomeric genome in Herpesvirus is tightly linked to its replication and recombination process [98, 101, 102], having access to isomer frequency along with full genome characterization will allow in the near future a better understanding of the complex life cycle as well as of the evolutionary processes of this type of virus.

Finally, and even if this was not the primarily goal of this study, TFF allowed to co-purify multiple DNA viruses along with OsHV-1. For this specific application of viral metagenomics, Illumina allowed, due to the huge sequencing depth, to taxonomically identify ten times more virus than Nanopore sequencing for contigs with size larger than 10 kb. The great majority of identified DNA virus belong to bacteriophage. It is interesting to note that some of them are phages infecting *Vibrio*, bacteria, which have been described as colonizing moribund oysters during POMS outbreak [25, 103], others are Pseudoalteromonas and Aeromonas phages, pathogens of bacteria considered as beneficial for oysters [104, 105]. Of course, our data does not disentangle the potential role of these phages in POMS but raised questions about the necessity of investigating the virome of healthy and diseased oysters as it has been done for the bacteriome [25].

CONCLUSION

This first evaluation of Nanopore sequencing for OsHV-1 genome characterization is extremely promising. Coupled with the appropriate bioinformatics pipeline this long-read sequencing approach enables identification of structural variation not observed with the short-read approach and allows characterization and quantification of genomic isomers. As shown in the present study, Nanopore sequencing still struggles to accurately sequence homopolymeric region. However, Nanopore is currently updating software for base calling as well as flow cell and chemistry to improve the resolution of the homopolymeric region [93, 106]. Illumina as well as Nanopore present difficulties to accurately resolve very long homopolymer (>20 bp) in GC-rich regions. It is thus of prime importance to focus efforts on methodological development for a better resolution of those regions that likely play a role for the diversity, replication and the stability of the viral genome [35, 107, 108].

The next challenge lies in combining TFF purification of virus populations harboured by a single individual and long-read sequencing to resolve viral haplotype phasing and better understand intra individual viral dynamics. Indeed, most of the whole-genome studies conducted on Herpesviruses to date only generate a consensus view of the viral population rather than a reconstruction of the viral haplotypes [96]. However, recent studies have shown that viral haplotype phasing could prove to be crucial in understanding the genetic determinants and viral evolution underpinning virulence, pathogenesis and drug resistance [109–111].

Funding information
J.D. was supported by a grant from the University of Montpellier. C.P. was supported by a grant from the Ifremer Scientific Board and from the Nouvelle-Aquitaine region. A.D.M. was financially supported by Ifremer, by EU DG SANTE through the European Reference Laboratory for mollusc diseases, and by DGAL through the National Reference Laboratory for Mollusc Diseases. This work benefited from support from the 'Laboratoire d'excellence' (LabEx) CeMEB, through the exploratory research project HaploFit and the use of the 'environmental genomics' facility (http://www.labex-cemeb.org/fr/genomique-environnementale-1). This work was also supported by the Ifremer Scientific Board, through the HemoVir and GT-oyster projects and by the 'Fond Européen pour les Affaires Maritimes et la Pêche' (FEAMP, GESTINNOV project n°PFEA470020FA1000007). The authors acknowledge the Pôle de Calcul et de Données Marines (PCDM; https://wwz.ifremer.fr/en/Research-Technology/Research-Infrastructures/Digital-infrastructures/Computation-Centre) for providing DATARMOR computing and storage resources. Some data were produced through the GenSeq technical facilities of the « Institut des Sciences de l'Evolution de Montpellier » with the support of LabEx CeMEB, an ANR 'Investissements d'avenir' programme (ANR-10-LABX-04–01). This study is set within the framework of the 'Laboratoires d'Excellences (LABEX)' TULIP (ANR-10-LABX-41). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Acknowledgements
We thank the staff of the Ifremer station at La Tremblade (ASIM), Frédéric Girardin and his team (Plateforme des Mollusques Marins de La Tremblade PMMLT) and Christophe Stavrakakis and his team (Plateforme des Mollusques Marins de Bouin PMMB). We also thank Louis Boismorand and Marc Leroy for technical assistance. We are also grateful to Jean-Christophe Avarre and Marie-ka Tilak (ISEM) for fruitful discussions and advice concerning Nanopore sequencing technology.

Author contributions
J.M.E. and B.M. conceptualized the study; B.M., N.F., M.B. and J.M.E. collected and processed the samples, J.D. and J.M.E. performed Nanopore sequencing, A.D.M. and G.C. did the bioinformatic analysis for Nanopore data, C.P. and G.C. did the bioinformatic analysis of the Illumina data, J.M.E. and G.C. drafted the manuscript and M.J. and B.M. corrected it. All authors read and approved the final manuscript.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
Not applicable, experimental works with bivalves are not currently subjected to ethical approval by Ifremer.

References

1. **Hall RJ**, **Wang J**, **Todd AK**, **Bissielo AB**, **Yen S**, ***et al.*** Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J Virol Methods* 2014;195:194–204.
2. **Kumar A**, **Murthy S**, **Kapoor A**. Evolution of selective-sequencing approaches for virus discovery and virome analysis. *Virus Res* 2017;239:172–179.
3. **Gaafar YZA**, **Ziebell H**, **Pappu HR**. Comparative study on three viral enrichment approaches based on RNA extraction for plant virus/viroid detection using high-throughput sequencing. *PLoS ONE* 2020;15:e0237951.
4. **Deng X**, **Achari A**, **Federman S**, **Yu G**, **Somasekar S**, ***et al.*** Metagenomic sequencing with spiked primer enrichment for viral diagnostics and genomic surveillance. *Nat Microbiol* 2020;5:443–454.
5. **Simner PJ**, **Miller S**, **Carroll KC**. Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clin Infect Dis* 2018;66:778–788.
6. **Gardy JL**, **Loman NJ**. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat Rev Genet* 2018;19:9–20.
7. **Davison AJ**, **Trus BL**, **Cheng N**, **Steven AC**, **Watson MS**, ***et al.*** A novel class of herpesvirus with bivalve hosts. *J Gen Virol* 2005;86:41–53.
8. **EFSA**. Oyster mortality. *EFSA Journal* 2015:1.
9. **Azéma P**, **Lamy J-B**, **Boudry P**, **Renault T**, **Travers M-A**, ***et al.*** Genetic parameters of resistance to vibrio aestuarianus, and oshv-1 infections in the pacific oyster, crassostrea gigas, at three different life stages. genetics selection evolution. *BioMed Central* 2017:1–16.
10. **Dégremont L**, **Morga B**, **Maurouard E**, **Travers M-A**. Susceptibility variation to the main pathogens of *Crassostrea gigas* at the larval, spat and juvenile stages using unselected and selected oysters to OsHV-1 and/or *V. aestuarianus*. *J Invertebr Pathol* 2021;183:107601.
11. **Xia J**, **Bai C**, **Wang C**, **Song X**, **Huang J**. Complete genome sequence of *Ostreid herpesvirus*-1 associated with mortalities of *Scapharca broughtonii* broodstocks. *Virol J BioMed Central* 2015;12:110–119.
12. **Cáceres-Martínez J**, **Vásquez-Yeomans R**, **Danigo P**, **Reyes-Roel C**. Histological alterations in pacific oysters crassostrea gigas that survived a summer mortality event in Baja California, Mexico. *J Aquat Anim Health* 2018;30:31–38.
13. **Hwang JY**, **Park JJ**, **Yu HJ**, **Hur YB**, **Arzul I**, ***et al.*** Ostreid herpesvirus 1 infection in farmed Pacific oyster larvae Crassostrea gigas (Thunberg) in Korea. *J Fish Dis* 2013;36:969–972.
14. **Jenkins C**, **Hick P**, **Gabor M**, **Spiers Z**, **Fell SA**, ***et al.*** Identification and characterisation of an ostreid herpesvirus-1 microvariant (OsHV-1 µ-var) in Crassostrea gigas (Pacific oysters) in Australia. *Dis Aquat Organ* 2013;105:109–126.
15. **Mortensen S**, **Strand Å**, **Bodvin T**, **Alfjorden A**, **Skår CK**, ***et al.*** *Summer mortalities* and detection of ostreid herpesvirus microvariant in Pacific oyster *Crassostrea gigas* in Sweden and Norway. *Dis Aquat Organ* 2016;117:171–176.
16. **Pernet F**, **Barret J**, **Le Gall P**, **Corporeau C**, **Dégremont L**, ***et al.*** Mass mortalities of pacific oysters *Crassostrea gigas* reflect infectious diseases and vary with farming practices in the Mediterranean Thau lagoon, France. *Aquacult Environ Interact* 2012;2:215–237.
17. **Prado-Alvarez M**, **Darmody G**, **Hutton S**, **O'Reilly A**, **Lynch SA**, ***et al.*** Occurrence of oshv-1 in crassostrea gigas cultured in ireland during an exceptionally warm summer. *Selection of Less Susceptible Oysters Front Physiol* 2016;7:492.
18. **Segarra A**, **Pépin JF**, **Arzul I**, **Morga B**, **Faury N**, ***et al.*** Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. *Virus Res* 2010;153:92–99.
19. **Garcia C**, **Thébault A**, **Dégremont L**, **Arzul I**, **Miossec L**, ***et al.*** Ostreid herpesvirus 1 detection and relationship with *Crassostrea gigas* spat mortality in France between 1998 and 2006. *Vet Res* 2011;42:73.
20. **Renault T**, **Faury N**, **Barbosa-Solomieu V**, **Moreau K**. Suppression substractive hybridisation (SSH) and real time PCR reveal differential gene expression in the Pacific cupped oyster, *Crassostrea gigas*, challenged with *Ostreid herpesvirus* 1. *Dev Comp Immunol* 2011;35:725–735.
21. **Petton B**, **Pernet F**, **Robert R**, **Boudry P**. Temperature influence on pathogen transmission and subsequent mortalities in juvenile Pacific oysters *Crassostrea gigas*. *Aquacult Environ Interact* 2013;3:257–273.
22. **Petton B**, **Boudry P**, **Alunno-Bruscia M**, **Pernet F**. Factors influencing disease-induced mortality of Pacific oysters *Crassostrea gigas*. *Aquacult Environ Interact* 2015;6:205–222.
23. **Le Roux F**, **Wegner KM**, **MPTI**. Oysters and vibrios as a model for disease dynamics in wild animals. *Trends Microbiol* 2016;24:568–580.
24. **Delisle L**, **Pauletto M**, **Vidal-Dupiol J**, **Petton B**, **Bargelloni L**, ***et al.*** High temperature induces transcriptomic changes in *Crassostrea gigas* that hinders progress of Ostreid herpesvirus (OsHV-1) and promotes survival . *J Exp Biol* 2020.
25. **de Lorgeril J**, **Lucasson A**, **Petton B**, **Toulza E**, **Montagnani C**, ***et al.*** Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. *Nat Commun* 2018;9:1–14.
26. **Rubio T**, **Oyanedel D**, **Labreuche Y**, **Toulza E**, **Luo X**, ***et al.*** Species-specific mechanisms of cytotoxicity toward immune cells determine the successful outcome of *Vibrio* infections. *Proc Natl Acad Sci U S A* 2019;116:14238–14247.
27. **Petton B**, **Destoumieux-Garzón D**, **Pernet F**, **Toulza E**, **de Lorgeril J**, ***et al.*** The Pacific Oyster Mortality Syndrome, a Polymicrobial and Multifactorial Disease: State of Knowledge and Future Directions. *Front Immunol* 2021;12:630343.
28. **Barbosa Solomieu V**, **Renault T**, **Travers M-A**. Mass mortality in bivalves and the intricate case of the Pacific oyster, *Crassostrea gigas*. *J Invertebr Pathol* 2015;131:2–10.
29. **Renault T**, **Moreau P**, **Faury N**, **Pépin J-F**, **Segarra A**, ***et al.*** Analysis of clinical ostreid herpesvirus 1 (malacoherpesviridae) specimens by sequencing amplified fragments from three virus genome areas.2nd ed. *J Virol* 2012;86(10):5942–5947.
30. **Renault T**, **Tchaleu G**, **Faury N**, **Moreau P**, **Segarra A**, ***et al.*** Genotyping of a microsatellite locus to differentiate clinical *Ostreid herpesvirus* 1 specimens. *Vet Res* 2014;45:3.
31. **Burioli EAV**, **Prearo M**, **Houssin M**. Complete genome sequence of *Ostreid herpesvirus* type 1 µVar isolated during mortality events in the Pacific oyster *Crassostrea gigas* in France and Ireland. *Virology* 2017;509:239–251.
32. **Abbadi M**, **Zamperin G**, **Gastaldelli M**, **Pascoli F**, **Rosani U**, ***et al.*** Identification of a newly described OsHV-1 µvar from the North Adriatic Sea (Italy). *J Gen Virol* 2018;99:693–703.
33. **Morga B**, **Jacquot M**, **Pelletier C**, **Chevignon G**, **Dégremont L**, ***et al.*** Genomic diversity of the ostreid herpesvirus type 1 across time and location and among host species. *Front Microbiol* 2021;12:711377.
34. **Delmotte J**, **Pelletier C**, **Morga B**, **Galinier R**, **Petton B**, ***et al.*** Genetic diversity and connectivity of the *Ostreid herpesvirus* 1

populations in France: a first attempt to phylogeographic inference for A marine mollusc disease. *Virus Evol* 2022;8:veac039.

35. **Kuny CV**, **Szpara ML**. Alphaherpesvirus genomics: past, present and future. *Curr Issues Mol Biol* 2021;42:41–80.
36. **Riaz N**, **Leung P**, **Barton K**, **Smith MA**, **Carswell S**, *et al*. Adaptation of Oxford Nanopore technology for hepatitis C whole genome sequencing and identification of within-host viral variants. *BMC Genomics* 2021;22:148–12.
37. **Rios G**, **Lacoux C**, **Leclercq V**, **Diamant A**, **Lebrigand K**, *et al*. Monitoring SARS-CoV-2 variants alterations in Nice neighborhoods by wastewater nanopore sequencing. *Lancet Reg Health Eur* 2021;10:100202.
38. **Toh X**, **Wang Y**, **Rajapakse MP**, **Lee B**, **Songkasupa T**, *et al*. Use of nanopore sequencing to characterize african horse sickness virus (AHSV) from the African horse sickness outbreak in thailand in 2020. *Transbound Emerg Dis* 2022;69:1010–1019.
39. **Chan WS**, **Au CH**, **Lam HY**, **Wang CLN**, **Ho DN-Y**, *et al*. Evaluation on the use of Nanopore sequencing for direct characterization of coronaviruses from respiratory specimens, and a study on emerging missense mutations in partial RdRP gene of SARS-CoV-2. *Virol J* 2020;17:183–13.
40. **King J**, **Harder T**, **Beer M**, **Pohlmann A**. Rapid multiplex MinION nanopore sequencing workflow for Influenza A viruses. *BMC Infect Dis* 2020;20:648–8.
41. **Batista FM**, **Stapleton T**, **Lowther JA**, **Fonseca VG**, **Shaw R**, *et al*. Whole genome sequencing of hepatitis a Virus Using A PCR-Free single-molecule nanopore sequencing approach. *Front Microbiol* 2020;11:874.
42. **Faria NR**, **Quick J**, **Claro IM**, **Thézé J**, **de Jesus JG**, *et al*. Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature* 2017;546:406–410.
43. **Eltom KH**, **Althoff AC**, **Hansen S**, **Böhlken-Fascher S**, **Yousif A**, *et al*. Differentiation of capripox viruses by nanopore sequencing. *Vaccines* 2021;9:351.
44. **Radukic MT**, **Brandt D**, **Haak M**, **Müller KM**, **Kalinowski J**. Nanopore sequencing of native adeno-associated virus (AAV) single-stranded DNA using a transposase-based rapid protocol. *NAR Genomics and Bioinformatics* 2020;2.
45. **Kulikov EE**, **Golomidova AK**, **Babenko VV**, **Letarov AV**. A simple method for extraction of the horse feces virome DNA, suitable for oxford nanopore sequencing. *Microbiology* 2020;89:246–249.
46. **Ben Chehida S**, **Filloux D**, **Fernandez E**, **Moubset O**, **Hoareau M**, *et al*. Nanopore sequencing is a credible alternative to recover complete genomes of geminiviruses. *Microorganisms* 2021;9:903.
47. **Karamitros T**, **Harrison I**, **Piorkowska R**, **Katzourakis A**, **Magiorkinis G**, *et al*. De novo assembly of human herpes virus type 1 (HHV-1) genome, mining of non-canonical structures and detection of novel drug-resistance mutations using short- and long-read next generation sequencing technologies. *PLoS One* 2016;11:e0157600.
48. **Karamitros T**, **van Wilgenburg B**, **Wills M**, **Klenerman P**, **Magiorkinis G**. Nanopore sequencing and full genome de novo assembly of human cytomegalovirus TB40/E reveals clonal diversity and structural variations. *BMC Genomics* 2018;19:577–12.
49. **Vereecke N**, **Carnet F**, **Pronost S**, **Vanschandevijl K**, **Theuns S**, *et al*. Genome sequences of equine herpesvirus 1 strains from a European outbreak of neurological disorders linked to a horse gathering in Valencia, Spain, in 2021. *Microbiol Resour Announc* 2021;10:e00333-21.
50. **Fuandila NN**, **Gosselin-Grenet A-S**, **Tilak M-K**, **Bergmann SM**, **Escoubas J-M**, *et al*. Structural variation turnovers and defective genomes: key drivers for the in vitro evolution of the large double-stranded DNA koi herpesvirus (KHV). Cold Spring Harbor Laboratory. *bioRxiv* 2022;2.
51. **Vinaya Kumar K**, **Shekhar MS**, **Otta SK**, **Karthic K**, **Ashok Kumar J**, *et al*. First report of a complete genome sequence of White spot syndrome virus from India. *Genome Announc* 2018;6:e00055-18.
52. **Gallagher MD**, **Matejusova I**, **Nguyen L**, **Ruane NM**, **Falk K**, *et al*. Nanopore sequencing for rapid diagnostics of salmonid RNA viruses. *Sci Rep* 2018;8:16307–16309.
53. **Liem M**, **Regensburg-Tuïnk T**, **Henkel C**, **Jansen H**, **Spaink H**. Microbial diversity characterization of seawater in a pilot study using Oxford Nanopore Technologies long-read sequencing. *BMC Res Notes* 2021;14:42–47.
54. **Beaulaurier J**, **Luo E**, **Eppley JM**, **Uyl PD**, **Dai X**, *et al*. Assembly-free single-molecule sequencing recovers complete virus genomes from natural microbial communities. *Genome Res* 2020;30:437–446.
55. **Urban L**, **Holzer A**, **Baronas JJ**, **Hall MB**, **Braeuninger-Weimer P**, *et al*. Freshwater monitoring by nanopore sequencing. *Elife* 2021;10:e61504.
56. **Wickramasinghe SR**, **Kalbfuss B**, **Zimmermann A**, **Thom V**, **Reichl U**. Tangential flow microfiltration and ultrafiltration for human influenza A virus concentration and purification. *Biotechnol Bioeng* 2005;92:199–208.
57. **Michalsky R**, **Passarelli AL**, **Pfromm PH**, **Czermak P**. Purification of the baculovirus Autographa californica M nucleopolyhedrovirus by tangential flow ultrafiltration. *Desalination* 2009;245:694–700.
58. **Eilts F**, **Harnischfeger J**, **Loewe D**, **Wolff MW**, **Salzig D**, *et al*. Production of *Baculovirus* and Stem Cells for Baculovirus-Mediated Gene Transfer into Human Mesenchymal Stem Cells. *Methods Mol Biol* 2021;2183:367–390.
59. **Negrete A**, **Pai A**, **Shiloach J**. Use of hollow fiber tangential flow filtration for the recovery and concentration of HIV virus-like particles produced in insect cells. *J Virol Methods* 2014;195:240–246.
60. **Loewe D**, **Grein TA**, **Dieken H**, **Weidner T**, **Salzig D**, *et al*. Tangential Flow Filtration for the Concentration of Oncolytic Measles Virus: The Influence of Filter Properties and the Cell Culture Medium. *Membranes* 2019;9:E160.
61. **Shi H**, **Pasco EV**, **Tarabara VV**. Membrane-based methods of virus concentration from water: a review of process parameters and their effects on virus recovery. *Environ Sci: Water Res Technol* 2017;3:778–792.
62. **Cordier C**, **Stavrakakis C**, **Morga B**, **Degrémont L**, **Voulgaris A**, *et al*. Removal of pathogens by ultrafiltration from sea water. *Environ Int* 2020;142:105809.
63. **Beaulaurier J**, **Luo E**, **Eppley JM**, **Uyl PD**, **Dai X**, *et al*. Assembly-free single-molecule sequencing recovers complete virus genomes from natural microbial communities. *Genome Res* 2020;30:437–446.
64. **Azéma P**, **Lamy J-B**, **Boudry P**, **Renault T**, **Travers M-A**, *et al*. Genetic parameters of resistance to Vibrio aestuarianus, and OsHV-1 infections in the Pacific oyster, *Crassostrea gigas*, at three different life stages. *Genet Sel Evol* 2017;49:23.
65. **Dégremont L**, **Bédier E**, **Soletchnik P**, **Ropert M**, **Huvet A**, *et al*. Relative importance of family, site, and field placement timing on survival, growth, and yield of hatchery-produced Pacific oyster spat (*Crassostrea gigas*). *Aquaculture* 2005;249:213–229.
66. **Schikorski D**, **Renault T**, **Saulnier D**, **Faury N**, **Moreau P**, *et al*. Experimental infection of Pacific oyster Crassostrea gigas spat by ostreid herpesvirus 1: demonstration of oyster spat susceptibility. *Vet Res* 2011;42:13–27.
67. **Webb SC**, **Fidler A**, **Renault T**. Primers for PCR-based detection of ostreid herpes virus-1 (OsHV-1): Application in a survey of New Zealand molluscs. *Aquaculture* 2007;272:126–139.
68. **Leruste A**, **Bouvier T**, **Bettarel Y**. Enumerating viruses in coral mucus. *Appl Environ Microbiol* 2012;78:6377–6379.
69. **Patel A**, **Noble RT**, **Steele JA**, **Schwalbach MS**, **Hewson I**, **Fuhrman JA**. Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. Nat Protoc. Nature Publishing Group; 2007;2(2):269–76.
70. **Suquet M**, **de Kermoysan G**, **Araya RG**, **Queau I**, **Lebrun L**, *et al*. Anesthesia in pacific oyster, crassostrea gigas. *Aquat Living Resour EDP Sciences* 2009;22(1):29–34.

71. **Dupont S**, **Lokmer A**, **Corre E**, **Auguet J-C**, **Petton B**, ***et al***. Oyster hemolymph is a complex and dynamic ecosystem hosting bacteria, protists and viruses. *Anim Microbiome* 2020;2:12.
72. **Chen S**, **Zhou Y**, **Chen Y**, **Gu J**. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–i890.
73. **Langmead B**, **Salzberg SL**. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–359.
74. **Li H**, **Handsaker B**, **Wysoker A**, **Fennell T**, **Ruan J**, ***et al***. The sequence alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
75. **Nurk S**, **Meleshko D**, **Korobeynikov A**, **Pevzner PA**. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 2017;27:824–834.
76. **Altschul SF**, **Gish W**, **Miller W**, **Myers EW**, **Lipman DJ**. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
77. **Cunningham C**, **Gatherer D**, **Hilfrich B**, **Baluchova K**, **Dargan DJ**, ***et al***. Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens. *J Gen Virol* 2010;91:605–615.
78. **Morse AM**, **Calabro KR**, **Fear JM**, **Bloom DC**, **McIntyre LM**. Reliable detection of herpes simplex virus sequence variation by high-throughput resequencing. *Viruses* 2017;9:E226.
79. **Li H**. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 2018;34:3094–3100.
80. **Kolmogorov M**, **Bickhart DM**, **Behsaz B**, **Gurevich A**, **Rayko M**, ***et al***. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods* 2020;17:1103–1110.
81. **Koren S**, **Walenz BP**, **Berlin K**, **Miller JR**, **Bergman NH**, ***et al***. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res* 2017;27:722–736.
82. **Vaser R**, **Sović I**, **Nagarajan N**, **Šikić M**. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 2017;27:737–746.
83. **Danecek P**, **Bonfield JK**, **Liddle J**, **Marshall J**, **Ohan V**, ***et al***. Twelve years of SAMtools and BCFtools. *Gigascience* 2021;10:giab008.
84. **Arndt D**, **Grant JR**, **Marcu A**, **Sajed T**, **Pon A**, ***et al***. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 2016;44:W16–21.
85. **Buchfink B**, **Reuter K**, **Drost H-G**. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 2021;18:366–368.
86. **Katoh K**, **Standley DM**. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.
87. **Darling ACE**, **Mau B**, **Blattner FR**, **Perna NT**. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 2004;14:1394–1403.
88. **Lemoine F**, **Correia D**, **Lefort V**, **Doppelt-Azeroual O**, **Mareuil F**, ***et al***. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res* 2019;47:W260–W265.
89. **Criscuolo A**, **Gribaldo S**. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol* 2010;10:210–221.
90. **Darriba D**, **Taboada GL**, **Doallo R**, **Posada D**. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 2012;9:772–2.
91. **Guindon S**, **Dufayard J-F**, **Lefort V**, **Anisimova M**, **Hordijk W**, ***et al***. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;59:307–321.
92. **Ross MG**, **Russ C**, **Costello M**, **Hollinger A**, **Lennon NJ**, ***et al***. Characterizing and measuring bias in sequence data. *Genome Biol* 2013;14:20–R51.
93. **Delahaye C**, **Nicolas J**. Sequencing DNA with nanopores: Troubles and biases. *PLoS One* 2021;16:e0257521.
94. **Stoler N**, **Nekrutenko A**. Sequencing error profiles of Illumina sequencing instruments. *NAR Genom Bioinform* 2021;3:lqab019.
95. **Houldcroft CJ**, **Beale MA**, **Breuer J**. Clinical and biological insights from viral genome sequencing. *Nat Rev Microbiol* 2017;15:183–192.
96. **Renner DW**, **Szpara ML**. Impacts of genome-wide analyses on our understanding of human herpesvirus diversity and evolution. *J Virol* 2018;92:126–13.
97. **Sozhamannan S**, **Holland MY**, **Hall AT**, **Negrón DA**, **Ivancich M**, ***et al***. Evaluation of signature erosion in ebola virus due to genomic drift and its impact on the performance of diagnostic assays. *Viruses* 2015;7:3130–3154.
98. **López-Muñoz AD**, **Rastrojo A**, **Kropp KA**, **Viejo-Borbolla A**, **Alcamí A**. Combination of long- and short-read sequencing fully resolves complex repeats of herpes simplex virus 2 strain MS complete genome. *Microb Genom* 2021;7.
99. **Hage E**, **Wilkie GS**, **Linnenweber-Held S**, **Dhingra A**, **Suárez NM**, ***et al***. Characterization of human cytomegalovirus genome diversity in immunocompromised hosts by whole-genome sequencing directly from clinical specimens. *J Infect Dis* 2017;215:1673–1683.
100. **Mahiet C**, **Ergani A**, **Huot N**, **Alende N**, **Azough A**, ***et al***. Structural variability of the herpes simplex virus 1 genome in vitro and in vivo. *J Virol* 2012;86:8592–8601.
101. **Chou J**, **Roizman B**. Isomerization of herpes simplex virus 1 genome: identification of the cis-acting and recombination sites within the domain of the a sequence. *Cell* 1985;41:803–811.
102. **McVoy MA**, **Ramnarain D**. Machinery to support genome segment inversion exists in a herpesvirus which does not naturally contain invertible elements. *J Virol* 2000;74:4882–4887.
103. **Le Roux F**, **Wegner KM**, **Polz MF**. Oysters and Vibrios as a Model for Disease Dynamics in Wild Animals. *Trends Microbiol* 2016;24:568–580.
104. **Desriac F**, **El Harras A**, **Simon M**, **Bondon A**, **Brillet B**, ***et al***. Alterins Produced by Oyster-Associated *Pseudoalteromonas* Are Antibacterial Cyclolipopeptides with LPS-Binding Activity. *Mar Drugs* 2020;18:E630.
105. **Gibson LF**, **Woodworth J**, **George AM**. Probiotic activity of Aeromonas media on the Pacific oyster, Crassostrea gigas, when challenged with Vibrio tubiashii. *Aquaculture* 1998;169:111–120.
106. **Wang Y**, **Zhao Y**, **Bollas A**, **Wang Y**, **Au KF**. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol* 2021;39:1348–1365.
107. **Olson AT**, **Wang Z**, **Rico AB**, **Wiebe MS**. A poxvirus pseudokinase represses viral DNA replication via A pathway antagonized by its paralog kinase. *PLoS Pathog* 2019;15:e1007608.
108. **Castillo-Lizardo M**, **Henneke G**, **Viguera E**. Replication slippage of the thermophilic DNA polymerases B and D from the Euryarchaeota Pyrococcus abyssi. *Front Microbiol* 2014;5:403.
109. **Wu NC**, **De La Cruz J**, **Al-Mawsawi LQ**, **Olson CA**, **Qi H**, ***et al***. HIV-1 quasispecies delineation by tag linkage deep sequencing. *PLoS One* 2014;9:e97505.
110. **Monaco DC**, **Zapata L**, **Hunter E**, **Salomon H**, **Dilernia DA**. Resistance profile of HIV-1 quasispecies in patients under treatment failure using single molecule, real-time sequencing. *AIDS* 2020;34:2201–2210.
111. **Della Bartola M**, **Byrne S**, **Mullins E**. Characterization of potato virus Y isolates and assessment of nanopore sequencing to detect and genotype potato viruses. *Viruses* 2020;12:E478.

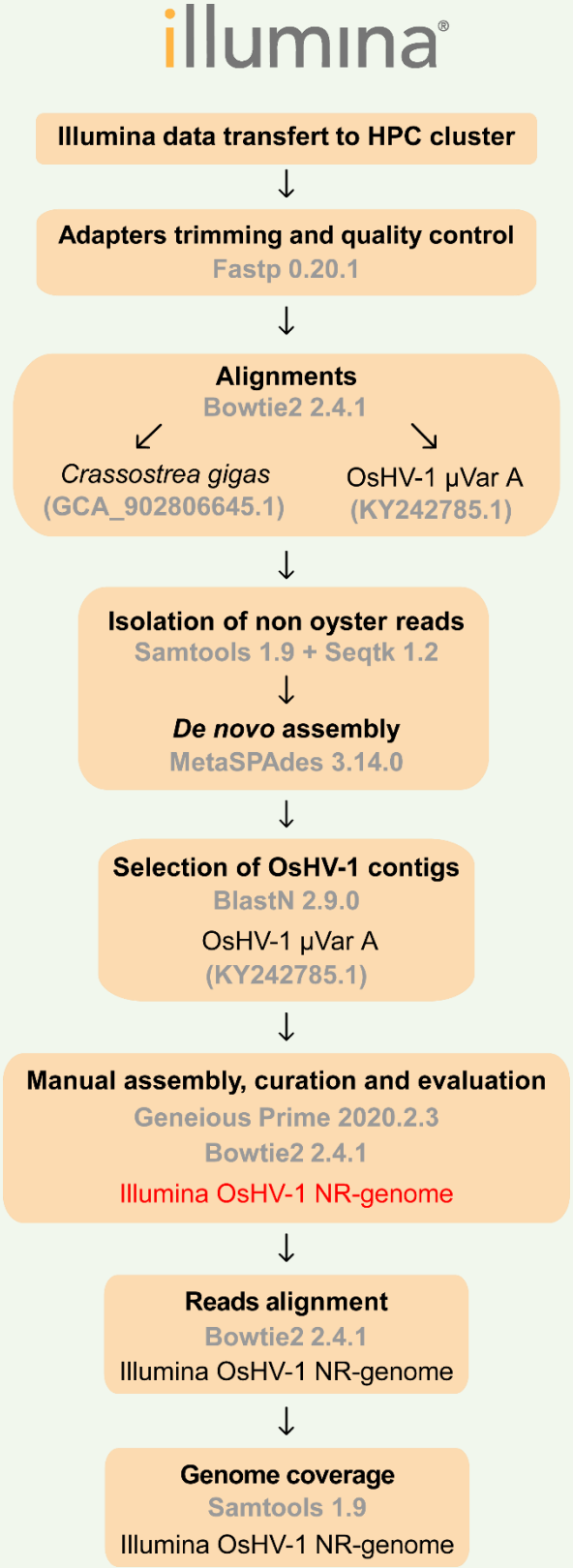
Messages clés :

- La filtration tangentielle est une bonne méthode pour purifier le virus OsHV-1. Elle permet **d'enrichir jusqu'à 10 fois** les particules virales présentes dans une suspension virale.

- Le séquençage *via* la technologie **Illumina** permet d'obtenir des séquences courtes d'ADN, avec un **faible taux d'erreurs** et une **profondeur optimale** pour les analyses de diversité virale. Au vu de ces caractéristiques, c'est la **technologie que j'ai sélectionné** pour séquencer l'ensemble des échantillons étudiés durant ma thèse.

- La technologie de séquençage **ONT** qui produit des longues séquences ADN est efficace pour l'assemblage OsHV-1 et permet de **s'affranchir des problèmes liés à l'architecture génomique** du virus. A l'heure actuelle le taux d'erreurs de cette technologie est trop élevé pour envisager des études de diversité du virus OsHV-1, mais c'est une **perspective prometteuse**.

- Le **pipeline bioinformatique** développé pour l'assemblage *de novo* de génomes d'OsHV-1 et l'analyse de sa diversité fonctionne.



Chapitre 3 – Etude de l’impact de l’espèce hôte sur la diversité virale

C. Pelletier

3.1 Résumé de l'article 2

En 1972, Farley décrivait pour la première fois, en Californie, aux Etats-Unis, un virus ayant une morphologie similaire à celle des herpesvirus, infectant une espèce de mollusque bivalve : l'huître Américaine *Magallana virginica*. Ce virus a été caractérisé comme appartenant aux Herpesviridae par microscopie électronique à transmission et nommé Ostreid Herpsvirus type 1 (OsHV-1). Très rapidement, OsHV-1 a été détecté au Royaume-Uni chez une autre espèce d'huître : l'huître plate Européenne *Ostrea edulis*. Ce virus a ensuite été découvert dans diverses espèces de bivalves marins à travers le monde, notamment l'huître du Pacifique *M. gigas*, l'huître plate européenne *O. edulis*, l'huître plate d'Antipodes *Ostrea angasi*, l'huître chilienne *Tiostrea chilensis*, les palourdes *Ruditapes decussatus* et *Ruditapes philippinarum* et la coquille Saint-Jacques *Pecten maximus*.

En France, deux espèces d'huîtres ont été infectées par le virus OsHV-1 : l'huître plate Européenne *O. edulis*, espèce native des côtes Européennes, et l'huître du Pacifique *M. gigas*, introduite depuis le Japon et le Canada au milieu du XX^{ème} siècle. Des évènements de transmission inter-espèce ont été suspecté, mais n'ont jamais pu être démontré. Actuellement aucune étude épidémiologique ne s'est intéressée à l'histoire évolutive des génotypes viraux retrouvés chez ces deux espèces. Peu d'informations sont disponibles sur le degré d'isolement génétique des génotypes viraux qui infectent ces deux espèces, de même que sur la fréquence des événements de transmission inter-espèces.

Cette étude utilise le séquençage haut-débit pour analyser la diversité génétique de l'OsHV-1 chez les huîtres Européennes et les huîtres du Pacifique en France et en Suède entre 2005 et 2010. Quarante nouveaux génomes d'OsHV-1 ont été séquencés et assemblés ce qui a permis *(i)* de décrire la diversité génomique de l'OsHV-1 dans ces deux espèces, *(ii)* d'évaluer l'isolement génétique entre les virus chez les deux espèces et *(iii)* de déterminer si la transmission inter-espèces de l'OsHV-1 s'est produite et dans quelle mesure. L'assemblage de génomes viraux a été effectuée grâce au pipeline bioinformatique développé dans le premier article de cette thèse et expliqué dans le chapitre 2. L'analyse de la diversité a été étudiée par une combinaison d'approches génétiques, génomiques et phylodynamiques.

Les résultats révèlent que les génomes de l'OsHV-1 présentent des variations structurales, y compris des délétions, des translocations génomiques et des inversions, suggérant que la recombinaison génétique pourrait jouer un rôle dans l'apparition de ces variations. De plus, les résultats indiquent que les génomes de l'OsHV-1 diffèrent en fonction de l'espèce hôte, avec une plus grande diversité génétique chez les huîtres du Pacifique, probablement en raison de leur plus grande population et des pratiques culturales. L'étude suggère que le virus OsHV-1 a évolué et s'est adapté induisant l'apparition

de deux lignées distinctes, l'une infectant les huîtres du Pacifique et l'autre infectant les huîtres européennes, avec une origine commune originaire des huîtres du Pacifique. Cependant, il n'y a que peu d'événements de transmission inter-espèces observés entre ces deux lignées.

Article 2

Genetic differentiation and host specialization among OsHV-1 infecting two oyster species in France

Camille Pelletier, Germain Chevignon, Nicole Faury, Isabelle Arzul, Céline Garcia, Bruno Chollet, Tristan Renault, Benjamin Morga, Maude Jacquot.

Genetic differentiation and host specialization among OsHV-1 infecting two oyster species in France

Camille Pelletier[1], Germain Chevignon[1], Nicole Faury[1], Isabelle Arzul[1], Céline Garcia[1], Bruno Chollet[1], Tristan Renault[2], Benjamin Morga[1#], Maude Jacquot[1#].

Affiliations

[1] Ifremer, RBE-ASIM, Unité Adaptations Santé des Invertébrés Marins, Avenue de Mus de Loup, 17390 La Tremblade, France.

[2] Ifremer, Département Ressources Biologiques et Environnement, Nantes, France

Running Head: Genetic diversity of OsHV-1 within two oyster species

#Address correspondence to Maude Jacquot, maude.jacquot@ifremer.fr, or Benjamin Morga, Benjamin.morga@ifremer.fr

Submitted in *Journal of Virology*.

Abstract

Cross-species transmission is one of the most significant causes of disease emergence in humans and other species. The Ostreid Herpesvirus type 1 has mostly been detected in the Pacific oyster *Magallana gigas* in the context of mortality events. However, it has been detected in other mollusc species including the flat oyster Ostrea edulis which raise questions about OsHV-1 host specialization. This study looked at the genetic differentiation of OsHV-1 in Pacific and European flat oysters, and the mechanisms that might underlie this genetic differentiation. To this end, high-throughput deep sequencing was used to analyze forty OsHV-1 genomes to characterize the genetic diversity and isolation of OsHV-1 viruses from both *O. edulis* and *M. gigas* in France as well as to assess the frequency of the virus cross-species transmission. To achieve this, comparative and population genomic, phylogenetic, and phylodynamic approaches were employed. Our findings highlight the significant influence of host species on the genetic diversity and differentiation among OsHV-1. The emergence of OsHV-1 in France seems to be traced back to the introduction of *M. gigas*, and the virus has since evolved into two distinct lineages following a cross-species event. Moreover, our analysis revealed selection signals in regions responsible for crucial biological functions and domains, indicating the potential for OsHV-1 to adapt and specialize in response to its respective host species. To get further insights into evolutionary processes and host-virus interactions, future research should focus on investigating the coevolution between OsHV-1 and a wider range of host species using phylogenetic approaches.

Importance

In 1972, a herpes-like virus was found infecting Crassostrea virginica in the USA during oyster experiments. Since the 1990s, Ostreid herpesvirus type 1 (OsHV-1) caused massive Pacific oyster (*M. gigas*) deaths, spreading to scallops, clams, and flat oysters. A recent study explored OsHV-1 lineages across hosts, questioning interspecies transmission frequency and host specialization impact on co-culture. Advanced genetics and computational tools were used to determine the frequency and impact of cross-species transmission between Pacific and European flat oysters. Results suggest rare but significant transmission events, generating a unique OsHV-1 lineage for flat oysters, indicating host specialization. This research unveils insights into viral evolution and interactions with diverse hosts.

Introduction

The study of multi-host-pathogen systems is crucial to improve our understanding of the spread and emergence of infectious diseases, and to develop effective strategies to address animal health problems. Many pathogens are able of multiplying themselves in multiple host species, and multiple pathogens genotypes may exhibit different phenotypic traits (1, 2). Understanding the partitioning of pathogens genetic diversity among host species is particularly important for disease transmission when pathogens have the ability to infect both wildlife and livestock (2, 3). Disentangling epidemiological cycles in such systems is often challenging, yet investigating how the genetic diversity of pathogens is partitioned among hosts species is key to properly assess and manage disease risk (1, 4). Pathogen genotype frequencies within host populations can vary depending on, for example, their transmission efficiency within and between host populations (5, 6), while barriers to transmission may result in population-specific patterns of infection for different genotypes (7–9).

Herpesviruses are an order of double-stranded DNA viruses commonly found across multiple species of invertebrates (*e.g.* mollusks) and vertebrates (*e.g.* fish, birds, pigs, horses, humans). However, they are characterized by their narrow host range and are known to coevolve with their hosts, leading to a long-standing relationship between virus and host species (10–12). Nevertheless, there have been several reports of cross-species transmission (CST) events in both vertebrates (*e.g.* transmission of *cercopithecine herpesvirus 1* from primates to humans (13)) and invertebrates such as mollusks (*e.g.* transmission of *ostreid herpesvirus 1* from Pacific oysters to scallops or flat oysters (10, 11)). Herpesviruses affecting mollusks are restricted to the *Malacoherpesviridae* family which currently includes two species, the *abalone herpesvirus* (AbHV) infecting gastropods and the *ostreid herpesvirus 1* (OsHV-1) infecting bivalves. Several OsHV-1 variants have been reported infecting a wide range of host species (14–17). Indeed, these variants can infect and cause mortalities in many bivalve species, including the scallops *Pecten maximus* (15) and *Chlamys farreri* (18), the clams *Anadara broughtonii* (19), *Ruditapes philippinarum* (20, 21) and *Ruditapes decussatus* (20, 21), as well as the oysters *Magallana angulata* (16), *Crassostrea virginica* (22), *Ostrea edulis* (16, 23, 24) and *Magallana gigas* (25–27).

In France, two oyster species have been shown to be infected by OsHV-1: the native European flat oyster *O. edulis* and the introduced Pacific oyster *M. gigas*. Pacific oysters were introduced in France at the end of the 1960s in response to the dramatic decline of the Portuguese oyster (*M. angulata*) populations (28). *M. gigas* is currently the most widely cultivated marine mollusk species, in France, in Europe and worldwide (29). While herpes like viruses were first observed in *C. virginica* adults in the United States of America in 1970, it has been detected in France in *M. gigas*, in a hatchery,

approximately 20 years after the introduction of *M. gigas* (27). OsHV-1 induces a major threat to juvenile Pacific oysters and has significant economic impacts in many production areas all around the world (18, 27, 30–33). On the other hand, OsHV-1 does not impacted the flat oyster production currently, as it only leads to rare symptoms in few cases accompanied by mortalities.

Traditionally, OsHV-1 genotypes have been characterized by targeted amplicon sequencing (15, 15, 24, 33–35). This approach allowed the characterization of OsHV-1 Var detected in clams *R. philippinarum* in 1997 (16) and scallops *P. maximus* in 2000 (15). This genotype has never been detected in wild clams or scallops as very limited detection of OsHV-1 have been carried out on these species in natural environment. Targeted approaches also allowed the characterization of OsHV-1 µVar, the genotype mostly detected in Europe currently. This genotype was first detected in spring 2008 and has been involved in mass mortalities of the Pacific oyster juveniles (35).

According to these observations, and because herpesviruses generally have a limited ability to infect a wide range of hosts, the assumption of OsHV-1 host specialization cannot be ruled out (16) despite the fact that several viral genotypes have been detected sporadically in various bivalve host species worldwide (15, 16, 24, 36). Indeed, the genetic information obtained so far suggests that different host species may be infected by distinct genotypes of OsHV-1 (19, 37), indicating a complex epidemiological cycle that involves a wide range of host species that still needs to be clarified.

One significant drawback of targeted approaches is that they cannot comprehensively capture the entire genome, leading to an incomplete interpretation of the genomic information (38, 39). Advances in sequencing technologies allowed for the acquisition of the first OsHV-1 genome obtained from *M. gigas* larvae collected in 1995. This genome has since then been used as a reference genome (12). To date, 52 OsHV-1 additional genomes have been sequenced from four host species: *C. farreri*, *A. broughtonii*, *O. edulis* and *M. gigas* (18, 19, 25, 37, 40, 40, 41). The genome of the virus is linear and its size ranges from 199 kb to 211 kb. It contains from 123 to 128 predicted ORFs, of which only 57 have putative functions (12, 18, 19, 25, 40, 41). OsHV-1 genomic architecture is similar to some other herpesvirus genomes (*i.e.* Herpesvirus group E such as HSV-1) with two invertible unique regions (U_L for Unique Long and U_S for Unique Short) surrounded by inverted repeats (TR_L/IR_L and TR_S/IR_S). An additional unique sequence (called X) can be found between IR_L and IR_S, resulting in the overall genomic organization that can be resumed as TR_L-U_L-IR_L-X-IR_S-U_S-TR_S (12). Preliminary analyses of a few number of OsHV-1 genomes, isolated from different bivalve species, revealed a relatively high genotypic diversity (18, 19, 37, 41). More specifically, comparative genomic analyses of viruses infecting *O. edulis* and *M. gigas* showed the presence of two large deletion areas in the UL region as well as host-specific polymorphisms along the genome (37). These results suggested that OsHV-1

evolution is strongly influenced by the host species (37). However, the low sampling size of previous studies did not allow the application of population genetic and phylodynamic analyses and still little is known about the level of genetic isolation of OsHV-1 genotypes infecting *O. edulis* and *M. gigas* species along with about the frequency of cross-species transmission events (37).

In the present study, high-throughput deep sequencing was used to characterize the genetic diversity of OsHV-1 found in both *O. edulis* and *M. gigas* in France and Sweden. To this end, 40 new OsHV-1 were sequenced and assembled. Using these data and previously gathered information about genotypes circulating in different host species, our aims were to *(i)* describe OsHV-1 genomic diversity patterns in both species, *(ii)* estimate the genetic isolation between viruses infecting both species and *(iii)* determine if OsHV-1 cross-species transmission occurred and at which frequency. To do so, a combination of comparative and population genomic, phylogenetic and phylodynamic approaches were applied and discussed results in the context of current knowledge of OsHV-1 molecular epidemiology.

Results

Detection of OsHV-1 in *O. edulis* and *M. gigas* collection samples

The number of viral copies varied from $1.47x10^3$ to $3.07x10^8$ copies per ng of DNA, with an average quantity of $1.34x10^7$ copies per ng of DNA for the samples collected from *O. edulis* (Fig. S1, Table S1). For samples collected from *M. gigas*, the viral copy numbers ranged from $2.21x10^4$ to $4.21x10^7$ copies per ng of DNA, with an average of $2.81x10^6$ copies per ng of DNA (Fig. S1, Table S1).

As the minimum threshold for efficient Illumina sequencing of virus particles is 103 copies, all samples were found to be suitable for sequencing (37, 42, 43).

In addition to qPCR quantification, the lesions related to the viral infection in flat oyster tissues has been highlighted because few studies of lesions exist, particularly in flat oyster larvae. Histological observations revealed the presence of apparently healthy flat oyster larvae (Fig. S2A), larvae with hemocytes infiltration and necrosis lesions (Fig. S2B to S2D) and also dead larvae (Fig. S2E). On moribund larvae, hemocyte infiltrations were light and generally located in the mantle (Fig. S2B); concerning necrosis, it was the main lesion observed and essentially located in the mantle and velum (Fig. 2C). It was associated with the presence of cells with abnormal nuclei; these abnormalities were characterized by the presence of perinuclear chromatin and an amorphous matrix in the central area of nucleus or abnormal enlarged nuclei (Fig. S2C). In some disease larvae, necrosis affected all tissues

(Fig. S2D). Dead larvae appeared in decomposition state with a lysis of all tissues and a loss of tissue architecture (Fig. S2E).

Sequencing statistics

Sequencing depth and the viral genome coverage were correlated to the total number of reads for the different libraries and to the number of copies of OsHV-1 copies (Fig. S1). The Illumina DNA sequencing step produced from 18.9 to 290.9 million reads for a given sample (average 112.2 M reads ± 60.3 SD, Fig. S1 and Table S1). Host reads represented from 5.2% to 91.2% of the total reads. The percentage of reads mapped to OsHV-1 varied from 0.013% to 38.51%, and the genome coverage varied accordingly from 12 to 26,790 reads depth. Using this sequencing dataset, 40 OsHV-1 NR-genomes were assembled *de novo*. The length of the assembled NR-genomes varied from 185.9 kbp to 190.2 kbp.

To verify that sequencing errors were not a major driver of intra-individual diversity and investigate the degree to which sequencing effort affected the number of intra-individual variations detected, a rarefaction analysis on the number of detected variations was performed. Rarefaction curves reached a plateau at 150,000 reads corresponding to an average of 110 reads depth (Fig. S3). Therefore, this value was used as the minimum number of reads necessary to accurately detect SNPs in our samples. 12 libraries did not reach this threshold and were excluded from all analyses focusing on intra-sample diversity. For others libraries, when a high number of sequences were used, no additional variations were detected, which would have been the case if there was an accumulation of sequencing errors.

Intra-sample OsHV-1 diversity

A total of 986 "intra-sample" polymorphisms, having an allelic frequency greater than 10%, was detected across 776 polymorphic sites along genomes (Table S2, Fig. 1A). In the genome group sampled from *O. edulis*, 536 polymorphisms were detected at 430 variable positions, out of which 67 positions carry more than two alleles. On the other hand, the genome group sampled from *M. gigas* exhibited 434 polymorphisms found at 327 variable positions, with 61 positions carrying more than two alleles. Additionally, 92 polymorphisms located on 51 variable positions were shared between genome groups of both species.

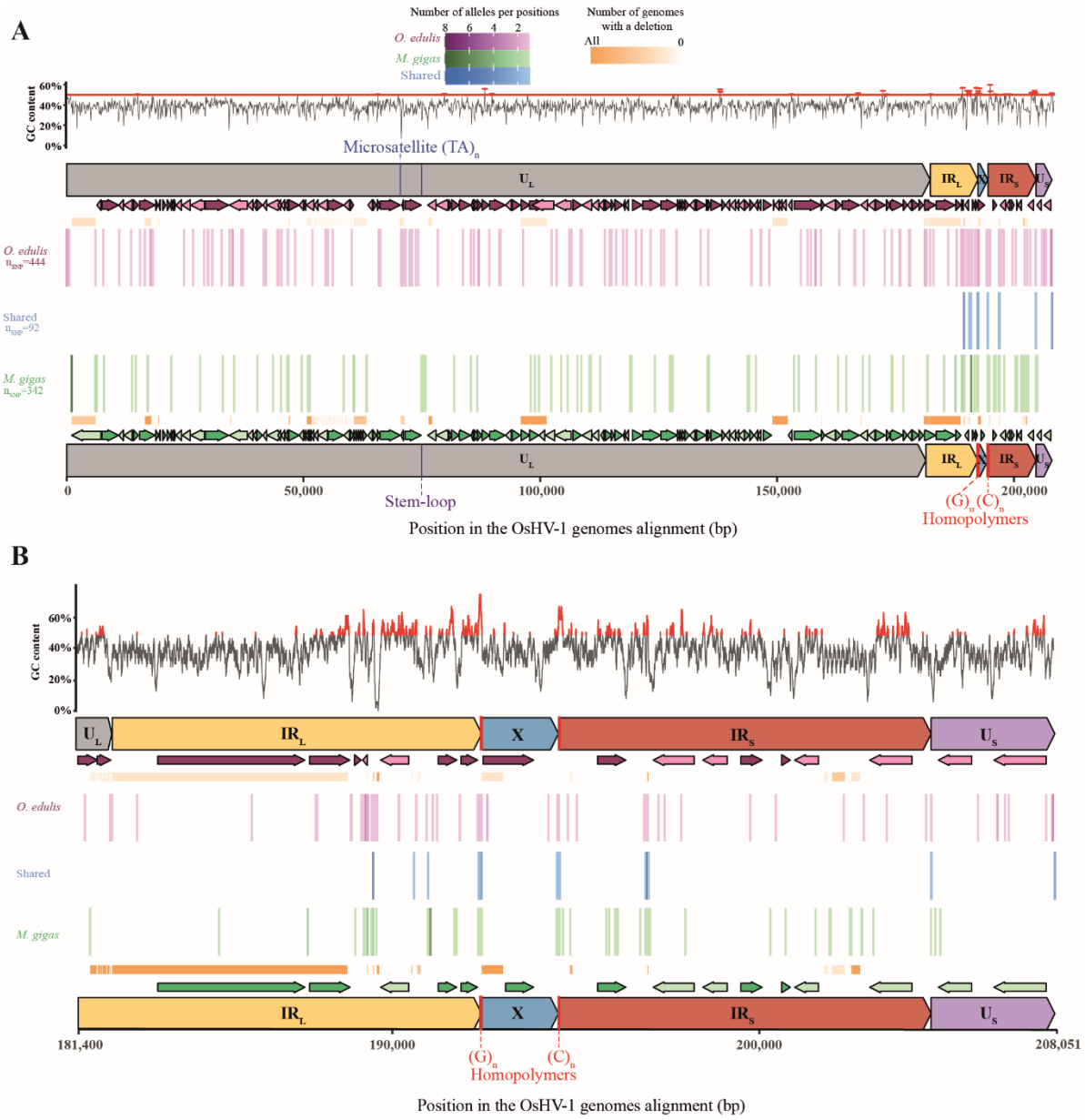


Fig. 1. distribution of SNPs within the OsHV-1 genome. *(A) Distribution of single nucleotide polymorphisms (SNPs) along the alignment of OsHV-1 de novo assembled genomes and (B) zoom on the repeated regions of OsHV-1 genomes. SNPs detected in the M. gigas population are represented in green, SNPs detected in the O. edulis population are represented in purple and SNPs detected in both populations are represented in blue. The intensity of the point varies according to the number of alleles detected for each position. Locations of SNPs are plotted against regions and positions along the OsHV-1 multiple alignment. Variations in GC content are represented in grey and when greater than 50% highlighted in red.*

Out of the 536 polymorphisms detected in *O. edulis* libraries, 336 were located within 58 predicted ORFs (Table S3). More precisely, 161 polymorphisms were found in ORFs coding for proteins with unknown functions, while the remaining polymorphisms were identified in ORFs coding for proteins with known domains or functions (Table S3). In detail, 105 polymorphisms were detected in ORF coding for putative transmembrane proteins (in ORFs 22, 41, 54, 57, 59, 65, 72, 77, and 111), 26 polymorphisms in ORF coding for putative DNA primase (ORFs 66 and 7), 15 SNPs in ORF coding for putative DNA packaging protein (ORF 109), 6 polymorphisms in ORF coding for DNA polymerase (in ORF 100), or 6 polymorphisms in ORF coding for Zinc-finger RING type protein, (in ORFs 9, 38, 97, and 117).

In viral genotypes sampled from *M. gigas*, 203 out of the 434 detected polymorphisms were located within 44 predicted ORFs. 57 polymorphisms were found in 21 ORFs coding for proteins with unknown functions, while the remaining 148 polymorphisms were identified in ORFs coding for proteins with known domains or functions. More precisely, 81 polymorphisms were detected in ORFs coding for transmembrane proteins (in ORFIN4, ORF 22, 41, 65, 68, and 77), 48 polymorphisms were detected in ORFs coding for membrane proteins (in ORF 16, 36, and 63) and 6 polymorphisms in ORFs coding for Zinc-finger, RING type. The remaining polymorphisms were detected in ORFs coding for DNA primase, DNA packaging, or DNA polymerase (Table S3).

In proportion to the size of the regions, the U_L region exhibits fewer polymorphisms (0.14% for *M. gigas* samples and 0.19% for *O. edulis* samples) compared to the other regions (Table S4). In the *M. gigas* samples, the genomic regions with the highest SNP density are the IR_L and IR_S regions, containing 0.81% and 0.57% of the SNPs, respectively, while 0.53% of the X and U_S regions contain SNPs. In the *O. edulis* samples, the U_S, IR_L, and X regions exhibit the highest SNP density, containing 1.03%, 0.99%, and 0.72% of the SNPs, respectively, while only 0.36% of the IRS region contains SNPs.

Interestingly, the polymorphisms shared by samples from both oyster species are located in the repeated regions IR_L and IR_S, particularly in regions that exhibit a GC content greater than 50% (Fig. 1B). No polymorphisms are detected in the X region, except in the homopolymers at the junctions with the repeated regions, which demonstrate an accumulation of SNPs.

The multidimensional scaling (MDS) analysis of allelic frequencies of polymorphic sites showed alleles frequencies in samples from *O. edulis* are different from those in samples from *M. gigas*, and indicate a clear structuring of intra-sample variations according to host species (Fig. 2A). However, 92 allelic frequencies were more similar within both species compared to others, indicating that some positions are variable independently of the species studied (Fig. 2A).

The genotype association graph, coupled with an MDS analysis, summarized the two-dimensional comparison of the co-occurrence frequency of each polymorphism inside and between both genotype groups (Fig. 2B). The modularity values of a network indicate the strength of connections within and between communities, and thus, provide information about the quality of network division. The analysis of alleles co-occurrence revealed the presence of three distinct groups based on modularity measurements. Out of these three groups, two clusters corresponding to *M. gigas* samples (represented by the dark and light green clusters on Fig. 2B) and one cluster corresponding to *O. edulis* (represented by the purple cluster on Fig. 2B). The dark green cluster of *M. gigas* samples comprises samples collected from various developmental stages between 2005 and 2010 in Charente-Maritime, the Breton inlet, and Normandie. On the other hand, the light green cluster of *M. gigas* samples consists of samples collected between 2006 and 2008 in Palavas and Thau lagoons, and the Breton inlet, encompassing all developmental stages.

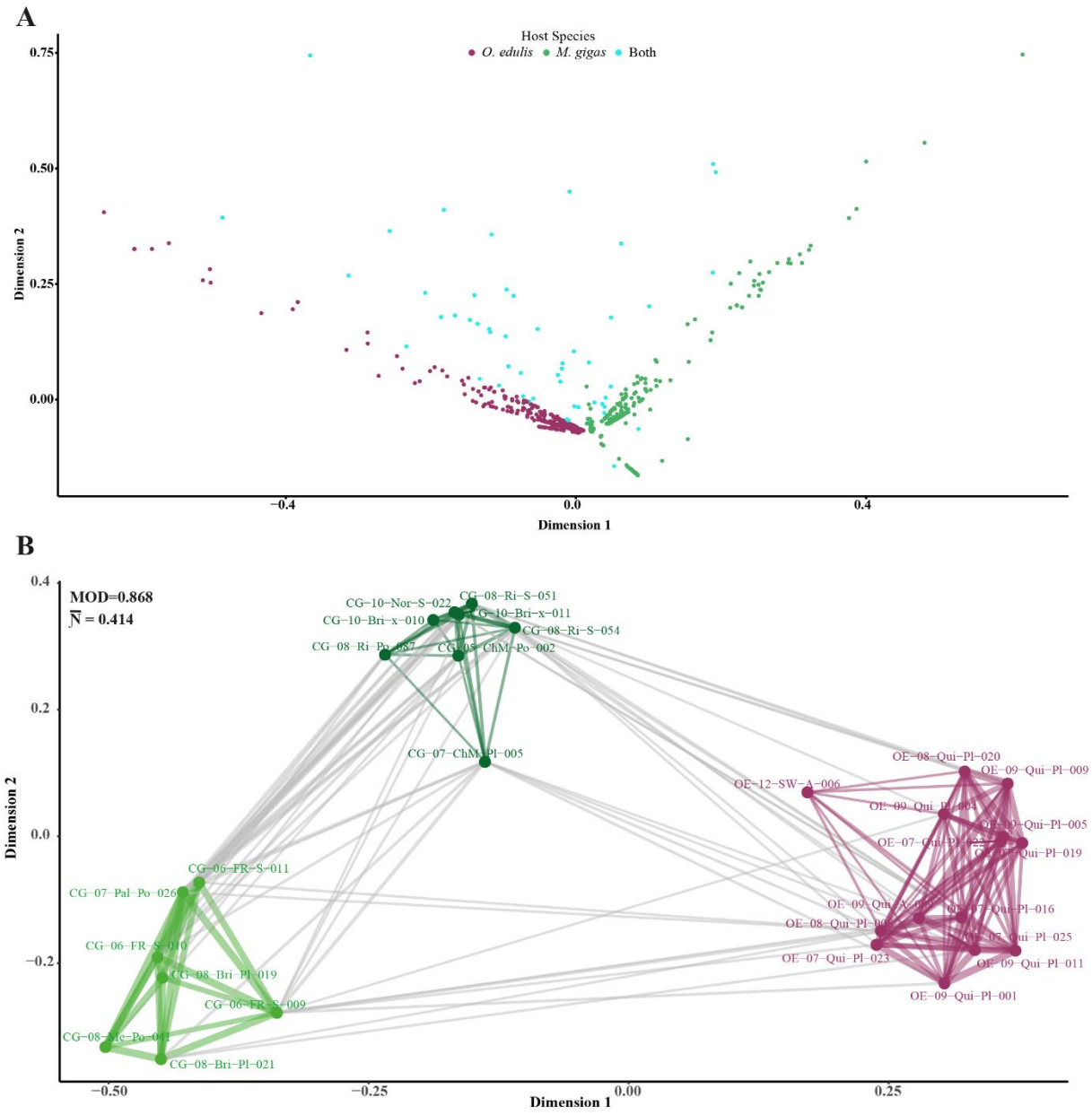


Fig. 2. From intra sample population diversity to host specialization. *A) Multidimensional scaling (MDS) plot showing allelic frequencies distribution at polymorphic sites. Points are colored according to the host species alleles were detected in. B) MDS representing oyster individual's similarity according to allelic frequencies of polymorphisms they harbored. The MDS analysis provides a visual representation of this relationship in a two-dimensional space and the genotype association graph shows the co-occurrence of alleles within individuals. Samples from O. edulis are shown in purple, while those from M. gigas are shown in green and are divided into two clusters. Links between samples/genotypes individuals are represented by lines, which vary in thickness based on the similarity between the two genotypes, and the color of the lines indicates if the individuals are from the same species or not.*

Structural variations

Using a whole genome alignment approach, a structural analysis of all OsHV-1 genomes assembled from *M. gigas* and *O. edulis* samples has been performed (Fig. S4). The whole genome comparison revealed six types of genomic architecture (Fig. 3). The first one includes a group of four genomes obtained from *M. gigas* and similar to the reference genome described by (12), characterized by NR-genomes of approximately 190 kbp. The five other structural variants were identified and described in comparison with the OsHV-1 reference genome (12). First, seven NR-genomes of 190 kbp collected from *M. gigas* differ from OsHV-1 reference genome by a reversion of 8,530 bp in the U_L region between ORF 36 which is disrupted (potentially not functional) and ORF 43. Second, a group of nine NR-genomes of 204 kbp sampled from *M. gigas* were analogous to the µVar genotype described by (25) which contains four deletions in the U_L region (ranging from 606 bp to 3,550 bp), one in the IRL (725 bp) region, and one insertion of 2,671 bp in the U_L region between ORF 43 and 44, for a total of approximately 10 kbp insertions or deletions. Third, one NR-genome of 192 kbp, sampled from *O. edulis*, contains a reverse translocation of the first part (between ORF 6 and ORF 4 which is disrupted) of the TR_L region (7 kbp) and an insertion of 2,204 bp upstream to the U_L region. In addition, this genome contains multiple large deletions (6 deletions for a total of 5,303 bp) and insertions (5 insertions for a total of 4,017 bp) for an average of 13 kbp insertions or deletions. Fourth, a group of twenty NR-genomes of 185 kbp and sampled from *O. edulis*, contains a reverse translocation of the first part of the TR_L region (4 kbp). In addition, these genomes contain multiple large deletions (11 deletions for a total of 6,423 bp) and insertions (4 insertions for a total of 1,133 bp) for an approximate 11 kbp InDels in total. Finally, a group of two NR-genomes contains the same reverse translocation, deletions, and insertions detected as those detected in the fourth group of twenty NR-genomes assembled from *O. edulis*, but in addition, an inversion of 1,470 bp was detected in the U_L region (ORF 50 disrupted) for a total of approximately 12 kbp variations.

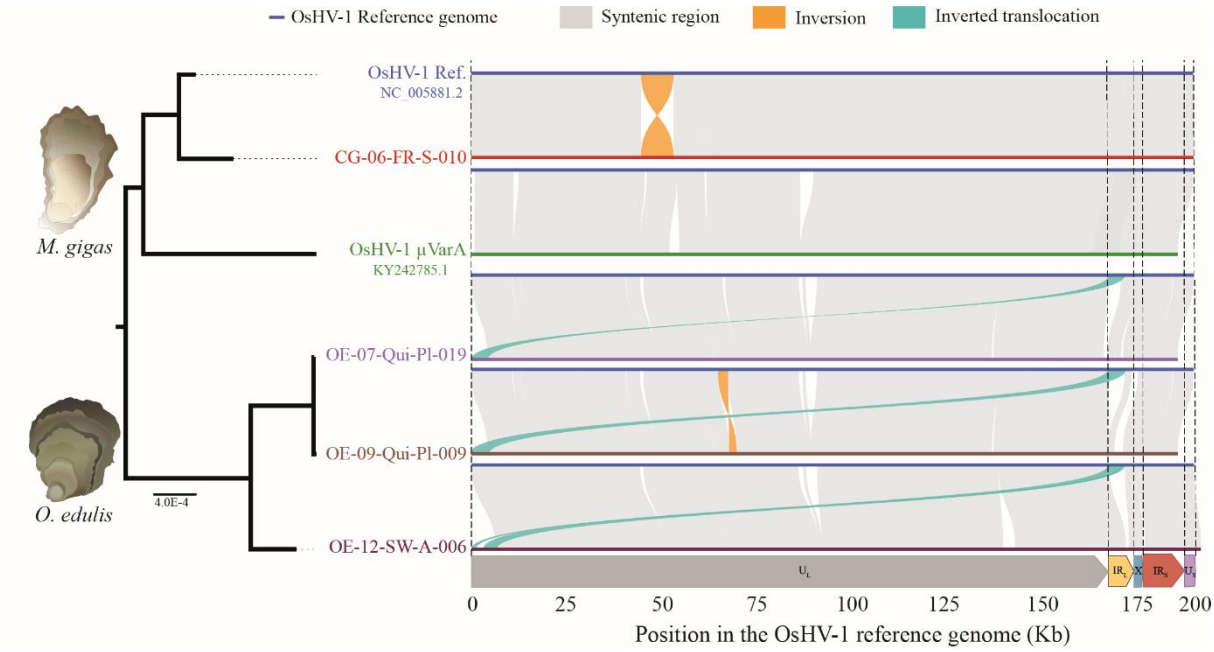


Fig. 3. Comparison of the six structural variants detected in the dataset. *On the right side of the structural variation graph, the phylogenetic relationships between the six representative individuals are plotted (estimated by a maximum likelihood approach [80]. The description of genomic structural variations was based on a pairwise comparison to the OsHV-1 reference genome (NC_005881.2, Davison et al. 2005). In each pairwise comparison, the reference genome is represented by a dark blue bar. Syntenic regions are represented in light grey, inversions in orange, and inverted translocations in light blue. Below the graph of structural variations is the architecture and position in the OsHV-1 reference genome.*

To ensure that homologous regions are all in phase, and that multiple alignment and phylogenetic analyses can be carried out, genomic regions of the forty-three OsHV-1 NR-genomes have been manually rearranged. The recombination analysis identified the occurrence of 17 potential recombination events along OsHV-1 genomes, that may have impacted two genomic regions (Fig. S5). The first potential recombination event was located in the U_L region between 61,474pb and 68,405bp, specifically between a microsatellite $(TA)_n$ present in the viral genome collected from *O. edulis* s and a stem loop predicted to be the probable origin of replication of OsHV-1. The other four recombination events detected were located in the repeated regions. These potential recombination events have been detected respectively between 173,860 and 175,357 bp at the junction between U_L and IR_L regions, between 178,966 and 180,440 bp and from 184,844 to 185,190 bp both in the IR_L region, and finally from 194,785 to 195,422 bp located in the IR_S region.

D'² values were mostly close to 1, indicating an absence of linkage disequilibrium (Fig. 4). However, one position close to the start of the genome within the ORF 6 (position 3,108), as well as two other positions at the end of the U_L respectively within ORF 88 and ORF 104 (positions 137,500 and 164,760), one position in the IR_L within ORF 114 (position 182,122), and one position in the IR_S (position 195,034) regions had very low linkage disequilibrium values, close to 0. None of the detected low D'² values was in a homopolymer region but two low D'² positions (positions 182,122 and 195,034) correspond to recombination breakpoints detected by Gubbins (Table S5).

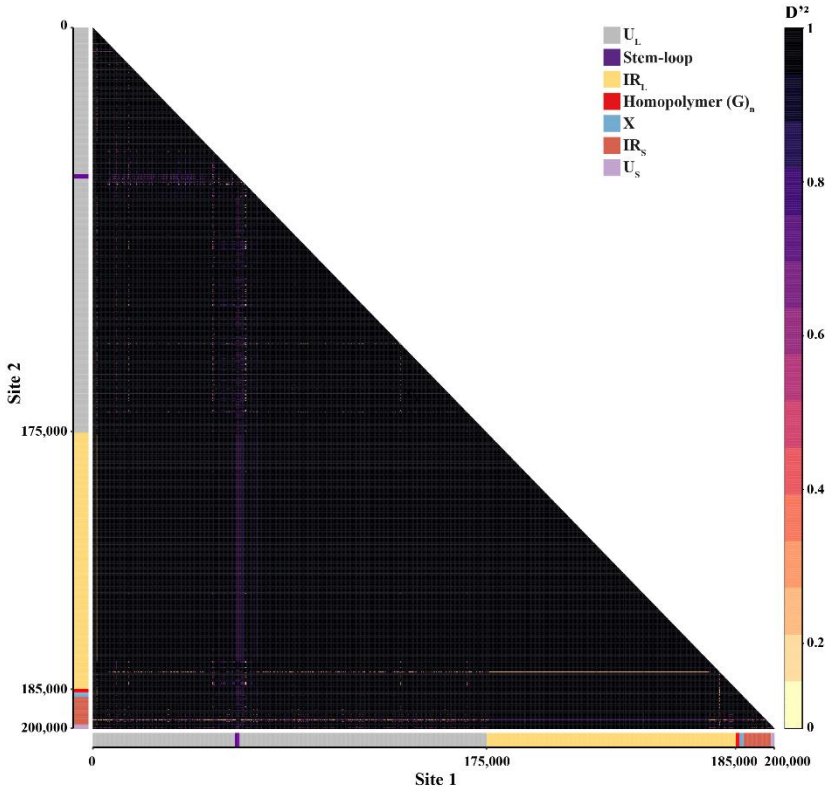


Fig. 4. Recombination hotspots. *Linkage disequilibrium (LD) heatmap between pairs of SNPs (site 1 and 2) along the alignment of 40 de novo assembled OsHV-1 genomes and the three OsHV-1 reference genomes: OsHV-1 Ref. (NC_005881.2; Davison et al. 2005), µVarA (KY242785; Burioli et al. 2017) and µVarB (KY271630; Burioli et al. 2017). LD values vary between 0 and 1, consequently, the color scale varies respectively from beige to black. At the bottom, the architecture of OsHV-1 genomes and position along the multiple alignment are represented.*

A total of 11,184 bp distributed within 13 regions potentially involved in recombination were removed from the alignment for the rest of the analysis. The suppressed regions are referenced in the table S5. More precisely, 541 SNPs have been removed from consensus-scale analyses.

A maximum likelihood phylogeny was then generated using a HKY+G+I model using PhyML (44). This tree showed that OsHV-1 genomes sampled from *M. gigas* were genetically distant from those sampled from *O. edulis* (Fig. S5). In the present study, the viral genotypes were clearly structured according to host species, with two main clusters identified (Fig. S5). Within each cluster, two subclusters were identified. In the *M. gigas* cluster, subcluster 1.1 corresponded to samples related to the OsHV-1 reference genome sampled between 1994 and 2008, while subcluster 1.2 corresponded to samples corresponding to µVar genotype sampled between 2008 and 2010. In the *O. edulis* cluster, subcluster 2.2 corresponded to samples collected between 2007 and 2009 in the Quiberon bay, while subcluster 2.1 corresponded to the sample collected in Sweden in 2012. The clustering was found to be primarily based on host species and then on the sampling date (Fig. S5). The multiple alignment of the forty-three genomes enabled us to determine the similarity percentages between every pair of genomes (Fig. S6, Table S6). To provide an overview of results, the average percentage of identities for each cluster and between clusters has been computed (Table 1). Identity percentages within clusters were close to 100%, indicating a high similarity between genomes within a cluster. In contrast, the average identity percentage varied between 96% (cluster 2.2 compared to 2.1) and 91% (cluster 1.1 compared to 2.1) when comparing pairs of genomes belonging to different clusters.

TABLE 1. Average percentage identities for each cluster and between clusters based on pairwise distances computing using MAFFT.

	Cluster 1.1	**Cluster 1.2**	**Cluster 2.1**	**Cluster 2.2**
Cluster 1.1	99,98%	94,66%	91,07%	92,51%
Cluster 1.2	94,66%	99,90%	92,96%	94,58%
Cluster 2.1	91,07%	92,96%	100%	96,03%
Cluster 2.2	92,51%	94,58%	96,03%	99,52%

Consensus-based diversity analysis

An AMOVA has been performed to test the structuring factors independently of the observed genetic differences (Table 2). The factors have been tested separately because of missing metadata. The

results indicating that 94.5% and 95.3% of the variance are explained respectively by host species and sampling origin while sampling year and host developmental stage explain only 40% and 66.4 % of the variance observed. In addition, 36% of the molecular variance is explained by the sequencing technology, there is therefore not much impact of sequencing technology on the consensus scale.

TABLE 2. Analysis of Molecular Variance (AMOVA) to determine the impact of host species, origin, year of sampling and host developmental stage on the OsHV-1 diversity at the genomic scale.

	DF	*SS*	*MS*	*Est. var.*	*%*	*PhiPT*
Host species	1	2.87e-04	2.87e-04	1.34e-05	94.5%	0.945
Sampling year	8	0.000152	1.89e-05	3.29e-06	40%	0.4
Origin	11	0.000304	2.76e-05	9.88e-06	95.3%	0.953
Host developmental stage	3	0.000196	6.54e-05	6.23e-06	66.4%	0.664
Sequencer	4	9.58e-05	2.39e-05	3.31e-06	36%	0.36

DF: degrees of freedom, SS: sum of square deviation, MS: mean square deviation, Est. var.: estimated variance, %: percentage of differentiation explained by the factor, PhiPT: population differentiation statistics.

The diversity analysis based on every NR-consensus genome was completed with the computation of commonly used measures in population genetics, including the fixation index G_{ST} and Tajima's D (Fig. 5). The genetic differentiation between species was estimated based on SNPs, and the global pairwise G_{ST} was computed for the alignment of both OsHV-1 genomes from *O. edulis* and *M. gigas* (Fig. 5A). The differentiation value between the two host species computed with a pairwise Nei G_{ST} was 0.745, indicating the presence of different fixed alleles in both groups. To determine regions that were highly differentiated between groups, G_{ST} was computed for 1-kb windows along the alignment of OsHV-1 genomes. G_{ST} values varied from 0.005 to 1 along genomes and were, on average, equal to 0.545. A threshold of G_{ST} = 0.6 was used to distinguish between highly differentiated and less-differentiated regions. Results show 34 out of 200 windows of 1-kb having a G_{ST} higher than 0.6 indicating high level of genetic differentiation of these regions (Fig. 5A). Among these 34 windows, 29 were located within the UL region and 5 were located in the IR_L region (Table S7). X, IR_S and U_S regions seemed to be less differentiated according to host species than the rest of the genome. Additionally, 100-bp windows were computed within these 1-kb windows to define more precisely positions of SNPs that were

differentiated. 87 100-bp windows out of 340 were detected as highly differentiated (G_{ST} > 0.6). These windows were located within 2 intergenic regions and 34 ORFs coding for proteins with mainly unknown functions (18 ORFs, see table S5), secreted proteins (ORFs 68, 88), proteins containing domains such as Zing-finger, Ring type proteins (ORFs 53, 97, 106), proteins with transmembrane domain (ORFs 22, 54, 59, 65.1, 68, 77, 84, 88), inhibitor of apoptosis (ORFs 87, 106), ribonucleotide reductase (ORF 20), DNA replication origin binding (ORF 115), large subunit DNA-packaging terminase (ORF 109).

Tajima's D values were calculated for 1-kb windows along the alignment to identify selection patterns along OsHV-1 genomes. At the overall group level, most of the Tajima's D values were positive (Fig. 5B), ranging from -1.87 to 4.44 with an average of 0.673. Later, Tajima's D was computed for each viral group individually. Within *O. edulis* OsHV-1 Tajima's D values were mostly negative, ranging from -2.02 to 0.54, with an average of -0.851. Within *M. gigas*, Tajima's D values were predominantly positive, varying from -1.97 to 2.92, with an average of 0.627, which was comparable to the values observed for the overall population.

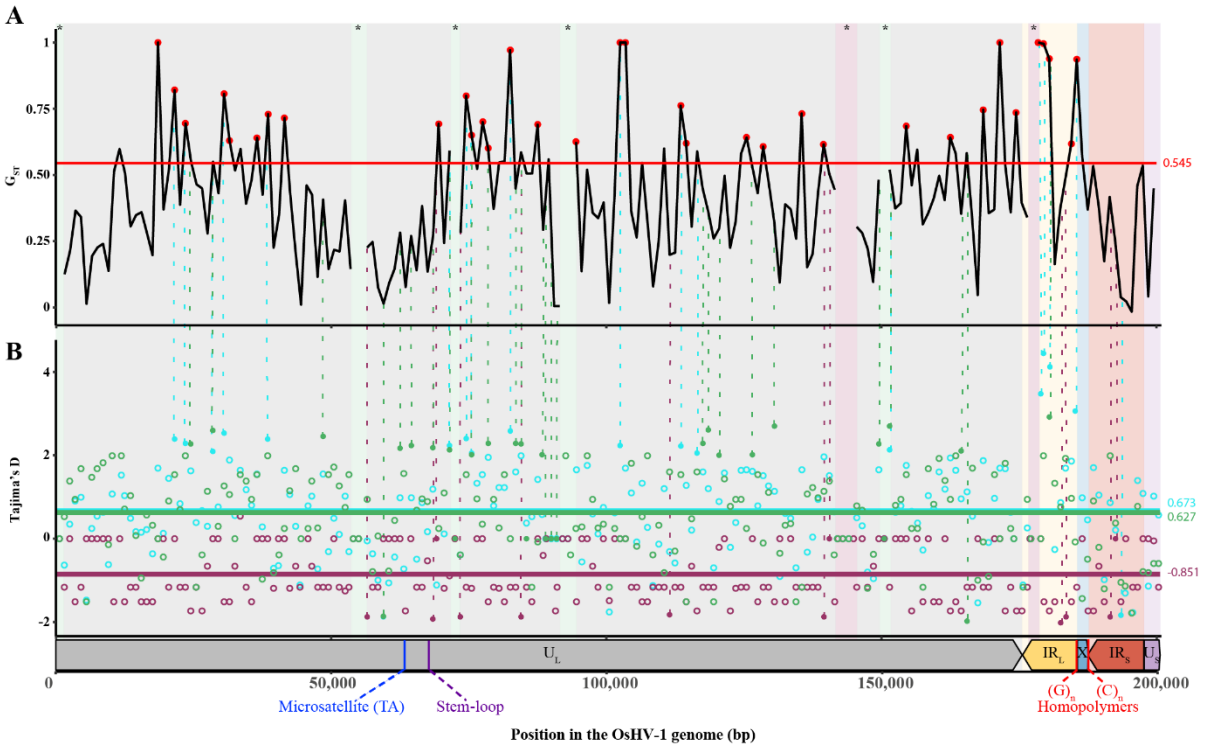


Fig 5. Consensus-based population genomics of OsHV-1. *(A) Genetic differentiation between samples isolates from both species computed using the mean GST values within 1-kb windows along the multiple alignment of OsHV-1 genomes. The red line at y=0.545 represent the mean global GST, which gives an*

overall estimate of the genetic differentiation of OsHV-1 between species. Red dots on the curve represent windows having a mean GST higher than 0.6. (B) Distribution of Tajima's D values obtained from 1-kb windows for all genomes included in this study (in blue), for OsHV-1 genomes sampled from M. gigas (in green) and O.edulis (in purple) If values of Tajima's D were significant, dots were filled, while if not significant, dots are empty. The blue, green and purple lines represent the averaged Tajima's D value for each genome data-set. For a better comparison of the neutrality and differentiation measures, significant Tajima's D values are linked by dotted lines to GST values. Below the graph is represented the architecture and position in the OsHV-1 genome multiple alignment. Genomic regions are reported on the two graphs above. Additionally, regions present only in M. gigas and O. edulis OsHV-1 subpopulations are represented with green and purple backgrounds respectively.

Tajima's D values for the *O. edulis* viral group suggest a selective sweep might have recently occurred along OsHV-1 genome. This was confirmed by RaISD results. In the *O. edulis* viral group, 168 regions with μ-statistic values ranging from 1 to 8 were found to have been impacted by selective sweeps. For *M. gigas* viral group, 587 regions with μ-statistic values ranging from 3 to 8 were affected (Fig. 6A). Regions impacted by selective sweeps in both groups were mainly located within ORFs coding for proteins with unknown functions. Furthermore, 8.3% to 9.9% of the selective sweep regions were located within intergenic regions (Table S8). The main domain impacted by selective sweep appear to be transmembrane protein (18.5% for *O. edulis* and 17.5% for *M. gigas*), Zinc-finger, RING-type (5.9% and 5.5% of the detected selective sweep regions), and viral maintenance functions such as DNA primase, polymerase, or packaging (Fig. 6B).

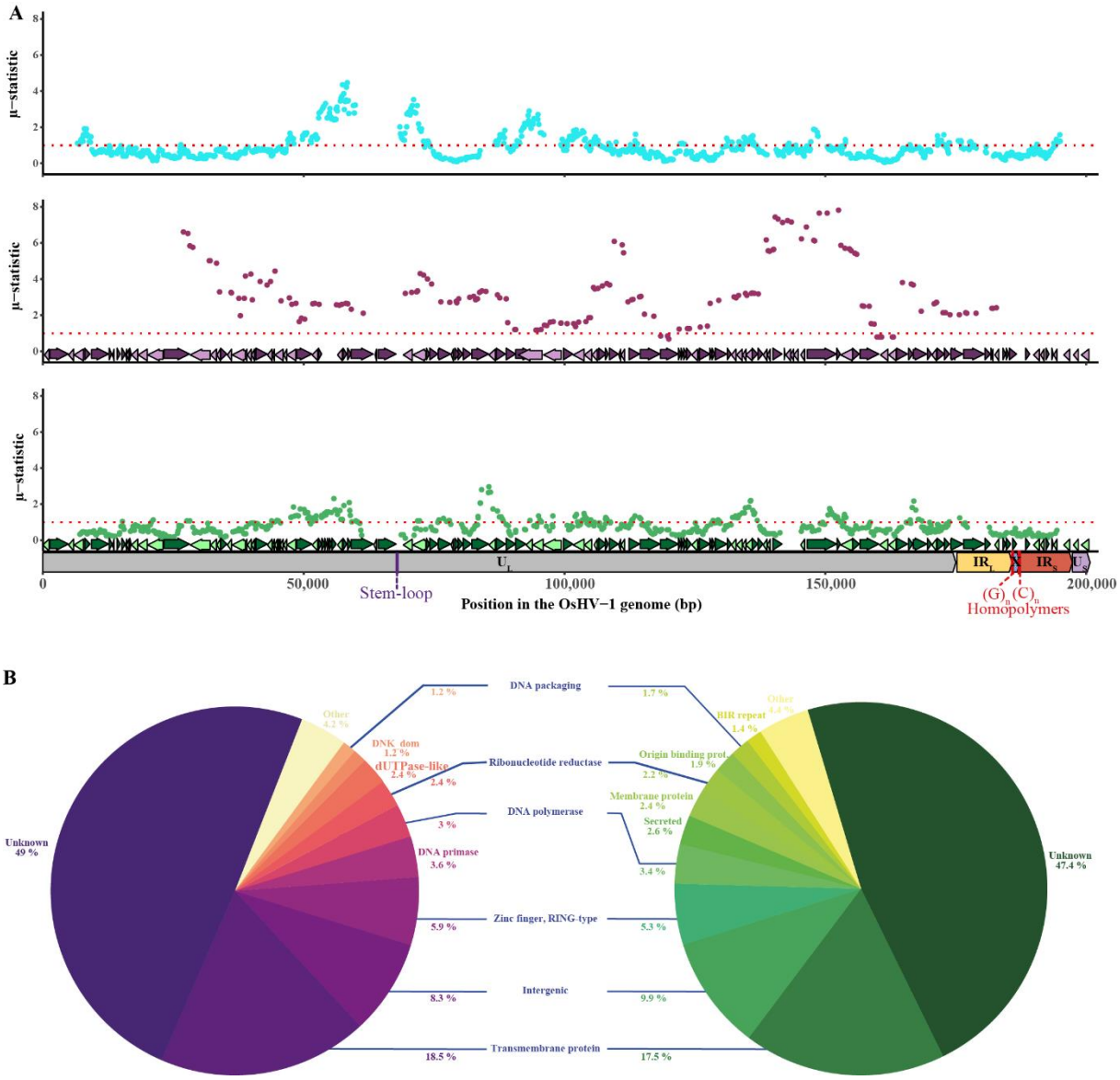


Fig 6. Detection of selection patterns. *(A) Detection of selective sweep signatures along the multiple alignment of OsHV-1 genomes in the global group (blue points), in the O. edulis group (purple points) and in the M. gigas group (green points). Below those graphs are represented the ORFs predicted on ancestral sequence of each group and the architecture and position in the OsHV-1 genomes multiple alignment. (B) Pie charts representing the proportion of selective sweep signals detected within ORFs coding for specific function or domain for both species.*

Ancestral states reconstruction

R^2 values of regression of root-to-tip genetic distance against sampling times are provided as a measure of how OsHV-1 evolution follows a clock-like pattern. According to R^2 values in best fitting root models,

OsHV-1 harbors a sufficient temporal signal to perform molecular clock-based approaches to estimate evolutionary rates over time (Fig. S7).

The comparison of MLE values, with higher values reflecting the most appropriate models, has led to the selection of an uncorrelated relaxed molecular clock and a constant size coalescent model (Table 3).

TABLE 3: maximum likelihood estimation values used to compare the different demographic and clock models.

	Constant size		*GMRF*		*Exponential growth*	
	Strict clock	Uncorrelated relaxed clock	Strict clock	Uncorrelated relaxed clock	Strict clock	Uncorrelated relaxed clock
PS	-1293,121731	**-1279,27**	-1312,775139	-1312,436377	-1294,032584	-1293,632567
SSS	-1293,270332	**-1279,441704**	-1312,798034	-1312,384202	-1294,529875	-1294,145544

Values in bold correspond to the maximum MLE used to select the model for ancestral state. reconstruction with Beast.

Moreover, the evolutionary rate has been computed for the global phylogeny and for each host subcluster (Fig. S7). For the global phylogeny, the evolutionary rate has been estimated at 1.50E-02 nucleotide substitutions per site per year. Within the cluster from *M. gigas*, the evolutionary rate was 2.37E-02 nucleotide substitutions per site per year. In comparison, within the genotypes of OsHV-1 infecting *O. edulis*, the evolutionary rate was slightly lower with a value of 1.20E-02 nucleotide substitutions per site per year.

The most recent common ancestor (MRCA) of the OsHV-1 genotypes was estimated to have existed in the 1990's (95% HPD interval [1989,1995]) and is associated with *M. gigas* (Fig. 7A). Then, the actual genotype µVar detected in France seems to have emerged around 2007 (95% HPD interval [2006,2008]). In parallel, one event of CST from *M. gigas* to *O. edulis* has been inferred to have occurred between 1998 (95% HPD interval [1993,2004]) and 2004 (95% HPD interval [2001,2008]) follow-up by a differentiation and a specialization of the lineages infecting *O. edulis*.

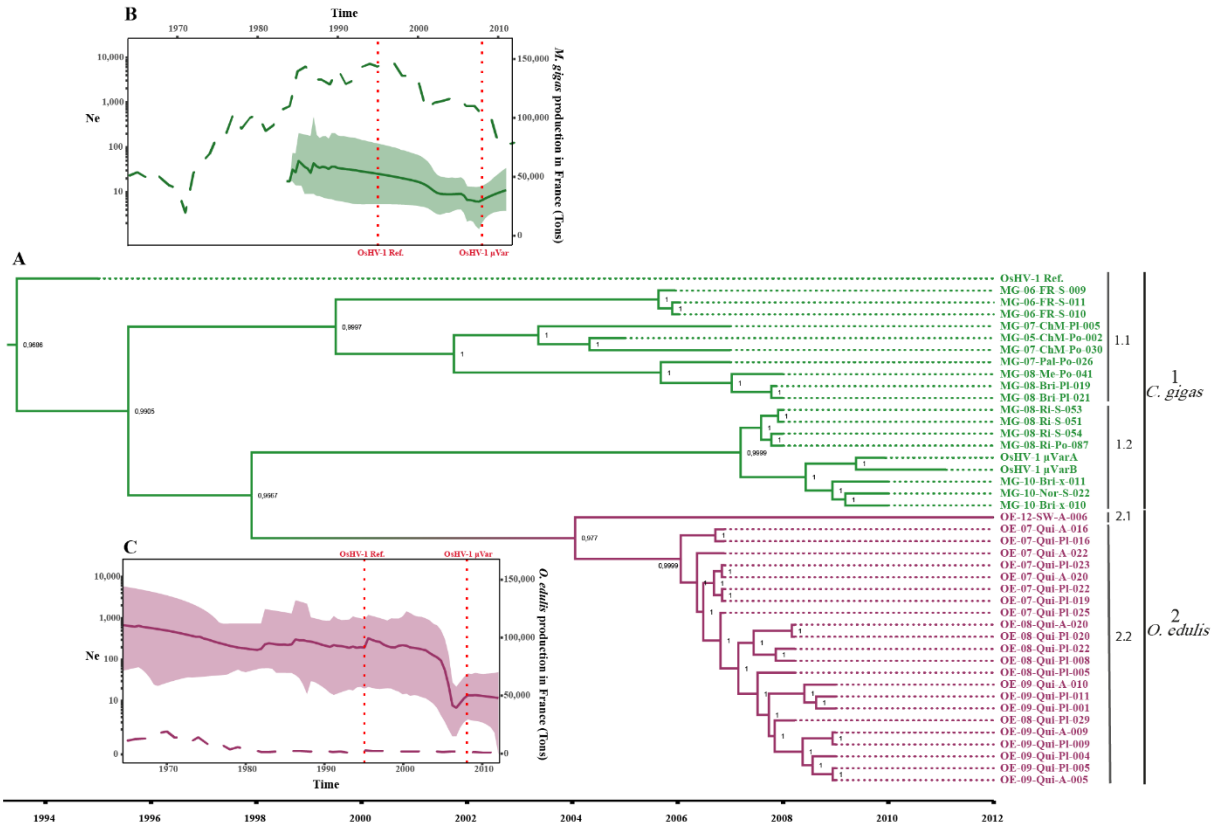


Fig 7. Evolution and demographic history of OsHV-1 within both host species. *A) Time-scaled phylogeny of OsHV-1 genomes and ancestral state reconstruction of the discrete host species trait. Demographical reconstruction of OsHV-1 population size and observed population within B) M. gigas and C) O. edulis. Similarly, to other figures, color of the branches of the Maximum clade credibility (MCC) tree represents the host species from which the virus was isolated (green for samples from M. gigas and purple for samples from O. edulis). Numbers at branch nodes represent the posterior probabilities of the host species inference.*

Demographic evolution over time within both viral groups

A model of the effective size of the viral group over time was constructed using output trees from ancestral state reconstruction. The viral populations sampled from *O. edulis* showed a constant decrease in "population" size from 1965 to 2002 (Fig. 7C). Between approximately 2002 and 2007, the decrease in "population" size was drastic and was quickly followed by a rapid expansion step. At present, the viral group is still experiencing a slight, constant decrease in "population" size. In contrast, the *M. gigas* viral group exhibited a fluctuating step of "population" size between 1982 and 1990, followed by a constant decrease in "population" size until 2008, which corresponds to the detection of the µVar genotype (Fig. 7B). After 2008, this group began to expand. Moreover, these results suggest

that the viral populations sampled from both oyster species was present before the first detection of OsHV-1 made in 1991 in France (26).

Discussion

Several approaches have been used to either detect or characterize OsHV-1 genotypes infecting *M. gigas*, including histology and microscopy (16, 36, 45, 46), PCR (16, 36) as well as restriction length polymorphisms (RFLP) (16). However, this is a recent genomic study that revealed a potential structuring of OsHV-1 lineages according to host species (37). This raised questions about OsHV-1 host specificity, transmission cycles and cross-species transmission (CST). Here, the diversity of OsHV-1 genomes between *M. gigas* and *O. edulis* collected from 2005 to 2012 in a dataset were compared, at both the intra-sample and global genomic scale, and gained insights into evolutionary history of OsHV-1 and CST frequency between both host species.

OsHV-1 population structure and differentiation driven by host species

Previous studies have provided preliminary evidences that the genetic diversity of OsHV-1 is influenced by various factors, including geographical location (37, 42), time (25, 37, 47), and host species (37).

Our analysis of structural variations identified six distinct variant genotype structures in our dataset, differentiated by large Indels, genomic translocations, and reversions mainly in the U_L and IR_L regions, which resulted in total size variations of approximately 6 kbp between genotypes. Our findings are consistent with those reported by Morga-Jacquot *et al.* (2021), who previously reported on the evolution of OsHV-1 and highlighted the prevalence of genomic deletions as a driving force (37). Consistent with their previous observations, our dataset also exhibited deletions in certain regions, including ORFs 11, 35 to 37, 48, and 62 to 63, and probably underestimated because of genomes non-redundancy. The exact mechanisms underlying the generation of these structural variants are not yet understood. However, it is plausible that recombination or rearrangement events, which are common in herpesviruses, play a crucial role in their appearance. Similarly, studies based on both Sanger sequencing (12) and recently Nanopore long reads sequencing (43) have demonstrated the presence of four major genomic isomers of the OsHV-1 genome, considered as structural variations, that were present in stoichiometric proportions. In herpesviruses, and especially in the herpes simplex virus 1, the presence of isomeric genomes is closely related to the replication and recombination mechanisms (48–51) and occurs early during the herpesvirus life cycle. During DNA replication, numerous random double-strand breaks efficiently initiate recombination events and give rise to new genomic rearrangements (52). Importantly, these structural variants appear to be significant and not detrimental to the maintenance and replication stages of the viral life cycle. On the contrary, these

variations may even be an advantage, as they potentially allow the emergence of new viral lineages and the adaptation of genotypes to new hosts (53).

Interestingly, OsHV-1 lineages sampled within the same species were genetically very close to each other (respectively 94% and 96% identities for the genomes collected from *M. gigas* and *O. edulis*), while lineages sampled from *O. edulis* were significantly different from those detected in *M. gigas* (between 91% and 94% identities). The genetic differentiation of OsHV-1 observed here is consistent with previous studies that have examined the genetic distances between viral genomes infecting closely related host species (54, 55). Similarly, Kolb *et al.* described the phenomenon of host shift in herpes simplex virus infecting old world monkeys, followed by co-evolution with its specific host, and demonstrated a strong correlation between viral genotype distance and phylogenetic host distance (54). However, within the samples collected from *M. gigas*, there is a strong clustering of genotypes according to sampling date, corresponding to the reference (pre-2008) and µVar (post-2008) genotypes. Additionally, the µVar genotypes were not detected in *O. edulis* samples investigated in this study and collected in the absence of OsHV-1 mortality episodes.

Host shifts have been identified as a key driver of genomic rearrangements and recombination events that facilitate host adaptation and specialization. An experimental evolution study has shown that structural rearrangements of the 3'-UTR of the Dengue virus genome were observed after 10 generations, leading to differential viral groups within mosquitoes and human (56). These rearrangements could potentially confer a high capacity for rapid adaptation and evolution in response to host shifts. Although our analysis of OsHV-1 identified potential recombination events, only a limited number of them were detected. Specifically, five recombination events were identified between the stem-loop regions and a microsatellite $(TA)_n$ located in the U_L region, as well as in the homopolymer regions at the junctions of the X with repeated regions. Notably, it has been demonstrated that a bias in the detection of recombination breakpoints could occur in repetitive regions, intergenic regions, and areas with high G+C content (57). This finding provides a potential explanation for the occurrence of only five recombination events observed in our study. Moreover, the frequency of recombination events in herpesviruses is not well-known (58). However, studies have demonstrated numerous recombination events between Herpes simplex virus 1 and 2 (59, 60) as well as between strains of Human cytomegalovirus (HCMV). These recombination events were detected throughout the genome and mainly in unique long regions (61, 62) as observed in our dataset. The genome of OsHV-1 may have undergone evolutionary processes that have led to these rare occurrences of recombination and potentially contributed to the adaptation of each lineage to its respective host (63). In order to fully understand the implications of the observed rearrangements and recombination events within the genome of OsHV-1, additional investigations are required. Specifically, it is crucial to examine

additional OsHV-1 samples sourced from diverse geographic locations and host species. This would help to determine whether similar patterns of genomic rearrangement and selection exist across different populations. Additionally, conducting experimental infections studies aiming at elucidating the functional consequences of these genetic modifications would provide valuable insights into their biological significance.

Among OsHV-1 clusters, contrasted differentiation and selection patterns are observed along the genome. In particular, regions encoding for essential biological functions or domains involved in viral particle binding to host cells, DNA synthesis, replication and packaging, as well as trans- and membrane proteins may be critical for adapting to new host species. Indeed, viral entry proteins play a crucial role in enhancing viral fitness, and a single amino acid mutation within these proteins can alter the virus's host range (64, 65). Furthermore, it has been shown that domains, such as envelope domains (including membrane glycoproteins and transmembrane receptors), auxiliary domains (such as Zinc-finger, RING type, dinucleoside kinase) and modulation and control domains (such as Interleukin, Interferon-regulatory factor or Zinc-finger) have been gained, duplicated, or lost during the evolution of Herpesviridae (66). Many of these acquired domains allowed viruses to bind specifically to host cells and to evade or modulate the host immune system. In our study, most of the selection signals have been detected within these domains, suggesting a potential adaptation of essential proteins leading to a specialization of viral genotypes to their host species. Furthermore, the host species was found to play a significant role in driving viral diversity through selection at the consensus scale (56). The selective pressure imposed by different host species has resulted, for example, in a diversification of Dengue viral lineages in mosquitoes and a reduction of viral population diversity in mammals (56).

Previous studies have not specifically investigated the intra-sample genetic diversity between infected host species but were more interested in comparing viral diversity in different organs or different individuals (*e.g.* vaccinated or unvaccinated) (67, 68). However, it is interesting to study within-host viral diversity and compare it to better understand minor diversity, differentiation and selection pressure. To investigate OsHV-1 diversity and differentiation according to host species, global scale genomic analyses were completed by within samples and within clusters. At the intra-sample scale, a clear contrast in the polymorphisms detected between host species was observed. Only 92 SNPs out of 1,051 were shared between both groups, and they were located within repeated regions and homopolymers at the junction between X and the repeated regions of the OsHV-1 genome. The accumulation of polymorphisms within the repeated regions at low frequencies has already been observed (37). There are several possible explanations for these results. Firstly, these regions are duplicated in the virus genome. Repeated regions are known, in herpesviruses, to allow rearrangements of single regions by pairing between these regions (52). In addition, these regions

contain multiple small repeated sequences that can vary in size and number from one genome to another. Some of the polymorphisms detected in these regions can be explained by the loss or gain of some of these repeats. Secondly, the use of a non-redundant genome to study the virus nucleotide diversity may induce a bias. Indeed, both the (TR_L/IR_L and IR_S/TR_S) repeat regions may contain polymorphism, which is detected here in a single region. Finally, GC-rich regions or regions containing various repeats, hairpin structures or long homopolymer stretches are notoriously difficult to sequence (69). Repeated regions combine several of these problematic structures such as GC-rich, multiple repeats, and long homopolymer stretches. Polymorphisms detected in these regions are therefore possibly difficult to interpret and their accumulation may be linked to sequencing errors due to the difficult template. Indeed, these regions are difficult to access by the polymerases used during sequencing steps due to DNA methylations or folds. Sequencing technologies that do not use polymerases, such as Nanopore, could potentially solve these sequencing problems.

The genotype association graph used to compare the co-occurrence frequency of each polymorphism within and between both genotype groups revealed three distinct groups: two corresponding to *M. gigas* and one corresponding to *O. edulis* population. These results suggested that genotypes present in *M. gigas* displayed a higher level of genetic diversity compared to those originating from *O. edulis*, that could be explained by host populations size and species culture. *M. gigas* is currently the most widely cultivated oyster in the world, with enormous population sizes, both locally within oyster beds or in the various production basins, and worldwide. Moreover, this species is often transferred from one basin to another, particularly in France due to breeding practices (70). In the case of the flat oyster, populations are small in France, mainly located in Brittany for a long time, and are transferred locally between North and South Brittany during their growth. It is particularly true for the flat oyster samples used in this study, because they were collected from natural populations which are not transferred. Moreover, the populations are distributed in "patches", which does not allow direct connections between the different populations and limit viral transmission (71).

The fine-scale diversity observed within OsHV-1 provides further support to findings obtained from consensus-based analyses and confirms the hypothesis that the virus is differentiated based on host species and has undergone coevolution with its respective new host following adaptation. An important avenue for future research would be to investigate the coevolution between OsHV-1 and multiple host species using phylogenetic approaches. This would involve sampling different host species from various locations to obtain a comprehensive understanding of the coevolutionary dynamics between the virus and its hosts and give further insights into the evolutionary processes and host-virus interactions shaping the viral diversity pattern.

As mentioned earlier, our sample selection aimed to maximize heterogeneity to explore the potential effects of various factors on viral diversity. However, it is crucial to acknowledge that the variability in sampling host stage, dates, and origin may have introduced biases in our results (72, 73). For example, this variability could have led to the overrepresentation or underrepresentation of specific viral genotypes or subgroups associated with particular factors such as host developmental stage variability or sampling time after infection, thereby compromising the accuracy of our representation of the viral diversity or have significant impact on ancestral states reconstruction (74). Furthermore, the heterogeneity of our dataset introduces challenges in disentangling confounding effects among factors, making it difficult to differentiate between clusters parameters and establish clear causal relationships between variables and their impact on the viral diversity. This complexity can hinder our ability to draw definitive conclusions about the specific factors influencing viral dynamics. Moreover, comparisons between samples may be limited due to inherent differences, such as temporal dynamics in viral spread or diversity across species, host stages, or sampling years. These differences introduce additional complexities and may hinder direct comparisons between samples, potentially impacting the generalizability of our findings.

CST frequency and OsHV-1 evolutionary history

Herpes like viruses particles in bivalves were first observed by Farley *et al.* in 1972 infecting adults of the *C. virginica* oyster species in USA during laboratory experiments (22). Shortly after, infections of another oyster species, *O. edulis*, were detected in the United Kingdom (12). Since these initial detections, OsHV-1 infections have been reported in various marine bivalve species, including *M. gigas* (25–27), *P. maximus* (15), *R. philippinarum* (20, 21) and *R. decussatus* (20, 21).

OsHV-1 is a measurably evolving virus (37). At the global genomic scale, the mean evolution rate inferred in this study was 7.19E-05 nucleotide substitutions per site per year for the global phylogeny which is comparable to the one inferred previously (37). Within the *M. gigas* cluster, mutation rates vary from 5.99E-05 to 3.29E-05 nucleotide substitutions per site per year while within the *O. edulis* cluster, mutation rate is slightly lower, estimated at 1.16E-05 nucleotide substitutions per site per year. These results were expected, since the effective population size of *M. gigas* is larger than that of *O. edulis*, which are now in decline. Nevertheless, the evolution rate of OsHV-1 is much higher than the evolution rate estimated for members of the Herpesvirales order under co-speciation respectively, 3.0E-09 nucleotide substitutions per site per year for mammalian and a*via*n herpesviruses (75), 4.4E-09 nucleotide substitutions per site per year for herpesviruses collected only from mammals (76) and 4.37E-09 nucleotide substitutions per site per year for cyprinivirus (77). These differences in evolution rates can be explained by the oyster farming conditions and by wild population sizes. In fact, oysters

live in close proximity to one another, in large population densities, which may facilitate virus transmission between oysters. It is well-known that the easier viral transmission is, the more viruses evolve and generate diversity because of increased replication rates (78). However, in this case, *M. gigas* virus groups are much larger than those of *O. edulis*, so the rate of evolution should be much higher for viruses infecting *M. gigas*, unless a second host with large populations is infected by the *O. edulis* genotypes and maintains a high mutation rate. A second hypothesis to explain the higher mutation rates in OsHV-1 flat oyster cluster than in Pacific oyster cluster is closely related to the reproduction cycle of the flat oyster, which stores larvae in its shell before releasing them to fix on a substrate (79). During this period, when the larvae are contained within the mother's shell, larval densities are extremely high, which facilitates viral transmission and influences the mutation rate of the viral genome.

The analysis of ancestral traits indicated that OsHV-1 has evolved and adapted to two distinct lineages, one infecting *M. gigas* and the other infecting *O. edulis*. These two lineages clustered together based on host species and year of sampling. The MRCA of these two lineages was traced back to *M. gigas*, suggesting that this host species may have played a crucial role in the evolution and emergence of the reference genotype. Intriguingly, the reconstruction of ancestral traits using genomes with and without recombinant regions produced two different tree topologies, indicating the possibility of coinfection of a single oyster by the two genotypes and the occurrence of a recombination event between these genotypes (80), which could have contributed to the evolution and emergence of the virus infecting *O. edulis*. Furthermore, although the first event of CST led to the emergence of two highly distinct genotypes, no subsequent CST or inter-species recombination events have been observed while the two host species *O. edulis* and *M. gigas* are living in nearby environment. As the only three other OsHV-1 genomes assembled from infected *C. farreri* (MF509813 and GQ153938.1) and *A. broughtonii* (KP412538.1) were not collected in France and are genetically distant from the genotypes sampled from *M. gigas* and *O. edulis* (37, 43), they were not included in the study. Moreover, no other samples collected from different host species were available to complete the dataset. These results provide an initial insight into the evolutionary history of OsHV-1, but need to be supplemented by the contribution of other viral genomes sampled from different species to understand the involvement of different host species in viral diversity and infer the full range of CST events.

The frequency of CST occurrences is determined by the presence of facilitators or barriers encountered by the virus (65, 81). CST events are particularly frequent in RNA viruses because of their rapid evolution (63). One of the most well-known viruses that can cross-species barriers is the Influenza virus, which is capable of rapidly adapting to different host species. The success of these viruses in

infecting new hosts has been linked to their ability to adapt to a wide range of host species, notably by changing their virus surface proteins (82) and by using different strategies to evade the host's immune system, such as the production of non-structural proteins that interfere with the host immune system (83). Moreover, their segmented genomes provide advantages for antigenic drift, antigenic shift, and recombination between strains infecting the same host, thus producing new combinations of genes that potentially favor the emergence of new viruses and increase their ability to infect and spread in new and unexpected host species (83–86). By contrast, CST events in herpesviruses are extremely rare due to intrinsic genomic properties that lead to a persistent or latency stage rather than CST. Firstly, herpesviruses use specific receptors, such as Class III membrane protein or entry glycoprotein, that are specific to the herpesvirus family to enter host cells (87). In the case of OsHV-1, only one Class III membrane protein was predicted in OsHV-1 genome, mainly involved in catalyzing fusion of viral and cellular membranes, suggesting a high specificity of the viral genes for a particular host species. In addition, herpesviruses have co-evolved with their specific host species over a long time, leading to specific adaptations that enhance their ability to infect and persist within their natural host rather than infecting multiple hosts (88–90). A herpesvirus that is adapted to one species may not be able to effectively be transmitted to or evade the immune system of a different species, reducing its ability to establish an infection (91). Furthermore, herpesviruses have a low mutation rate, because of error correcting activity of exonuclease-polymerase, which does not allow a rapid evolution and adaptation to the new host species.

In the specific case of OsHV-1, CST events between oysters may be potentially limited by ecological and physiological barriers. In fact, only a few oyster production sites (Brittany, Normandy and Languedoc-Roussillon, Marennes-Oléron and Arcachon) bring the two oyster species together. In addition, given the phylogeny of oyster species and the genetic distance between them, physiological barriers are inevitably present, notably due to the immune defenses of the different host species (92). In our study, high selective sweep signals were detected within the ORFs coding for transmembrane and membrane proteins, as well as for other functions involved in the maintenance of the replication process and packaging of the virus suggesting adaptation of the different OsHV-1 genotypes to a specific host species. Moreover, a previous study involved injecting flat oysters with OsHV-1 virus isolated from naturally moribund Pacific oysters, resulting in lower viral loads and reduced mortality rates compared to *M. gigas* infections. These findings suggest the presence of specific host-virus interactions and potentially during the immune response (93). Based on the results, the co-cultivation of the two oyster species *M. gigas* and *O. edulis* does not stand for a major infection hazard, but it could increase genetic recombination and cross-species transmission events opportunities thus potentially promoting the emergence of variants that could be problematic.

Based on our findings, it seems that the French ancestral OsHV-1 genotype did originate from *M. gigas* but rather was differentiated into at least two genotypes that subsequently adapted to different marine bivalve host species such as *O. edulis* and *M. gigas*. The adaptation process continued through coevolution, which was probably facilitated by host-specific factors such as immunity, living environment (temperature, salinity, light, host density), and food availability, as well as viral genomic properties such as mutation rates, host entry receptor specificity, and low genetic diversity.

Considering the obtained results and the differentiation among various viral lineages, it would be prudent to improve the classification of Malacoherpesviridae by incorporating the observed genotypes found in different host species.

To deepen our understanding of the evolutionary processes and host-virus interactions, research should focus on investigating the coevolution between OsHV-1 and a wider range of host species using phylogenetic approaches. This would involve sampling from diverse host species across different geographical regions. By conducting such studies, we can gain valuable insights into the diversification, adaptation, and coevolutionary dynamics of OsHV-1 with its various hosts.

Materials and Methods

OsHV-1 sampling

To study the diversity of OsHV-1 in different host species, archived samples from Ifremer National NRL and EURL collection were used. OsHV-1 viruses were sampled from wild Pacific oysters *M. gigas* (seventeen samples) collected in the field during OsHV-1 infection outbreaks between 2005 and 2010 from individuals of different developmental stages (larvae, spats and adults) at five sites along the French coast (Antioche inlet, Palavas lagoon, Thau lagoon, Brittany inlet, Pirou Agon-Coutainville). Wild flat oysters *O. edulis* (twenty-three samples) collected in the field from April to August between 2007 and 2009 in the Quiberon bay as describe by (94) and one sample has been collected in Sweden (Table S1, Fig. 8A). This last sample was used to compare the viral genotype from France and Sweden within *O. edulis* individuals to widen the geographical area studied. Criterion used to select *M. gigas* samples was i) the geographic proximity to *O. edulis* samples and ii) the year of samples collection, which covers the *O. edulis* samples collection period. Before DNA extraction, animals have been stored at -80°C or in ethanol from 6 to 14 years.

In addition, 22 *O. edulis* pools of larvae were sampled from adults' gills and were preserved in Davidson's fixative for histological analyses.

Samples were named according to the nomenclature recommended by (95) which is made of the virus name, the sampling location, the date of collection, the host species, and the sample number (*e.g.* OsHV-1/Rivedoux.FRA/2008/Cg-S-053 for the sample number 053 collected in 2008, in Rivedoux, from *M. gigas* spat). However, hereafter, shortened names have been used to facilitate reading (*e.g.* CG-08-Ri-S-053 for the same sample, see Table S1 for correspondence).

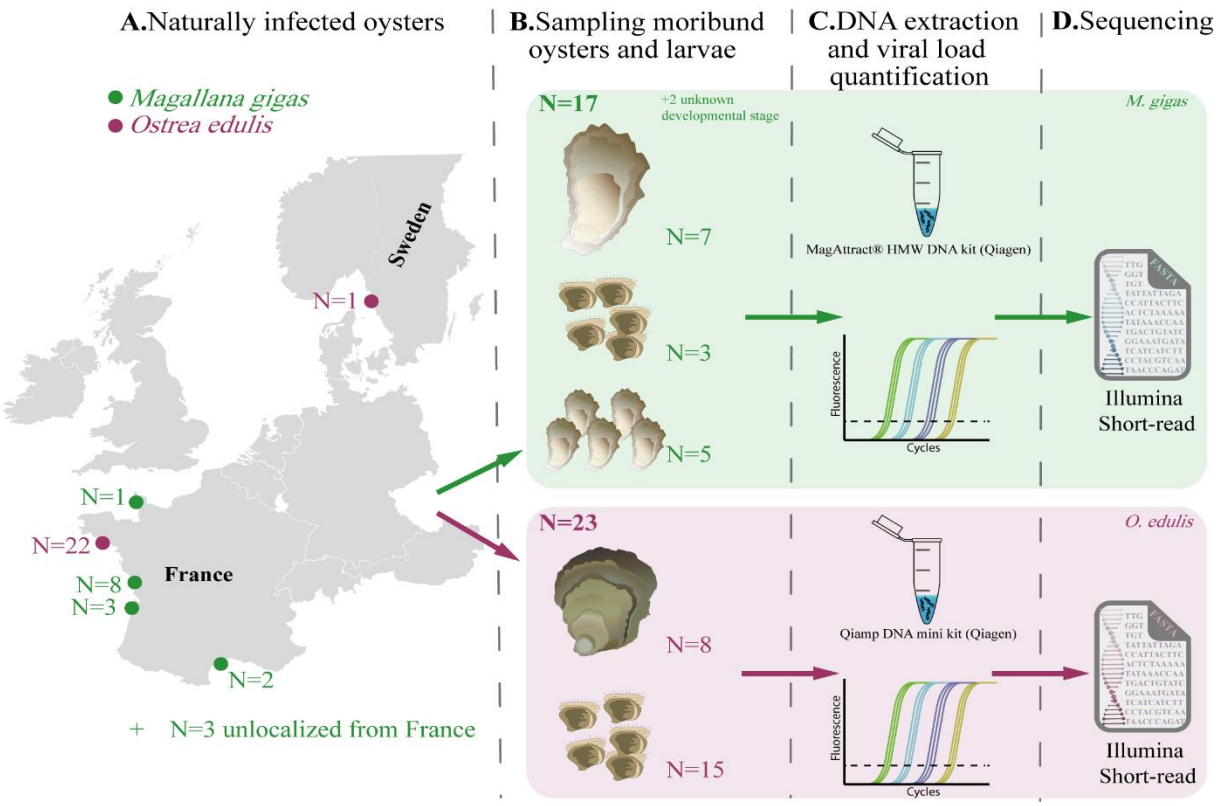


Fig 8. Schematic representation of the experimental sampling design. *(A-B) Larvae, spat, and adults naturally infected with the virus were collected from M. gigas and O. edulis in France and Sweden. (C) DNA was extracted from individual tissues and the viral copy number was assessed by qPCR. D) DNA samples were sequenced using the Illumina short-read technology. M. gigas and O. edulis samples and treatments were represented respectively in green and purple.*

Histological assessment of tissue lesions during OsHV-1 infection

OsHV-1 is responsible for tissue lesions that have already been described in Pacific oysters (96). Lesions in *O. edulis* tissues related to OsHV-1 infection have been rarely described by histology (45, 46, 93). In order to detect tissue lesions and assess the presence of OsHV-1 within flat oyster tissues, a histological analysis has been performed. Six pools of Ostrea edulis larvae collected in the Quiberon bay and previously stored in Davidson's fixative (see above) were transferred to 70% ethanol after 24-48h of

fixation before processing standard histological protocol of paraffin inclusion (97). Embedded paraffin blocks were then sectioned (2-3µm) and stained with hematoxylin and eosin. Histological sections were examined for pathological conditions under light microscopy (Olympus BX63) and photography were taken with an associate camera (Olympus DP74).

DNA extraction, viral copy numbers quantification and sequencing

Depending on host species, two different methods were used for samples DNA extraction (Fig. 1C). For *M. gigas*, DNA was extracted from mantle of moribund individual spats, from pool of larvae or from pool of shred oysters using the Qiamp DNA Mini kit (Qiagen) according to the manufacturer's protocol. For *O. edulis*,DNA was extracted from gills of individual adults or from pool of larvae using the MagAttract® HMW DNA kit (Qiagen) according to the manufacturer's protocol.

DNA purity and concentration were assessed using a Nano-Drop ND-1000 spectrometer (Thermo Scientific) and Qubit® dsDNA BR assay kits (Molecular Probes Life Technologies), respectively. Then quantification of viral copy numbers was carried out in duplicate by quantitative PCR using a Mx3005 P Thermocycler (Agilent) as described by (98). Each reaction contained 1µl of Milli-Q Water, 2µl of each primer at 5,5 µM, OsHVDP For (forward) 5'-ATTGATGATGTGGATAATCTGTG-3' and OsHVDP Rev (reverse) 5'-GGTAAATACCATTGGTCTTGTTCC-3', 10µl of Brilliant III Ultra-Fast SYBR® Green PCR Master Mix (Agilent), and finally 25 ng of DNA sample from infected oysters or Milli-Q water (non-template control) in a total volume of 20µl. The qPCR program consisted in 3 min at 95 °C followed by 40 cycles of amplification at 95 °C for 5s and 60 °C for 20s. Results were expressed as the log-transformed copy number of viral DNA per nanogram of extracted DNA (cp.ng-1).

Illumina DNA-Seq library preparation and sequencing were performed at four sequencing platforms (see Table S1 for details about libraries preparation and sequencing technologies). Libraries were prepared using (i) True seq DNA PCR free kit, (ii) Shotgun PCR-free library preparation kit (Lucigen) or (iii) Illumina genomic PCR-free kit. The sequencing step was performed using (i) HiSeqTM 2500 device (paired-ends, 100 pb) by GenoToul (Toulouse, France), (ii) HiSeqTM 4000 device (paired-ends, 150 bp) by the Lille Integrated Genomics Advanced Network for personalized medicine (LIGAN-PM) sequencing platform (UMR 8199, CNRS, Lille, France) or (iii) NovaSeqTM 6000 device (paired-ends, 150 bp) by the Genome Quebec Company (Genome Quebec Innovation Center, McGill University, Montreal, Canada) or by Fasteris Life Science Genesupport (Plan-les-Ouates, Switzerland).

De novo assembly of OsHV-1 genomes

De novo OsHV-1 genome assemblies were obtained by processing sequencing reads as previously described (37,38 and see 48 for details on the bioinformatic pipeline). Briefly, for each library, raw

reads were quality filtered (PHRED score > 31) and adapters trimmed using Fastp v0.20.1 (with parameter -q 31; (99). In order to estimate the number of OsHV-1 reads and coverage in each sample, filtered reads were aligned to the reference genome of OsHV-1 (accession number: NC_005881.2; (12) using Bowtie2 v2.4.1 (parameter --local and --no-unal; 50). Meanwhile, reads were aligned to the reference genome of *M. gigas* (accession number: GCF_902806645.1, (101)) or *O. edulis* (accession number: GCA_024362755.1, (102)) according to the species from which the sample was collected, using Bowtie 2 v2.4.1 (parameter --local). Non-oyster reads were selected with SAMtools v1.9 (option view -f 4,(103)) and Seqtk v1.2 subseq (default parameters; (104)). Non-oyster reads were then *de novo* assembled with metaSPAdes v3.14.0 (parameter --meta; (105)). OsHV-1 contigs were selected using BLASTn v2.12.0 (options -max_target_seqs 1, -evalue "1e-10"; (106)) against the OsHV-1 reference genome. OsHV-1 contigs were then manually re-ordered in Geneious Prime v2022.2.2 by visualizing reads alignment on the *de novo* assembled contigs to identify overlaps between trimmed reads at each contig extremities. To avoid inverted repeat regions reconstruction issues for all subsequent analyses, every OsHV-1 genome was reconstructed as non-redundant (NR) genome containing only one copy of each duplicated regions (IRL and IRS) as described by (107), (42) and (43). NR-genomes were validated and corrected by similarity matrix visualization of each genome to itself using the dotplot tool (based on the EMBOSS v6.5.7 tool dottup) implemented in Geneious Prime v2022.2.2. To verify and clean the final assemblies as well as to assess coverage uniformity, quality filtered reads were aligned back to the new constructed OsHV-1 NR-genomes using Bowtie 2 v2.4.1 (parameter --end-to-end).

Genomic architecture and ORF predictions

The genomic annotation in this study has been carried out at both genome and loci scales, including the identification of genomic regions (U_L, IR_L, X, IR_S, U_S), ORF prediction and functional annotation, signal peptides and transmembrane domain prediction.

Genomic annotations

At the genome scale, the structural genomic regions annotation, was performed using a combination of tools. Firstly, BLASTn v2.12.0 (106) implemented in Geneious Prime v2022.2.2 was used to compare each OsHV-1 NR-genome assembled in this study with the OsHV-1 reference genome (NC_005881.2) and transfer raw annotation to assembled NR-genome. Secondly, the structural annotations were corrected using coverage plots generated after remapping reads to control and clean the assembly. By observing where the coverage levels experienced changes or abrupt shifts, it allows the identification and correction of the exact boundaries between repeated and unique regions.

ORF, signal peptides and transmembrane domains predictions

To ensure consistency in the predictions across different genomes and to facilitate comparability of polymorphisms and gene positions, an ancestral sequence of OsHV-1 for each host species was reconstructed. To this end, first multiple alignment of genomes sampled from *M. gigas* (twenty genomes) or from *O. edulis* (twenty-three genomes) was performed using MAFFT v1.4.0 (108). Analyses were performed using the model of nucleotide substitution HKY+G+I based on Akaike Information Criterion (AIC) obtained using MEGA v11.0.11 (109). Next, ancestral sequences of studied OsHV-1 genomes for both host species were reconstructed using Beast v1.10.4 (110). Different combinations of coalescent models (*i.e.* constant size, GMRF Bayesian skyride, expansion growth, and exponential growth) and molecular clocks (*i.e.* strict clock and uncorrelated relaxed clock) were tested to find out the most appropriate model. The maximum likelihood estimator (MLE) calculated by path sampling (PS) and stepping-stone sampling (SSS) were compared to select the most appropriate model. According to results obtained, ancestral sequences were reconstructed using a strict clock and a constant size coalescent model for the *M. gigas* group, and an uncorrelated relaxed clock and an expansion growth model for the *O. edulis* group and MCMC chains of 10,000,000 steps were ran and sub-sampled every 1,000,000 generations.

Subsequently, local predictions of genes, signal peptides and transmembrane domain on the two reconstructed ancestral sequences were performed. ORFs were predicted with Prodigal v2.6.3 (111) with default parameters, a tool already used for double-stranded DNA viruses infecting eukaryotes (112). Functional annotations of predicted proteins were performed with EggNOG-mapper v2.1.8 (113). Signal peptides were predicted using SignalP v6.0 (114) and transmembrane domains using DeepTMHMM v1.0.24 (115).

Within individual's genomic diversity

Analysis of polymorphisms rarefaction

To verify that sequencing errors was not a major driver of intra-individual diversity and investigate the degree to which sequencing effort affected the number of intra-individual variations detected, a rarefaction analysis on the number of detected variations was performed for each sequencing library (42, 116). To this end, an iterative variant calling analysis was performed on subsampled reads that were mapped on the previously assembled genomes. In order to accelerate analyses, a nonlinear subsampling incrementation was performed based on 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 50%, 80% and 100% of the total number of reads of libraries. After each

subsampling/variant calling step, variants were classified as SNPs (single nucleotide polymorphism) and InDels (insertion or deletions) and counted in order to increment rarefaction curves.

Variant calling

Then, to determine the genomic diversity within samples, variations were called against each of the 40 genomes previously assembled, using Freebayes v1.3.5, (117) with the following settings: --use-mapping-quality, --min-repeat-entropy 1, --haplotype-length 0, --min-alternate-count 5, --pooled-continuous, --hwe-priors-off, --allele-balance-priors-off and --gvcf. All the variations were then normalized, multiallelic variants and biallelic blocks were decomposed and then sorted and deduplicates (using programs normalize, decompose, decompose_blocksub, sort and uniq) with vt v2015.11.10 (118) using default parameters. Polymorphisms detected by Freebayes were mapped on the corresponding consensus, then all the forty-one consensuses with their corresponding annotated polymorphisms were aligned using MAFFT v1.4.0 (108) implemented in Geneious Prime to re-coordinate all polymorphisms. Re-coordinated polymorphisms were then used to create unique identifier as described by (42). Briefly, unique identifiers are composed of (i) the position of the polymorphism along the multiple alignment, (ii) the reference allele (REF) and (iii) the alternative allele (ALT) (*i.e.* position_REF >ALT).

Statistical description of the intra-individual diversity

A MultiDimensional Scaling (MDS) plots was used to examine the impact of host species on the genetic diversity of OsHV-1. The first MDS analysis summarized the allelic frequencies of each SNP within the sequencing libraries, allowing to explore the distances between pairs of SNPs in terms of their allelic frequencies within each library. This information was linked to those of host species for the visualization step. The second MDS was performed using a spearman correlation matrix based on individuals' genetic similarities. The number of clusters of individuals were searched using the kmeans method and were represented on the MDS plot. Then, a genotype association analysis was performed using the co-occurrence matrix of genotypes within individuals, as well as the adjacency and incidence of vertices on each other. Connections between edges and their weights were used to represent the strength of interaction between individuals in the MDS graph.

Both MDS and plotting were performed in R v4.2.2 using stats v4.2.2 (119), igraph v1.3.5 (120) and ggplot2 v3.4.0 (121) packages.

To examine the density of SNPs within each genomic region (*i.e.* U_L, IR_L, X, IR_S, U_S), the ratio of the number of SNPs detected within a region to the length of the region ((Number of SNPs)/(Genomic region length)) has been calculated.

Whole genome alignment and recombination detection

Identification and description of OsHV-1 structural variations

To compare the genomic architecture between viral genotypes and determine large-scale evolutionary events, the forty reconstructed NR-genomes were aligned with three published OsHV-1 genomes (OsHV-1 µVarA, *M. gigas*, 2010, France: KY242785, OsHV-1 µVarB, *M. gigas*, 2011, Ireland: KY271630, and OsHV-1 Ref., *M. gigas*, 1995, France: NC_005881.2) using Mauve v1.1.3 (progressive Mauve algorithm, with default parameters, (122)). These three genomes were available in public databases and were also assembled from Pacific oysters. Furthermore, the genotypes are very similar to those detected in our libraries. However, the three published genomes assembled from clams or scallops were not include as they are too distant from the genotypes studied in this research, and not enough genomes are available for these species. Our multiple alignment of the forty-three OsHV-1 genomes allowed us to identify typical genomic structures based on rearranged blocks and to classify genomes according to their structure. Specifically, five distinct genomic structures were identified in this study.

To plot the five types of genomic structures observed in our samples, pairwise alignments of the five representative genotypes against the OsHV-1 reference genome (NC_005881.2) were conducted using minimap2 v2.24 (parameters -ax asm5 --eqx; (123)). Alignments were then sorted and indexed using SAMtools v1.14 (103). Structural rearrangements were characterized using Syri v1.5.4 (124) and plot with the python package plotsr v0.5.3 (125).

Manual restructuring of OsHV-1 genomes and multiple alignment

As some genomes contained large rearrangements, thirty assembled genomes were manually modified, from the Mauve multiple alignment, in Geneious Prime to ensure that the genomic architecture was identical for every genome, making them comparable to each other. Genomes were then aligned using MAFFT v1.4.0 (108) that allowed the computation of nucleotide identity percentages between pair of genomes.

Recombination events detection

Gubbins v3.3 (126) with default settings was used to identify regions that might be impacted by recombination events among the assembled OsHV-1 NR-genomes. Additionally, to confirm results obtained with Gubbins, linkage disequilibrium values (D') were computed to test for random association of allelic distribution and then explore the possibility of recombination in our dataset. D' values were computed in DNAsp6 (127) based on informative sites and significance was assessed using a Fisher's exact test. $D^{2\prime}$ values were then plotted as a heatmap in R using the heatmap function from

the ggplot package v3.4.0 (121). All genomic regions detected as potentially recombinant were removed from the downstream analyses (Table S5).

Consensus based diversity computation

Population structure and evolution analysis through phylogenetics

First, a phylogenetic analysis has been performed to explore the population structure and evolutionary relationships among OsHV-1 genotypes. A phylogenetic search that applied a maximum-likelihood approach was performed on the alignment of the forty-three genomes using PhyML v3.3.2 (44) with 1000 bootstraps replications. The best fitting model of evolution for the samples was found using MEGA v11.0.11 (109) based on AIC. Therefore, a HKY+G+I model was used. Results were visualized using Fig.Tree v1.4.4 (available at http://tree.bio.ed.ac.uk/software/Fig.tree/).

Investigating genetic differentiation between groups

Then, description of consensus-based diversity was done using population genetic statistics such as the de*via*tion of allelic frequencies from the neutrality equilibrium (Tajima's D) and the differentiation level of viruses infecting the two host populations (GST).

To obtain insights into de*via*tion from demographic equilibrium and/or selective neutrality, Tajima's D values were computed in DNAsp6 v6.12.03 (127) for 1kb-windows along the alignment in order to standardize the amount of available information among analyses. Tajima's D values were calculated for the overall population (forty-three genomes) as well as for samples from each host species separately. Under the neutral evolutionary model, Tajima's D is expected to be null. At a locus scale, positive Tajima's D values can be induced by balancing selection, due to the maintenance of highly divergent variants and, negative Tajima's D values can be the result of purifying selection or of a recent selective sweep, which both result in an excess of weakly divergent alleles. At the whole genome scale, positive Tajima's D values reflect a decrease of the group effective size thus favoring intermediate frequency alleles, while negative Tajima's D values suggest an increase of the population effective size implying expansion of rare alleles.

Then differentiation measures were computed between groups as they reflect the genetic variability among groups of individuals relative to the variability of a whole sample. The pairwise Nei GST method was used and, per locus GST values were computed for 1-kb windows with the option diff_stats in R using the package mmod v1.3.3 (128). GST values were also computed for 100-bp windows within 1-kb windows that were detected as differentiated to define precisely regions involved in differentiation and facilitated the link with ORF potential function for biological hypothesis assessment.

An AMOVA was conducted to investigate the possible factors contributing to the structuring of OsHV-1 diversity at the consensus level. The global alignment of genomes was used for the AMOVA, following the same procedure as described in the previous paragraph (see "Within individual's genomic diversity" for more details).

Selective sweeps detection

To explore the potential effects of natural selection, signals of selective sweeps were searched in our datasets. A selective sweep occurs when a new beneficial mutation increases in frequency and becomes fixed, resulting in a reduction or elimination of genetic variation near the mutation. RAiSD v2.9 (129) was used to detect selective sweep signatures along OsHV-1 genomes.

Ancestral traits reconstruction

In order to analyze the temporal signal and the 'clocklikeness' of the ML tree and first explore the evolution rate of OsHV-1, TempEst v1.5.3 (130) has been used to plot the divergence of each tip from the root over time. The root-to-tip divergence (RTTD) plot has been performed separately for *O. edulis* group (twenty-three genomes) and for the two subclusters of the OsHV-1 *M. gigas* group (respectively nine and eleven genomes) using a midpoint rooted tree and a best fitting root approaches for comparisons.

Analysis were performed using a HKY+G+I model of nucleotide substitution based on AIC obtained using MEGA v11.0.11 (109). As before, different combinations of coalescent models (*i.e.* constant size, GMRF Bayesian skyride, and exponential growth) and molecular clocks (*i.e.* strict clock and uncorrelated relaxed clock) were tested to find the most appropriate model for our dataset. The maximum likelihood estimator (MLE) calculated by path sampling (PS) and stepping-stone sampling (SSS) for each model were compared. According to those, an uncorrelated relaxed molecular clock and a constant size coalescent model has been used accordingly to MLE and a MCMC chain of 10,000,000 steps was run and sub-sampled every 1,000,000 generations using Beast v1.10.4 (110). The effective sample size, representing the mixing of the population (Ne > 200), and the stationary distribution of the trace were checked in Tracer v1.7.2 (131) for all parameters, using an initial burn-in of 10%. Finally, maximum clade credibility (MCC) trees were generated from tree output files from BEAST in TreeAnnotator v1.10.4 (110), and annotated in Fig.Tree v1.4 (http://tree.bio.ed.ac.uk/software/Fig.tree/).

Modeling of demographic evolution

To assess the historical population dynamics of OsHV-1 in both host species, a demographic analysis was conducted by reconstructing past changes in population size over time. To do so, the log files of the trees that were previously generated to reconstruct ancestral sequences of OsHV-1 in both host species were used. The demographic history was then reconstructed using a GMRF Skyride model in Tracer v1.7.2 (131).

Data availability

Raw data as well as reconstructed genomes have been deposited on the NCBI database under Bioproject PRJNA999485 accession numbers SAMN36736021 to SAMN36736060 for future reference and accessibility (Table S1).

Acknowledgments

The present study was supported by the French Institute for Exploitation of the Sea (Ifremer), by DGAL (French General Directorate for Food) through the National Reference Laboratory for Mollusc Diseases and by the European-Union Reference Laboratory for Mollusc Diseases, Ifremer, La Tremblade. CP was supported by grant from the region Nouvelle-Aquitaine and from the scientific direction of Ifremer. We also thank the SeBiMER from Ifremer for their technical assistance in bioinformatic.

We declare that we have no competing interests.

References

1. Roche B, Benbow ME, Merritt R, Kimbirauskas R, McIntosh M, Small PLC, Williamson H, Guégan J-F. 2013. Identifying the Achilles' heel of multi-host pathogens: The concept of keystone "host" species illustrated by Mycobacterium ulcerans transmission. Environ Res Lett 8:045009.
2. Woolhouse MEJ, Taylor LH, Haydon DT. 2001. Population Biology of Multihost Pathogens. Science 292:1109–1112.
3. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008. Global trends in emerging infectious diseases. 7181. Nature 451:990–993.
4. Archie EA, Luikart G, Ezenwa VO. 2009. Infecting epidemiology with genetics: a new frontier in disease ecology. Trends in Ecology & Evolution 24:21–30.
5. Handel A, Rohani P. 2015. Crossing the scale from within-host infection dynamics to between-host transmission fitness: a discussion of current assumptions and knowledge. Philos Trans R Soc Lond B Biol Sci 370:20140302.
6. VanderWaal KL, Ezenwa VO. 2016. Heterogeneity in pathogen transmission: mechanisms and methodology. Functional Ecology 30:1606–1622.
7. Geoghegan JL, Senior AM, Holmes EC. 2016. Pathogen population bottlenecks and adaptive landscapes: overcoming the barriers to disease emergence. Proceedings of the Royal Society B: Biological Sciences 283:20160727.

8. Kuiken T, Holmes EC, McCauley J, Rimmelzwaan GF, Williams CS, Grenfell BT. 2006. Host Species Barriers to Influenza Virus Infections. Science 312:394–397.

9. Park M, Loverdo C, Schreiber SJ, Lloyd-Smith JO. 2013. Multiple scales of selection influence the evolutionary emergence of novel pathogens. Philosophical Transactions of the Royal Society B: Biological Sciences 368:20120333.

10. Azab W, Dayaram A, Greenwood AD, Osterrieder N. 2018. How Host Specific Are Herpesviruses? Lessons from Herpesviruses Infecting Wild and Endangered Mammals. Annu Rev Virol 5:53–68.

11. Bandín I, Dopazo CP. 2011. Host range, host specificity and hypothesized host shift events among viruses of lower vertebrates. Vet Res 42:67.

12. Davison AJ, Trus BL, Cheng N, Steven AC, Watson MS, Cunningham C, Deuff R-ML, Renault T. 2005. A novel class of herpesvirus with bivalve hosts. 1. Journal of General Virology 86:41–53.

13. Huff JL, Barry PA. 2003. B-Virus (Cercopithecine herpesvirus 1) Infection in Humans and Macaques: Potential for Zoonotic Disease. Emerg Infect Dis 9:246–250.

14. Arzul I, Corbeil S, Morga B, Renault T. 2017. Viruses infecting marine molluscs. J Invertebr Pathol 147:118–135.

15. Arzul I, Nicolas J-L, Davison AJ, Renault T. 2001. French Scallops: A New Host for Ostreid Herpesvirus-1. 2. Virology 290:342–349.

16. Arzul I, Renault T, Lipart C, Davison AJ. 2001. Evidence for interspecies transmission of oyster herpesvirus in marine bivalves. 4. Journal of General Virology 82:865–870.

17. Carnegie RB, Arzul I, Bushek D. 2016. Managing marine mollusc diseases in the context of regional and international commerce: policy issues and emerging concerns. Philosophical Transactions of the Royal Society B: Biological Sciences 371:20150215.

18. Bai C-M, Morga B, Rosani U, Shi J, Li C, Xin L-S, Wang C-M. 2019. Long-range PCR and high-throughput sequencing of Ostreid herpesvirus 1 indicate high genetic diversity and complex evolution process. Virology 526:81–90.

19. Xia J, Bai C, Wang C, Song X, Huang J. 2015. Complete genome sequence of Ostreid herpesvirus-1 associated with mortalities of Scapharca broughtonii broodstocks. 1. Virol J 12:110.

20. Renault, Arzul. 2001. Herpes-like virus infections in hatchery-reared bivalve larvae in Europe: specific viral DNA detection by PCR. 3. J Fish Diseases 24:161–167.

21. Renault T. 1998. Infections herpétiques chez les invertébrés : détection de virus de type herpès chez les mollusques bivalves marins. 2. Virologie 401–403.

22. Farley CA, Banfield WG, Kasnic G, Foster WS. 1972. Oyster Herpes-Type Virus. 4062. Science 178:759–760.

23. Comps M. 1988. Epizootic diseases of oysters associated with viral infections. 18. Am Fish Soc Spe Publ 23–37.

24. Renault T, Lipart C, Arzul I. 2001. A herpes-like virus infects a non-ostreid bivalve species: virus replication in Ruditapes philippinarum larvae. Diseases of Aquatic Organisms 45:1–7.

25. Burioli EAV, Prearo M, Houssin M. 2017. Complete genome sequence of Ostreid herpesvirus type 1 µVar isolated during mortality events in the Pacific oyster Crassostrea gigas in France and Ireland. Virology 509:239–251.

26. Nicolas J-L, Comps M, Cochennec N. 1992. Herpes-like virus infecting Pacific oyster larvae. 1. Bull Eur Ass Fish Pathol 12:11.

27. Renault T, Cochennec R, Le Deuff RM, Chollet B. 1994. HERPES-LIKE VIRUS INFECTING JAPANESE OYSTER (CRASSOSTREA GIGAS) SPAT 3.

28. Grizel H, Heral M. 1991. Introduction into France of the Japanese oyster (Crassostrea gigas). 47. Journal du Conseil - Conseil International pour l'Exploration de la Mer 399–403.

29. FAO. 2023. Production mondiale par source de production de la pêche et aquaculture - Quantité (1950 - 2021).

30. Burge CA, Griffin FJ, Friedman CS. 2006. Mortality and herpesvirus infections of the Pacific oyster Crassostrea gigas in Tomales Bay, California, USA. Dis Aquat Organ 72:31–43.

31. EFSA Panel on Animal Health and welfare (AHAW). 2010. Scientific Opinion on the increased mortality events in Pacific oysters, Crassostrea gigas: Oyster mortality. 11. EFSA Journal 8:1894.

32. Paul-Pont I, Dhand NK, Whittington RJ. 2013. Spatial distribution of mortality in Pacific oysters Crassostrea gigas: reflection on mechanisms of OsHV-1 transmission. Dis Aquat Organ 105:127–138.

33. Renault T, Moreau P, Faury N, Pepin J-F, Segarra A, Webb S. 2012. Analysis of Clinical Ostreid Herpesvirus 1 (Malacoherpesviridae) Specimens by Sequencing Amplified Fragments from Three Virus Genome Areas. Journal of Virology 86:5942–5947.

34. Davison AJ. 2002. Evolution of the herpesviruses. 1–2. Veterinary Microbiology 86:69–88.

35. Segarra A, Pépin JF, Arzul I, Morga B, Faury N, Renault T. 2010. Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, Crassostrea gigas, in France in 2008. 1. Virus Research 153:92–99.

36. Arzul I, Renault T, Lipart C. 2001. Experimental herpes-like viral infections in marine bivalves: demonstration of interspecies transmission. Dis Aquat Org 46:1–6.

37. Morga B, Jacquot M, Pelletier C, Chevignon G, Dégremont L, Biétry A, Pepin J-F, Heurtebise S, Escoubas J-M, Bean TP, Rosani U, Bai C-M, Renault T, Lamy J-B. 2021. Genomic Diversity of the Ostreid Herpesvirus Type 1 Across Time and Location and Among Host Species. Front Microbiol 12:711377.

38. Maiden MCJ, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND. 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol 11:728–736.

39. Sunde J, Yıldırım Y, Tibblin P, Forsman A. 2020. Comparing the Performance of Microsatellites and RADseq in Population Genetic Studies: Analysis of Data for Pike (Esox lucius) and a Synthesis of Previous Studies. Frontiers in Genetics 11.

40. Abbadi M, Zamperin G, Gastaldelli M, Pascoli F, Rosani U, Milani A, Schivo A, Rossetti E, Turolla E, Gennari L, Toffan A, Arcangeli G, Venier P. 2018. Identification of a newly described OsHV-1 µvar from the North Adriatic Sea (Italy). 5. Journal of General Virology 99:693–703.

41. Ren W, Chen H, Renault T, Cai Y, Bai C, Wang C, Huang J. 2013. Complete genome sequence of acute viral necrosis virus associated with massive mortality outbreaks in the Chinese scallop, Chlamys farreri. Virology Journal 10:110.

42. Delmotte J, Pelletier C, Morga B, Galinier R, Petton B, Lamy J-B, Kaltz O, Avarre J-C, Jacquot M, Montagnani C, Escoubas J-M. 2022. Genetic diversity and connectivity of the Ostreid herpesvirus 1 populations in France: A first attempt to phylogeographic inference for a marine mollusc disease. Virus Evolution 8:veac039.

43. Dotto-Maurel A, Pelletier C, Morga B, Jacquot M, Faury N, Dégremont L, Bereszczynki M, Delmotte J, Escoubas J-M, Chevignon G. 2022. Evaluation of tangential flow filtration coupled to long-read sequencing for ostreid herpesvirus type 1 genome assembly. Microbial Genomics 8.

44. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321.

45. da Silva P, Renault T, Fuentes J, Villalba A. 2008. Herpesvirus infection in European flat oysters Ostrea edulis obtained from brood stocks of various geographic origins and grown in Galicia (NW Spain). Dis Aquat Org 78:181–188.

46. Comps M, Cochennec N. 1993. A Herpes-like Virus from the European Oyster Ostrea edulis L. Journal of Invertebrate Pathology 62:201–203.

47. Burioli EAV, Prearo M, Riina MV, Bona MC, Fioravanti ML, Arcangeli G, Houssin M. 2016. Ostreid herpesvirus type 1 genomic diversity in wild populations of Pacific oyster Crassostrea gigas from Italian coasts. Journal of Invertebrate Pathology 137:71–83.

48. Mahiet C, Ergani A, Huot N, Alende N, Azough A, Salvaire F, Bensimon A, Conseiller E, Wain-Hobson S, Labetoulle M, Barradeau S. 2012. Structural Variability of the Herpes Simplex Virus 1 Genome In Vitro and In Vivo. 16. J Virol 86:8592–8601.

49. López-Muñoz AD, Rastrojo A, Kropp KA, Viejo-Borbolla A, Alcamí A. 2021. Combination of long- and short-read sequencing fully resolves complex repeats of herpes simplex virus 2 strain MS complete genome. Microb Genom 7:000586.

50. Chou J, Roizman B. 1985. Isomerization of herpes simplex virus 1 genome: identification of the cis-acting and recombination sites within the domain of the a sequence. Cell 41:803–811.

51. Chaitanya KV. 2019. Structure and Organization of Virus Genomes. Genome and Genomics 1–30.

52. Pérez-Losada M, Arenas M, Galán JC, Palero F, González-Candelas F. 2015. Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. Infect Genet Evol 30:296–307.

53. Turner PE, Elena SF. 2000. Cost of Host Radiation in an RNA Virus. Genetics 156:1465–1470.

54. Kolb AW, Brandt CR. 2020. Genomic nucleotide-based distance analysis for delimiting old world monkey derived herpes simplex virus species. BMC Genomics 21:436.

55. Longdon B, Day JP, Alves JM, Smith SCL, Houslay TM, McGonigle JE, Tagliaferri L, Jiggins FM. 2018. Host shifts result in parallel genetic changes when viruses evolve in closely related species. PLoS Pathog 14:e1006951.

56. Villordo SM, Filomatori CV, Sánchez-Vargas I, Blair CD, Gamarnik AV. 2015. Dengue Virus RNA Structure Specialization Facilitates Host Adaptation. PLOS Pathogens 11:e1004604.

57. Lee K, Kolb AW, Sverchkov Y, Cuellar JA, Craven M, Brandt CR. 2015. Recombination Analysis of Herpes Simplex Virus 1 Reveals a Bias toward GC Content and the Inverted Repeat Regions. J Virol 89:7214–7223.

58. Renner DW, Szpara ML. 2017. Impacts of Genome-Wide Analyses on Our Understanding of Human Herpesvirus Diversity and Evolution. Journal of Virology 92:e00908-17.

59. Burrel S, Boutolleau D, Ryu D, Agut H, Merkel K, Leendertz FH, Calvignac-Spencer S. 2017. Ancient Recombination Events between Human Herpes Simplex Viruses. Molecular Biology and Evolution 34:1713–1721.

60. Koelle DM, Norberg P, Fitzgibbon MP, Russell RM, Greninger AL, Huang M-L, Stensland L, Jing L, Magaret AS, Diem K, Selke S, Xie H, Celum C, Lingappa JR, Jerome KR, Wald A, Johnston C. 2017. Worldwide circulation of HSV-2 × HSV-1 recombinant strains. Sci Rep 7:44084.

61. Sijmons S, Thys K, Mbong Ngwese M, Van Damme E, Dvorak J, Van Loock M, Li G, Tachezy R, Busson L, Aerssens J, Van Ranst M, Maes P. 2015. High-throughput analysis of human cytomegalovirus genome diversity highlights the widespread occurrence of gene-disrupting mutations and pervasive recombination. 15. J Virol 89:7673–7695.

62. Lassalle F, Depledge DP, Reeves MB, Brown AC, Christiansen MT, Tutill HJ, Williams RJ, Einer-Jensen K, Holdstock J, Atkinson C, Brown JR, Van Loenen FB, Clark DA, Griffiths PD, Verjans GMGM, Schutten M, Milne RSB, Balloux F, Breuer J. 2016. Islands of linkage in an ocean of pervasive recombination reveals two-speed evolution of human cytomegalovirus genomes. Virus Evol 2:vew017.

63. Parrish CR, Holmes EC, Morens DM, Park E-C, Burke DS, Calisher CH, Laughlin CA, Saif LJ, Daszak P. 2008. Cross-Species Virus Transmission and the Emergence of New Epidemic Diseases. Microbiology and Molecular Biology Reviews 72:457–470.

64. Thakur S, Sasi S, Pillai SG, Nag A, Shukla D, Singhal R, Phalke S, Velu GSK. 2022. SARS-CoV-2 Mutations and Their Impact on Diagnostics, Therapeutics and Vaccines. Front Med (Lausanne) 9:815389.

65. Webby R, Hoffmann E, Webster R. 2004. Molecular constraints to interspecies transmission of viral pathogens. Nat Med 10:S77–S81.

66. Brito AF, Pinney JW. 2020. The evolution of protein domain repertoires: Shedding light on the origins of the Herpesviridae family. Virus Evolution 6:veaa001.

67. Al-Khatib HA, Smatti MK, Ali FH, Zedan HT, Thomas S, Ahmed MN, El-kahlout RA, Al Bader MA, Elgakhlab D, Coyle PV, Abu-Raddad LJ, Al Thani AA, Yassine HM. 2022. Comparative analysis of within-host diversity among vaccinated COVID-19 patients infected with different SARS-CoV-2 variants. iScience 25:105438.

68. Gu H, Quadeer AA, Krishnan P, Ng DYM, Chang LDJ, Liu GYZ, Cheng SMS, Lam TTY, Peiris M, McKay MR, Poon LLM. 2023. Within-host genetic diversity of SARS-CoV-2 lineages in unvaccinated and vaccinated individuals. Nat Commun 14:1793.

69. Kieleczawa J. 2006. Fundamentals of Sequencing of Difficult Templates—An Overview. J Biomol Tech 17:207–217.

70. Lupo C, Ezanno P, Arzul I, Garcia C, Jadot C, Joly J-P, Renault T, Bareille N. 2016. How network analysis of oyster movements can improve surveillance and control programs of infectious diseases? Front Vet Sci 3.

71. Nørgaard LS, Phillips BL, Hall MD. 2019. Infection in patchy populations: Contrasting pathogen invasion success and dispersal at varying times since host colonization. Evolution Letters 3:555–566.

72. Chikhi L, Sousa VC, Luisi P, Goossens B, Beaumont MA. 2010. The Confounding Effects of Population Structure, Genetic Diversity and the Sampling Scheme on the Detection and Quantification of Population Size Changes. Genetics 186:983–995.

73. Sul JH, Martin LS, Eskin E. 2018. Population structure in genetic studies: Confounding factors and mixed models. PLOS Genetics 14:e1007309.

74. Liu P, Song Y, Colijn C, MacPherson A. 2022. The impact of sampling bias on viral phylogeographic reconstruction. PLOS Glob Public Health 2:e0000577.

75. McGeoch DJ, Dolan A, Ralph AC. 2000. Toward a Comprehensive Phylogeny for Mammalian and A*via*n Herpesviruses. Journal of Virology 74:10401–10406.

76. McGeoch DJ, Cook S, Dolan A, Jamieson FE, Telford EAR. 1995. Molecular Phylogeny and Evolutionary Timescale for the Family of Mammalian Herpesviruses. Journal of Molecular Biology 247:443–458.

77. Donohoe O, Zhang H, Delrez N, Gao Y, Suárez NM, Davison AJ, Vanderplasschen A. 2021. Genomes of Anguillid Herpesvirus 1 Strains Reveal Evolutionary Disparities and Low Genetic Diversity in the Genus Cyprinivirus. Microorganisms 9:998.

78. Peck KM, Lauring AS. 2018. Complexities of Viral Mutation Rates. Journal of Virology 92:10.1128/jvi.01031-17.

79. Helmer L, Farrell P, Hendy I, Harding S, Robertson M, Preston J. 2019. Active management is required to turn the tide for depleted Ostrea edulis stocks from the effects of overfishing, disease and invasive species. PeerJ 7:e6431.

80. Feng C, Wang J, Liston A, Kang M. 2023. Recombination Variation Shapes Phylogeny and Introgression in Wild Diploid Strawberries. Molecular Biology and Evolution 40:msad049.

81. Tian J, Sun J, Li D, Wang N, Wang L, Zhang C, Meng X, Ji X, Suchard MA, Zhang X, Lai A, Su S, Veit M. 2022. Emerging viruses: Cross-species transmission of coronaviruses, filoviruses, henipaviruses, and rotaviruses from bats. Cell Reports 39:110969.

82. Long JS, Mistry B, Haslam SM, Barclay WS. 2019. Host and viral determinants of influenza A virus species specificity. 2. Nat Rev Microbiol 17:67–81.

83. Shao W, Li X, Goraya MU, Wang S, Chen J-L. 2017. Evolution of Influenza A Virus by Mutation and Re-Assortment. Int J Mol Sci 18:1650.

84. Kim H, Webster RG, Webby RJ. 2018. Influenza Virus: Dealing with a Drifting and Shifting Pathogen. Viral Immunology 31:174–183.

85. Both GW, Sleigh MJ, Cox NJ, Kendal AP. 1983. Antigenic drift in influenza virus H3 hemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. J Virol 48:52–60.

86. Webster RG, Laver WG, Air GM, Schild GC. 1982. Molecular mechanisms of variation in influenza viruses. Nature 296:115–121.

87. Krummenacher C, Carfí A, Eisenberg RJ, Cohen GH. 2013. Entry of Herpesviruses into Cells: The Enigma Variations, p. 178–195. In Pöhlmann, S, Simmons, G (eds.), Viral Entry into Host Cells. Springer New York, New York, NY.

88. McGeoch DJ, Rixon FJ, Davison AJ. 2006. Topics in herpesvirus genomics and evolution. 1. Virus Research 117:90–104.

89. Kaján GL, Doszpoly A, Tarján ZL, Vidovszky MZ, Papp T. 2020. Virus–Host Coevolution with a Focus on Animal and Human DNA Viruses. J Mol Evol 88:41–56.

90. Waltzek T, Kelley G, Alfaro M, Kurobe T, Davison A, Hedrick R. 2009. Phylogenetic relationships in the family Alloherpesviridae. Dis Aquat Org 84:179–194.

91. O'Connor CM, Sen GC. 2021. Innate Immune Responses to Herpesvirus Infection. 8. Cells 10:2122.

92. Danic-Tchaleu G, Heurtebise S, Morga B, Lapègue S. 2011. Complete mitochondrial DNA sequence of the European flat oyster Ostrea edulis confirms Ostreidae classification. BMC Res Notes 4:400.

93. López Sanmartín M, Power DM, de la Herrán R, Navas JI, Batista FM. 2016. Experimental infection of European flat oyster Ostrea edulis with ostreid herpesvirus 1 microvar (OsHV-1µvar): Mortality, viral load and detection of viral transcripts by in situ hybridization. Virus Research 217:55–62.

94. Arzul I, Langlade A, Chollet B, Robert M, Ferrand S, Omnes E, Lerond S, Couraleau Y, Joly J-P, François C, Garcia C. 2011. Can the protozoan parasite Bonamia ostreae infect larvae of flat oysters Ostrea edulis? Veterinary Parasitology 179:69–76.

95. Kuny CV, Szpara ML. 2022. Alphaherpesvirus Genomics: Past, Present and Future. Current Issues in Molecular Biology 41–80.

96. Schikorski D, Renault T, Saulnier D, Faury N, Moreau P, Pépin J-F. 2011. Experimental infection of Pacific oyster Crassostrea gigas spat by ostreid herpesvirus 1: demonstration of oyster spat susceptibility. 1. Vet Res 42:27.

97. Howard D W. 2004. Histological techniques for marine bivalve mollusks and crustaceans. 5. NOAA Technical Memorandum NOS NCCOS.

98. Pepin JF, Riou A, Renault T. 2008. Rapid and sensitive detection of ostreid herpesvirus 1 in oyster samples by real-time PCR. 2. Journal of Virological Methods 149:269–276.

99. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. 17. Bioinformatics 34:i884–i890.

100. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. 4. Nat Methods 9:357–359.

101. Peñaloza C, Gutierrez AP, Eöry L, Wang S, Guo X, Archibald AL, Bean TP, Houston RD. 2021. A chromosome-level genome assembly for the Pacific oyster Crassostrea gigas. GigaScience 10:giab020.

102. Boutet I, Alves Monteiro HJ, Baudry L, Takeuchi T, Bonnivard E, Billoud B, Farhat S, Gonzales-Araya R, Salaun B, Andersen AC, Toullec J-Y, Lallier FH, Flot J-F, Guiglielmoni N, Guo X, Li C, Allam B, Pales-Espinosa E, Hemmer-Hansen J, Moreau P, Marbouty M, Koszul R, Tanguy A. Chromosomal assembly of the flat oyster (Ostrea edulis L.) genome as a new genetic resource for aquaculture. Evolutionary Applications n/a.

103. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. 16. Bioinformatics 25:2078–2079.

104. Li H. 2013. Seqtk: a fast and lightweight tool for processing FASTA or FASTQ sequences.

105. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. 5. Journal of Computational Biology 19:455–477.

106. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. 3. Journal of Molecular Biology 215:403–410.

107. Morse AM, Calabro KR, Fear JM, Bloom DC, McIntyre LM. 2017. Reliable Detection of Herpes Simplex Virus Sequence Variation by High-Throughput Resequencing. Viruses 9:226.

108. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. 14. Nucleic Acids Res 30:3059–3066.

109. Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution 38:3022–3027.

110. Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.

111. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119.

112. González-Tortuero E, Krishnamurthi R, Allison HE, Goodhead IB, James CE. 2021. Comparative analysis of gene prediction tools for viral genome annotation. preprint. Bioinformatics.

113. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. bioRxiv https://doi.org/10.1101/2021.06.03.446934.

114. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. 7. Nat Biotechnol 40:1023–1025.

115. Hallgren J, Tsirigos KD, Pedersen MD, Armenteros JJA, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv https://doi.org/10.1101/2022.04.08.487609.

116. Cameron ES, Schmidt PJ, Tremblay BJ-M, Emelko MB, Müller KM. 2021. Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. 1. Sci Rep 11:22302.

117. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv:12073907 [q-bio].

118. Tan A, Abecasis GR, Kang HM. 2015. Unified representation of genetic variants. 13. Bioinformatics 31:2202–2204.

119. R Development Core Team. 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

120. Csardi G, Nepusz T. 2006. The igraph software package for complex network research. InterJournal Complex Systems:1695.

121. Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer New York, New York, NY. https://link.springer.com/10.1007/978-0-387-98141-3. Retrieved 10 May 2023.

122. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. Genome Res 14:1394–1403.

123. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100.

124. Goel M, Sun H, Jiao W-B, Schneeberger K. 2019. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. Genome Biology 20:277.

125. Goel M, Schneeberger K. 2022. plotsr: visualizing structural similarities and rearrangements between multiple genomes. Bioinformatics 38:2922–2926.

126. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Research 43:e15.

127. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and Evolution 34:3299–3302.

128. Winter DJ. 2012. MMOD : an R library for the calculation of population differentiation statistics. Mol Ecol Resour 12:1158–1160.

129. Alachiotis N, Pavlidis P. 2018. RAiSD detects positive selection based on multiple signatures of a selective sweep and SNP vectors. 1. Commun Biol 1:1–11.

130. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol 2:vew007.

131. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Systematic Biology 67:901–904.

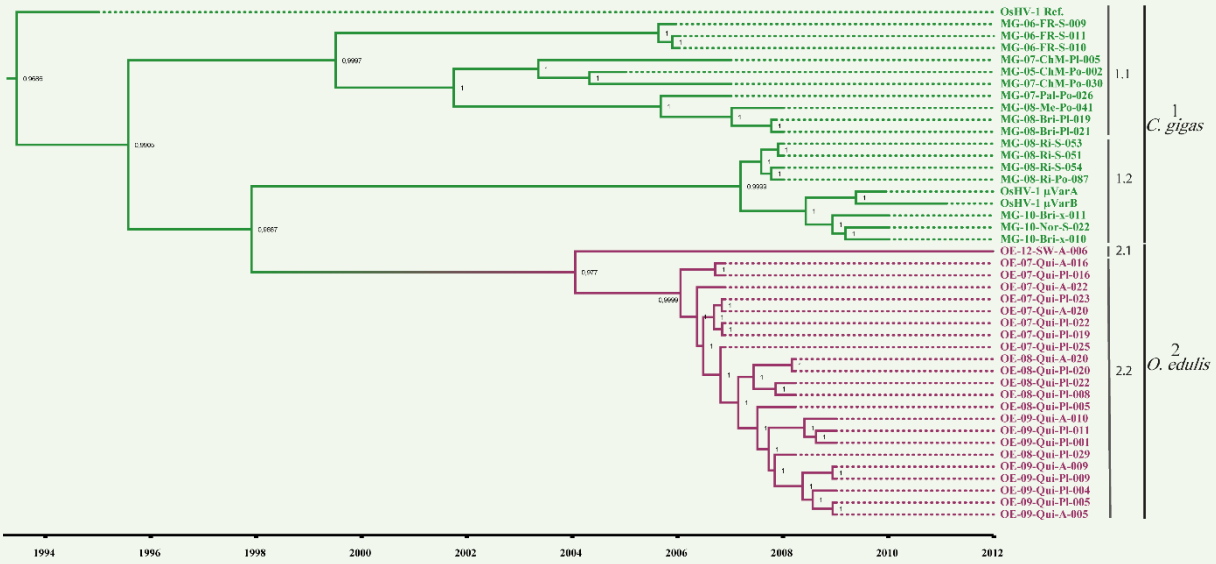
Messages clés :

- La **diversité génétique** du virus OsHV-1 est **influencée par l'espèce hôte**. Des variations structurales, telles que de grandes insertions et délétions, ont été identifiées dans le génome d'OsHV-1, indiquant la prévalence des **réarrangements génomiques comme moteur de son évolution**.

- Les lignées d'OsHV-1 infectant *M. gigas* et *O. edulis*, ont montré une **différenciation génétique significative**. La proximité génétique des lignées au sein de la même espèce hôte suggère une forte association entre OsHV-1 et son hôte, soutenant l'idée d'une **adaptation spécifique à l'hôte et de la coévolution**.

- L'analyse des polymorphismes au sein des génotypes d'OsHV-1 a révélé que **les génotypes collectés chez *M. gigas* présentaient un niveau plus élevé de diversité génétique par rapport à ceux d'*O. edulis*.** Cette différence pourrait être attribuée aux tailles de population plus grandes et aux transferts fréquents d'huîtres entre les bassins, facilitant la transmission virale et l'évolution.

- Bien que les événements de CST soient fréquents chez certains virus, les événements de **CST du virus OsHV-1 entre les espèces d'huîtres sont relativement rares** en raison de barrières écologiques et physiologiques. Il est probable que le génotype ancestral d'OsHV-1 ait été initialement originaire de *M. gigas* mais se soit différencié en génotypes adaptés à différentes espèces hôtes.



Chapitre 4 – Caractérisation spatio-temporelle de la diversité du virus OsHV-1 chez *M. gigas*

4.1 Résumé de l'article 3

Le virus OsHV-1 est le premier *herpesvirus* isolé chez les invertébrés et il est responsable de mortalités chez l'huître creuse *Magallana gigas* au cours des dernières décennies, causant d'importants dommages économiques. Le virus infecte principalement les jeunes huîtres, mais il peut également infecter d'autres espèces de bivalves. Des variants d'OsHV-1, tels que l'OsHV-1 µVar, ont été associés à des mortalités massives chez les huîtres dans différentes régions du monde.

Actuellement, très peu d'études portant sur la diversité du virus OsHV-1 sont disponibles. Les études précédentes sur l'OsHV-1 se sont principalement concentrées sur des régions génomiques spécifiques, et peu d'informations sont disponibles sur la diversité génomique spatio-temporelle du virus. Ainsi, l'étude propose d'utiliser une approche de séquençage complet du génome à partir d'échantillons historiques pour évaluer l'évolution de l'OsHV-1 à l'échelle épidémiologique et caractériser la diversité génomique entre différentes espèces hôtes et régions géographiques. L'objectif de l'étude était *(i)* de mieux comprendre l'évolution du virus OsHV-1 au cours du temps et son potentiel d'adaptation en fonction des espèces hôtes, *(ii)* de déterminer si les populations virales chez des huîtres du Pacifique infectées provenant de régions différentes du monde, présentent des structures génomiques ou des polymorphismes similaires.

Dans cette étude, des haplotypes locaux sur les régions codantes (CDS) et des consensus globaux ont pu être reconstruits à partir d'un alignement des séquences sur le génome de référence d'OsHV-1, qui ont ensuite servi à réaliser une analyse phylogénétique. En parallèle, une analyse d'appel de variant a pu être effectué, permettant ainsi d'étudier la diversité virale intra-individuelle de chaque échantillon.

Cette étude a montré que l'OsHV-1 présente une diversité génétique importante, avec des haplotypes viraux multiples au sein d'un même isolat viral. Certaines régions du génome viral semblent héberger plus d'haplotypes que d'autres, ce qui indique la présence de régions conservées et de régions hypervariables. De larges délétions ont été détectées, certaines correspondent aux délétions détectées dans le génome d'OsHV-1 µVar, tandis que d'autres n'ont encore jamais été rapportées dans la littérature. L'analyse phylogénétique montre que la diversité génétique actuelle des isolats d'OsHV-1 a émergé au cours des trois dernières décennies. Les isolats de virus provenant d'hôtes spécifiques et de différentes régions géographiques se sont regroupés de manière distincte dans l'arbre phylogénétique, suggérant une évolution spécifique liée à la géographie et aux hôtes.

Cet article s'intitule « Genomic diversity of the Ostreid Herpesvirus Type 1 across time and location and among host species » (Morga-Jacquot *et al.*, 2021) a été placé en annexe 1 de ce manuscrit.

4.2 Résumé de l'article 4

En 2019, seuls six génomes complets du virus étaient disponibles dans les bases de données publiques. Cette carence est très probablement liée au fait que le virus est difficilement purifiable et les séquençages, encore chers, résultaient alors en une majorité de séquences de l'hôte. Ce dernier point rendait difficile l'analyse bio-informatique nécessaire à l'assemblage des génomes viraux. De plus, dans le contexte précédemment décrit, il était nécessaire de caractériser les processus évolutifs et épidémiologiques qui agissent sur la diversité du virus OsHV-1. Un premier article de Morga et collaborateurs (article 3), présenté dans ce même chapitre, est la première étude portant sur la diversité du virus OsHV-1 qui s'intéresse au génome complet du virus OsHV-1. Cependant, plusieurs limitations à cette analyse apparaissaient comme importante à résoudre pour les futures études de diversité. Cette étude a utilisé des haplotypes viraux reconstruits par alignement de séquences courtes sur le génome de référence d'OsHV-1 plutôt qu'une approche d'assemblage *de novo*. Ce type de reconstruction introduit des biais significatifs dans l'analyse de la diversité, en particulier dans l'évaluation des variations structurales de type insertion, qui sont souvent ignorées. Cela affecte également l'étape d'appel de variants, pouvant potentiellement conduire à l'identification de SNP qui pourraient ne pas être réellement présents dans la population. Dans ce contexte, il était important d'analyser la diversité du virus OsHV-1 après avoir reconstruit *de novo* des génomes viraux afin de limiter les biais potentiels.

Pour mieux comprendre la diversité génétique d'OsHV-1 et ses liens avec sa distribution géographique, cette étude a utilisé une approche de séquençage complet du génome sur des échantillons provenant de différentes régions d'élevage d'huîtres en France (la Rade de Brest, le bassin de Marennes-Oléron et l'étang de Thau) prélevés lors d'épisodes de mortalités d'huîtres creuses *M. gigas*. Au total, 21 nouveaux génomes complets d'OsHV-1 µVar ont pu être assemblés à partir de ces échantillons et du pipeline développé dans le chapitre 2. Les analyses d'appel de variant ont permis de détecter 400 SNP dans les librairies dont 51 positions variables étaient partagées par l'ensemble des échantillons, et les librairies construites à partir des échantillons du bassin de Marennes-Oléron contiennent une diversité plus importante que les autres. Les analyses phylogénétiques et génomiques ont montré une grande diversité génétique du virus non seulement entre les différentes régions d'élevage d'huîtres, mais aussi entre les individus au sein des mêmes régions. Ces résultats suggèrent que la diversité génétique d'OsHV-1 est influencée par des facteurs épidémiologiques et peut jouer un rôle dans la propagation du virus entre les régions productrices d'huîtres. En utilisant des modèles phylodynamiques et phylogéographiques, nous avons pu proposer un scénario expliquant les relations phylogénétiques entre les populations virales rencontrées dans les différentes régions d'élevage d'huîtres.

Article 4

Genetic diversity and connectivity of the *Ostreid herpesvirus* 1 populations in France: a first attempt to phylogeographic inference for marine mollusc diseases

Jean Delmotte, Camille Pelletier, Benjamin Morga, Richard Galinier, Bruno Petton, Jean-Baptiste Lamy, Oliver Kaltz, Jean-Christophe Avarre, Maude Jacquot, Caroline Montagnani, Jean-Michel Escoubas

VIRUS EVOLUTION

Virus Evolution, 2022, **8(1)**, 1–14
DOI: https://doi.org/10.1093/ve/veac039
Advance access publication date: 23 April 2022
Research Article

Genetic diversity and connectivity of the Ostreid herpesvirus 1 populations in France: A first attempt to phylogeographic inference for a marine mollusc disease

Jean Delmotte,[1,†] Camille Pelletier,[2,†] Benjamin Morga,[2] Richard Galinier,[3] Bruno Petton,[4] Jean-Baptiste Lamy,[2] Oliver Kaltz,[5] Jean-Christophe Avarre,[5,‡] Maude Jacquot,[2,*,§] Caroline Montagnani,[1] and Jean-Michel Escoubas[1,*,**]

[1]IHPE, CNRS, Ifremer, UPVD, University of Montpellier, Montpellier F-34095, France, [2]Ifremer, RBE-ASIM, Station La Tremblade, La Tremblade F-17390, France, [3]IHPE, CNRS, Ifremer, UPVD, University of Montpellier, Perpignan F-66000, France, [4]Ifremer, CNRS, IRD, Ifremer, LEMAR UMR 6539 Université de Bretagne Occidentale, Argenton-en-Landunvez F-29840, France and [5]ISEM, IRD, CNRS, University of Montpellier, Montpellier F-34095, France
[†]Authors contributed equally to this work.
[‡]https://orcid.org/0000-0001-6899-7052
[§]https://orcid.org/0000-0001-5945-4689
[**]https://orcid.org/0000-0002-6138-1148
*Corresponding authors: E-mail: Maude.Jacquot@ifremer.fr; Jean.Michel.Escoubas@ifremer.fr

Abstract

The genetic diversity of viral populations is a key driver of the spatial and temporal diffusion of viruses; yet, studying the diversity of whole genomes from natural populations still remains a challenge. Phylodynamic approaches are commonly used for RNA viruses harboring small genomes but have only rarely been applied to DNA viruses with larger genomes. Here, we used the Pacific oyster mortality syndrome (a disease that affects oyster farms around the world) as a model to study the genetic diversity of its causative agent, the Ostreid herpesvirus 1 (OsHV-1) in the three main French oyster-farming areas. Using ultra-deep sequencing on individual moribund oysters and an innovative combination of bioinformatics tools, we *de novo* assembled twenty-one OsHV-1 new genomes. Combining quantification of major and minor genetic variations, phylogenetic analysis, and ancestral state reconstruction of discrete traits approaches, we assessed the connectivity of OsHV-1 viral populations between the three oyster-farming areas. Our results suggest that the Marennes-Oléron Bay represents the main source of OsHV-1 diversity, from where the virus has dispersed to other farming areas, a scenario consistent with current practices of oyster transfers in France. We demonstrate that phylodynamic approaches can be applied to aquatic DNA viruses to determine how epidemiological, immunological, and evolutionary processes act and potentially interact to shape their diversity patterns.

Key words: herpesvirus; OsHV-1; Pacific oyster; phylogeography; non-redundant genomes; genetic diversity; viral spread; minor variations.

1. Introduction

Viruses are disease agents that have high levels of genetic diversity. This high diversity often means that they lack shared genetic markers, such as ribosomal DNA sequences that are common to all prokaryotes and eukaryotes (Nkili-Meyong et al. 2016), making it difficult to characterize viruses genetically. Furthermore, many viruses—especially those with RNA genomes—are not stable genetic entities but exist as clouds of many phylogenetically related genetic variations, known as viral quasispecies. This genetic organization currently encumbers our understanding of viral diseases and their evolution, and it impedes the straightforward characterization of virus population structure (Vignuzzi et al. 2006; Lauring and Andino 2010). Studies have shown that the level of genetic diversity within these viral populations likely influences viral pathogenicity, dissemination, and host immune evasion. Full-length genome analyses are required to identify intra-host viral population structure, reveal molecular traits with epidemiological significance, or detect low-frequency, but nonetheless relevant viral variants. These difficulties apply to RNA viruses, but also to certain large DNA viruses, such as herpesviruses, whose genomic variability rivals that of many RNA viruses (Renzette et al. 2013; Hammoumi et al. 2016; Renner and Szpara 2018). For example, human cytomegalovirus shows considerable inter-host and intra-host genetic divergence across tissue compartments and times of infection (Renzette et al. 2014, 2015). In addition, the evolution of a disease is sometimes not fully explained by intrinsic host factors, but by the genotypic diversity of herpesviruses (Akhtar et al. 2019).

Aquaculture is one of the fastest-growing food-producing sectors, representing an important animal protein supply for human

© The Author(s) 2022. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

consumption, with an expanding role in global food security. Today, the biggest threat arising due to the intensification and globalization of aquaculture is infectious diseases. The management and mitigation of the emergence and spread of these infectious diseases are key issues to address to ensure the sustainability of this industry (Stentiford et al. 2012; Pernet et al. 2016; Burge, Shore-Maggio, and Rivlin 2017). One illustration is the Pacific oyster mortality syndrome (POMS), which threatens global *Crassostrea gigas* oyster production, a main sector in aquaculture worldwide (reviewed in Petton et al. 2021). Since 2008, this syndrome has caused mass mortality in cultivated oysters around the world, from Europe to America and Asia (Friedman et al. 2005; Moss et al. 2007; Vasquez-Yeomans, Garcia-Ortega, and Caceres-Martinez 2010; Peeler et al. 2012; Jenkins et al. 2013; Barbosa Solomieu, Renault, and Travers 2015; Gittenberger, Voorbergen-Laarman, and Engelsma 2016; Abbadi et al. 2018; Barbieri et al. 2019). In 2010, the causative agent of these massive mortalities was identified: it is an emerging genotype of a herpes-like virus named Ostreid herpesvirus 1 (OsHV-1) (Segarra et al. 2010; Pernet et al. 2012; Renault et al. 2012; Paul-Pont, Dhand, and Whittington 2013; Petton et al. 2013). Two major genetic factors seem to affect the severity of POMS: the genetic background of the oysters (Dégremont, Nourry, and Maurouard 2015; de Lorgeril et al. 2018, 2020) and OsHV-1 genetic diversity (Martenot et al. 2011; Delmotte et al. 2020; Friedman et al. 2020). Since the characterization of the emergent genotype OsHV-1 μVar, associated with the 2008 high-mortality events (Segarra et al. 2010), several variants of the μVar genotype have been described (reviewed in Barbosa Solomieu, Renault, and Travers 2015). Interestingly, in 2016, a survey of OsHV-1 genetic diversity carried out on wild *C. gigas* populations along the coasts of Italy demonstrated the high diversity of this virus in natural oyster populations (Burioli et al. 2016). However, most studies on OsHV-1 genetic diversity have been based on polymerase chaine reaction (PCR) molecular markers focusing on a few variable regions of the viral genome, which may not reflect the whole genomic diversity. For instance, in 2017, sequencing of the whole genome of OsHV-1 μVar and its comparison with the reference genome published in 2005 (Davison et al. 2005) showed that the two genomes also differed by the loss or addition of several open reading frames (ORFs), indicating that whole-genome sequencing is necessary to fully reveal viral diversity and to better understand the virus' origin and evolution (Burioli, Prearo, and Houssin 2017). To date, no study has specifically looked at the links between the geographic distribution of OsHV-1 and its genetic diversity at the whole-genome level. As far as we know, the entire assembled genomes of only four OsHV-1 infecting *C. gigas* oysters are available, namely the reference genome (AY509253), two OsHV-1 µVar (μVar A KY242785.1 and μVar B KY271630.1), and more recently OsHV-1-PT (MG561751.2) (Davison et al. 2005; Burioli, Prearo, and Houssin 2017; Abbadi et al. 2018). These four assemblages show significant genomic diversity; however, they tell us very little about the geographic distribution and dissemination of the virus.

As shown, for instance, with the H1N1 influenza virus (Rambaut et al. 2008), the dengue virus (Allicock et al. 2012), the Zika virus (Theze et al. 2018), the Acquired Immune Deficiency Syndrome virus (Faria et al. 2014), and more recently Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Cov-2) (Martin, VanInsberghe, and Koelle 2021), molecular epidemiology based on full genomes can be the key to understanding the epidemics of emerging viruses. Therefore, the purpose of the present study was to characterize the genomic diversity in viral populations of OsHV-1 μVar encountered in oyster-farming areas both within and between host individuals, to better understand the dissemination of OsHV-1 μVar populations, and to assess the feasibility of applying phylodynamic and phylogeographic models to OsHV-1 whole-genome sequences. Using a deep-sequencing approach conducted at the individual level and novel bioinformatics analyses, we identified twenty-one new complete genome sequences of OsHV-1 μVar from the three most important oyster-farming areas in France. Phylogenetic analyses combined with comparative genomics and inspection of variant frequencies showed that viral genetic diversity differs greatly not only between areas but also between individuals within areas. An ancestral state reconstruction of discrete trait approach allows us to propose a scenario explaining the phylogenetic relationships between viral populations encountered in the three farming areas. This novel combination of approaches, together with epidemiological information, holds promise for understanding OsHV-1 evolution and epidemic dynamics and for identifying transmission patterns between oyster-producing regions. These data can be highly useful for developing novel disease management strategies.

2. Results

2.1 Variability in sequencing depth does not impair the analysis of OsHV-1 diversity

We used specific-pathogen-free (SPF) *C. gigas* juveniles to sample OsHV-1 diversity during natural infections in the three farming areas (Fig. 1A). The cumulative mortalities for the three batches of oysters were 85, 81, and 70 per cent for Thau (Th), Marennes-Oléron (MO), and Brest (Br), respectively. No mortality was observed in the control group.

For each individual oyster, deep sequencing produced from 86.2 to 145.9 M reads per sample (average 109.7 M reads ±16.7 SD; Fig. 1B; Supplementary Table S1). To evaluate OsHV-1 sequencing depth in each sample, sequencing reads were mapped on the OsHV-1 μVar A genome (KY242785.1) (Burioli, Prearo, and Houssin 2017). The number of reads mapping to the OsHV-1 μVar A genome greatly varied from one individual to the other, independently of the number of reads obtained for each sample: from 0.65 M reads (0.66 per cent) for MO ind3 to 11.94 M reads (13.21 per cent) for Th ind10 (Fig. 1B; Supplementary Table S1). Therefore, OsHV-1 genome coverage varied from 480× to 8,745× for MO ind3 and Th ind10, respectively.

A rarefaction analysis was run using OsHV-1 μVar A as a reference genome to verify that OsHV-1 coverage was sufficient to accurately quantify viral genomic diversity in each sample. For single-nucleotide polymorphisms (SNPs) and for insertions-deletions (InDels), rarefaction curves reached a plateau around 300,000 reads, corresponding to an OsHV-1 coverage of 220× (Fig. 1C). Since the number of mapped reads of each sample was clearly above this threshold, sequencing depth was considered sufficient to capture OsHV-1 genomic diversity. Rarefaction curves indicate that viral diversity was higher in infected MO oysters than in the other two areas. Comparison of the average number of individual polymorphisms deduced from rarefaction analysis between the three environments showed that (1) the number of SNPs in viruses from MO (170 ± 27.7) was significantly higher than that of viruses from Br (136.4 ± 5.9; $P \leq 0.001$) and Th (143.8 ± 5.8; $P \leq 0.05$) and (2) the number of InDels in viruses from MO (35.9 ± 3.1 SD) was also significantly higher than that of viruses from Th (25.7 ± 0.9; $P \leq 0.0001$) (Fig. 1D; Supplementary Table S2). In addition, in MO, viral diversity greatly varied from one individual to the other; for example, MO ind3 contained 222 SNPs and 42 InDels, whereas MO ind4 contained only 144 SNPs and 33 InDels.

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

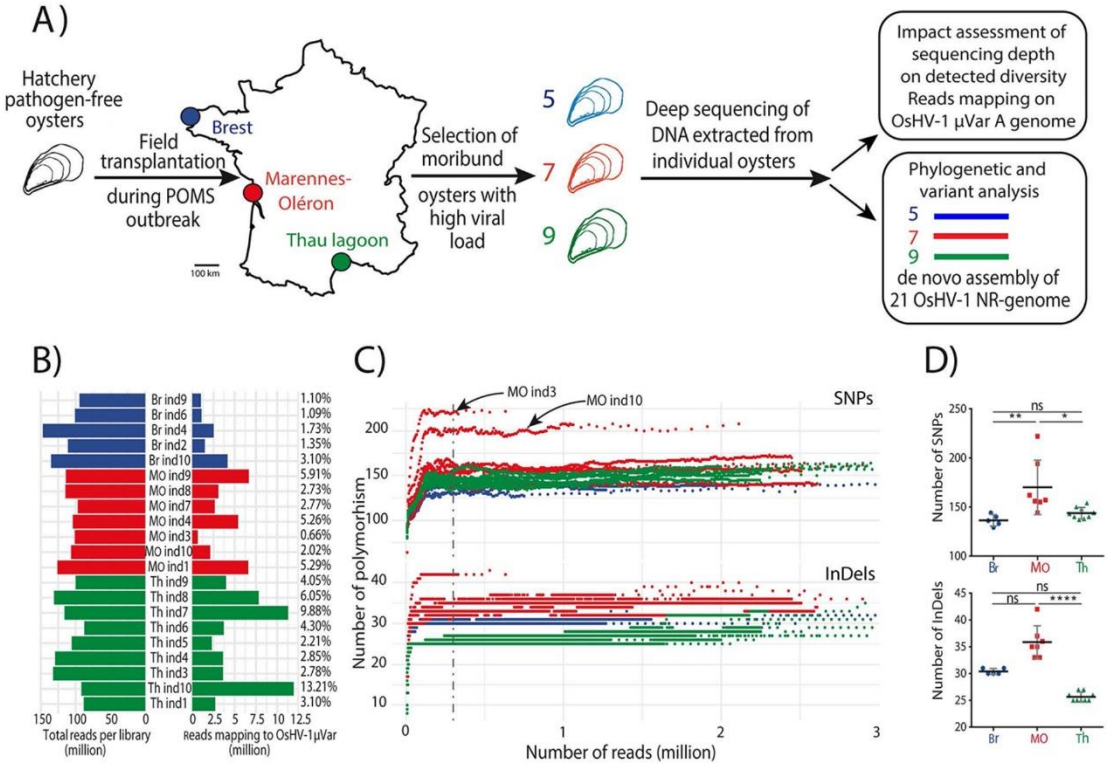


Figure 1. Variability in sequencing depth does not impair the analysis of OsHV-1 diversity. (A) Schematic representation of the sampling of OsHV-1 genetic diversity in the three main French oyster-farming areas during the POMS event in 2018 (see Section 4 for more details). Each color represents a farming area, with Br in blue (five individual samples), MO in red (seven individual samples), and Th Lagoon (Th) in green (nine individual samples). (B) Quantification of the number of reads in the library and the number of reads mapping to the OsHV-1 μVar A genome. Numbers on the right side of the figure indicate the percentages of reads mapping to the OsHV-1 μVar A genome. (C) Rarefaction curves showing the accumulation of polymorphisms (SNPs and InDels top and bottom panels, respectively) during iterative variant calling analysis of the twenty-one libraries on the OsHV-1 μVar A genome. The dashed vertical line indicates 0.3 M reads, corresponding to the threshold at which the rarefaction curves reach a plateau. (D) Comparison of the number of SNPs (top) and InDels (bottom) between the three farming areas. Asterisks indicate the *P*-value according to Kruskal–Wallis and Dunn's multiple comparison tests (ns: non-significant, $*P \leq 0.05$, $**P \leq 0.01$ $****P \leq 0.0001$).

2.2 Phylogenetic analysis of non-redundant OsHV-1 genomes reveals that all three OsHV-1 populations belong to the µVar genotype

We assembled non-redundant (NR) genomes so as to reduce the complexity of the OsHV-1 genome and keep only one copy of each repeated region in the virus sequence. The OsHV-1 μVar A genome contains two unique regions and three repeats: the unique long (UL) region (164,268 bp), flanked by two inverted repeats (repeat long (RL), 7,338 bp), and the unique short (US) region (3,370 bp), flanked by two inverted repeats (repeat short (RS), 9,777 bp), and two X regions (1,510 bp) (Fig. 2A, upper panel). Considering the high sequence identity within each set of repeats (between 99.9 and 100 per cent), removing one of the repeats makes it possible to reliably map the reads from these repeated regions without affecting the overall genome composition. Removing a copy of each repeated region led to an NR OsHV-1 μVar A genome (NR-genome) of 186,262 bp, which represents 91 per cent of the full-length genome and contains all the coding sequences (Fig. 2A, lower panel).

Using this approach, twenty-one NR-genomes were assembled *de novo* from the sequencing data generated during this study, as well as six other OsHV-1 (belonging to the *Malacoherpesviridae* family and *Ostreavirus* genus) published genomes. A phylogenetic analysis of these twenty-seven NR-genomes revealed that the twenty-one new OsHV-1 isolates were phylogenetically closer to OsHV-1 μVar A and B isolates (Supplementary Fig. S1) than to the OsHV-1 reference genome (AY509253.2), suggesting that they belong to the μVar genotype.

Among newly sequenced isolates, those originating from Th and Br were grouped into two distinct clusters reflecting their geographic origin (Fig. 2B, Clusters 1 and 2). Conversely, MO OsHV-1 isolates were distributed over these two clusters: five in Cluster 1 and two in Cluster 2. The phylogenetic distance among Th or Br isolates (average $2.47 \times 10^{-5} \pm 4.68 \times 10^{-5}$ SD and $1.43 \times 10^{-4} \pm 7.91 \times 10^{-5}$ SD, respectively) was lower than that observed among the MO isolates (average $5.00 \times 10^{-3} \pm 4.21 \times 10^{-3}$), indicating greater viral diversity in MO than in the other two farming areas.

2.3 Comparative genomics of the twenty-one NR genomes confirms that OsHV-1 diversity is higher in MO than in the other two farming areas

To compare the genetic diversity at the whole-genome scale, we used the twenty-one NR-genomes to construct an average NR genome (Avg-genome) (see Section 4.5 for more details).

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

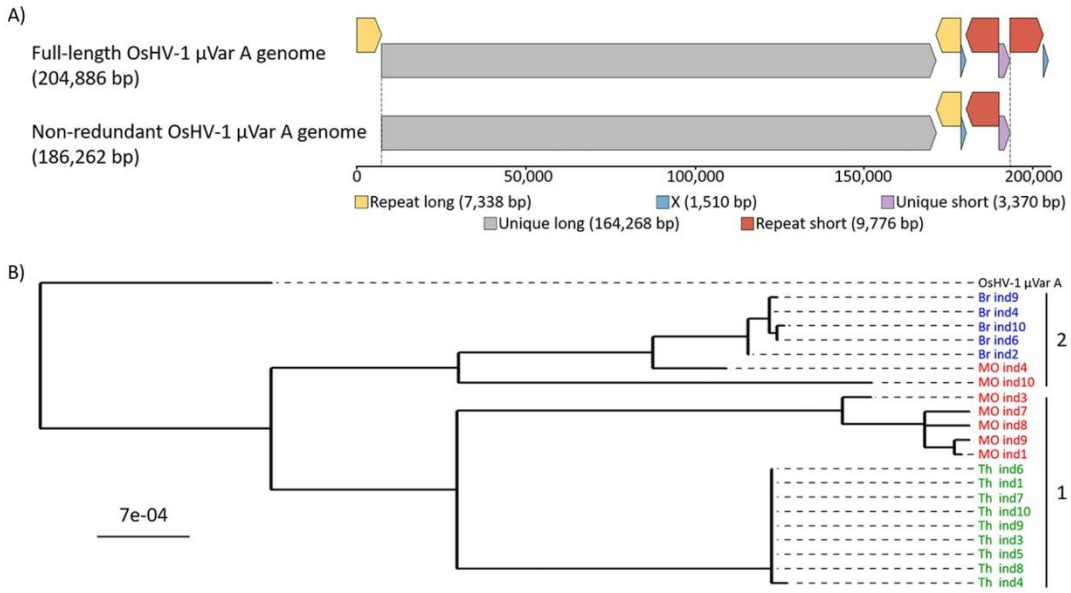


Figure 2. Phylogenetic relationship between the twenty-one *de novo* assembled and other OsHV-1 NR genomes. (A) Full-length (upper panel) and NR (lower panel) genomes of OsHV-1 µVar A (KY242785). The OsHV-1 µVar A genome consists of two unique regions, referred to as UL (in grey) and US (in purple), and three repeat regions, referred to as RL (in yellow), RS (in orange), and X region (X, in blue). The first two repeat regions each have copies at a terminal locus (TRL and TRS) and an internal locus (IRL and IRS) within the genome. The NR-genome consists of the linear fragment starting at the beginning of the UL and ending at the end of the US region. This strategy includes all the genetic information without duplication. All the regions within the genome and the NR genome are to scale. (B) Shown is a maximum-likelihood phylogenetic tree of the twenty-two OsHV-1 NR-genomes rooted in OsHV-1 µVar A. The twenty-one isolates from this study are colored according to their geographic origin, Br in blue, MO in red, and Th Lagoon in green. Within those, two clusters were defined based on genetic distances (numbered on the right).

The comparison of the twenty-one NR-genomes to the Avg-genome identified a total of 399 genetic variations (SNPs and InDels, hereafter called variations), and Spearman's correlation coefficient (–0.382; $P = 0.087$) indicates that there is no relationship between the viral load and the number of variations per sample (Supplementary Fig. S2).

The 399 variations spread over 310 positions, and twenty-two of them contained from two to seven allelic variations (Supplementary Table S3). Prediction of the possible functional effects of the 399 variations revealed that 216 (54.1 per cent) were distributed over 94 ORFs (summarized in Supplementary Fig. S3 and detailed in Supplementary Table S3). The twenty-one variations predicted to have a high impact (frameshift or start lost) were distributed over eleven ORFs. Unfortunately, these results are not very informative, given that six ORFs were annotated 'unknown' due to the absence of homologies with any known proteins and the five others were annotated according to domain homologies suggesting a protein localization or a putative function.

The average number of variations per genome was significantly lower in samples from Th (63.7 ± 2.7) than from MO (92.1 ± 25.8; $P < 0.05$) and Br (120.8 ± 2.8; $P < 0.001$), whereas the average number of variations in MO was not significantly different from Br (Fig. 3A, Supplementary Table S3). However, viruses from MO showed the highest diversity, with 290 variations accounting for 72.7 per cent of the total diversity. Among the variations, 187 (64.5 per cent) were specific to MO, and 114 were observed in one sample only (singletons) (Fig. 3B, Supplementary Table S3). None of these 187 variations were shared by all the samples from this area. The majority of MO singleton variations were observed in two samples: ind10 (fifty-six) and ind4 (twenty-three) (Fig. 3C). For Th and Br farming areas, 82 and 132 variations were identified, among which 70 and 38 were area-specific, respectively (Fig. 3B, Supplementary Table S3). In contrast within MO, 57.9 per cent (22/38) and 75.5 per cent (53/70) of the variations were common to all the samples of Br and Th, respectively (Fig. 3B, Supplementary Table S3). Among the ninety-two variations shared by Br and MO, eighty-seven (94.6 per cent) were shared by two MO individuals (ind10 and ind4) and all Br individuals (Fig. 3C, purple arrow). These two MO individuals also shared five variations with all Th individuals (Fig. 3C, red arrows).

Interestingly, these major variations were only characterized in a small number of individuals within the samples analyzed; indeed, 129 (32 per cent) of them are found in only one individual, and only 2 (0.5 per cent) are shared by more than 50 per cent of individuals.

2.4 Within-individuals diversity of OsHV-1: the minor variations are more frequent in MO than in the other two farming areas

To fully characterize OsHV-1 genetic diversity, we performed variant calling analyses, using the Avg-genome as the reference, to quantify both major (frequency >50 per cent) and minor (frequency <50 per cent) variations within each of the twenty-one individual samples. Among the 399 variations identified as major variations, 132 had variable frequencies (above or below 50 per cent) across samples; moreover, 208 additional minor variations were identified (Fig. 4A, Supplementary Table S4). These 208 minor variations are spread over 164 positions of the Avg-genome, and 17 of these positions are showing 2–12 allelic variations (Supplementary Table S4). SnpEff predicts that 72 (18.0 per cent) of the 208

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

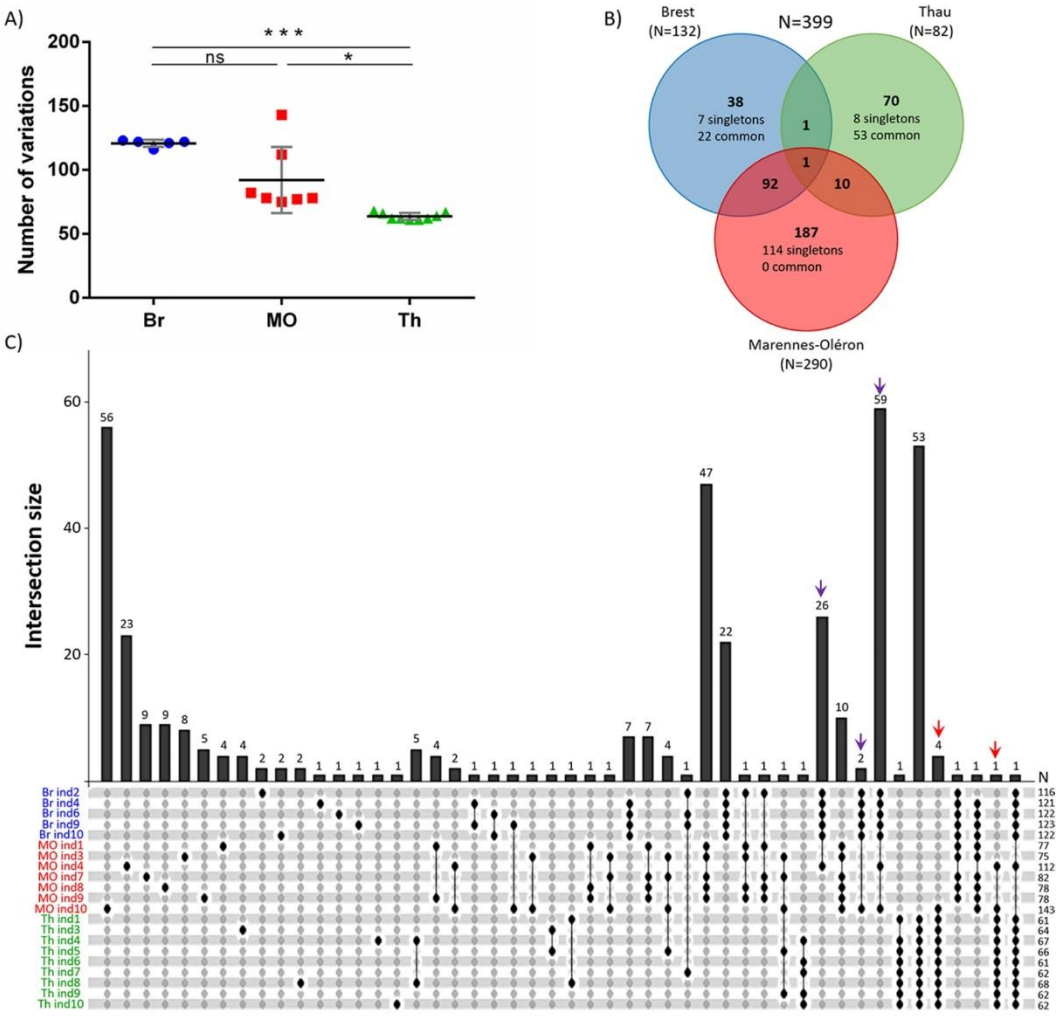


Figure 3. Comparative genomics of the twenty-one NR genomes of OsHV-1 isolated from three oyster-farming areas. (A) Distribution of the 399 variations between the twenty-one individuals and the three farming areas. Significance was calculated according to Kruskal–Wallis and Dunn's multiple comparison tests (ns: non-significant, $*P < 0.05$, $***P < 0.001$). (B) Venn diagram summarizing the distribution of variations between the three farming areas. Color depends on the farming area with Br in blue, MO in red, and Th Lagoon in green. The numbers of variations per farming area are indicated in parentheses. Singletons are variations that occur in only one sample, and the term 'common' indicates variations present in all the sampled individuals at a given location. (C) UpSet plot of variation distribution among the twenty-one NR-genomes. The bottom panel shows the specific combinations (or intersections), and the vertical bars indicate the number of variations within these combinations. Colored arrows on the top indicate variations shared by two individuals from MO (ind4 and ind10) and all individuals from Br (purple arrows) or all individuals from Th (red arrows). Numbers (*N*) on the right side indicate the number of variations per sample.

variations were distributed over 34 predicted ORFs (summarized in Supplementary Fig. S4 and detailed in Supplementary Table S4). The seven variations predicted to have a high impact (frameshift or start gained) were distributed in six ORFs. But, again, these results were not very informative, given that four of these ORFs were annotated 'unknown' and the two others were annotated according to domain homologies suggesting a membrane location of the predicted protein.

Frequencies of the 208 strictly minor variations ranged from 5.0 to 48.7 per cent (mean = 9.57 ± 6.52 per cent SD). The number of minor variations was lower in oysters from Br (88) than in oysters from Th (98) or MO (114) (Fig. 4B, Supplementary Table S4). Fifty-five of these 208 minor variations (24.5 per cent) were found in all three farming areas. Moreover, all site-specific variations were found in only one individual oyster (singletons).

A hierarchical clustering analysis applied to the minor variations showed that individuals did not cluster according to their geographic origin (Fig. 4C). In some cases, we even observed two individuals from different farming areas clustering together (MO ind10 and Br ind2; Fig. 4C), suggesting that some OsHV-1 found in these individuals had a common origin. Figure 4C also revealed that the fifty-one variations shared by the three farming areas (common) were found in a majority of sampled individuals (mean = 13.9 ± 5.7 SD), suggesting that they had not arisen within

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

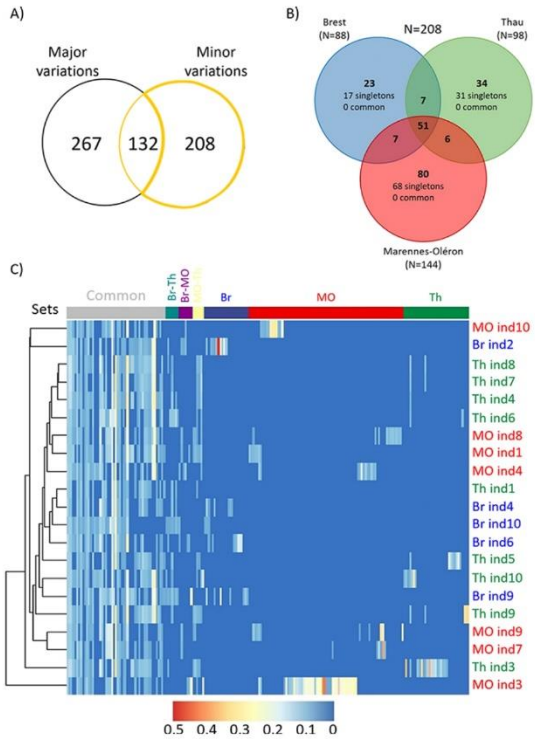


Figure 4. Variant calling analysis and minor variation distribution in the three farming areas. (A) Quantification of both major and minor variations using a variant calling analysis using the Avg-genome as the reference. (B) Venn diagram summarizing the distribution of the 208 strictly minor variations (frequency <50 per cent) between the three farming areas (Br in blue, MO in red, and Th Lagoon in green). The number of minor variations within each farming area is indicated in parentheses. Singletons correspond to variations that occur in only one oyster sample, whereas the term 'common' corresponds to variations present in all the oyster samples within the dataset. (C) Heatmap of the allelic frequency of strictly minor variations across the twenty-one samples with a gradient from 0 (blue) to 0.5 (red). Hierarchical clustering using the Euclidean distance of the allele frequency is displayed on the top. Each line of the heatmap corresponds to a strictly minor variation. Variation sets on the top are derived from the Venn diagram. Samples are indicated on the right and colored according to their geographic origin.

individual oysters, but had been transmitted horizontally to other oysters.

2.5 Distribution of within and between individual genetic variations among samples reveals oyster-farming areas' connectivity

The 132 variations that had variable frequencies across samples (Fig. 4A) were present as minor variations in only four samples: three from MO (ind3, ind4, and ind10) and one from Br (ind2) (Fig. 5A). Among them, samples MO ind3 and MO ind10 contained 125 of these minor variations. Sample MO ind10 contained thirty minor variations (with a frequency ranging from 5 to 7.2 per cent; Fig. 5B; Supplementary Table S5) that were major variations in samples MO ind1/3/7/8/9. For instance, the variation at position 65,003 (T>A) of the Avg-genome had a frequency of 7 per cent in MO ind10, whereas its frequency varied from 76 to 99 per cent in MO ind1/3/7/8/9 (Fig. 5C). Similar fluctuations in variation frequency were observed for the other twenty-nine variations (Supplementary Table S5). These thirty variations highlighted a clear genetic connectivity between the MO ind10 sample and the five MO samples (ind1/3/7/8/9) located in the other phylogenetic clade (Fig. 5D).

Furthermore, MO ind3 contained ninety-five minor variations that corresponded to major variations in MO ind10, forty-six that corresponded to major variations in MO ind4 and all Br samples, and four that corresponded to major variations in all Th samples (Fig. 5A and E). For instance, the variation at position 57,080 (T>C) in the Avg-genome had a frequency of 24 per cent in MO ind3, whereas its frequency was 93 per cent in MO ind10 and over 99 per cent in all Th samples (Fig. 5F; Supplementary Table S5). Another example is the variation at position 65,610 (C>T) in the Avg-genome: its frequency grew from 24 per cent in MO ind3 to 92 per cent in MO ind10, and 99 per cent in MO ind4 and all Br samples (Fig. 5F, Supplementary Table S5). Again, minor variations highlighted genetic connectivity between samples from different farming areas and from phylogenetic Clusters 1 and 2 (Fig. 5G; Supplementary Table S5).

2.6 Ancestral state reconstruction of discrete trait

The analysis of minor variations highlighted links not only between MO samples and samples from the other two farming areas, but also between MO samples, which segregated into two different clusters according to whole-genome phylogenetic analyses. These results indicate that all the OsHV-1 variations characterized in this study originated from the MO farming area. This is congruent with ancestral state reconstruction of the discrete locations (Fig. 6A, Supplementary Fig. S5) that showed evidence for MO being the ancestral location of the twenty-one OsHV-1 isolates sequenced. Moreover, the Bayesian stochastic search variable selection approach suggested significant transition rates of OsHV-1 between the three sampling sites regardless of the alignment used to perform the analysis, i.e. full NR-genomes or NR-genomes with regions potentially impacted by recombination events removed (Supplementary Table S6). Indeed, two rates were inferred with decisive supports: from MO to Th (Bayes factor (BF) = 3,683) and from MO to Br (BF = 11,051) (Fig. 6B).

3. Discussion

The accurate description of viral genetic diversity (major and minor variations) depends on several parameters, such as the quantity and quality of viral genetic material, sequencing quality, and depth, as well as the bioinformatics tools used for data analyses (Sanjuán and Domingo-Calap 2019). In the present study, we characterized OsHV-1 genetic diversity in three oyster-farming areas during POMS outbreaks. To do so, oysters from the same cohort were exposed to POMS outbreaks in the field, and OsHV-1 genetic diversity was investigated in individual infected moribund oysters. To optimize the characterization of OsHV-1 variations, we selected moribund oysters with high viral loads. Then, DNA extracted from these oysters was used to prepare sequencing libraries using a PCR-free kit to remove genomic coverage biases associated with PCR amplification. Using this approach, we obtained an average OsHV-1 genome coverage of ≈3,000×. However, because OsHV-1 coverage varied among samples (ranging from 480× to 8,745×), it was necessary to find a way to verify that coverage was sufficient to accurately quantify viral genomic diversity in each sample. Using a rarefaction analysis, we determined the threshold at which sequencing depth was sufficient to capture all of the genetic diversity present in the studied samples. Although this analysis is commonly used in ecology to assess

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

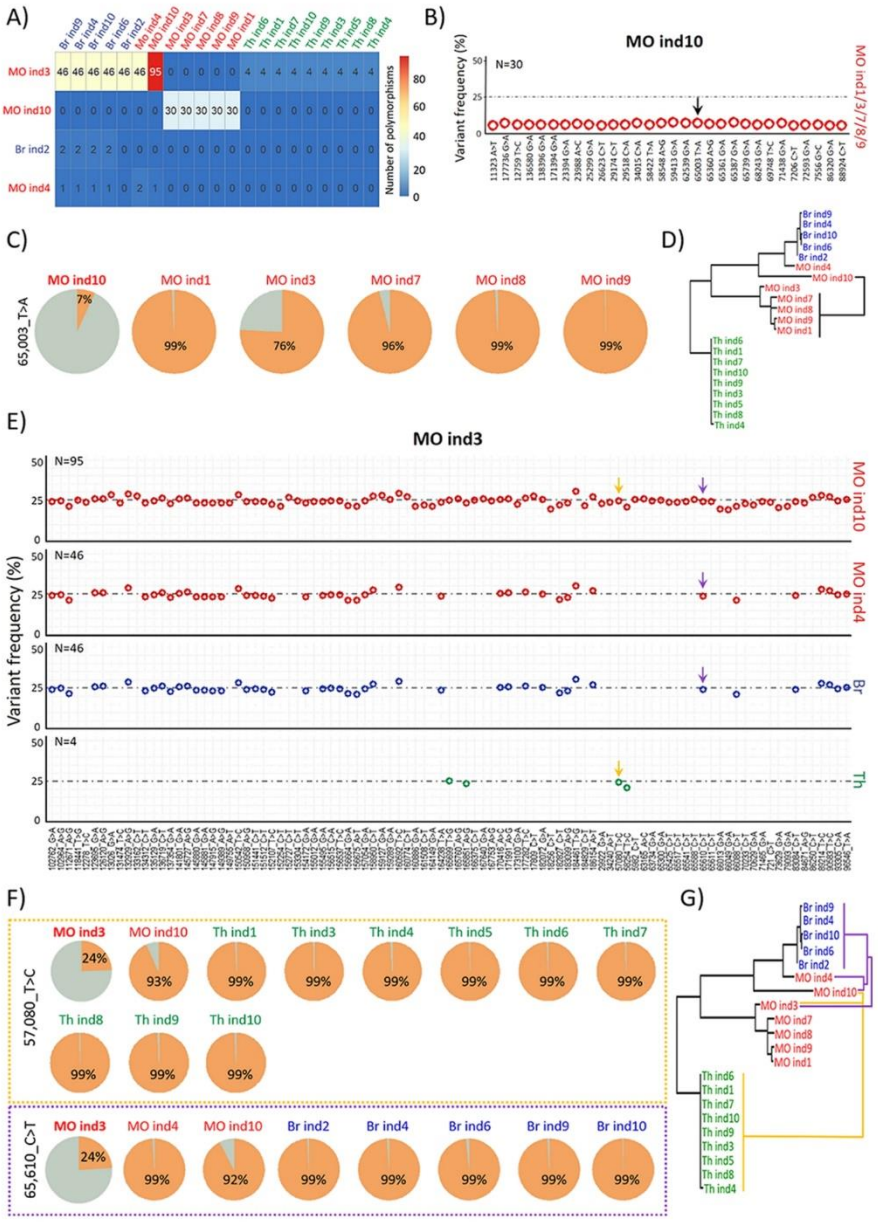


Figure 5. Some minor OsHV-1 variations found in the MO farming area are major variations in the other two farming areas. (A) Matrix of the 132 variations, which were major or minor, according to the sample. Rows display the four samples in which the 132 variations were minor, and columns show all twenty-one samples. Intersections give the number of shared variations with a minor (row) and major (column) frequency. For self-comparisons, no number is given. (B) Dot plot of the 30 minor variations of MO ind10, which were major variations in MO ind1/3/7/8/9. Positions on the Avg-genome and the type of SNPs are indicated on the *x*-axis. Variation frequency (in per cent) is indicated in ordinate. The black arrow indicates the variation that is detailed in (C). (C) Pie charts of the frequency of the variation found at position 65,003 of the NR-genome in MO ind1/3/7/8/9/10. (D) Phylogenetic tree of the twenty-one *de novo* assembled NR-genomes rooted in the midpoint. The black line shows the genetic connectivity established between MO ind10 and MO ind1/3/7/8/9 from variation frequencies at the thirty positions indicated in (B). (E) Dot plot showing the position and type of SNPs (*x*-axis) and variation frequency (*y*-axis) of the ninety-five minor variations of MO ind3 that were present as major variations in MO ind10 (ninety-five variations), in MO ind4 (forty-six variations), as well as in all individuals from Br (forty-six variations) and Th (four variations). Yellow and purple arrows indicate the two variations (57,080_T > C and 65,610_C > T, respectively) that are detailed in (F). (F) The minor variation found in MO ind3 at position 57,080 of the Avg-genome corresponds to major variations in MO ind10 and all Th samples (upper panel, framed in yellow). The minor variation found in MO ind3 at position 65,610 of the Avg-genome corresponds to major variations in MO ind4/10 and all Br samples (lower panel, framed in purple). (G) Phylogenetic tree of the twenty-one *de novo* assembled NR-genomes rooted in the midpoint. The yellow lines show the genetic connectivity established between MO ind3 and MO ind10 plus all Th samples based on the study of variation frequencies at the four positions indicated in (E). The purple lines show the genetic connectivity established between MO ind3 and MO ind4, MO ind10 plus all Br samples based on the study of variation frequencies at the ninety-five (MO ind4) and forty-six (MO ind10 and all Br samples) positions indicated in (E).

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

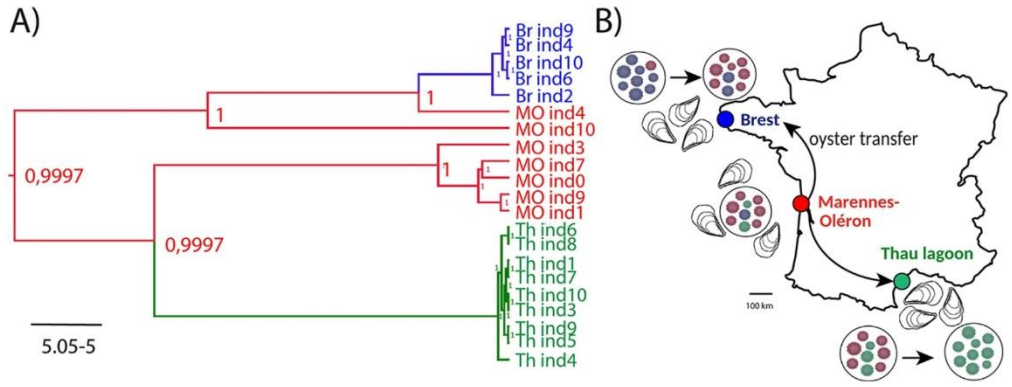


Figure 6. Graphical representation of OsHV-1 population connectivity in France. (A) MCC phylogeny of NR-genomes of OsVH-1 isolates and ancestral state reconstruction of the discrete locations: Br, MO and Th lagoon. The color of the branches represents the most probable location state of their descendent nodes. The color-coding is consistent with other figures: Br in blue, MO in red, and Th in green. Numbers at branch nodes indicate the posterior location probabilities. Given that we did not incorporate any timing information into our analysis (no calibration prior), the scale beneath the tree is expressed in the average number of substitutions per time unit. (B) Bayesian stochastic search variable selection was used to infer dispersal routes with non-zero rates using BFs. Routes that were inferred to have decisive support (BF > 100) are plotted on map (black arrows).

the species richness in a community (Chase et al. 2018), it is, to our knowledge, rarely applied to study viral genetic diversity. This information can be crucial when studying viral diseases in which the high diversity of the virus may potentially increase the fitness of the viral population, making it hard to eradicate (Domingo, Sheldon, and Perales 2012; Al Khatib et al. 2020).

Having a great sequencing depth was an advantage for OsHV-1 genome assembly; however, the task was complicated by the presence of large repeated sequences, corresponding to 9.1 per cent of the genome of OsHV-1 μVar A. To solve this problem, we used an approach commonly used for herpesvirus genome analyses (Cunningham et al. 2010; Morse et al. 2017): contigs generated by *de novo* assembly were ordered using a reference-based approach to ultimately assemble NR genomes in which only one copy of each repeated element was kept. These NR genomes constituted the keystone of our bioinformatics analyses. Because each of them contained the entire genomic sequence of the virus, we were able to carry out an exhaustive search for intra-individual (sample) genetic variations and characterize the viral population structure. These genomes were also aligned with each other to perform whole-genome phylogeny and genome-wide comparative genomics.

Our results revealed that viral populations of OsHV-1 are a heterogeneous set of genomes rather than a single dominating genome, as revealed for another aquatic herpesvirus, the Cyprinid herpesvirus 3 (Hammoumi et al. 2016). They also suggest that all OsHV-1 characterized in our study share a common ancestor with OsHV-1 μVar characterized by Burioli et al. in Normandy (France), reinforcing the idea that genotypes closely related to μVar have replaced the reference genotype and are predominant in oyster-farming areas along the French coasts (Burioli, Prearo, and Houssin 2017; Burioli et al. 2018). Our findings also showed that OsHV-1 μVar populations are much more diverse than previously thought. This is congruent with results from Delmotte et al., showing that OsHV-1 μVar from the Bay of Br and Th Lagoon constitute two distinct viral populations (Delmotte et al. 2020) as well as with the recent relatively high estimate of OsHV-1 evolutionary rate (Morga et al. 2021). We therefore argue it is now necessary to combine common approaches relying on the generation and the analysis of consensus genomes for intra-host diversity characterization in order to accurately study the epidemiology of DNA viruses (Renner and Szpara 2018).

Here, we identified major and minor variations within viral populations that substantially vary in frequency among sampled oyster individuals. Thus, ultra-deep sequencing can support the reconstruction of the true structure of a viral population and inform on relationships between individuals from the same farming area or between individuals from the three farming areas (Fig. 5D and G, respectively). Ancestral state reconstruction, used to infer the dynamics of dispersal of OsHV-1 populations across the three farming areas (Fig. 6A), along with the distribution of major and minor variations, suggests that the MO farming area is the main source of OsHV-1 viral diversity, from where variations are dispersed to the other farming areas. This hypothesis is in agreement with what is known about oyster transfers. Indeed, MO is the main area of *C. gigas* spat collection, and juvenile oysters are then transferred to growing sites along the coasts of Brittany, Normandy, and also along the Mediterranean (Buestel et al. 2009) (Fig. 6B). Importantly, the three studied sites are separated by long distances (>500 km) and therefore are not directly connected through ocean currents. OsHV-1 thus likely disseminates via the transportation of infected oysters, similar to the spread of other diseases via livestock transfer (e.g. Keeling et al. 2001). Effectively several studies showed that wild-caught spats or oysters surviving an OsHV-1 infection contain the virus and are therefore able to transmit it when cohabiting with uninfected oysters (Schikorski et al. 2011; Dégremont and Benabdelmouna 2014; Agnew et al. 2020).

OsHV-1 diversity found in the MO Bay was higher than in the Bay of Br and Th Lagoon. This may be related to the fact that the MO Bay has become one of the largest French oyster-farming areas, with intensive shellfish cultivation and chronic overstocking (Buestel et al. 2009). While we acknowledge the potential bias resulting from small sampling size, this is coherent with large viral population size and increased opportunities for recombination. Mutation and recombination rates being good prognostic markers for the possible emergence of phenotypically different viruses, studying them at the different farming areas would help to better prevent future new OsHV-1 variation outbreaks. Moreover, a network analysis of oyster movements revealed that oysters are

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

moved up to nine times during their production cycle, with peaks of transfers in spring and autumn. These aquacultural practices likely contribute to the spread of OsHV-1 and in turn increase its genetic diversity (Lupo et al. 2016).

One hypothesis to explain the lower diversity in the Bay of Br and Th Lagoon samples could be that only few oysters were infected during field transplantation, and then the infection would be spread in the experimental tanks. However, this hypothesis is unlikely: first because dead oysters were observed in the batches as soon as they were recovered from the field, which clearly shows that they were infected, and secondly because genetic variations present only in one individual (singletons) were characterized in samples from Br and Th. Given that the fate of new genetic variations is largely determined by host selection, another hypothesis to explain the lower diversity in the Bay of Br and Th Lagoon is that oyster populations exert a bottleneck on viral diversity imported from MO. Indeed, it can be speculated that some OsHV-1 variations imported from MO have not found oysters with a genetic background compatible with their replication and survival. However, several studies conducted in Europe to document oyster genetic variability and population structure have revealed a high level of genetic diversity, but no genetic differentiation between French populations (Rohfritsch et al. 2013; Vendrami et al. 2019; Lapegue et al. 2020). This lack of differentiation suggests that selection pressure exerted by oysters is not sufficient to explain differences in viral genetic diversity between farming areas. Genetic drift also influences the probability and rate by which alleles increase or decrease in frequency in a viral population; drift results in the loss of genetic diversity because only a subset of the population contributes to the next generation (Sanjuán and Domingo-Calap 2019). Furthermore, the relationship between environmental factors and mortality events in Pacific oysters has been well documented (reviewed in Alfaro, Nguyen, and Merien 2018). Environmental factors, such as temperature, food availability, water quality, and salinity or UV radiation have been identified as risk factors that directly or indirectly affect POMS dynamics. Given the variation in these factors among the studied oyster-farming areas, identifying the effects of these factors on the structure of OsHV-1 μVar populations may shed new light on the processes of variation selection. Therefore, a larger OsHV-1 study should focus on characterizing the epidemiological and evolutionary processes that shape OsHV-1 genetic diversity patterns and on identifying the ecological, environmental, and anthropogenic factors influencing OsHV-1 dispersal using a combination of phylogeography and spatial epidemiology approaches (Pybus et al. 2012; Dellicour, Rose, and Pybus 2016).

Combining data on OsHV-1 viral population structure using an innovative combination of bioinformatics tools with robust epidemiological information from the different farming areas can offer valuable insight into the dynamics of POMS infection (routes of transmission and potential reservoirs) and thus ultimately help implement effective integrated viral disease management strategies (Avarre 2017).

4. Materials and methods

4.1 OsHV-1 field infection and oyster sampling

All *C. gigas* oysters used in the present study were produced in August 2017 in the Ifremer experimental facilities located in Argenton (Brittany, France), as SPF juveniles (Petton et al. 2015). This SPF oyster cohort was generated from 132 genitors (108 females and 24 males) derived from wild-caught spat collected between 2012 and 2015 (33 genitors for each year). To sample OsHV-1 diversity, SPF oysters originating from the same batch were transplanted into three oyster-farming areas (~1,000 individuals per farm; average weight 1.5 g/oyster) during a disease outbreak when seawater temperature was above 16°C: (1) the Th Lagoon (Mèze, lat.: 43.379 and long.: 3.571) in May 2018, (2) the MO Bay (La Floride, lat.: 45.803 and long.: −1.153) in June 2018, and (3) the Bay of Br (Logonna-Daoulas, lat.: 48.335 long.: −4.318) in July 2018 (Fig. 1A). It has been established that massive mortalities can occur as early as 5 days after oyster transplantation in the natural environment (Petton et al. 2015; Dupont et al. 2020). For this reason, oysters were transferred back to laboratory facilities 5 days after transplantation. Dead oysters were removed from each batch, and the remaining oysters were then maintained in seawater tanks at 20°C (density of 20–25 oysters/l), and moribund oysters were collected daily and stored at −80°C until analysis. A control group was composed of oysters that had not been transplanted into natural environment. Monitoring was stopped after three consecutive days without any oyster mortality. During transport, oysters were packed in polystyrene containers and kept moist by covering them with a damp cloth.

4.2 DNA extraction, viral load quantification, and sequencing

DNA was extracted from individual moribund oysters using the MagAttract® HMW DNA kit (Qiagen) according to the manufacturer's protocol. DNA purity and concentration were checked using a Nano-Drop ND-1000 spectrometer (Thermo Scientific) and Qubit® double-stranded DNA HS assay kits (Molecular Probes Life Technologies), respectively. Quantification of OsHV-1 was carried out using quantitative PCR (qPCR). Amplification reactions were performed using the Roche LightCycler 480 real-time thermocycler on three technical replicates (qPHD-Montpellier GenomiX platform, Montpellier University, France). The total qPCR reaction volume was 1.5 µl, comprising 0.5 µl of DNA (40 ng/µl) and 1 µl of LightCycler 480 SYBR Green I Master mix (Roche) containing 0.5 µM PCR primers (Eurogentec SA). The primers used were virus-specific and targeted the region of the OsHV-1 genome predicted to encode a catalytic subunit of DNA polymerase (ORF100, AY509253): forward: 5′-ATTGATGATGTGGATAATCTGTG-3′ and reverse: 5′-GGTAAATACCATTGGTCTTGTTCC-3′ (Webb, Fidler, and Renault 2007). The following program was applied: enzyme activation at 95°C for 10 min followed by forty cycles of denaturation (95°C, 10 s), annealing (60°C, 20 s), and elongation (72°C, 25 s). To check the specificity of the amplification, a subsequent melting step was applied. Twenty-one samples with a viral load greater than or equal to 10^5 genomic units (GU)/ng of DNA were selected for sequencing (five from Br, seven from MO, and nine from Th). DNA-Seq library preparation and sequencing were performed by the Genome Quebec Company (Genome Quebec Innovation Center, McGill University, Montreal, Canada) using the Shotgun PCR-free library preparation kit (Lucigen) and the NovaSeq™ 6000 Sequencing system (Illumina®) (paired ends, 150 bp).

4.3 Raw read quality control and selection of viral reads

Read quality was assessed using FastQC v0.11.8 (Andrews 2010), and sequence adapters were removed using trimmomatic 0.39 (Bolger, Lohse, and Usadel 2014). Reads shorter than 50 bp were discarded, and bases at the start and the end of a read were trimmed when their Phred quality score was below 30. Additionally, reads were clipped if the average quality within

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

a 4-bp sliding window fell below 15. PCR duplicates were removed using Picard v2.22.4 ('Picard toolkit', 2019 http://broadinstitute.github.io/picard/ accessed 21 April 2022). Then, viral reads were extracted using KrakenUniq 0.5.8 (Breitwieser, Baker, and Salzberg 2018) and Seqkit v0.12.0 (Shen et al. 2016) with the seq –name –only-id and grep pattern-file. To avoid any host contamination, extracted reads were aligned to a *C. gigas* reference genome (assembly version V9; Zhang et al. 2012) using Bowtie2 4.8.2 (Langmead and Salzberg 2012), and non-aligned viral reads were kept for subsequent analyses.

4.4 Rarefaction analysis

In order to verify that OsHV-1 sequencing depth was sufficient to accurately characterize whole viral genomic diversity, we performed a rarefaction analysis for each of the twenty-one libraries. To this end, we performed an iterative variant calling analysis by subsampling and mapping the reads on the OsHV-1 μVar A genome (GenBank: KY242785). In order to accelerate analyses, a nonlinear subsampling incrementation was performed. We started by subsampling 1,000 reads until reaching 10 per cent of the total number of reads of the library analyzed. Then, we subsampled 4,000, 11,000, 21,000, and 41,000 reads until reaching 30, 50, 80, and 100 per cent of the total number of reads of the library, respectively. After each step of subsampling/variant calling, variants were compared to the previously identified ones, and the new ones were classified in SNPs and InDels in order to increment rarefaction curves.

4.5 *De novo* assembly of NR OsHV-1 genomes

Viral reads were used for *de novo* assembly using SPAdes with the –meta and – only assembler options (Nurk et al. 2017). Assembled scaffolds were aligned to the OsHV-1 μVar A genome (GenBank No. KY242785.1) using BLAST 2.9.0 (Altschul et al. 1990) with an *e*-value of 0.00001. Scaffolds were then extended using SSPACE v3.0 (Boetzer et al. 2011). Considering that the OsHV-1 genome is composed of a combination of unique (U) and repeated (R) regions (TRL–UL–IRL–X–IRS–US–TRS-X; Davison et al. 2005) and that the latter have a size varying from 160 to 9,777 bp, most of these repeat sequences could not be resolved using a kmer approach.

Scaffolds were grouped according to their size using Seqkit v0.12.0 (Shen et al. 2016). The first group of scaffolds had a size between 2 and 4 kbp and made it possible to select the US region. The second group was composed of sequences with a size between 4 and 20 kbp, which made it possible to select the scaffold constituting IRL–X–IRS. Finally, the last group contained sequences greater than 20 kbp and made it possible to select the UL region. We used these three groups of scaffolds to construct NR-genomes that contain only one copy of each repeated region (UL–IRL–X–IRS–US) (Morse et al. 2017). Because the repeat regions are inverted complements of each other, several adjustments were made to avoid the assembly of sequences in the wrong orientation. Using this strategy, twenty-one OsHV-1 NR-genomes were assembled from the twenty-one individual oyster samples.

4.6 Whole-genome comparisons and phylogenetic analyses

The twenty-one NR-genomes were then aligned with the six NR-genomes derived from previously published OsHV-1 genomes (AY509253, KU096999, KP412538, KY242785, KY271630, MG561751 and MF509813) using MAFFT v7.475 with default parameters (Katoh et al. 2002). A maximum-likelihood phylogenetic inference was conducted on the twenty-seven NR-genomes using PhyML v3.1 (Guindon et al. 2009) with 100 bootstrap replications. We used a Hasegawa–Kishino–Yano nucleotide substitution model with invariant sites and gamma-distributed categories of rate variation (HKY+G+I), which was considered as the most appropriate model by jModelTest v2.1.10 (Darriba et al. 2012) based on Akaike's information criterion. Similarly, in order to get a finer picture of phylogenetic relationships among samples, the twenty-one NR-genomes were aligned with OsHV-1 μVar A (KY242785) genome only, and a maximum-likelihood tree was obtained using an HKY+G+I substitution model. The results were locally visualized using the R package ggTree v2.0.4 (Yu et al. 2016) and FigTree v1.4.4 (available at http://tree.bio.ed.ac.uk/software/figtree/ accessed 21 April 2022).

4.7 Genomic variability within and between farming areas

An Avg-genome of the twenty-one newly generated NR-genomes was built using the 'cons' argument from the EMBOSS suite v6.6.0.0 with the default settings (Rice, Longden, and Bleasby 2000). The NR-genomes were compared to the Avg-genome using MUMer4 4.0.0beta2 (Marcais et al. 2018) with the option 'nucmer -c 100 -l 15 –f' to identify all the variable positions between the twenty-one NR-genomes.

To assess the genomic variability between the farming areas, both minor and major variations of each sequencing library (twenty-one) were called on the Avg-genome previously generated using Freebayes v1.3.2-dirty (Garrison and Marth 2012), with the following settings: –use-mapping-quality, –min-repeat-entropy 1, –haplotype-length 0, –min-alternate-count 5, –pooled- continuous, –hwe-priors-off, –allele-balance-priors-off. The resulting variant calling outputs were normalized using BCFtools v- (Narasimhan et al. 2016), decomposed with vt v1.0.0 (Tan, Abecasis, and Kang 2015), and split with vcflib v1.0.0. Considering the very rare occurrence of multi-nucleotide polymorphisms, these latter were counted as InDels. Variable positions with a frequency of >50 per cent (major variations) were compared to those obtained with MUMer4 for comparative genomics, and all were validated. They were subsequently subtracted from the variant calling files to retain only minor variations (with a frequency of <50 per cent). A unique identifier composed of the position and sequence information (e.g. 128,737_A>G) of each minor variation was created. Set analyses of nucleotide variations were performed with a matrix approach using R (R Core Team 2018) operation and visualized with the Venn Diagram package (Chen and Boutros 2011), UpSetR v1.4.0 (Lex et al. 2014), or pheatmap v1.0.12.

All downstream analyses (tables, graphs, plot creation, and edition) were performed in R 3.6 (R Core Team 2018) on R studio IDE (R-Studio-Team 2015) with an extensive use of Dplyr v1.0.0 (Wickham et al. 2022) and ggplot2 (Wickham 2016) from the tidyverse v1.3.0 package (Wickham et al. 2019). Genome visualization was carried out using gggenes v0.4.0. For statistical analyses, samples were treated as three groups, representing the different farming areas (Br, MO, and Th). Kruskal–Wallis and Dunn's multiple comparison tests were used locally (GraphPad Prism 8.4.2) to compare the number of SNPs or InDels.

4.8 Genomic annotation and variation effect predictions

Genomic annotation was performed using five gene prediction softwares: Prodigal (Hyatt et al. 2010), FragGeneScan (Rho *et al.*, 2010; version 1.31), GeneMark with heuristic models

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

(Zhu, Lomsadze, and Borodovsky 2010; version 3.25), GeneMarkS for virus (Besemer, Lomsadze, and Borodovsky 2001; version 4.28), and finally ORF finder implemented in Geneious Prime (version 2022.0.1). All identified ORFs were manually sorted in Geneious Prime, and only those predicted by at least three gene predictors were retained for genomic annotation. Variant calling files (VCF) of major and minor variations were used to generate two global VCF containing the coordinates of the variations accorging to the Avg-genome. SnpEff (Cingolani et al. 2012; version 4.1 l) was then used to predict the effect of major and minor variations using a home-made database constructed with the Avg-genome and its annotations.

4.9 Ancestral state reconstruction of discrete trait

To infer the geographic origin of our sampled isolates, sampling sites (Th, MO and Br) were modeled as a discrete trait for each NR-genome sequence over the genealogy by ancestral state inference using a discrete asymmetric phylogenetic diffusion model (Lemey et al. 2014) in BEAST v1.10.4 (Suchard et al. 2018). This approach estimates the probability of the internal nodes and branches being associated with a specific sampling site, based on information about sampling site states of the samples at the branch tips. A Bayesian stochastic search variable selection procedure (Lemey et al. 2009) was employed to allow for diffusion between specific location pairs to be included or excluded from the model. Based on Akaike's information criterion ranking in jModelTest v2.1.10 (Darriba et al. 2012), analyses were performed under a GTR+I+G model of nucleotide substitution. A strict molecular clock and a constant coalescent model were used, and an Markov chain Monte Carlo chain of 10,000,000 steps was run and sub-sampled every 1,000 generations. Convergence to the stationary distribution and sufficient mixing (effective sample size >200) for all parameter estimates were checked in Tracer after removing the initial 10 per cent of the samples as burn-in. The BEAGLE library was used to increase computational speed (Ayres et al. 2012). BFs were computed in SpreaD3 (Bielejec et al. 2016) and used as a measure of support for identifying frequently invoked transition rates to explain the diffusion process. Support for a rate was considered substantial when BF > 3, strong if BF > 10, and decisive if BF > 100 (Jeffreys 1961).

To investigate the possible impact of recombination on the ancestral state reconstruction, the analysis was repeated on an alignment where genomic regions potentially impacted by recombination events were removed. To do so, we used recombination detection program (RDP) version 4.97 (Martin et al. 2015), with default settings, to identify any recombination events between the OsHV-1 NR-genomes we sequenced. As evidence for recombination, we took events detected by at least two of seven different recombination detection methods implemented in RDP: RDP, MAXCHI, and GENECONV methods in primary scanning mode and the Bootscan, CHIMAERA, SisScan, and 3SEQ methods in secondary scanning mode, each with a Bonferroni-corrected *P*-value cut-off of 0.05. Two genomic regions were detected by RDP4 as potentially impacted by three recombination events and removed from the alignment (213 SNPs remaining after pairwise deletion at gapped sites, 46 after complete deletion) before running jModelTest, BEAST, and SpreaD3 again.

Maximum clade credibility (MCC) trees were generated from tree output files from BEAST in TreeAnnotator v1.10.4 and annotated in the FigTree v1.4.4 graphical user interface (available at http://tree.bio.ed.ac.uk/software/figtree/ accessed 21 April 2022).

Data availability

The datasets generated from this study can be found in the Sequence Read Archive database BioProject accession number PRJNA681599 with submission ID SUB8642385. The twenty-one OsHV-1 genomes generated in this study are available in GenBank under the following accession numbers: Brest samples (from OM811577 to OM811581), Marennes-Oléron samples (from OM811582 to OM811588) and Thau samples (from OM811589 to OM811597). All the scripts generated in this study are freely available at https://github.com/propan2one/OshV-1-molepidemio accessed 21 April 2022. Complementary information is available from the corresponding authors upon reasonable request.

Supplementary data

Supplementary data is available at *Virus Evolution* online.

Acknowledgements

The present study was supported by the EU project VIVALDI (H2020 program, no. 678589) led by Ifremer, the CNRS, and the University of Montpellier. The authors acknowledge the Pôle de Calcul et de Données Marines (https://wwz.ifremer.fr/en/Research-Technology/Research-Infrastructures/Digital-infrastructures/Computation-Centre accessed 21 April 2022) for providing DATARMOR computing and storage resources. We thank the staff of the Ifremer stations at Argenton (Laboratoire de Physiologie des Invertébrés, Physiologie Fonctionnelle des Organismes Marins) and La Tremblade (Adaptation et Santé des Invertébrés Marins (ASIM)), Frédéric Girardin and his team (Plateforme des Mollusques Marins de La Tremblade PMMLT), and the Comité Régional Conchylicole de Méditerranée for technical support during transplantation experiments. We also thank Nicole Faury (Ifremer, ASIM) and Marc Leroy (CNRS, IHPE) for technical assistance. We are also grateful to Eric Rivals Laboratoire d'Informatique de Robotique et de Microelectronique de Montpellier for fruitful discussions.

Funding

University of Montpellier to J.D.; Ifremer and Nouvelle Aquitaine region for C.P. 'Laboratoire d'Excellence' (LabEx) CeMEB, through the exploratory research project HaploFit and the use of the 'environmental genomics' facility (http://www.labex-cemeb.org/fr/genomique-environnementale-1 accessed 21 April 2022); Ifremer Scientific Board, through the HemoVir and GT-oyster projects; 'Fond Européen pour les Affaires Maritimes et la Pêche' (GESTINNOV project no. PFEA470020FA1000007). This study is set within the framework of the 'LabEx' TULIP (ANR-10-LABX-41). The funders had no role in study design, data collection, and interpretation or the decision to submit the work for publication.

Conflict of interest: We declare that we have no competing interests.

Author's contribution

J.-C.A., O.K., B.M., J.D., and J.-M.E. were involved in the study conception and design. B.M., B.P., C.P., J.D., and J.-M.E. were involved in the collection of samples and in the experimental work. J.D., C.P., M.J., and J.-M.E. were involved in bioinformatics and statistical analyses. J.D., R.G., C.M., M.J., and J.-M.E. drafted the manuscript, and all authors revised and approved the final manuscript.

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

Importance

Phylogeography is a field of research that attempts to reconstruct the relationships between individual genotypes within a species and then correlate these genealogical relationships with their geographic and temporal origin. This field of research has become an essential step in the understanding of pandemics, in particular, to determine the origin, spread, and evolution of a pathogen as currently illustrated in studies on viral pandemics. However, because phylogeographic analyses are based on genome variation, stable genomes yield less information than labile genomes. Accordingly, viruses with double-stranded DNA genomes generally have lower nucleotide diversity than RNA viruses. In this study, by combining the use of both major and minor genetic variations with phylogeographic analyses of the oyster herpesvirus OsHV-1, we highlight genealogical relationships that are not depicted in phylogenetic trees based on consensus viral genomes only. These data offer a plausible scenario reflecting the origin and spread of OsHV-1 populations between oyster-farming sites.

References

Abbadi, M. et al. (2018) 'Identification of a Newly Described OsHV-1 Microvar from the North Adriatic Sea (Italy)', *The Journal of General Virology*, 99: 693–703.

Agnew, M. V. et al. (2020) 'Differential Mortality and High Viral Load in Naive Pacific Oyster Families Exposed to OsHV-1 Suggests Tolerance Rather Than Resistance to Infection', *Pathogens*, 9: 1057–75.

Akhtar, L. N. et al. (2019) 'Genotypic and Phenotypic Diversity of Herpes Simplex Virus 2 within the Infected Neonatal Population', *mSphere*, 4: e00590–18.

Al Khatib, H. A. et al. (2020) 'Within-Host Diversity of SARS-CoV-2 in COVID-19 Patients with Variable Disease Severities', *Frontiers in Cellular and Infection Microbiology*, 10: 575613.

Alfaro, A. C., Nguyen, T. V., and Merien, F. (2018) 'The Complex Interactions of Ostreid Herpesvirus 1, Vibrio Bacteria, Environment and Host Factors in Mass Mortality Outbreaks of *Crassostrea gigas*', *Reviews in Aquaculture*, 11: 1148–68.

Allicock, O. M. et al. (2012) 'Phylogeography and Population Dynamics of Dengue Viruses in the Americas', *Molecular Biology and Evolution*, 29: 1533–43.

Altschul, S. F. et al. (1990) 'Basic Local Alignment Search Tool', *Journal of Molecular Biology*, 215: 403–10.

Andrews, S. (2010) *FastQC: A Quality Control Tool for High Throughput Sequence Data*. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ accessed 21 Apr 2022.

Avarre, J. C. (2017) 'Editorial: Molecular Tracing of Aquatic Viruses: Where Epidemiology Needs to Meet Genomics', *Frontiers in Microbiology*, 8: 1498.

Ayres, D. L. et al. (2012) 'BEAGLE: An Application Programming Interface and High-performance Computing Library for Statistical Phylogenetics', *Systematic Biology*, 61: 170–3.

Barbieri, E. S. et al. (2019) 'First Detection of Ostreid Herpesvirus 1 in Wild *Crassostrea gigas* in Argentina', *Journal of Invertebrate Pathology*, 166: 107222.

Barbosa Solomieu, V., Renault, T., and Travers, M. A. (2015) 'Mass Mortality in Bivalves and the Intricate Case of the Pacific Oyster, *Crassostrea gigas*', *Journal of Invertebrate Pathology*, 131: 2–10.

Besemer, J., Lomsadze, A., and Borodovsky, M. (2001) 'GeneMarkS: A Self-training Method for Prediction of Gene Starts in Microbial Genomes. Implications for Finding Sequence Motifs in Regulatory Regions', *Nucleic Acids Research*, 29: 2607–18.

Bielejec, F. et al. (2016) 'SpreaD3: Interactive Visualization of Spatiotemporal History and Trait Evolutionary Processes', *Molecular Biology and Evolution*, 33: 2167–9.

Boetzer, M. et al. (2011) 'Scaffolding Pre-assembled Contigs Using SSPACE', *Bioinformatics*, 27: 578–9.

Bolger, A. M., Lohse, M., and Usadel, B. (2014) 'Trimmomatic: A Flexible Trimmer for Illumina Sequence Data', *Bioinformatics*, 30: 2114–20.

Breitwieser, F. P., Baker, D. N., and Salzberg, S. L. (2018) 'KrakenUniq: Confident and Fast Metagenomics Classification Using Unique K-mer Counts', *Genome Biology*, 19: 198.

Buestel, D. et al. (2009) 'History, Status, and Future of Oyster Culture in France', *Journal of Shellfish Research*, 28: 813–20.

Burge, C. A., Shore-Maggio, A., and Rivlin, N. D. (2017) 'Ecology of Emerging Infectious Diseases of Invertebrates'. In: Hajek, A. E. (ed.) *Ecology of Invertebrate Diseases*, pp. 587–612. John Wiley & Sons Ltd: Hoboken, New Jersey.

Burioli, E. A. et al. (2016) 'Ostreid Herpesvirus Type 1 Genomic Diversity in Wild Populations of Pacific Oyster *Crassostrea gigas* from Italian Coasts', *Journal of Invertebrate Pathology*, 137: 71–83.

Burioli, E. A. V., Prearo, M., and Houssin, M. (2017) 'Complete Genome Sequence of Ostreid Herpesvirus Type 1 microVar Isolated during Mortality Events in the Pacific Oyster *Crassostrea gigas* in France and Ireland', *Virology*, 509: 239–51.

Burioli, E. A. V. et al. (2018) 'A Novel Divergent Group of Ostreid Herpesvirus 1 muVar Variants Associated with a Mortality Event in Pacific Oyster Spat in Normandy (France) in 2016', *Journal of Fish Diseases*, 41: 1759–69.

Chase, J. M. et al. (2018) 'Embracing Scale-dependence to Achieve a Deeper Understanding of Biodiversity and Its Change across Communities', *Ecology Letters*, 21: 1737–51.

Chen, H., and Boutros, P. C. (2011) 'VennDiagram: A Package for the Generation of Highly-Customizable Venn and Euler Diagrams in R', *BMC Bioinformatics*, 12: 35.

Cingolani, P. et al. (2012) 'A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of Drosophila Melanogaster Strain W(1118); Iso-2; Iso-3', *Fly*, 6: 80–92.

Cunningham, C. et al. (2010) 'Sequences of Complete Human Cytomegalovirus Genomes from Infected Cell Cultures and Clinical Specimens', *The Journal of General Virology*, 91: 605–15.

Darriba, D. et al. (2012) 'jModelTest 2: More Models, New Heuristics and Parallel Computing', *Nature Methods*, 9: 772.

Davison, A. J. et al. (2005) 'A Novel Class of Herpesvirus with Bivalve Hosts', *The Journal of General Virology*, 86: 41–53.

de Lorgeril, J. et al. (2018) 'Immune-suppression by OsHV-1 Viral Infection Causes Fatal Bacteraemia in Pacific Oysters', *Nature Communications*, 9: 4215.

——— et al. (2020) 'Differential Basal Expression of Immune Genes Confers *Crassostrea gigas* Resistance to Pacific Oyster Mortality Syndrome', *BMC Genomics*, 21: 63.

Dégremont, L., and Benabdelmouna, A. (2014) 'Mortality Associated with OsHV-1 in Spat *Crassostrea gigas*: Role of Wild-Caught Spat in the Horizontal Transmission of the Disease', *Aquacultural International*, 22: 1767–81.

Dégremont, L., Nourry, M., and Maurouard, E. (2015) 'Mass Selection for Survival and Resistance to OsHV-1 Infection in *Crassostrea gigas* Spat in Field Conditions: Response to Selection after Four Generations', *Aquaculture*, 446: 111–21.

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

Dellicour, S., Rose, R., and Pybus, O. G. (2016) 'Explaining the Geographic Spread of Emerging Epidemics: A Framework for Comparing Viral Phylogenies and Environmental Landscape Data', *BMC Bioinformatics*, 17: 82.

Delmotte, J. et al. (2020) 'Contribution of Viral Genomic Diversity to Oyster Susceptibility in the Pacific Oyster Mortality Syndrome', *Frontiers in Microbiology*, 11: 1579.

Domingo, E., Sheldon, J., and Perales, C. (2012) 'Viral Quasispecies Evolution', *Microbiology Molecular Biology Reviews*, 76: 159–216.

Dupont, S. et al. (2020) 'Oyster Hemolymph Is a Complex and Dynamic Ecosystem Hosting Bacteria, Protists and Viruses', *Animal Microbiome*, 2: 12.

Faria, N. R. et al. (2014) 'HIV Epidemiology. The Early Spread and Epidemic Ignition of HIV-1 in Human Populations', *Science*, 346: 56–61.

Friedman, C. S. et al. (2005) 'Herpes Virus in Juvenile Pacific Oysters *Crassostrea gigas* from Tomales Bay, California, Coincides with Summer Mortality Episodes', *Diseases of Aquatic Organisms*, 63: 33–41.

——— et al. (2020) 'Unraveling Concordant and Varying Responses of Oyster Species to Ostreid Herpesvirus 1 Variants', *The Science of the Total Environment*, 739: 139752.

Garrison, E., and Marth, G. (2012) *Haplotype-based Variant Detection from Short-read Sequencing*. arXiv:1207.3907 [q-bio.GN].

Gittenberger, A., Voorbergen-Laarman, M. A., and Engelsma, M. Y. (2016) 'Ostreid Herpesvirus OsHV-1 Mu Var in Pacific Oysters *Crassostrea gigas* (Thunberg 1793) of the Wadden Sea, a UNESCO World Heritage Site', *Journal of Fish Diseases*, 39: 105–9.

Guindon, S. et al. (2009) 'Estimating Maximum Likelihood Phylogenies with PhyML', *Methods in Molecular Biology (Clifton, N.J.)*, 537: 113–37.

Hammoumi, S. et al. (2016) 'Targeted Genomic Enrichment and Sequencing of CyHV-3 from Carp Tissues Confirms Low Nucleotide Diversity and Mixed Genotype Infections', *PeerJ*, 4: e2516.

Hyatt, D. et al. (2010) 'Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification', *BMC Bioinformatics*, 11: 119.

Jeffreys, H. (1961) *Theory of Probability*, 3rd edn. Clarendon Press: Oxford.

Jenkins, C. et al. (2013) 'Identification and Characterisation of an Ostreid Herpesvirus-1 Microvariant (Oshv-1 Micro-var) in *Crassostrea gigas* (Pacific Oysters) in Australia', *Diseases of Aquatic Organisms*, 105: 109–26.

Katoh, K. et al. (2002) 'MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform', *Nucleic Acids Research*, 30: 3059–66.

Keeling, M. J. et al. (2001) 'Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape', *Science*, 294: 813–7.

Langmead, B., and Salzberg, S. L. (2012) 'Fast Gapped-Read Alignment with Bowtie 2', *Nature Methods*, 9: 357–9.

Lapegue, S. et al. (2020) 'Genetic Characterization of Cupped Oyster Resources in Europe Using Informative Single Nucleotide Polymorphism (SNP) Panels', *Genes*, 11: 451–68.

Lauring, A. S., and Andino, R. (2010) 'Quasispecies Theory and the Behavior of RNA Viruses', *PLoS Pathogens*, 6: e1001005.

Lemey, P. et al. (2014) 'Unifying Viral Genetics and Human Transportation Data to Predict the Global Transmission Dynamics of Human Influenza H3N2', *PLoS Pathogens*, 10: e1003932.

——— et al. (2009) 'Bayesian Phylogeography Finds Its Roots', *PLoS Computational Biology*, 5: e1000520.

Lex, A. et al. (2014) 'UpSet: Visualization of Intersecting Sets', *IEEE Transactions on Visualization and Computer Graphics*, 20: 1983–92.

Lupo, C. et al. (2016) How Network Analysis of Oyster Movements Can Improve Surveillance and Control Programs of Infectious Diseases? AquaEpi (Aquatic Animal Epidemiology), Oslo, Norway.

Marcais, G. et al. (2018) 'MUMmer4: A Fast and Versatile Genome Alignment System', *PLoS Computational Biology*, 14: e1005944.

Martenot, C. et al. (2011) 'Detection of Different Variants of Ostreid Herpesvirus 1 in the Pacific Oyster, *Crassostrea gigas* between 2008 and 2010', *Virus Research*, 160: 25–31.

Martin, D. P. et al. (2015) 'RDP4: Detection and Analysis of Recombination Patterns in Virus Genomes', *Virus Evolution*, 1: vev003.

Martin, M. A., VanInsberghe, D., and Koelle, K. (2021) 'Insights from SARS-CoV-2 Sequences', *Science*, 371: 466–7.

Morga, B. et al. (2021) 'Genomic Diversity of the Ostreid Herpesvirus Type 1 across Time and Location and among Host Species', *Frontiers in Microbiology*, 12: 711377.

Morse, A. M. et al. (2017) 'Reliable Detection of Herpes Simplex Virus Sequence Variation by High-Throughput Resequencing', *Viruses*, 9: 226.

Moss, J. A. et al. (2007) 'Pathogens in *Crassostrea ariakensis* and Other Asian Oyster Species: Implications for Non-native Oyster Introduction to Chesapeake Bay', *Diseases of Aquatic Organisms*, 77: 207–23.

Narasimhan, V. et al. (2016) 'BCFtools/RoH: A Hidden Markov Model Approach for Detecting Autozygosity from Next-generation Sequencing Data', *Bioinformatics*, 32: 1749–51.

Nkili-Meyong, A. A. et al. (2016) 'Contribution of Next-Generation Sequencing to Aquatic and Fish Virology', *Intervirology*, 59: 285–300.

Nurk, S. et al. (2017) 'metaSPAdes: A New Versatile Metagenomic Assembler', *Genome Research*, 5: 824–34.

Paul-Pont, I., Dhand, N. K., and Whittington, R. J. (2013) 'Spatial Distribution of Mortality in Pacific Oysters *Crassostrea gigas*: Reflection on Mechanisms of OsHV-1 Transmission', *Diseases of Aquatic Organisms*, 105: 127–38.

Peeler, E. J. et al. (2012) 'Investigation of Mortality in Pacific Oysters Associated with Ostreid Herpesvirus-1 Mu Var in the Republic of Ireland in 2009', *Preventive Veterinary Medicine*, 105: 136–43.

Pernet, F. et al. (2012) 'Mass Mortalities of Pacific Oysters *Crassostrea gigas* Reflect Infectious Diseases and Vary with Farming Practices in the Mediterranean Thau Lagoon, France', *Aquaculture Environmental Interactions*, 2: 215–37.

——— et al. (2016) 'Infectious Diseases in Oyster Aquaculture Require a New Integrated Approach', *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371: 1689.

Petton, B. et al. (2015) '*Crassostrea gigas* Mortality in France: The Usual Suspect, a Herpes Virus, May Not Be the Killer in This Polymicrobial Opportunistic Disease', *Frontiers in Microbiology*, 6: 686.

——— et al. (2021) 'The Pacific Oyster Mortality Syndrome, a Polymicrobial and Multifactorial Disease: State of Knowledge and Future Directions', *Frontiers in Immunology*, 12: 630343.

——— et al. (2013) 'Temperature Influence on Pathogen Transmission and Subsequent Mortalities in Juvenile Pacific Oysters *Crassostrea gigas*', *Aquaculture Environment Interactions*, 3: 257–73. Inter-reseach science publisher.

Pybus, O. G. et al. (2012) 'Unifying the Spatial Epidemiology and Molecular Evolution of Emerging Epidemics', *Proceedings of the National Academy of Sciences of the United States of America*, 109: 15066–71.

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.

Rambaut, A. et al. (2008) 'The Genomic and Epidemiological Dynamics of Human Influenza A Virus', *Nature*, 453: 615–9.

Renault, T. et al. (2012) 'Analysis of Clinical Ostreid Herpesvirus 1 (*Malacoherpesviridae*) Specimens by Sequencing Amplified Fragments from Three Virus Genome Areas', *Journal of Virology*, 86: 5942–7.

Renner, D. W., and Szpara, M. L. (2018) 'Impacts of Genome-Wide Analyses on Our Understanding of Human Herpesvirus Diversity and Evolution', *Journal of Virology*, 92: e00908–17.

Renzette, N. et al. (2013) 'Rapid Intrahost Evolution of Human Cytomegalovirus Is Shaped by Demography and Positive Selection', *PLoS Genetics*, 9: e1003735.

——— et al. (2014) 'Human Cytomegalovirus Intrahost Evolution-A New Avenue for Understanding and Controlling Herpesvirus Infections', *Current Opinion in Virology*, 8: 109–15.

——— et al. (2015) 'Limits and Patterns of Cytomegalovirus Genomic Diversity in Humans', *Proceedings of the National Academy of Sciences of the United States of America*, 112: E4120–8.

Rho, M., Tang, H., and Ye, Y. (2010) 'FragGeneScan: predicting genes in short and error-prone reads', *Nucleic Acids Research*, 38: e191.

Rice, P., Longden, I., and Bleasby, A. (2000) 'EMBOSS: The European Molecular Biology Open Software Suite', *Trends in Genetics: TIG*, 16: 276–7.

Rohfritsch, A. et al. (2013) 'Population Genomics Shed Light on the Demographic and Adaptive Histories of European Invasion in the Pacific Oyster, *Crassostrea gigas*', *Evolutionary Applications*, 6: 1064–78.

R-Studio-Team (2015) *RStudio: Integrated Development Environment for R*. RStudio Inc: Boston, MA.

Sanjuán, R., and Domingo-Calap, P. (2019) 'Genetic Diversity and Evolution of Viral Populations', *Encyclopedia of Virology*, 2021: 53–61.

Schikorski, D. et al. (2011) 'Experimental Ostreid Herpesvirus 1 Infection of the Pacific Oyster *Crassostrea gigas*: Kinetics of Virus DNA Detection by q-PCR in Seawater and in Oyster Samples', *Virus Research*, 155: 28–34.

Segarra, A. et al. (2010) 'Detection and Description of a Particular Ostreid Herpesvirus 1 Genotype Associated with Massive Mortality Outbreaks of Pacific Oysters, *Crassostrea gigas*, in France in 2008', *Virus Research*, 153: 92–9.

Shen, W. et al. (2016) 'SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation', *PLoS One*, 11: e0163962.

Stentiford, G. D. et al. (2012) 'Disease Will Limit Future Food Supply from the Global Crustacean Fishery and Aquaculture Sectors', *Journal of Invertebrate Pathology*, 110: 141–57.

Suchard, M. A. et al. (2018) 'Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10', *Virus Evolution*, 4: vey016.

Tan, A., Abecasis, G. R., and Kang, H. M. (2015) 'Unified Representation of Genetic Variants', *Bioinformatics*, 31: 2202–4.

Theze, J. et al. (2018) 'Genomic Epidemiology Reconstructs the Introduction and Spread of Zika Virus in Central America and Mexico', *Cell Host & Microbe*, 23: 855–64 e857.

Vasquez-Yeomans, R., Garcia-Ortega, M., and Caceres-Martinez, J. (2010) 'Gill Erosion and Herpesvirus in *Crassostrea gigas* Cultured in Baja California, Mexico', *Diseases of Aquatic Organisms*, 89: 137–44.

Vendrami, D. L. J. et al. (2019) 'Detailed Insights into pan-European Population Structure and Inbreeding in Wild and Hatchery Pacific Oysters (*Crassostrea gigas*) Revealed by Genome-wide SNP Data', *Evolutionary Applications*, 12: 519–34.

Vignuzzi, M. et al. (2006) 'Quasispecies Diversity Determines Pathogenesis through Cooperative Interactions in a Viral Population', *Nature*, 439: 344–8.

Webb, S. C., Fidler, A., and Renault, T. (2007) 'Primers for PCR-based Detection of Ostreid Herpes Virus-1 (Oshv-1): Application in a Survey of New Zealand Molluscs', *Aquaculture*, 272: 126–39.

Wickham, H. (2016) *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag: New York.

——— et al. (2019) 'Welcome to the Tidyverse', *Journal of Open Source Software*, 4: 1686.

——— et al. (2022) *Dplyr: A Grammar of Data Manipulation*. RStudio.

Yu, G. et al. (2016) 'GGTREE: An R Package for Visualization and Annotation of Phylogenetic Trees with Their Covariates and Other Associated Data', *Methods in Ecology and Evolution*, 8: 28–36.

Zhang, G. et al. (2012) 'The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation', *Nature*, 490: 49–54.

Zhu, W. H., Lomsadze, A., and Borodovsky, M. (2010) 'Ab Initio Gene Identification in Metagenomic Sequences', *Nucleic Acids Research*, 38: e132.

Downloaded from https://academic.oup.com/ve/article/8/1/veac039/6572772 by Ifremer, Bibliothèque La Pérouse user on 29 July 2023

4.3 Résumé de l'article 5

Au travers des deux précédents articles (article 3 et article 4), nous avons vu l'importance d'étudier la diversité génétique du virus OsHV-1 et son évolution, d'une part pour comprendre les facteurs structurant sa diversité, et d'autre part pour comprendre les processus épidémiologiques qui influencent sa transmission dans les bassins ostréicoles du monde entier.

Les deux études précédemment présentées suggèrent une structuration spatio-temporelle des populations du virus OsHV-1. Cependant, ces deux études comportent des biais dans l'analyse de la diversité, pouvant être expliqué par la méthode de reconstruction des génotypes viraux (comme expliqué pour l'article 3) ou par l'utilisation d'un consensus global pour effectuer les étapes d'appel de variant. En effet, dans l'article 4, l'ensemble des analyses de diversité intra-individuelles sont basées sur la reconstruction d'un consensus global qui fusionne la diversité observée dans les 21 échantillons utilisés. Cette approche a été utilisée pour permettre une correspondance et une comparaison des positions présentant des variations. Le biais majeur induit par ce type d'appel de variant est l'identification de SNP qui n'existe pas dans la population. De plus dans ces deux études, la taille d'échantillonnage est très limitée ce qui ne permet pas de tirer des conclusions précises sur les processus évolutifs et épidémiologiques qui agissent sur le génome du virus OsHV-1. Les résultats obtenus dans ces deux études soulignent tout de même l'importance d'étudier la diversité génomique complète avec un nombre important d'échantillons pour comprendre les structures de population virale et les routes de dispersion, des éléments essentiels pour élaborer des mesures de sécurité contre la propagation du virus.

Dans cette étude, nous avons caractérisé la structure spatio-temporelle de la population d'OsHV-1 à l'échelle mondiale. Nous avons séquencé et assemblé *de novo* 274 nouveaux génomes d'OsHV-1 prélevés chez *M. gigas* dans différentes zones de production d'huîtres à travers le monde entre 1993 et 2022. Des analyses de variations structurales ont permis d'étudier les réarrangements génomiques d'OsHV-1. Une approche phylogéographique combinée avec des analyses permettant de déterminer les populations ancestrales composant chaque individu a été utilisée pour reconstituer l'origine ancestrale d'OsHV-1 dans le monde et les voies de dispersion du virus en France. Cette approche globale a permis d'obtenir des informations essentielles sur la diversité et l'évolution d'OsHV-1, ainsi que sur les facteurs influençant sa propagation à l'échelle mondiale et nationale.

Les résultats mettent en évidence une évolution temporelle dynamique de l'architecture génomique d'OsHV-1, avec l'émergence du génotype d'OsHV-1 de référence, puis de l'OsHV-1 µVar en France en 2008, et des génomes intermédiaires dans le Pacifique. De plus, cette étude retrace l'histoire évolutive

mondiale du virus OsHV-1, liée à l'introduction de *M. gigas* dans le monde pour l'aquaculture, tout en soulignant l'importance de la région de Marennes-Oléron en France en tant que source clé de diversité virale. Ce bassin est actuellement le plus gros producteur d'huîtres d'Europe, ce qui influence potentiellement la diversité virale. Cette étude offre un aperçu complet de la diversité génomique et de l'histoire évolutive d'OsHV-1 à l'échelle mondiale, fournissant ainsi des bases essentielles pour la gestion ciblée de ce virus et son impact sur les huîtres et l'industrie de conchylicole.

Article 5

Global diversity and dispersal routes of the *Ostreid herpesvirus type 1*

Camille Pelletier, Germain Chevignon, Maude Jacquot, Benjamin Morga.

Global diversity and dispersal routes of the Ostreid herpesvirus type 1

Camille Pelletier[1], Germain Chevignon[1], Maude Jacquot[1], Benjamin Morga[1].

Affiliations

[1] Ifremer, RBE-ASIM, Unité Adaptations Santé des Invertébrés Marins, Avenue de Mus de Loup, 17390 La Tremblade, France.

Corresponding authors:

Benjamin Morga

E-mail address: Benjamin.morga@ifremer.fr,

Tel +33 5 46 76 26 49

Fax +33 5 46 76 26 11

Maude Jacquot

E-mail address: maude.jacquot@ifremer.fr,

Tel +33 5 46 76 26 82

Prepared and formatted for submission to *Virus evolution*

Abstract

The Herpesviridae family, comprising double-stranded DNA viruses, poses significant health threats to both humans and animals, causing various infectious diseases. One of the major challenges in oyster production involves Ostreid herpesvirus 1 (OsHV-1), a virus that has caused mass mortalities in Pacific oysters *Magallana gigas*. The rapid spread of the virus OsHV-1 is linked to the global movement of oysters for farming purposes. Recent research has focused on understanding population structure and evolution of OsHV-1 using phylogeographic approaches. These approaches help trace the dispersal routes of the OsHV-1 virus. The study at hand employed high-throughput sequencing to analyze 274 new OsHV-1 genomes collected from various oyster production areas worldwide from 1994 to 2022. Through a synergy of comparative genomics and phylodynamic methods, we have delved into the virus's evolutionary history on a global scale, with a specific emphasis on tracing the routes of viral dispersal within France. Deletions, and recombination within the viral genome are highlighted as driving forces in its evolution. The study also traces the evolution history of OsHV-1, identifying the reference genotype's emergence subsequent evolution into the different OsHV-1 µVar variant detected wolrdwide. In France, the Marennes-Oléron farming region serves as a primary source of viral diversity, with the virus disseminating to other farming areas. Oyster movement practices play a crucial role in the virus's spread, with oysters transferred multiple times during their production cycle. This study contributes to a deeper understanding of the history of OsHV-1 and informs potential control measures against the disease.

Keywords: Herpesvirus; OsHV-1; Pacific oyster; phylogeography; genetic diversity; viral spread

Introduction

Herpesviridae, a family of double-stranded DNA viruses, pose a significant threat to human and animal health, causing a wide range of infectious diseases. In humans, notable examples include *Herpes simplex virus* (HSV), which is involved in oral and genital herpes (Whitley *et al.*, 2007); *Epstein Barr virus* (EBV), associated with infectious mononucleosis (Fugl & Andersen, 2019); and *Varicella Zoster virus* (VZV) responsible for varicella during the primary infection and for herpes zoster after reactivation (Ayoade & Kumar, 2023). Herpesviruses can also induce significant diseases in various animal species. For instance, *Gallid Herpesvirus 2*, also referred to as Marek's disease virus, is accountable for tumor formation on multiple organs, immune suppression, and paralysis in chickens (Gimeno *et al.*, 1999). Similarly, *Cyprinid herpesvirus 3* leads to extensive mortalities in common carp and koi fishes, accompanied by skin lesions, gill damage, and internal organ failure (Michel *et al.*, 2010). Additionally, *Haliotid herpesvirus 1* contributes to substantial mortality among wild and farmed abalones, and is associated with nerve system lesions (Chang *et al.*, 2005; Corbeil, 2020). Impact of Herpesvirus is particularly pronounced in aquaculture, a vital industry for food production, where they have been responsible for devastating mortalities in farms, leading to substantial economic losses (Matsui *et al.*, 2008; Renault, 2011).

One prominent sector within the marine aquaculture industry is oyster production, with the Pacific oyster *Magallana gigas* (previously called *Crassostrea gigas*) being the dominant oyster species cultivated worldwide, accounting for 98% of oyster production (FAO, 2023). The pacific oyster *M. gigas* has suffered from numerous mass mortality events at the larval and spat developmental stages due to a viral disease caused by the *Ostreid herpesvirus 1* (Arzul *et al.*, 2017; Davison *et al.*, 2005; Farley *et al.*, 1972; Nicolas *et al.*, 1992; Segarra *et al.*, 2010). OsHV-1 belongs to the Malacoherpesviridae family. The first detection of a herpes-like virus infecting a bivalve mollusk species was reported in 1970, in California, in a hatchery of American cupped oyster *Magallana virginica* (Farley *et al.*,1972). Since this initial detection, the virus has been rapidly detected in Europe (Batista *et al.*, 2015; Lynch *et al.*, 2012; Nicolas *et al.*, 1992; Peeler *et al.*, 2012; Renault *et al.*, 1994; Roque *et al.*, 2012), Asia (Boudry *et al.*, 1998; Hwang *et al.*, 2013; Moss *et al.*, 2007; Ren *et al.*, 2013; Xia *et al.*, 2015), Oceania (Cochennec *et al.*, 2011; Hine & Thorne, 1997; Jenkins *et al.*, 2013; M. J. Keeling *et al.*, 2001; Paul-Pont *et al.*, 2014; Webb *et al.*, 2007), and even North Africa (Fabioux *et al.*, 2002). Its swift dispersion is potentially associated with the movement of oysters among global oyster basins, particularly during the introduction of the Pacific oyster, *M. gigas*, and during various steps of oyster farming (Delmotte *et al.*, 2022). Pacific oysters were initially imported from Japan to Canada and the coasts of the USA in the 1920s, and later to Australia and New Zealand in the 1950s (Nell, 2001; Quayle, 1988). In the early

1970s, Pacific oyster spat and broodstock were introduced in large quantities to France and Argentina from Japan and Canada (Grizel & Heral, 1991; Wörner *et al.*, 2019).

Since this initial detection, the virus has been associated with mortalities in multiple host species including the scallops *Pecten maximus* (Arzul, Nicolas, *et al.*, 2001) and *Azumapecten farreri* (Bai *et al.*, 2019), the clams *Anadara broughtonii* (Xia *et al.*, 2015), *Ruditapes philippinarum* (Renault & Arzul, 2001; Renault, 1998, p. 199) and *Ruditapes decussatus* (Renault & Arzul, 2001; Renault, 1998, p. 199), as well as the oysters *Magallana angulata* (Arzul, Renault, *et al.*, 2001), *Ostrea edulis* (Arzul, Renault, *et al.*, 2001; Comps, 1988; Renault *et al.*, 2001), and *M. gigas* (Burioli *et al.*, 2017; Nicolas *et al.*, 1992; Renault *et al.*, 1994).

Early studies employing targeted techniques such as PCR, microsatellites, and amplicon sequencing revealed the presence of different genotypes (*i.e.* allelic composition of an individual) that emerged subsequently to the introduction of the reference genotype (Davison *et al.*, 2005). The first genotype to be identified was found in clam and oyster larvae in 1997 (Arzul, Renault, *et al.*, 2001) and later in scallop larvae in 2001 (Arzul, Nicolas, *et al.*, 2001). This genotype was designated as OsHV-1 Var and is characterized by a deletion of at least 2.8 kilobase pairs (kbp) in the TR_L region in comparison to the OsHV-1 reference genome (Arzul, Nicolas, *et al.*, 2001). In 2008, mass mortalities of Pacific oyster spat were reported in Europe, coinciding with the emergence of a new genotype initially detected in France (Martenot *et al.*, 2011; Renault *et al.*, 2012; Segarra *et al.*, 2010), followed by its identification in Ireland (Peeler *et al.*, 2012) and UK (Lynch *et al.*, 2012). This genotype was named µVar due to genetic variations observed at a microsatellite located upstream of ORF4, and single nucleotide polymorphisms (SNPs) within ORF4 and ORF42/43 (OIE, 2019). Multiple genotypes of the µVar variants (*i.e.* subtype of the virus that becomes genetically distinct from a main strain because of one or more mutations in its genome over time) have since been described (Barbosa Solomieu *et al.*, 2015).

OsHV-1 genome was first sequenced, assembled and characterized in 2005 from infected larvae *M. gigas* sampled in a French commercial hatchery in 1992 (Davison *et al.*, 2005). This genome has a length of 207 kbp and is composed of unique regions (U_L/U_S: unique long or short) surrounded by repeated regions (TR_L/TR_S: terminal repeat long or short and IR_L/IR_S: inverted repeat long or short). Its genomic architecture can be summarized as TR_L-U_L-IR_L-X-IR_S-U_S-TR_S (Davison *et al.*, 2005). Prior to 2017, the detection and examination of OsHV-1 diversity primarily involved targeting genomic sequencing of regions bellow 1 kbp (Renault *et al.*, 2011, 2014). However, in 2017, the genomes of two micro-variants of OsHV-1: OsHV-1 µVarA and µVarB were published (Burioli *et al.* 2017). The first comprehensive comparisons between OsHV-1 reference genome (Davison *et al.*, 2005) and these two micro-variants revealed significant differences, including the loss of 6.4 kbp and the acquisition of approximately 3

kbp (Burioli *et al.*, 2017). Recently, new sequenced genomes of OsHV-1 collected from various host species, including *M. gigas*, *A. farreri* and *S. broughtonii*, and locations such as Italy, China or France have been deposited in public databases (Abbadi *et al.*, 2018; Bai *et al.*, 2019; Delmotte *et al.*, 2022; Hwang *et al.*, 2013; Xia *et al.*, 2015). These genomes also exhibit significant differences with the reference genome of OsHV-1. These findings highlight the importance of studying genome variability at the whole-genome level to accurately characterize viral population structures among individuals and detect low-frequency genotypes. This approach enables the inference of evolutionary and epidemiological histories and can provide further insights into dispersal routes (Bloomquist *et al.*, 2010; Houldcroft *et al.*, 2017; Nahata *et al.*, 2021).

Currently, there are 52 complete genomes of OsHV-1 available in public databases. However, only two studies have explored the spatio-temporal diversity of OsHV-1 (Delmotte-Pelletier *et al.*, 2022; Morga-Jacquot *et al.*, 2021). A preliminary study focused on the global and temporal genomic diversity and evolution of OsHV-1 across multiple host species and suggested differentiation of OsHV-1 genotypes based on the host species and sampling year (Morga-Jacquot *et al.*, 2021). However, this study used viral haplotypes reconstructed by read mapping on the OsHV-1 reference genome rather than a *de novo* assembly approach. This type of reconstruction introduces significant biases in the analysis of diversity, particularly in the assessment of insertion-type structural variations, which are often ignored. It also impacts the variant calling step, potentially leading to the identification of SNPs that may not actually be present in the population. Conversely, a more recent study investigated the diversity of the OsHV-1 virus within three main oyster production regions in France in 2018 using *de novo* assembled OsHV-1 genomes (Delmotte-Pelletier *et al.*, 2022). This analysis demonstrated a clustering of samples based on their geographical origin within France, with greater diversity in the Marennes-Oléron basin. It suggests virus transfers between oyster basins, potentially linked to oyster movements. Nevertheless, the smaller sample size used in this study means that no precise conclusions can be drawn about the epidemiological and evolutionary history of the virus.

These initial studies raised questions about the factors influencing the structure of OsHV-1 populations around the world, as well as about its evolutionary history, from its introduction in Europe to the present day. These informations are crucial to propose measures to control the disease worldwide.

Recently developed phylogeographic models provide a powerful framework for gaining insights into the diffusion processes of pathogens from time-stamped and geo-referenced sequences. One approach to infer viral lineages movements is based on phylogeographic reconstructions in continuous space (Lemey *et al.*, 2010; Pybus *et al.*, 2012). However, no standard method has been developed to deal with uncertainty of locations and samples heterogeneity (Kalkauskas *et al.*, 2021). A second

phylogeographic method estimate locations of samples' ancestors from a diffusion model in discrete space (Lemey *et al.*, 2009). This offers particularly interesting opportunities to infer viral dispersal routes, especially for viruses that propagate by transferring between distinct demes. Furthermore, this approach is less influenced by heterogenic sampling where intermediate locations remain unsampled (Lemey *et al.*, 2009).

In the present study, we used short-reads high-throughput deep sequencing to characterize the OsHV-1 population structure worldwide across time. To this end, we have sequenced and *de novo* assembled 274 new OsHV-1 genomes sampled in different oyster production areas worldwide from 1994 to 2022. Using these data and previous information about genotypes circulation in the field, our objectives were to study OsHV-1 genomic diversity patterns according to geographical origin and time. Additionally, we aimed at identifying the main dispersal routes of the virus in France. In order to achieve this, we used a combination of comparative genomic approaches, along with phylogenetic and phylodynamic approaches. We then examined the findings in the light of current knowledge of OsHV-1 history.

Results

Structural variations

Based on the whole genome alignment of the 274 *de novo* assembled genomes from samples collected worldwide with the four published genomes, a whole genomic structural analysis was conducted. This comprehensive comparison yielded six distinct types of genomic architecture (Figure 1). The first group comprises 10 genomes sampled in France between 1993 and 2008 that are closely related to the reference genome (NC_005881.2; Davison *et al.* 2005). These genomes are characterized by non-redundant (NR) genomes of approximately 190 kbp. This whole genome alignment also revealed five other structural variants, which were further characterized and compared to the OsHV-1 reference genome. These genotypes exhibit distinct genomic arrangements. First, 10 NR-genomes (group 2) of 190 kbp collected in France between 2005 and 2008 differed from the OsHV-1 reference genome by an 8,530 bp reversion in the U_L region located between ORF 36 and ORF 43 which encodes respectively membrane protein, Zinc-finger, Ring-type protein, secreted protein, transmembrane protein and protein coding for unknown function. Secondly, a group of 219 NR-genomes (group 3) of 204 kbp collected in Europe (France, Germany, Spain, Ireland) between 2008 and 2022, were found to be similar to the µVar genotype described by (Burioli *et al.*, 2017). This genotype is characterized by a deletion of 12 bp within a microsatellite located in the TR_L/IR_L region (respectively between 4487-4510 bp and 178547-178570 bp) upstream to the ORF4 (Segarra *et al.*, 2010). Additionally, the µVar genotype exhibited four deletions in the U_L region, ranging from 606 bp to 3,550 bp in length, one

deletion of 725 bp at the junction between U_L and IR_L region within ORF114 (unknown function), and one insertion of 2,671 bp in the U_L region between ORF 43 and 44 that encodes Zinc-finger, Ring-type protein and protein coding for unknown function. In total, the µVar genotype showed approximately 10 kbp of insertions or deletions. Thirdly, 1 genome (group 4), 186 kbp collected in New-Zealand in 2012 was similar to the OsHV-1 µVar genotype but with an additional deletion of 6 kbp at the end of the IR_S and beginning of US regions. Fourthly, 7 NR-genomes (group 5) of 184 kbp collected in USA on the pacific coast (San Diego), in New-Zealand and in Korea between 2011 and 2018 were characterized by an inverted-translocation of 4 kpb at the end of IR_S region at the right extremity of the genome. Consequently, the U_S region is localized in the middle of the IR_S region. Moreover, in comparison with the OsHV-1 reference genome, one large insertion of 2.6 kpb was observed in the U_L region and six large deletions ranging from 603 bp to 3.6 kbp were detected in the U_L and a deletion of 446 pb was detected in the IR_L region. These genotypes do not have the previously characterized 12 bp deletion in the microsatellite region of OsHV-1 µVar genotype, but instead has a 6 bp deletion in this region. In total, these genotypes showed approximately 5.5 kbp of insertions or deletions. Finally, 15 NR-genomes (group 6) collected in USA on the pacific coast (Tomales Bay), Australia, and New-Zealand between 2006 and 2018 of 190 kpb exhibit similar variations as the other genomes collected in the Pacific Ocean, but with an additional deletion of 267 pb and an insertion of 510 pb in the IR_S region (Figure 1).

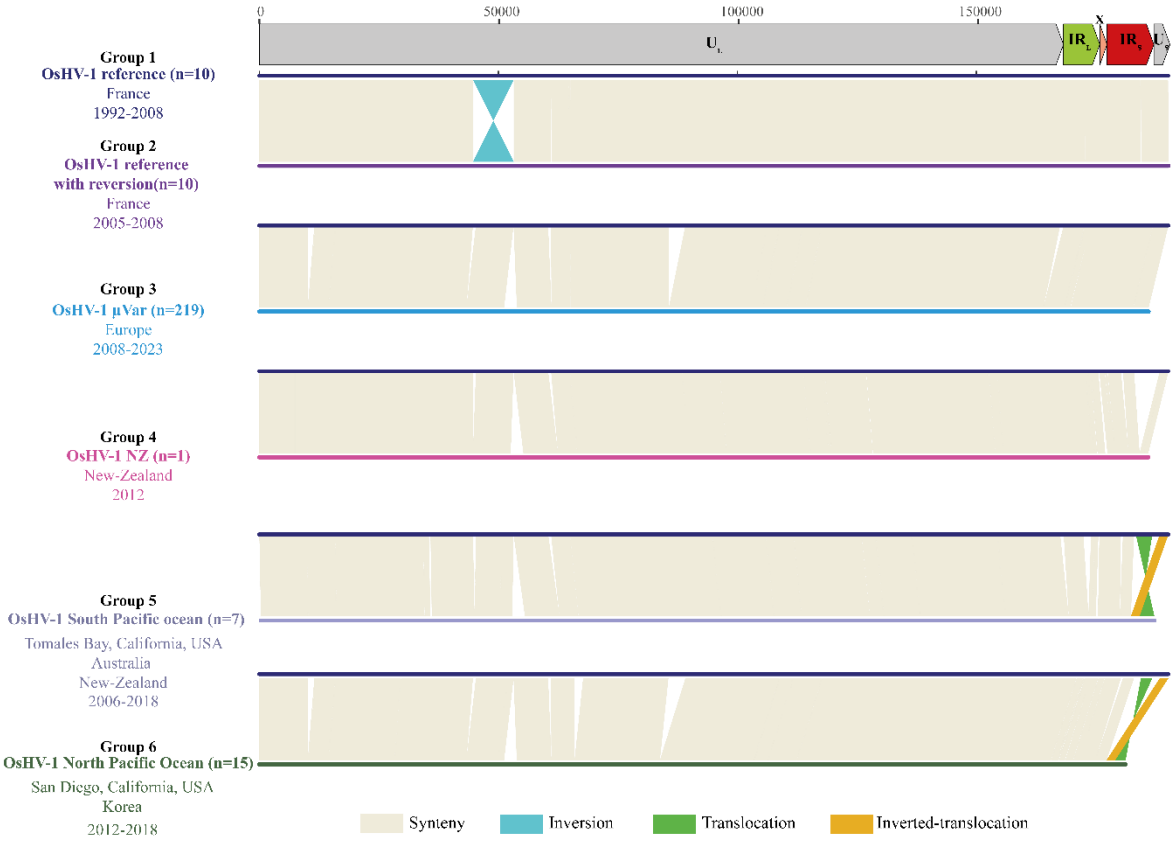


Figure 1: Comparative genomic showing the six OsHV-1 genomic structure type. *The description of genomic structural variations was based on a pairwise comparison to the OsHV-1 reference genome (NC_005881.2, Davison et al. 2005) performed using Mummer. In each pairwise comparison, the reference genome is represented by a dark blue bar. Syntenic regions are represented in beige, inversions in light blue, translocation in green and inverted translocations in orange. Above the graph of structural variations is the architecture and position in the OsHV-1 reference genome.*

Populations admixture

Population admixture is a commonly used analysis to determine the ancestral population composition of each individual genome. The population's genetic structure was assessed through the utilization of STRUCTURE software. Following multiple independent runs of the Gibbs sampler, it was determined that the optimal number of ancestral populations, assumed to be present in the dataset employed for this study, is K=9. Based on population admixture and ancestral populations similarity, we have defined four arbitrary groups of OsHV-1 genomes (Figure 2). The first group (Figure 2) is composed uniquely of samples collected within the Marennes-Oléron basin between 2021 and 2022 excepted for one sample collected in the Brest harbor in 2022. The second group (Figure 2) is composed of samples collected in Europe (Spain, Germany, Ireland and France) between 2008 and 2021. The third group of individuals (Figure 2) is composed by samples collected exclusively in France between 2018 and 2022. Finally, the last group (Figure 2) is composed of individuals collected in the Pacific Ocean from USA, New-Zealand, Australia and Korea and by samples collected in France before 2008, corresponding to the reference genotype.

Interestingly, three samples from USA collected in San Diego and two samples collected in Korea was partly related to ancestral population 2 (Figure 2, green bars), ancestral population 4 (Figure 2, dark purple bars), ancestral population 5 (Figure 2, salmon bars) and ancestral population 6 (Figure 2, orange bars). One individual (2008-CA-040) collected in the Thau lagoon in 2008 present no admixture and is the unique sample related to ancestral population 7 (Figure 2, yellow bars).

Additionally, five distinct pattern of population structure was observed in France. Firstly, 169 OsHV-1 genotypes were characterized by no admixture of ancestral populations. Then 64 French OsHV-1 genotypes were characterized by admixture between two or three ancestral populations from ancestor 1 (red bars), 2 (green bars) or 3 (pink bars).

Finally, four genotypes collected in Thau lagoon and in Breton inlet in 2008 were characterized by admixture between ancestor 1 (red ancestral population), ancestor 2 (green ancestral population) and ancestor 4 (dark purple ancestral population).

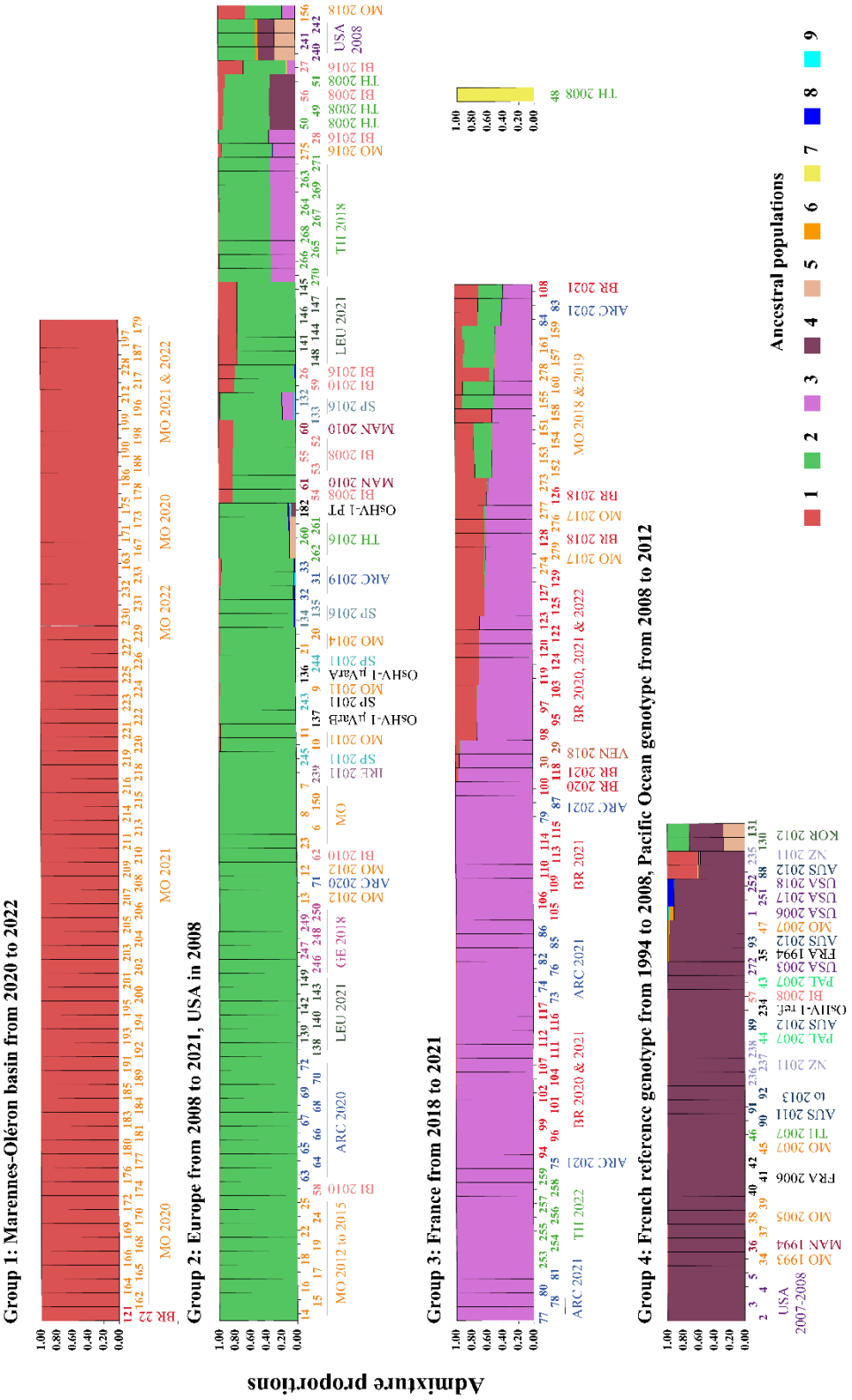


Figure 2: Population admixture and assignation of ancestral population to each of the 279 individuals. *Bar plots representing the assignation of each individual OsHV-1 genome to an ancestral population based on admixture analysis. The analysis of population admixture has been performed using the STRUCTURE software, assuming nine ancestral populations (represented as red, green, pink, dark purple, salmon, orange, yellow, dark blue, light blue). The x axis represents individuals OsHV-1 genomes from populations sorted according to their reported ancestries (defined by Q-Matrix) and four groups of individuals OsHV-1 genome were defined arbitrary based on individuals OsHV-1 genome similarities regarding ancestral populations. Each individual OsHV-1 genome is represented by a vertical stacked column of color-coded admixture proportions that reflects genetic contributions from putative ancestral populations.*

From the Gubbins recombination analysis (Croucher *et al.*, 2015), 16 potential recombination events were identified along OsHV-1 genomes, which might have impacted six specific genomic regions (Figure S1). Eleven potential recombination events occurred in the UL region and five in the repeated regions (Figure S1). All these regions have been removed from the alignment for subsequent analyses.

Maximum clade credibility tree and ancestral states inference

R^2 value of regression of root-to-tip genetic distance against sampling times are provided as a measure of how OsHV-1 evolution follows a clock-like pattern (Figure 3A). According to R^2 value ($R^2 = 0.344$) obtained by the best-fitting root tree, OsHV-1 harbors a sufficient temporal signal to use phylogenetic molecular clock analysis to estimate evolutionary rates over time.

Comparison of different discrete phylogeographic models of OsHV-1 dispersal showed statistical support for the combination of an uncorrelated relaxed molecular clock and a GMRF coalescent model (Table S3).

The time-scaled phylogeny distinguished six phylogenetic groups (Figure 3B). The first cluster correspond to genotypes collected in France between 1993 and 2008 and includes the reference OsHV-1 genotype (Davison *et al.*, 2005). The second cluster correspond to OsHV-1 genotypes collected in the South-West coasts of Pacific Ocean (New-Zealand and Australia) between 2011 and 2013. Samples collected from the North-East coasts of Pacific Ocean (California, USA) between 2003 and 2018 grouped together in a third cluster. Both OsHV-1 genotypes collected from Korea in 2012 clustered in a forth phylogenetic group. The fifth cluster grouped three OsHV-1 genotypes collected from USA in

2018. Finally, the last cluster grouped genotypes collected from European coasts (Germany, Italy, Spain, Ireland and France) between 2008 and 2022.

Furthermore, the mean substitution rate along the overall phylogeny was estimated to be 3.6093E-03 nucleotide substitutions per site per year (subs/site/year) (95% HPD interval [3.2291E-03 - 3.09444E-03]). The most recent common ancestor (MRCA) of OsHV-1 sampled genotypes was estimated to have existed in 1991 (95% HPD interval [1990-1992]) in the Marennes-Oléron basin. Subsequently, it can be inferred that this genotype might have evolved differently in Europe and the Pacific Ocean. In Europe, the first detected genotype was the OsHV-1 reference genotype sampled in 1992 (Davison *et al.*, 2005). In parallel, from this reference genotype, another lineage might have emerged and could have been first detected in the Pacific Ocean between 1994 (95% HPD interval [1991-1998]) and 2018. Subsequently, in Europe, after the emergence of this lineage, it might have evolved into the OsHV-1 µVar genotype, which is currently the main genotype detected.

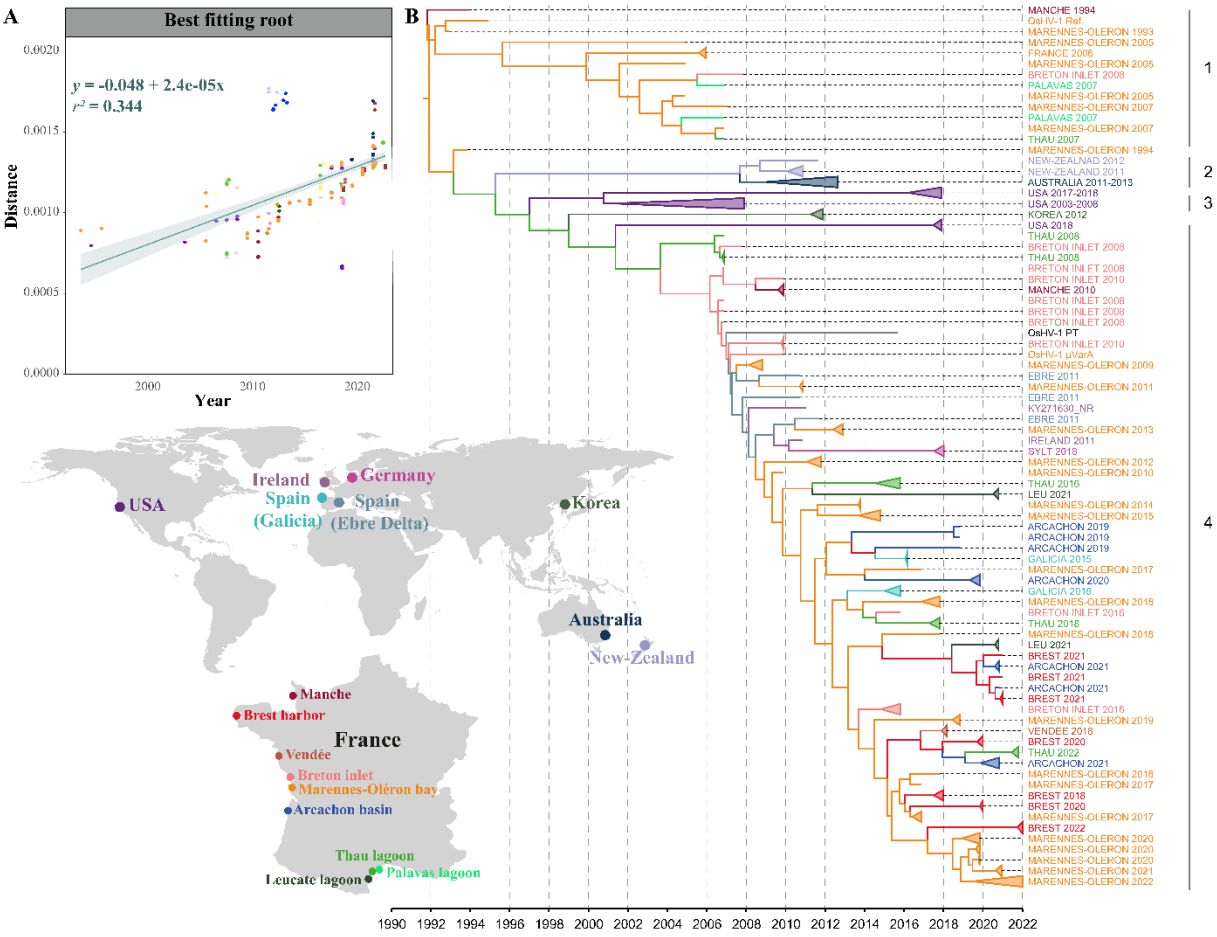


Figure 3: Ancestral state reconstruction of OsHV-1 specific origin between 1990 and 2022 worldwide. *A) Root-to-tip distance in maximum likelihood tree versus sampling year. B) Time-scaled Maximum clade credibility of 279 OsHV-1 genomes which was obtained using Bayesian phylogenetic inference in program BEAST under a HKY+G+I model of nucleotide substitution, uncorrelated relaxed clock and GMRF demographical models. Tip colors correspond to origin of sampling and branch colors correspond to most probable origin of the common ancestor as specified on the map. This figure was produced in FigTree v1.4.4 (Rambaut, 2019). Four clusters were defined (numbered on the right).*

Reconstruction of OsHV-1 dispersal routes in France

In order to explore the viral spread of OsHV-1 in France, an ancestral state analysis and an inference of dispersal routes has been performed specifically for the 238 OsHV-1 genomes sampled in France (Figure 4A). The Maximum Clade Credibility (MCC) tree showed two clusters: (i) the cluster A, corresponding to samples collected before 2008 and (ii) the cluster B corresponding to samples collected after 2008. Surprisingly, two samples collected in 1994 respectively in the Manche Channel and in the Marennes-Oléron basin did not cluster with samples from the first cluster. Within the cluster A, two subcluster has been highlighted. The cluster A.1 grouped samples collected in the Marennes-Oléron basin in 1994 and 1995. The cluster A.2 grouped samples collected in the Marennes-Oléron basin between 2005 and 2008. The cluster B was more diverse as the cluster A, and show 13 subclusters that gather samples according to geographical origin and sampling date (Figure 4A). All these subclusters are structured according to geographical origin and sampling dates. Moreover, genotypes collected in the Marennes-Oléron basin are distributed throughout the phylogeny, indicating high viral diversity in this basin.

The MRCA has been dated back to 1992 (95% HPD interval [1991-1993]). The ancestral state reconstruction of specific locations provides evidences supporting Marennes-Oléron as the ancestral location for the 239 OsHV-1 isolates that were sequenced (Figure 4A). Furthermore, the Bayesian stochastic search variable selection approach indicated significant transition rates of OsHV-1 among the nine sampling sites (Table 1, Figure 4B). In fact, two transition rates were inferred with decisive statistical support (Bayes Factor (BF) > 100): one from Marennes-Oléron to Breton inlet (BF = 2107) and another from Vendée to Arcachon bay (BF = 100). Additionally, six transition rates were inferred with strong support (BF > 10): from Marennes-Oléron to Vendée (BF = 96) and Thau (BF = 21), from Arcachon bay to Brest harbor (BF = 73) and Marennes-Oléron (BF = 19) and from Thau to the Breton inlet (BF = 64) and the Manche sea (BF = 16). Finally, eight transition rates were inferred with substantial statistical support (BF > 3): from Marennes-Oléron to Leucate (BF = 9), Arcachon (BF = 9)

and Palavas (BF = 4), from the Breton inlet to Palavas (BF = 5), Thau lagoon (BF = 4) and the Manche sea (BF = 3), from Vendée to Leucate (BF = 5) and from Brest harbor to Marennes-Oléron (BF = 4).

Table 1: Bayes factors computed to infer dispersal routes of OsHV-1 in France.

From	To	Bayes factor	Posterior probability	Significance
Marennes-Oléron	Breton inlet	2107,19	1,00	Decisive
Vendée	Arcachon	100,53	0,93	
Marennes-Oléron	Vendée	96,27	0,93	Strong
Arcachon	Brest	73,44	0,91	
Thau	Breton inlet	63,58	0,90	
Marennes-Oléron	Thau	21,03	0,74	
Arcachon	Marennes-Oléron	19,00	0,72	
Thau	Manche	16,01	0,69	
Marennes-Oléron	Leucate	9,17	0,56	Substantial
Marennes-Oléron	Arcachon	8,99	0,55	
Breton inlet	Palavas	4,74	0,39	
Vendée	Leucate	4,65	0,39	
Breton inlet	Thau	4,44	0,38	
Brest	Marennes-Oléron	4,13	0,36	
Marennes-Oléron	Palavas	3,94	0,35	
Breton inlet	Manche	3,00	0,29	

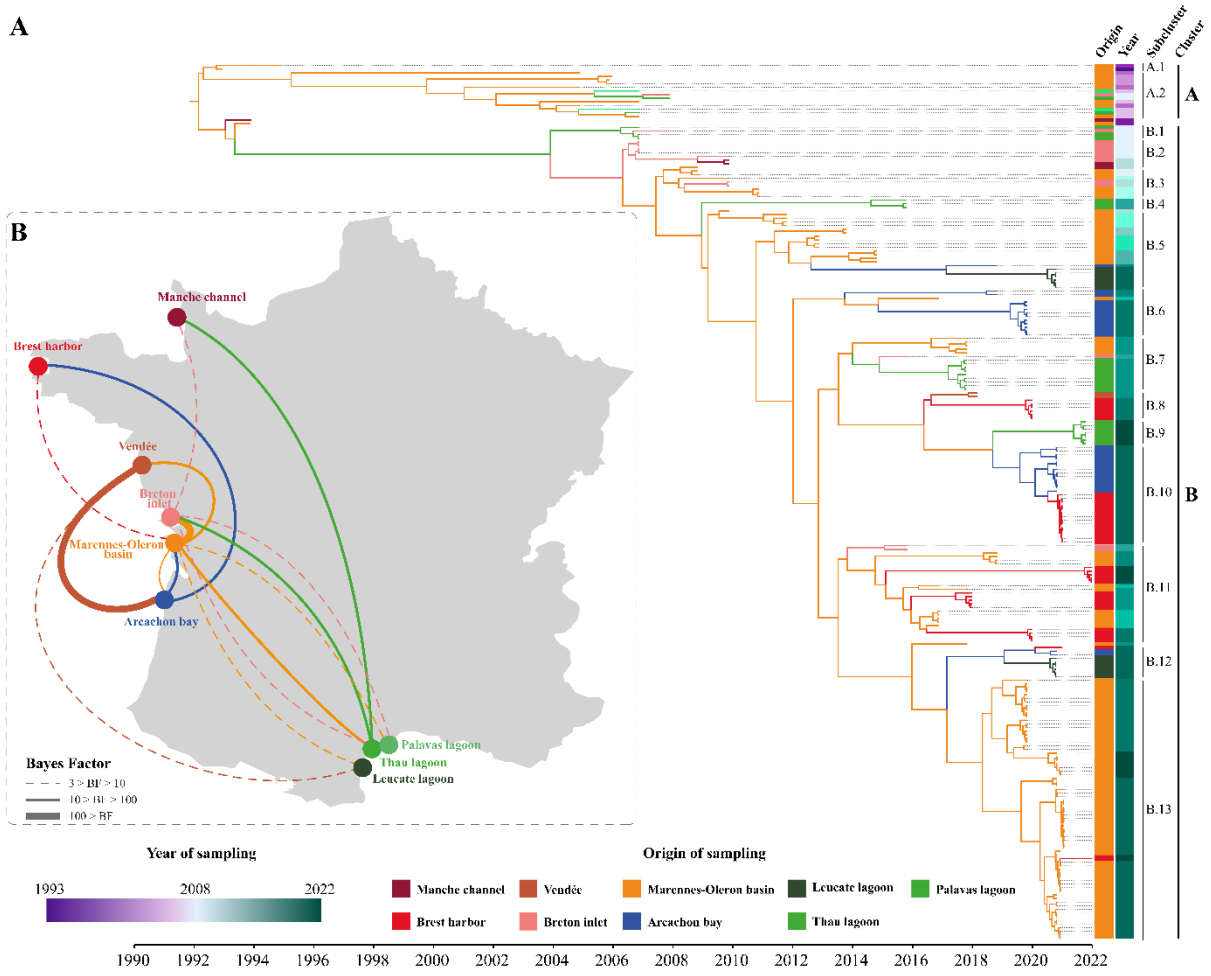


Figure 4: Ancestral state reconstruction of OsHV-1 specific origin between 1990 and 2022 in France. *A) Maximum clade credibility time-calibrated tree from the ancestral states' reconstruction generated in BEAST under a HKY+G+I model of nucleotide substitution, uncorrelated relaxed clock and GMRF demographical models. The tree is scaled in year with the final sampling year being 2022. Tip colors correspond to origin of sampling and branch colors correspond to most probable origin of the common ancestor referenced on the map. This figure was produced in FigTree v1.4.4 (Rambaut, 2019). On the right are the geographical origins associated with each sample and the sampling date as described in the legend panel. Two clusters and 15 subclusters were defined (numbered on the right). B) Bayesian stochastic search variable selection (BSSVS) was used to infer dispersal routes with none-zero rates using Bayes factors (BF). Routes that were inferred to have decisive support (BF > 100), strong support (BF > 10) and substantial support (BF > 3) are plotted on map. The thickness of the line represents the significance of the BF and the colour represents the origin of the transfer.*

Discussion

The study of OsHV-1 dispersal is crucial for mollusc health and the related economy. This virus seriously impacts Pacific oyster populations, making it a significant concern for global aquaculture. To better understand the consequence of OsHV-1 on its host, we need to decipher its epidemiology and phylogeography. This study provides the first comprehensive report on the global diversity of the OsHV-1 virus worldwide and complements the two previous studies on the diversity of the OsHV-1 virus that were conducted at the whole genome scale (Delmotte-Pelletier *et al.*, 2022; Morga-Jacquot *et al.*, 2021). The large sample size, global geographical coverage, and the time span covered by the samples have enabled the study of the spatio-temporal diversity of viral genotypes in various global production basins and to trace back a plausible evolutionary history of the virus worldwide from 1990 to 2022. Furthermore, the dense sampling in France has enabled the tracing of the virus's dispersal routes within French oyster production areas.

Evolution of OsHV-1 genomic architecture over time

Using a comprehensive whole-genome alignment strategy, the evolution of the OsHV-1 genomic architecture has been explored. The initial viral genotype identified in samples collected between 1990 and 2008 is the reference genotype OsHV-1. It is characterized by its 207 kb genome and a genomic architecture denoted as TR_L-U_L-IR_L-X-IR_S-U_S-TR_S. As previously documented certain variants exhibiting an inversion of approximately 4 kb within the U_L region (Davison *et al.*, 2005). This particular inversion was previously described in a minor part (25 %) of the population and is exclusive to genomes collected in France between 1993 and 2008. The ORFs affected by this inversion are located between ORF36 and ORF43. In these genomic variants, the ORF36 encoding a membrane protein is likely non-functional as it is disrupted. The other ORFs remain intact despite being inverted, so their functionality may only be affected by reversion. This type of rearrangement in the viral genome can lead to its inability to complete a full replication cycle in the absence of intact viruses that can complement the missing functions (Vignuzzi & López, 2019).

In France and more broadly in Europe, in 2008, this genotype primarily evolved through genetic material deletions. These variants are characterized by a genome of 204 kb and by a specific deletion of 12 base pairs within a microsatellite located upstream of ORF4 (protein coding for unknown function), as well as a deletion of over 600 base pairs between ORF 43 (protein coding for unknown function) and 44 (encoding a Deoxynucleoside kinase domain) (Burioli *et al.*, 2017; Segarra *et al.*, 2010). This genotype has been consequently named OsHV-1 µVar. In total, 10 kb of structural variations,

including insertions and deletions, have occurred in these genomes. The OsHV-1 µVar genotype is currently predominant in Europe and has replaced the reference genotype after 2008. Intriguingly, this genotype has been detected for the first time in France in 2008 and seem to have emerged rapidly with an increase of virulence (Burioli *et al.*, 2017). According to the phylogenetic analysis, four OsHV-1 samples obtained from the Thau Lagoon and the Breton Inlet exhibit a genetic admixture between ancestral populations corresponding to the OsHV-1 reference genotype and the OsHV-1 µVar genotype. Furthermore, when conducting recombination analysis on samples collected in France in 2008, we identified a combination of DNA segments from the OsHV-1 reference and OsHV-1 µVar genotypes within eight genomes.

Notably, two intergenic regions located around the stem-loop, one in the microsatellite upstream of ORF4, and four distinct genomic regions encoding unknown proteins, transmembrane proteins, and Zinc-finger, Ring-type proteins were identified as recombinant regions. The first hypothesis derived from these findings is that the OsHV-1 µVar genotype emerged through a recombination event between the OsHV-1 reference genotype and another genetic source, likely originated in one of these two basins. Since this genotype was detected in two widely separated production areas and outside any possible natural communications, it is plausible that the variant has emerge in one location and has been spread out in the other basin during oyster transfers (M. J. Keeling *et al.*, 2001; Lupo *et al.*, 2016; Morse, 2004). However, it is challenging to draw a definitive conclusion regarding this hypothesis due to the intricate nature of tracking the complex movements of oysters across the world. First, the limited number of samples collected in the year 2008 diminishes the robustness of the analysis. Second, the absence of precise sampling dates for these three samples adds a layer of uncertainty. Finally, it's worth noting that the sampling was exclusively conducted within these two basins, which might not provide a comprehensive representation of the broader context.

Meanwhile, in the USA (Tomales Bay), New Zealand, and Australia between 2006 and 2018, an intermediate genotype between the OsHV-1 reference and OsHV-1 µVar genotypes has been identified. These variants exhibit numerous rearrangements and deletions in the UL region similar to the µVar genotype, such as deletion within the ORF4 and ORF42/43. However, notably, these variants do not contain a full-length deletion in the microsatellite upstream of ORF 4, as seen in the reference genotype, and do not contain some deletion such as deletion of ORFs62/63. These observations are concordant with previous but partial description made using PCR approach (Jenkins *et al.*, 2013; S. Keeling *et al.*, 2014). Additionally, these genotypes are characterized by a translocation of the U_S region and a reverse translocation of the IRL region that have never been describe before. It is possible that these genotypes sampled in the Pacific Ocean between 2006 and 2011 represent an intermediate evolutionary genotype between the OsHV-1 reference genotype and the OsHV-1 µVar genotype

actually detected in Europe. Population admixture analyses confirm these findings. The results reveal genotypes that fall in between at least three ancestral populations. The first ancestral population consists of the reference OsHV-1 genotype, the second population is composed of the OsHV-1 µVar genotype, and finally, the third ancestral population corresponds to a genotype found in Korea. It is highly probable that a recombination event has occurred between the reference OsHV-1 genotype, OsHV-1 µVar, and the OsHV-1 genotype collected in Korea and led to a specific genotype detected in Australia, USA and New-Zealand.

Finally, another genomic architecture has been identified, corresponding to samples collected in Korea, the USA (San Diego), and New Zealand between 2011 and 2018. These genotypes are similar to the OsHV-1 µVar genotype; they contain the same deletions, but instead of a 12-base pair deletion in the microsatellite, there is only a 6-base pair deletion (Bai *et al.*, 2019; Burge *et al.*, 2021; Shimahara *et al.*, 2012). Microsatellites sometimes are more abundant in the untranslated regions or in regulatory regions, which can subsequently impact the expression of viral genes (Y.-C. Li *et al.*, 2004). In bacteria, it has already been demonstrated that microsatellite instability, meaning variations in the number of repetitions, can regulate the binding of transcriptional regulatory proteins and gene expression (Martin *et al.*, 2005). Certain microsatellites are associated with pathogenesis and virulence, and they can enhance antigenic diversity as a means to evade the host immune response (Y.-C. Li *et al.*, 2004). In the genome of osHV-1, the function of this microsatellite has not been elucidated, but it is located upstream of the ORF 4, which encodes a protein of unknown function. Consequently, drawing definitive conclusions about this microsatellite is challenging. In addition, the US region is translocated ahead of the IRS region. Our findings are consistent with those reported previously on the evolution of OsHV-1 and highlighted the prevalence of genomic deletions as a driving force (Burioli *et al.*, 2017; Morga *et al.*, 2021). Consistent with their previous observations, our dataset also exhibited deletions in certain regions previously described, including ORFs 35 to 37 (membrane protein and protein coding for unknown functions, Renault *et al.*, 2012), 48 (protein coding for unknown functions, Burioli *et al.*, 2017), and 62 to 63 (membrane protein and protein coding for unknown functions, Burioli *et al.*, 2017), and probably underestimated because of genomes non-redundancy. Interestingly, the ORF36 is deleted within genomes of all microvariants genotypes (Renault *et al.*, 2012). In certain reference OsHV-1 genomes, this ORF is disrupted by the occurrence of a genomic inversion, suggesting an initial evolution of this region in genomes containing the inversion of the UL region. Our understanding of the precise mechanisms responsible for the emergence of structural variants in viruses like OsHV-1 is currently incomplete, but these events could led to defective viral genomes (Vignuzzi & López, 2019). Nevertheless, it's plausible that common events like recombination or rearrangement, frequently observed in herpesviruses, play a pivotal role in giving rise to these genetic alterations. Recent

investigations employing sequencing techniques, including Sanger sequencing (Davison *et al.*, 2005) and more recently, Nanopore long-read sequencing (Dotto-Maurel *et al.*, 2022), have shed light on the existence of four major genomic isomers within the OsHV-1 genome. These isomers, regarded as structural variations, have been found to be present in consistent proportions. This phenomenon of isomeric genomes is not unique to OsHV-1 but is also seen in other herpesviruses, notably herpes simplex virus 1 (Mahiet *et al.*, 2012). It's closely linked to the replication and recombination mechanisms within these viruses, occurring early in their life cycles (Chaitanya, 2019; Chou & Roizman, 1985; Mahiet *et al.*, 2012). During DNA replication, the occurrence of numerous random double-strand breaks efficiently initiates recombination events, ultimately leading to the creation of new genomic rearrangements (Pérez-Losada *et al.*, 2015). What makes this even more intriguing is that these structural variants appear to not impede the virus's capacity for self-maintenance and replication. In fact, they might offer specific benefits. These genetic alterations hold the potential to promote the emergence of new viral lineages and enable genotypes to adapt to new host environments (Pérez-Losada *et al.*, 2015). Essentially, these variations may serve as a crucial factor in enhancing the virus's ability to adapt and persist in the face of changing environments.

Worldwide OsHV-1 population evolution history

Two previous studies have demonstrated the spatial structuration of OsHV-1 both in France (Delmotte *et al.*, 2022) and at a worldwide scale (Morga-Jacquot *et al.*, 2021), using small datasets of respectively 21 genomes and 22 haplotypes. In this study the time-scaled phylogeny reconstruction produced using 279 OsHV-1 genomes, collected worldwide between 1993 and 2022 confirms previous results about genomes clustering according to sampling location and to collection date. The MRCA of all these lineages was identified as the OsHV-1 reference genotype (Davison *et al.*, 2005) that has emerged around 1991 and subsequently evolved into *i)* the OsHV-1 µVar genotype (Burioli *et al.*, 2017) that has emerged in 2008 and is actually the only genotype detected in Europe and *ii)* into a specific OsHV-1 µVar lineage detected in the Pacific ocean (Burge *et al.*, 2021), which corresponds to a genome intermediate between the reference genotype and the French OsHV-1 µvar genotype, that has potentially emerged around 1995. These emergences may be linked to the introduction of *M. gigas* worldwide for its adaptability and ease of cultivation (Andrews, 1980), but they may also be linked to the intensification of Pacific oyster farming notably by exchanges of spats between hatcheries and between oyster production areas, because of high oysters' densities on farming areas or the use of particular genetic oyster background. Furthermore, pre-seeded oysters may have come into contact with the virus during their development without exhibiting any symptoms (Dégremont, 2013; Peeler

et al., 2012). Consequently, they can act as asymptomatic carriers of the OsHV-1 virus in a persistent phase. This can lead to the inadvertent spread of the virus during oysters' movements.

The mean evolutionary rate of OsHV-1 has been estimated to 3.6093E-03 subs/site/year. This value falls between the estimates from two prior studies (Morga-Jacquot *et al.*, 2021; Pelletier *et al.*, 2023). The first study estimated the virus evolution rate at 1.50E-02 subs/site/year, but it encompassed virus genomes collected from various host species, which can significantly influence the virus's evolution due to host-specific factors and regulatory mechanisms (Pelletier *et al.*, 2023). In the second study, the average evolutionary rate was 6.787E-05 subs/site/year (Morga-Jacquot *et al.*, 2021) but may be skewed by the haplotype reconstruction method, which does not facilitate the complete reconstruction of a genome and may overlook genetic material insertions. For various herpesviruses, including *Herpes Simplex Virus 1* (HSV1), *Varicella Zoster Virus* (VZV), and *Marek's Disease Virus* (MDV), the rates of evolution have been estimated at 8.21E-05 subs/site/year, 6.56E-06 subs/site/year, and 8.25E-06 subs/site/year, respectively (K. Li *et al.*, 2022; Sanjuán, 2012). These values are lower than the one estimated for OsHV-1 virus. These differences can be explained by the substantial but biased sample size in certain countries, as well as by oyster densities in the field or by the cohabitation of oysters with various cultured species, such as mussels, other oyster species, abalones, clams or cockles that potentially favored evolution through recombination between various genotypes infecting these host species (Casto *et al.*, 2020; Crespo-Bellido *et al.*, 2021).

Dispersal routes of OsHV-1 in France is linked to cultural practices

Reconstructing dispersal routes for viruses is of paramount importance in understanding and effectively manage infectious diseases. This effort provides insights into both the dynamics of disease transmission and the intricate process of virus evolution (Nahata *et al.*, 2021). To conduct this type of analysis, it is essential to have a substantial heterochronous dataset that encompasses relevant geographical areas (Lemey *et al.*, 2009). In the case of the OsHV-1 virus, it was crucial to explore the virus's dispersal routes in France. This dataset comprises 249 genomes collected from the seven largest French farming areas (including Manche Channel, Atlantic Ocean and Mediterranean Sea), spanning from the virus's initial detection in 1992 to the present day.

The use of ancestral state reconstruction to deduce patterns of OsHV-1 dispersal routes across nine distinct farming areas, indicates that the Marennes-Oléron farming region serves as the primary source of OsHV-1 viral diversity. From this location, OsHV-1 variants are disseminated to the remaining farming areas. These results are congruent with results obtained from a previous study conducted on

samples collected within Brest harbor, Marennes-Oléron basin and Thau lagoon exclusively in 2018 (Delmotte *et al.*, 2022), but our study adds robustness and provides additional insights into the involvement of different oyster farming areas. These results are in line with existing knowledge of oyster movement. (M. J. Keeling *et al.*, 2001; Lupo *et al.*, 2016; Morse, 2004). Specifically, Marennes-Oléron plays a pivotal role as the primary area for collecting *M. gigas* spat, and subsequently, young oysters are transferred to cultivation sites along the Brittany, Normandy, and the Mediterranean coasts (Buestel *et al.*, 2009). Notably, the nine examined sites are geographically distant and not directly connected through ocean currents. Consequently, OsHV-1 likely spreads through a helping mechanism such as the transmission dynamics observed in the movement of diseases during the transfer of livestock (M. J. Keeling *et al.*, 2001; Morse, 2004). Analysis of oyster movement unveiled that oysters are transferred up to nine times throughout their production cycle, particularly during spring and autumn, period during which OsHV-1-related mortalities appear (Lupo *et al.*, 2016). These aquaculture practices contribute greatly to the genetic diversity of OsHV-1 through its dissemination in the different farming areas (Lupo *et al.*, 2016). Moreover, it has already been demonstrated that the virus could be transmitted to healthy oysters by cohabitation with oysters that have survived to OsHV-1 infection, demonstrating the transmission impact of transferring healthy carrier oysters to uninfected oyster farms (Degremont and Benabdelmouna, 2014).

Results obtained from the time-scaled phylogeny reconstruction indicate that French OsHV-1 genotypes identified in this study trace their lineage back to a common ancestor shared with the well-known OsHV-1 reference genotype (Davison *et al.*, 2005). This lineage has subsequently evolved in the OsHV-1 µVar genotype (Burioli *et al.*, 2017), that has been detected for the first time in 2008 in France. This genotype has replaced the reference genotype in France and more widely in Europe as previously, reported in different studies (Burioli *et al.*, 2017, 2018; Delmotte *et al.*, 2022).

By integrating OsHV-1 architecture, viral population structure data through a combination of genetic and phylogeographic tools and robust epidemiological data derived from diverse worldwide farming regions, we can gain valuable insights into the dynamics of OsHV-1 infection. The present study establishes a crucial baseline understanding of OsHV-1 diversity, evolution and transmission, laying the groundwork for targeted management strategies aimed at mitigating the impact of this virus on oyster populations and the aquaculture industry.

Material and methods

Samples collection

To study the viral diversity of OsHV-1 worldwide across time, a total of 274 samples of *M. gigas* naturally infected in the wild were collected between 1993 and 2022 (Table S1, Figure 5). These samples were originated from different country around the world (Figure 1A): Ireland (n=1), Korea (n=2), New-Zealand (n=4), Germany (n=5), Australia (n=6), Spain (n=7), USA (n=11) and France (n=238). In France, samples were collected at nine different locations (eg Manche, Atlantic and mediterranean coast) (Figure 1B): Palavas lagoon (Hérault, n=2), in Bouin (Vendée, n=2), in the Manche channel (Normandie, n=3), in the Leucate lagoon (Aude, n=12), in the Breton inlet (Charente-Maritime and Vendée, n=12), in the Thau lagoon (n=24), in Arcachon basin (Gironde, n=28), in the Brest harbor (Finistère, n=38), and in the Marennes-Oléron bay (Charente-Maritime, n=113). Additionally, four samples have been sampled in France but the exact location is unknown.

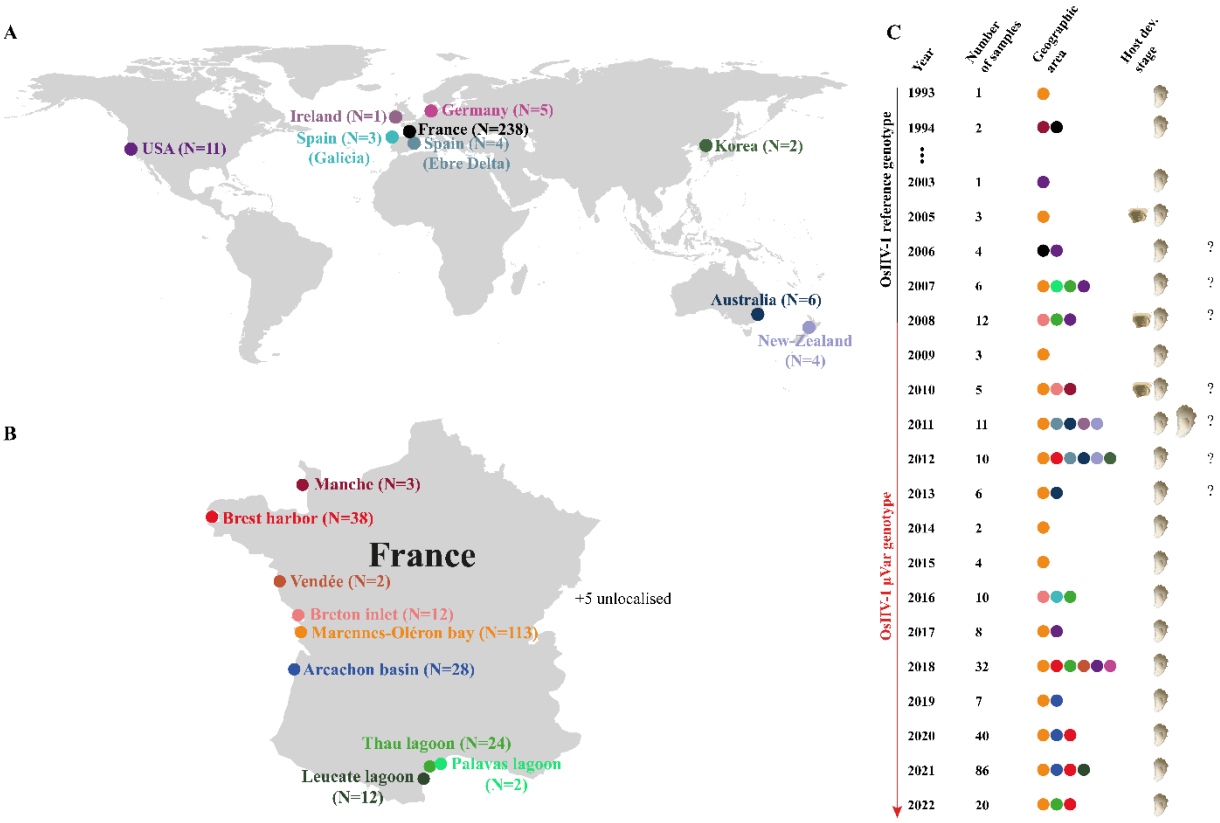


Figure 5: Samples origin and sampling metadata. *A) Overview of locations and number of samples collected around the world. B) Overview of locations and number of samples collected in France. C) Year of sampling and number of samples, geographic areas and host developmental stages associated.*

DNA extraction, viral copy numbers quantification and sequencing

DNA was extracted from mantle of individual moribund oysters using the MagAttract® HMW DNA kit or QIAamp DNA mini kit (Qiagen) according to manufacturer's protocol (Table S1). DNA purity and concentration were assessed by Nano-Drop ND-1000 spectrometer (Thermo Scientific) and Qubit® dsDNA BR assay kits (Molecular Probes Life Technologies), respectively. Then quantification of viral copies was carried out by quantitative PCR using a Mx3005 P Thermocycler (Agilent), as described in Pepin *et al.* (2008).

Illumina DNA-seq library were prepared and sequenced by six sequencing platforms (Table S1). One sample has been sequenced by the Genotoul (INRAE, Castanet-Tolosan, France), three samples have been sequenced by GATC biotech (Eurofins Scientific, Eberberg, Germany), 26 samples have been sequenced by LIGAN -PM Genomics (CNRS UMR 8199, Lille, France), 161 samples have been sequenced by the Genome Quebec Company (Genome Quebec Innovation Center, McGill University, Montreal, Canada), 24 samples have been sequenced by Fasteris Life Science Genesupport SA (Plan-les-Ouates, Switzerland) and 59 samples have been sequenced by IntegraGen (Evry, France).

De novo assembly of OsHV-1 genomes

The process of viral genome assembly plays a crucial role in studying viral diversity as it enables the reconstruction of complete genomes from samples. The *de novo* assembly of OsHV-1 genomes have been reconstructed using the framework previously developed (Delmotte-Pelletier *et al.* 2022 and Dotto-Maurel *et al.* 2022) with additional details on the bioinformatic pipeline provided by Dotto-Maurel *et al.* (2022).

To ensure the accuracy and uniformity of the final assemblies, quality-filtered reads were realigned to the newly constructed OsHV-1 NR-genomes using Bowtie 2 v2.4.1 (parameter --end-to-end). Raw data and reconstructed genomes have been deposited in the EBI database under accession numbers xx for future reference and accessibility (Table S1).

Whole genome alignment and recombination detection

Identification and description of OsHV-1 structural variations

At the genomic scale, variations between different genotypes of OsHV-1 can be attributed to structural variations (SV) such as insertions, deletions, inversions, translocations, and duplications. To analyze and compare the genomic architecture among viral genotypes and discern large-scale evolutionary

events, we aligned the 274 reconstructed NR-genomes with four published OsHV-1 genomes using the progressiveMauve algorithm implemented in Mauve v2.4.0 with default parameters (Darling *et al.*, 2004). These four genomes, namely OsHV-1 PT (2016, Italy: MG561751; Abbadi *et al.*, 2018), OsHV-1 µVarA (2010, France: KY242785; Burioli *et al.*, 2017), OsHV-1 µVarB (2011, Ireland: KY271630; Burioli *et al.*, 2017), and OsHV-1 Ref. (1992-1993, France: NC_005881.2; Davison *et al.*, 2005). They were selected because they were collected on the Pacific oyster.

From this alignment, we were able to identify distinct SV based on rearranged blocks and classify the genomes accordingly. Specifically, we observed five distinct SV in our study. To streamline the results, the focus is on these five specific structures.

To visualize these five types of SV, pairwise and multiple whole genomes alignments of the five representative genotypes were conducted against the OsHV-1 reference genome (NC_005881.2) using the progressiveMauve algorithm implemented in Mauve v2.4.0 with default parameters (Darling *et al.*, 2004).

Manual rearrangement of OsHV-1 genomes and multiple alignment

To perform the phylogenetic analysis, it was crucial to establish consistent genomic structures across all genomes. Since 34 genomes exhibited significant rearrangements in comparison to OsHV-1 reference genome (Davison *et al.*, 2005), manual modifications were performed on these assembled genomes. Using the Mauve multiple alignment in Geneious Prime, the genomic architecture of each genome was adjusted to ensure uniformity, thereby enabling meaningful comparisons among them. Subsequently, the genomes were aligned using MAFFT v1.4.0 (Katoh *et al.*, 2002).

OsHV-1 populations isolation

Genetic diversity within populations is influenced by subgroups and geographic isolation (Novembre, 2016). In order to assess the potential impact of geographic origin on the structure of the OsHV-1 population, a population structure analysis was conducted. First, Single Nucleotide Polymorphisms (SNP) data was obtained from the multiple alignment of all 279 NR-genomes previously generated and extracted using SNP-sites v2.5.1 (Page *et al.*, 2016). Next, the software STRUCTURE v 2.3.4 was employed to establish the population admixture and assigned each individual to ancestral populations (Pritchard *et al.*, 2000). To determine the optimal number of populations (K), independent runs of the Gibbs sampler were executed with K values ranging from 1 to 20. The Gibbs sampler utilized a Bayesian

approach applying Markov Chain Monte Carlo inference. Each run consisted of 50,000 iterations, with a burn-in period of 5,000 iterations. The STRUCTURE software assigned a posterior probability to each tested population structure, which was used to select the appropriate number of populations as described by Porras-Hurtado *et al.* (2013)(Table S3). According to this, K=9 was used to infer the population structure. To precisely determine the population structure, a last run with 100,000 iterations, and a burn-in period of 10,000 iterations was run for K=9. Finally, bar plots representing the estimated membership coefficients were generated with STRUCTURE using the output Q-matrix and groups of individuals were defined arbitrary based on their ancestral similarities.

Recombination events detection

The presence of recombinant regions along alignments can introduce biases in tree reconstruction due to the potentially distinct evolutionary histories of those regions compared to the rest of the genome (Lanier & Knowles, 2012). To address this issue, we employed Gubbins v3.3 (Croucher *et al.*, 2015) with default settings to detect regions in the assembled OsHV-1 NR-genomes that may have been affected by recombination events. Results obtained from Gubbins were further validated using linkage disequilibrium values (D') computed in DNAsp6 (Rozas *et al.*, 2017). The computation of D' was based on informative sites, and significance was assessed using a Fisher's exact test. To visualize the $D^{2'}$ values, a heatmap was generated in R using the ggplot package v3.4.0 (Wickham, 2009). Any genomic regions identified as potentially recombinant were subsequently excluded from the downstream analyses (Table S2).

Phylogeographic inferences

To reconstruct the worldwide ancestral origin of OsHV-1 and infer its dispersal routes in France, a discrete asymmetric phylogenetic diffusion model was used in BEAST v1.10.4 (Drummond & Rambaut, 2007). This involved modeling the host origin as a discrete trait for each NR-genome sequence over the phylogenetic tree and inferring ancestral states. To estimate the likelihood of internal nodes and branches being associated with a specific origin, ancestral state inference based on host information from the branch tips was employed. The best fitting model of evolution for the samples was found using MEGA v11.0.11 (Tamura *et al.*, 2021) based on Akaike Information Criterion (AIC). Our analysis used a HKY+G+I model of nucleotide substitution.

To identify the most suitable combination of coalescent (*i.e.* constant size, Gaussian Markov Random Field or GMRF) and molecular clock models (*i.e.* strict clock and uncorrelated relaxed clock) for our dataset, each possible combination has been tested. The maximum likelihood estimator (MLE) was calculated using path sampling (PS) and stepping-stone sampling (SSS) for each model. Based on these estimators, an uncorrelated relaxed molecular clock and a GMRF coalescent model was defined as most appropriate. A Markov Chain Monte Carlo (MCMC) with 10,000,000 steps, sampling every 1,000,000 generations using Beast v1.10.4 was ran. Additionally, a Markov jumps approach (Minin & Suchard, 2008a, 2008b) was used to determine the number of transitions between spatial locations. The BEAGLE library was used to increase computational speed (Ayres *et al.*, 2012).

To assess the mixing of the population (Ne > 200) and the stationary distribution of the trace, we examined the effective sample size in Tracer v1.7.2 (Rambaut *et al.*, 2018). This control was performed for all parameters, with an initial burn-in of 10%. Finally, we generated maximum clade credibility (MCC) trees using the tree output files from BEAST and annotated them in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

The same methodology has been applied to samples collected in France.

Bayes Factors (BF) were computed in SpreaD3 (Bielejec *et al.*, 2016) and used as a measure of support for identifying frequently invoked transition rates to explain the diffusion process. Support for a rate was considered substantial when BF > 3, strong if BF > 10, and decisive if >100 (Jeffreys, 1961).

Data availability

The datasets generated from this study can be found in the SRA database BioProject accession number xxx with submission ID x. Complementary information is available from the corresponding authors upon reasonable request.

Acknowledgement

This work was supported by the French Institute for Exploitation of the Sea (Ifremer), by DGAL (French General Directorate for Food) through the National Reference Laboratory for Mollusc Diseases and by the European-Union Reference Laboratory for Mollusc Diseases, Ifremer, La Tremblade. Additionally, it was supported by the EU project VIVALDI (H2020 program, no. 678589) led by Ifremer.

CP was supported by grant from the region Nouvelle-Aquitaine and from the scientific direction of Ifremer.

The study was conceptualized and designed by BM and CP. BM and CP were responsible for sample collection and conducting experimental procedures. CP, GC and MJ performed bioinformatics and statistical analyses. CP drafted the manuscript, with all authors contributing to its revision and approval.

The authors would like to express their gratitude to the staff at the Ifremer stations in Argenton (LPI, PFOM) and La Tremblade (PMMLT) for their technical support in producing standard (NSI) oysters and conducting transplantation experiments. Special thanks go to Nicole Faury (Ifremer, ASIM) for their technical assistance. Thank you to Jean-Michel Escoubas (IHPE, Ifremer, Montpellier), Bruno Petton (LEMAR, Ifremer, Argenton), and Johan Vieira (Capena, Arcachon) for their assistance during field monitoring activities. Thanks to all the worldwide collaborators: Colleen Burge (Coastal and Marine Sciences Institute, California), Mathias Wegner (AWI, Germany), Stein Mortensen (Institute of Marine Research, Norway), Yoshiko Shimahara (National Reasearch Institute of Aquaculture, Fisheries Research Agency, Japan).

References

Abbadi, M., Zamperin, G., Gastaldelli, M., Pascoli, F., Rosani, U., Milani, A., Schivo, A., Rossetti, E., Turolla, E., Gennari, L., Toffan, A., Arcangeli, G., & Venier, P. (2018). Identification of a newly described OsHV-1 µvar from the North Adriatic Sea (Italy). Journal of General Virology, 99(5), Article 5. https://doi.org/10.1099/jgv.0.001042

Andrews, J. D. (1980). A Review Of Introductions Of Exotic Oysters And Biological Planning For New Importations.

Arzul, I., Corbeil, S., Morga, B., & Renault, T. (2017). Viruses infecting marine molluscs. Journal of Invertebrate Pathology, 147, 118–135. https://doi.org/10.1016/j.jip.2017.01.009

Arzul, I., Nicolas, J.-L., Davison, A. J., & Renault, T. (2001). French Scallops: A New Host for Ostreid Herpesvirus-1. Virology, 290(2), Article 2. https://doi.org/10.1006/viro.2001.1186

Arzul, I., Renault, T., Lipart, C., & Davison, A. J. (2001). Evidence for interspecies transmission of oyster herpesvirus in marine bivalves. Journal of General Virology, 82(4), Article 4. https://doi.org/10.1099/0022-1317-82-4-865

Ayoade, F., & Kumar, S. (2023). Varicella-Zoster Virus (Chickenpox). In StatPearls. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK448191/

Bai, C.-M., Morga, B., Rosani, U., Shi, J., Li, C., Xin, L.-S., & Wang, C.-M. (2019). Long-range PCR and high-throughput sequencing of Ostreid herpesvirus 1 indicate high genetic diversity and complex evolution process. Virology, 526, 81–90. https://doi.org/10.1016/j.virol.2018.09.026

Barbosa Solomieu, V., Renault, T., & Travers, M.-A. (2015). Mass mortality in bivalves and the intricate case of the Pacific oyster, Crassostrea gigas. Journal of Invertebrate Pathology, 131, 2–10. https://doi.org/10.1016/j.jip.2015.07.011

Batista, F. M., López-Sanmartín, M., Grade, A., Morgado, I., Valente, M., Navas, J. I., Power, D. M., & Ruano, F. (2015). Sequence variation in ostreid herpesvirus 1 microvar isolates detected in dying and asymptomatic Crassostrea angulata adults in the Iberian Peninsula: Insights into viral origin and spread. Aquaculture, 435, 43–51. https://doi.org/10.1016/j.aquaculture.2014.09.016

Bloomquist, E. W., Lemey, P., & Suchard, M. A. (2010). Three roads diverged? Routes to phylogeographic inference. Trends in Ecology & Evolution, 25(11), 626–632. https://doi.org/10.1016/j.tree.2010.08.010

Boudry, P., Heurtebise, S., Collet, B., Cornette, F., & Gerard, A. (1998). Differentiation between populations of the Portuguese oyster, Crassostrea angulata (Lamark) and the Pacific oyster, Crassostrea gigas (Thunberg), revealed by mtDNA RFLP analysis (2). 226(2), Article 2.

Buestel, D., Ropert, M., Prou, J., & Goulletquer, P. (2009). History, Status, and Future of Oyster Culture in France. Journal of Shellfish Research, 28(4), Article 4. https://doi.org/10.2983/035.028.0410

Burge, C. A., Friedman, C. S., Kachmar, M. L., Humphrey, K. L., Moore, J. D., & Elston, R. A. (2021). The first detection of a novel OsHV-1 microvariant in San Diego, California, USA. Journal of Invertebrate Pathology, 184, 107636. https://doi.org/10.1016/j.jip.2021.107636

Burioli, E. A. V., Prearo, M., & Houssin, M. (2017). Complete genome sequence of Ostreid herpesvirus type 1 µVar isolated during mortality events in the Pacific oyster Crassostrea gigas in France and Ireland. Virology, 509, 239–251. https://doi.org/10.1016/j.virol.2017.06.027

Burioli, E. A. V., Varello, K., Lavazza, A., Bozzetta, E., Prearo, M., & Houssin, M. (2018). A novel divergent group of Ostreid herpesvirus 1 µVar variants associated with a mortality event in Pacific oyster spat in Normandy (France) in 2016. Journal of Fish Diseases, 41(11), 1759–1769. https://doi.org/10.1111/jfd.12883

Casto, A. M., Roychoudhury, P., Xie, H., Selke, S., Perchetti, G. A., Wofford, H., Huang, M.-L., Verjans, G. M. G. M., Gottlieb, G. S., Wald, A., Jerome, K. R., Koelle, D. M., Johnston, C., & Greninger, A. L. (2020). Large, Stable, Contemporary Interspecies Recombination Events in Circulating Human Herpes Simplex Viruses. The Journal of Infectious Diseases, 221(8), 1271–1279. https://doi.org/10.1093/infdis/jiz199

Chaitanya, K. V. (2019). Structure and Organization of Virus Genomes. Genome and Genomics, 1–30. https://doi.org/10.1007/978-981-15-0702-1_1

Chang, P. H., Kuo, S. T., Lai, S. H., Yang, H. S., Ting, Y. Y., Hsu, C. L., & Chen, H. C. (2005). Herpes-like virus infection causing mortality of cultured abalone Haliotis diversicolor supertexta in Taiwan. Diseases of Aquatic Organisms, 65(1), 23–27. https://doi.org/10.3354/dao065023

Chou, J., & Roizman, B. (1985). Isomerization of herpes simplex virus 1 genome: Identification of the cis-acting and recombination sites within the domain of the a sequence. Cell, 41(3), 803–811. https://doi.org/10.1016/S0092-8674(85)80061-1

Cochennec, N., Baud, J.-P., Pépin, J.-F., & et al. (2011). Les surmortalités des naissains d'huîtres creuses, Crassostrea gigas: Acquis des recherches en 2010. https://archimer.ifremer.fr/doc/00033/14423/11719.pdf

Comps, M. (1988). Epizootic diseases of oysters associated with viral infections. American Fisheries Society Special Publication, 18, Article 18.

Corbeil, S. (2020). Abalone Viral Ganglioneuritis. Pathogens, 9(9), 720. https://doi.org/10.3390/pathogens9090720

Crespo-Bellido, A., Hoyer, J. S., Dubey, D., Jeannot, R. B., & Duffy, S. (2021). Interspecies Recombination Has Driven the Macroevolution of Cassava Mosaic Begomoviruses. Journal of Virology, 95(17), e00541-21. https://doi.org/10.1128/JVI.00541-21

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., Parkhill, J., & Harris, S. R. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Research, 43(3), e15. https://doi.org/10.1093/nar/gku1196

Darling, A. C. E., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. Genome Research, 14(7), 1394–1403. https://doi.org/10.1101/gr.2289704

Davison, A. J., Trus, B. L., Cheng, N., Steven, A. C., Watson, M. S., Cunningham, C., Deuff, R.-M. L., & Renault, T. (2005). A novel class of herpesvirus with bivalve hosts. Journal of General Virology, 86(1), Article 1. https://doi.org/10.1099/vir.0.80382-0

Dégremont, L. (2013). Size and genotype affect resistance to mortality caused by OsHV-1 in Crassostrea gigas. Aquaculture, 416–417, 129–134. https://doi.org/10.1016/j.aquaculture.2013.09.011

Degremont, L., & Benabdelmouna, A. (2014). Mortality associated with OsHV-1 in spat Crassostrea gigas: Role of wild-caught spat in the horizontal transmission of the disease. Aquaculture International, 22(6), 1767–1781. https://doi.org/10.1007/s10499-014-9781-7

Delmotte, J., Pelletier, C., Morga, B., Galinier, R., Petton, B., Lamy, J.-B., Kaltz, O., Avarre, J.-C., Jacquot, M., Montagnani, C., & Escoubas, J.-M. (2022). Genetic diversity and connectivity of the Ostreid herpesvirus 1 populations in France: A first attempt to phylogeographic inference for a marine mollusc disease. Virus Evolution, 8(1), veac039. https://doi.org/10.1093/ve/veac039

Dotto-Maurel, A., Pelletier, C., Morga, B., Jacquot, M., Faury, N., Dégremont, L., Bereszczynki, M., Delmotte, J., Escoubas, J.-M., & Chevignon, G. (2022). Evaluation of tangential flow filtration coupled to long-read sequencing for ostreid herpesvirus type 1 genome assembly. Microbial Genomics, 8(11). https://doi.org/10.1099/mgen.0.000895

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7(1), 214. https://doi.org/10.1186/1471-2148-7-214

Fabioux, C., Huvet, A., Lapegue, S., Heurtebise, S., & Boudry, P. (2002). Past and present geographical distribution of populations of Portuguese (Crassostrea angulata) and Pacific (C. gigas) oysters along the European and north African Atlantic coats. Haliotis, 31, 33–44. https://archimer.ifremer.fr/doc/00000/2785/

Farley, C. A., Banfield, W. G., Kasnic, G., & Foster, W. S. (1972). Oyster Herpes-Type Virus. Science, 178(4062), Article 4062. https://doi.org/10.1126/science.178.4062.759

Fugl, A., & Andersen, C. L. (2019). Epstein-Barr virus and its association with disease—A review of relevance to general practice. BMC Family Practice, 20, 62. https://doi.org/10.1186/s12875-019-0954-3

Gimeno, I. M., Witter, R. L., & Reed, W. M. (1999). Four distinct neurologic syndromes in Marek's disease: Effect of viral strain and pathotype. Avian Diseases, 43(4), 721–737.

Grizel, H., & Heral, M. (1991). Introduction into France of the Japanese oyster (Crassostrea gigas). Journal Du Conseil - Conseil International Pour l'Exploration de La Mer, 47, Article 47. https://archimer.ifremer.fr/doc/00000/2760/

Hine, P., & Thorne, T. (1997). Replication of herpes-like viruses in haemocytes of adult flat oysters Ostrea angasi:an ultrastructural study. Diseases of Aquatic Organisms, 29, 189–196. https://doi.org/10.3354/dao029189

Houldcroft, C. J., Beale, M. A., & Breuer, J. (2017). Clinical and biological insights from viral genome sequencing. Nature Reviews Microbiology, 15(3), Article 3. https://doi.org/10.1038/nrmicro.2016.182

Hwang, J. Y., Park, J. J., Yu, H. J., Hur, Y. B., Arzul, I., Couraleau, Y., & Park, M. A. (2013). Ostreid herpesvirus 1 infection in farmed Pacific oyster larvae Crassostrea gigas (Thunberg) in Korea. Journal of Fish Diseases, n/a-n/a. https://doi.org/10.1111/jfd.12093

Jenkins, C., Hick, P., Gabor, M., Spiers, Z., Fell, S., Gu, X., Read, A., Go, J., Dove, M., O'Connor, W., Kirkland, P., & Frances, J. (2013). Identification and characterisation of an ostreid herpesvirus-1 microvariant (OsHV-1 µ-var) in Crassostrea gigas (Pacific oysters) in Australia. Diseases of Aquatic Organisms, 105(2), Article 2. https://doi.org/10.3354/dao02623

Kalkauskas, A., Perron, U., Sun, Y., Goldman, N., Baele, G., Guindon, S., & Maio, N. D. (2021). Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. PLOS Computational Biology, 17(1), e1008561. https://doi.org/10.1371/journal.pcbi.1008561

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research, 30(14), Article 14. https://doi.org/10.1093/nar/gkf436

Keeling, M. J., Woolhouse, M. E. J., Shaw, D. J., Matthews, L., Chase-Topping, M., Haydon, D. T., Cornell, S. J., Kappey, J., Wilesmith, J., & Grenfell, B. T. (2001). Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. Science, 294(5543), 813–817. https://doi.org/10.1126/science.1065973

Keeling, S., Brosnahan, C., Williams, R., Gias, E., Hannah, M., Bueno, R., McDonald, W., & Johnston, C. (2014). New Zealand juvenile oyster mortality associated with ostreid herpesvirus 1—An opportunistic longitudinal study. Diseases of Aquatic Organisms, 109(3), Article 3. https://doi.org/10.3354/dao02735

Lanier, H. C., & Knowles, L. L. (2012). Is Recombination a Problem for Species-Tree Analyses? Systematic Biology, 61(4), 691–701. https://doi.org/10.1093/sysbio/syr128

Lemey, P., Rambaut, A., Bedford, T., Faria, N. R., Bielejec, F., Baele, G., Russell, C. A., Smith, D. J., Pybus, O. G., Brockmann, D., & Suchard, M. A. (2012). The seasonal flight of influenza: A unified framework for spatiotemporal hypothesis testing (arXiv:1210.5877). arXiv. http://arxiv.org/abs/1210.5877

Lemey, P., Rambaut, A., Drummond, A. J., & Suchard, M. A. (2009). Bayesian Phylogeography Finds Its Roots. PLOS Computational Biology, 5(9), e1000520. https://doi.org/10.1371/journal.pcbi.1000520

Lemey, P., Rambaut, A., Welch, J. J., & Suchard, M. A. (2010). Phylogeography Takes a Relaxed Random Walk in Continuous Space and Time. Molecular Biology and Evolution, 27(8), 1877–1885. https://doi.org/10.1093/molbev/msq067

Li, K., Yu, Z., Lan, X., Wang, Y., Qi, X., Cui, H., Gao, L., Wang, X., Zhang, Y., Gao, Y., & Liu, C. (2022). Complete genome analysis reveals evolutionary history and temporal dynamics of Marek's disease virus. Frontiers in Microbiology, 13. https://www.frontiersin.org/articles/10.3389/fmicb.2022.1046832

Li, Y.-C., Korol, A. B., Fahima, T., & Nevo, E. (2004). Microsatellites Within Genes: Structure, Function, and Evolution. Molecular Biology and Evolution, 21(6), 991–1007. https://doi.org/10.1093/molbev/msh073

Lupo, C., Ezanno, P., Arzul, I., Garcia, C., Jadot, C., Joly, J.-P., Renault, T., & Bareille, N. (2016). How network analysis of oyster movements can improve surveillance and control programs of infectious diseases? Frontiers in Veterinary Science, 3. https://doi.org/10.3389/conf.FVETS.2016.02.00044

Lynch, S. A., Carlsson, J., Reilly, A. O., Cotter, E., & Culloty, S. C. (2012). A previously undescribed ostreid herpes virus 1 (OsHV-1) genotype detected in the pacific oyster, Crassostrea gigas, in Ireland. Parasitology, 139(12), Article 12. https://doi.org/10.1017/S0031182012000881

Mahiet, C., Ergani, A., Huot, N., Alende, N., Azough, A., Salvaire, F., Bensimon, A., Conseiller, E., Wain-Hobson, S., Labetoulle, M., & Barradeau, S. (2012). Structural Variability of the Herpes Simplex Virus 1 Genome In Vitro and In Vivo. Journal of Virology, 86(16), Article 16. https://doi.org/10.1128/JVI.00223-12

Martenot, C., Oden, E., Travaillé, E., Malas, J.-P., & Houssin, M. (2011). Detection of different variants of Ostreid Herpesvirus 1 in the Pacific oyster, Crassostrea gigas between 2008 and 2010. Virus Research, 160(1–2), Article 1–2. https://doi.org/10.1016/j.virusres.2011.04.012

Martin, P., Makepeace, K., Hill, S. A., Hood, D. W., & Moxon, E. R. (2005). Microsatellite instability regulates transcription factor binding and gene expression. Proceedings of the National Academy of Sciences, 102(10), 3800–3804. https://doi.org/10.1073/pnas.0406805102

Matsui, K., Honjo, M., Kohmatsu, Y., Uchii, K., Yonekura, R., & Kawabata, Z. (2008). Detection and significance of koi herpesvirus (KHV) in freshwater environments. Freshwater Biology, 53(6), 1262–1272. https://doi.org/10.1111/j.1365-2427.2007.01874.x

Michel, B., Fournier, G., Lieffrig, F., Costes, B., & Vanderplasschen, A. (2010). Cyprinid Herpesvirus 3. Emerging Infectious Diseases, 16(12), 1835–1843. https://doi.org/10.3201/eid1612.100593

Minin, V. N., & Suchard, M. A. (2008a). Counting labeled transitions in continuous-time Markov models of evolution. Journal of Mathematical Biology, 56(3), 391–412. https://doi.org/10.1007/s00285-007-0120-8

Minin, V. N., & Suchard, M. A. (2008b). Fast, accurate and simulation-free stochastic mapping. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1512), 3985–3995. https://doi.org/10.1098/rstb.2008.0176

Morga, B., Jacquot, M., Pelletier, C., Chevignon, G., Dégremont, L., Biétry, A., Pepin, J.-F., Heurtebise, S., Escoubas, J.-M., Bean, T. P., Rosani, U., Bai, C.-M., Renault, T., & Lamy, J.-B. (2021). Genomic Diversity of the Ostreid Herpesvirus Type 1 Across Time and Location and Among Host Species. Frontiers in Microbiology, 12, 711377. https://doi.org/10.3389/fmicb.2021.711377

Morse, S. S. (2004). Factors and determinants of disease emergence. 23(2), 443. https://doi.org/10.20506/rst.23.2.1494

Moss, J., Burreson, E., Cordes, J., Dungan, C., Brown, G., Wang, A., Wu, X., & Reece, K. (2007). Pathogens in Crassostrea ariakensis and other Asian oyster species: Implications for non-native oyster introduction to Chesapeake Bay. Diseases of Aquatic Organisms, 77, 207–223. https://doi.org/10.3354/dao01829

Nahata, K. D., Bollen, N., Gill, M. S., Layan, M., Bourhy, H., Dellicour, S., & Baele, G. (2021). On the Use of Phylogeographic Inference to Infer the Dispersal History of Rabies Virus: A Review Study. Viruses, 13(8), Article 8. https://doi.org/10.3390/v13081628

Nell, J. A. (2001). The History of Oyster Farming in Australia. Marine Fisheries Review.

Nicolas, J.-L., Comps, M., & Cochennec, N. (1992). Herpes-like virus infecting Pacific oyster larvae. (1). 12(1), Article 1.

Novembre, J. (2016). Pritchard, Stephens, and Donnelly on Population Structure. Genetics, 204(2), 391–393. https://doi.org/10.1534/genetics.116.195164

Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., & Harris, S. R. (2016). SNP-sites: Rapid efficient extraction of SNPs from multi-FASTA alignments. Microbial Genomics, 2(4). https://doi.org/10.1099/mgen.0.000056

Paul-Pont, I., Evans, O., Dhand, N. K., Rubio, A., Coad, P., & Whittington, R. J. (2014). Descriptive epidemiology of mass mortality due to Ostreid herpesvirus-1 (OsHV-1) in commercially farmed Pacific oysters (Crassostrea gigas) in the Hawkesbury River estuary, Australia. Aquaculture, 422–423, 146–159. https://doi.org/10.1016/j.aquaculture.2013.12.009

Peeler, E. J., Allan Reese, R., Cheslett, D. L., Geoghegan, F., Power, A., & Thrush, M. A. (2012). Investigation of mortality in Pacific oysters associated with Ostreid herpesvirus-1µVar in the Republic of Ireland in 2009. Preventive Veterinary Medicine, 105(1–2), Article 1–2. https://doi.org/10.1016/j.prevetmed.2012.02.001

Pelletier, C., Chevignon, G., Faury, N., Arzul, I., Garcia, C., Chollet, B., Renault, T., Morga, B., & Jacquot, M. (2023). Genetic differentiation and host specialization among OsHV-1 infecting two oyster species in France [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2023.08.23.554398

Pepin, J. F., Riou, A., & Renault, T. (2008). Rapid and sensitive detection of ostreid herpesvirus 1 in oyster samples by real-time PCR. Journal of Virological Methods, 149(2), Article 2. https://doi.org/10.1016/j.jviromet.2008.01.022

Pérez-Losada, M., Arenas, M., Galán, J. C., Palero, F., & González-Candelas, F. (2015). Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. Infection, Genetics and Evolution, 30, 296–307. https://doi.org/10.1016/j.meegid.2014.12.022

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., & Lareu, M. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. Frontiers in Genetics, 4. https://www.frontiersin.org/articles/10.3389/fgene.2013.00098

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. Genetics, 155(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Pybus, O. G., Suchard, M. A., Lemey, P., Bernardin, F. J., Rambaut, A., Crawford, F. W., Gray, R. R., Arinaminpathy, N., Stramer, S. L., Busch, M. P., & Delwart, E. L. (2012). Unifying the spatial epidemiology and molecular evolution of emerging epidemics. Proceedings of the National Academy of Sciences, 109(37), 15066–15071. https://doi.org/10.1073/pnas.1206598109

Quayle, D. B. (1988). Quayle, D. B. (1988). Pacific oyster culture in BC Canadian Bull. Fisheries Aquat. Sci, 218, 241.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Systematic Biology, 67(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Ren, W., Chen, H., Renault, T., Cai, Y., Bai, C., Wang, C., & Huang, J. (2013). Complete genome sequence of acute viral necrosis virus associated with massive mortality outbreaks in the Chinese scallop, Chlamys farreri. Virology Journal, 10(1), 110. https://doi.org/10.1186/1743-422X-10-110

Renault & Arzul. (2001). Herpes-like virus infections in hatchery-reared bivalve larvae in Europe: Specific viral DNA detection by PCR. Journal of Fish Diseases, 24(3), Article 3. https://doi.org/10.1046/j.1365-2761.2001.00282.x

Renault, T. (1998). Infection herpétiques chez les invertébrés: Détection de virus de type herpès chez les mollusques bivalves marins. Annales de l'Institut Pasteur., 2, Article 2. https://archimer.ifremer.fr/doc/00000/2903/

Renault, T., & Académie vétérinaire de France, P.-V., FRA (meeting organiser). (2011). LES VIRUS INFECTANT LES MOLLUSQUES MARINS: UN EXEMPLE D'ACTUALITÉ, LES HERPÈSVIRUS. Bulletin de l'Académie vétérinaire de France, 1, Article 1. https://doi.org/10.4267/2042/48108

Renault, T., Cochennec, R., Le Deuff, R. M., & Chollet, B. (1994). HERPES-LIKE VIRUS INFECTING JAPANESE OYSTER (CRASSOSTREA GIGAS) SPAT. 3.

Renault, T., Faury, N., Barbosa-Solomieu, V., & Moreau, K. (2011). Suppression substractive hybridisation (SSH) and real time PCR reveal differential gene expression in the Pacific cupped oyster, Crassostrea gigas, challenged with Ostreid herpesvirus 1. Developmental & Comparative Immunology, 35(7), Article 7. https://doi.org/10.1016/j.dci.2011.02.004

Renault, T., Lipart, C., & Arzul, I. (2001). A herpes-like virus infects a non-ostreid bivalve species: Virus replication in Ruditapes philippinarum larvae. Diseases of Aquatic Organisms, 45, 1–7. https://doi.org/10.3354/dao045001

Renault, T., Moreau, P., Faury, N., Pepin, J.-F., Segarra, A., & Webb, S. (2012). Analysis of Clinical Ostreid Herpesvirus 1 (Malacoherpesviridae) Specimens by Sequencing Amplified Fragments from Three Virus Genome Areas. Journal of Virology, 86(10), 5942–5947. https://doi.org/10.1128/JVI.06534-11

Renault, T., Tchaleu, G., Faury, N., Moreau, P., Segarra, A., Barbosa-Solomieu, V., & Lapègue, S. (2014). Genotyping of a microsatellite locus to differentiate clinical Ostreid herpesvirus 1 specimens. Veterinary Research, 45(1), Article 1. https://doi.org/10.1186/1297-9716-45-3

Roque, A., Carrasco, N., Andree, K. B., Lacuesta, B., Elandaloussi, L., Gairin, I., Rodgers, C. J., & Furones, M. D. (2012). First report of OsHV-1 microvar in Pacific oyster (Crassostrea gigas) cultured in Spain. Aquaculture, 324–325, 303–306. https://doi.org/10.1016/j.aquaculture.2011.10.018

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and Evolution, 34(12), 3299–3302. https://doi.org/10.1093/molbev/msx248

Sanjuán, R. (2012). From Molecular Genetics to Phylodynamics: Evolutionary Relevance of Mutation Rates Across Viruses. PLOS Pathogens, 8(5), e1002685. https://doi.org/10.1371/journal.ppat.1002685

Segarra, A., Pépin, J. F., Arzul, I., Morga, B., Faury, N., & Renault, T. (2010). Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, Crassostrea gigas, in France in 2008. Virus Research, 153(1), Article 1. https://doi.org/10.1016/j.virusres.2010.07.011

Shimahara, Y., Kurita, J., Kiryu, I., Nishioka, T., Yuasa, K., Kawana, M., Kamaishi, T., & Oseko, N. (2012). Surveillance of Type 1 Ostreid Herpesvirus (OsHV-1) Variants in Japan. Fish Pathology, 47(4), Article 4. https://doi.org/10.3147/jsfp.47.129

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution, 38(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Vignuzzi, M., & López, C. B. (2019). Defective viral genomes are key drivers of the virus–host interaction. Nature Microbiology, 4(7), Article 7. https://doi.org/10.1038/s41564-019-0465-y

Webb, S. C., Fidler, A., & Renault, T. (2007). Primers for PCR-based detection of ostreid herpes virus-1 (OsHV-1): Application in a survey of New Zealand molluscs. Aquaculture, 272(1–4), Article 1–4. https://doi.org/10.1016/j.aquaculture.2007.07.224

Whitley, R., Kimberlin, D. W., & Prober, C. G. (2007). Pathogenesis and disease. In A. Arvin, G. Campadelli-Fiume, E. Mocarski, P. S. Moore, B. Roizman, R. Whitley, & K. Yamanishi (Eds.), Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge University Press. http://www.ncbi.nlm.nih.gov/books/NBK47449/

Wörner, S., Dragani, W. C., Echevarria, E. R., Carrasco, M., & Barón, P. J. (2019). An Estimation of the Possible Migration Path of the Pacific Oyster (Crassostrea gigas) Along the Northern Coast of Patagonia. Estuaries and Coasts, 42(3), 806–821. https://doi.org/10.1007/s12237-018-00492-z

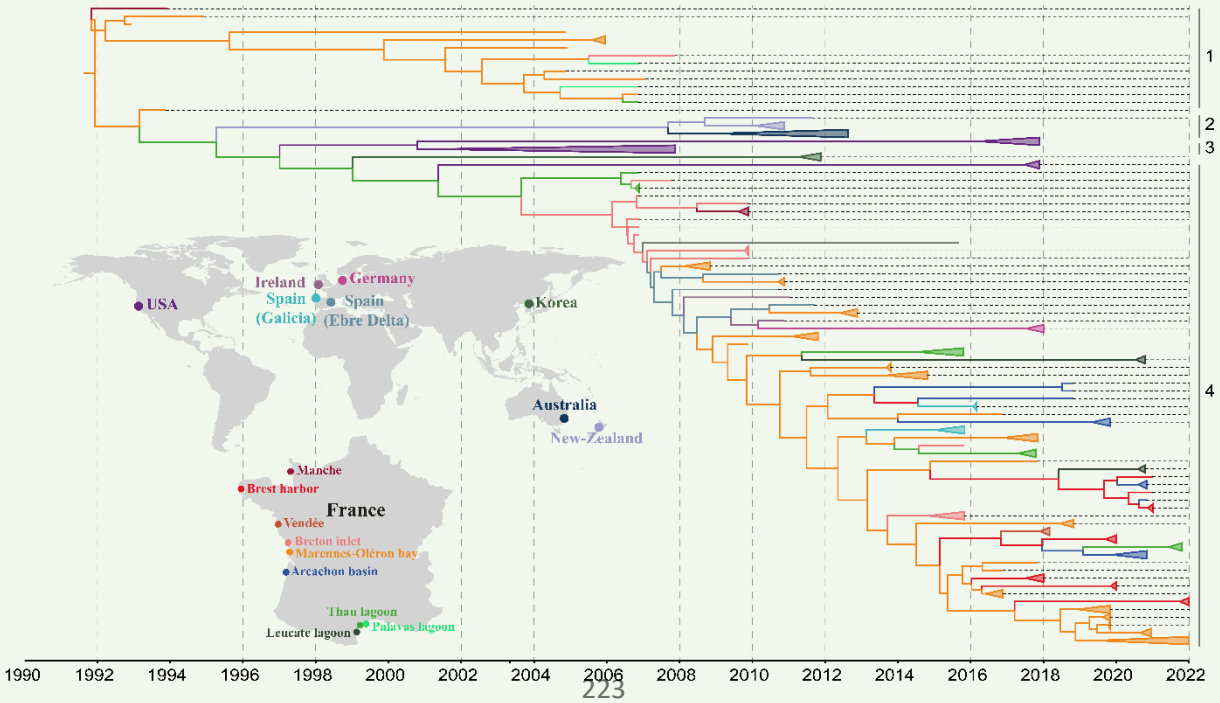
Xia, J., Bai, C., Wang, C., Song, X., & Huang, J. (2015). Complete genome sequence of Ostreid herpesvirus-1 associated with mortalities of Scapharca broughtonii broodstocks. Virology Journal, 12(1), Article 1. https://doi.org/10.1186/s12985-015-0334-0

Messages clés :

- OsHV-1 a évolué temporellement par l'intermédiaire **de délétions et de réarrangements** génomiques qui ont contribué à cette diversification. De plus, des génotypes intermédiaires ont été identifiés dans le Pacifique, suggérant une évolution continue du virus.

- Ces études ont permis de reconstituer **l'histoire évolutive** du virus OsHV-1 à l'échelle mondiale. Le virus a été retracé jusqu'à un ancêtre commun, le génotype de référence, apparu vers 1991. Ce génotype ancestral a donné naissance à plusieurs lignées virales, dont l'OsHV-1 µVar.

- L'étude montre que la région de **Marennes-Oléron en France joue un rôle central en tant que source principale de diversité du virus OsHV-1**, avec une dispersion ultérieure vers d'autres zones de culture d'huîtres en France. **Les pratiques aquacoles, telles que le transfert d'huîtres juvéniles, sont probablement responsables de la dispersion du virus entre les bassins de culture en France**.



Chapitre 5 – Identification des mécanismes évolutifs du virus OsHV-1 par une approche expérimentale

C. Pelletier

5.1 Résumé de l'article 6

Les deux précédents chapitres (chapitre 3 et 4) se sont concentrés principalement sur l'étude des processus épidémiologiques qui influencent la diversité et la dispersion du virus OsHV-1 dans le monde. Malgré d'importantes avancées dans la compréhension des processus épidémiologiques et évolutifs de l'OsHV-1, des lacunes subsistent quant à ses mécanismes évolutifs et à sa capacité à s'adapter aux caractéristiques génétiques de son hôte. Bien que traditionnellement les virus à ADN double brin ayant un grand génome sont connus pour avoir des taux de mutation plus faibles que les virus à ARN, le virus OsHV-1 montre une diversité génétique notable et quantifiable. Le virus évolue grâce à divers processus, notamment la mutation et la recombinaison génétique. Deux autres processus, la dérive génétique et la sélection naturelle, peuvent influencer l'évolution d'un virus. Ces deux processus n'ont à l'heure actuelle pas pu être mis en cause dans l'évolution du virus OsHV-1. Notre compréhension des processus évolutifs contribuant à la diversité des populations du virus OsHV-1 reste donc limitée.

Dans ce contexte, cette étude vise à approfondir les connaissances sur les mécanismes évolutifs qui conduisent à l'adaptation et à la diversification ou à la sélection des génotypes d'OsHV-1 en fonction de l'origine des hôtes face à l'infection virale. Pour ce faire, nous avons utilisé une approche d'évolution expérimentale du virus en conditions contrôlées chez trois familles d'huîtres du Pacifique ayant des niveaux de résistance contrastés face à l'infection virale. Au total, 90 échantillons ont été séquencés et ont permis d'étudier les fréquences alléliques des différents génotypes au cours des générations d'infection. Les fréquences alléliques des différents polymorphismes et leur fixation dans les populations ont pu être suivi par une méthode statistiques appliquées au traitement de données en séries temporelles.

Au cours de cette évolution expérimentale, une réduction de la diversité génétique virale a été observée, probablement due à des goulots d'étranglement de population. L'analyse des fréquences des SNP a révélé des clusters distincts correspondant aux différentes familles d'huîtres et aux générations d'infection, indiquant une évolution des génotypes au cours des générations d'infection. Des signatures de sélection positives ont été détectées exclusivement dans le génome d'OsHV-1 lorsqu'il était propagé dans la famille d'huître ayant le niveau de résistance le plus faible, suggérant une adaptation du virus aux hôtes plus sensibles. Ces signatures de sélection se trouvent dans diverses régions du génome viral, y compris les régions répétées et la région U_L, impliquant une adaptation de protéines essentielles du virus.

Article 6

Simulation of 30 years of *Magallana gigas* infection with the virus OsHV-1 revealed evolution through positive selection

Camille Pelletier, Nicole Faury, Mickaël Mege, Lionel Dégremont, Germain Chevignon, Maude Jacquot, Benjamin Morga.

Simulation of 30 years of *Magallana gigas* infection with the virus OsHV-1 revealed evolution through positive selection

Camille Pelletier[1], Nicole Faury[1], Mickaël Mege[1,#], Lionel Dégremont[1], Germain Chevignon[1], Maude Jacquot[1], Benjamin Morga[1].

Affiliations

[1] Ifremer, RBE-ASIM, Unité Adaptations Santé des Invertébrés Marins, Avenue de Mus de Loup, 17390 La Tremblade, France.

Current address: Oniris, INRAE, UMR BIOEPAR, site de la Chantrerie, Route de Gachet - CS40706, 44307 Nantes cedex 3

Corresponding authors:

Benjamin Morga

E-mail address: Benjamin.morga@ifremer.fr,

Tel +33 5 46 76 26 49

Fax +33 5 46 76 26 11

Maude Jacquot

E-mail address: maude.jacquot@ifremer.fr,

Tel +33 5 46 76 26 82

Prepared and formatted for submission to *PCI evolutionary biology*.

Abstract

Ostreid herpesvirus 1 (OsHV-1) poses a significant threat to the global oyster farming industry, causing substantial economic losses due to mortality outbreaks. While OsHV-1 primarily affects the Pacific oyster *Magallana gigas*, it has been linked to mortality events in various host species. Despite advancements in understanding OsHV-1 epidemiology, knowledge gaps persist regarding its evolutionary mechanisms and adaptation to host genetic backgrounds. This study employs experimental evolution and extensive genomic analysis to unravel the dynamics of OsHV-1 evolution in response to oyster host genetic variation. Genetic mutations, particularly transitions and transversions, played a significant role in shaping the viral population, leading to a trend toward genetic homogenization. Positive selection signals were observed in the oyster line with higher susceptibility, suggesting adaptation of viral genotypes to specific host genetic backgrounds. These findings shed light on the complex evolutionary dynamics of OsHV-1 and its interactions with oyster hosts. Understanding how this virus adapts to host genetic diversity is crucial for developing strategies to mitigate its impact on the oyster farming industry and provides valuable insights into the broader mechanisms of viral evolution in response to host variation.

Introduction

Ostreid herpesvirus 1 (OsHV-1), a double-stranded DNA virus belonging to the Malacoherpesviridae oyster line, poses a significant threat to the global oyster farming industry. In recent years, mortality outbreaks associated with OsHV-1 have caused substantial economic losses. This virus primarily affects the Pacific oyster *Magallana gigas* (previously called *Crassostrea gigas*), making it vulnerable from the larval to the adult stages in naive oysters (Azéma, Lamy, *et al.*, 2017; Burioli *et al.*, 2017; Nicolas *et al.*, 1992; Renault *et al.*, 1994). Meanwhile, oyster farmers report mortality associated with OsHV-1 mainly in spat, as survivors demonstrate genetic resistance upon subsequent exposure to the virus (Evans *et al.*, 2017). However, OsHV-1 has been linked to mortality events in various host species, including scallops such as *Pecten maximus* (Arzul, Nicolas, *et al.*, 2001), *Azumapecten farreri* (Bai *et al.*, 2019), clams like *Anadara broughtonii* (Xia *et al.*, 2015), *Ruditapes philippinarum* (Renault & Arzul, 2001; Renault, 1998, p. 199) and *Ruditapes decussatus* (Renault & Arzul, 2001; Renault, 1998, p. 199), as well as oysters such as *Crassostrea angulata* (Arzul, Renault, *et al.*, 2001), *Crassostrea virginica* (Farley *et al.*, 1972), and *Ostrea edulis* (Arzul, Renault, *et al.*, 2001; Comps, 1988; Renault *et al.*, 2001).

The first comprehensive characterization of the OsHV-1 genome was published in 2005 (Davison *et al.*, 2005). The genome of OsHV-1 spans approximately 207 kilobase pairs (kb) in length and exhibits a distinctive genomic organization comprising unique regions (U_L/U_S: unique long or short) flanked by repeated regions (TR_L/TR_S: terminal repeat long or short and IR_L/IR_S: inverted repeat long or short). The architecture of the OsHV-1 genome can be summarized as TR_L-U_L-IR_L-X-IR_S-U_S-TR_S (Davison *et al.*, 2005). Advances in high-throughput sequencing have facilitated the genomic characterization of several OsHV-1 genotypes, including OsHV-1 µVar (Burioli *et al.*, 2017), OsHV-1 PT (Abbadi *et al.*, 2018), and OsHV-1 SB (Xia *et al.*, 2015). These genomic datasets have revealed a spatiotemporal structuring of viral genotypes in France (Delmotte-Pelletier *et al.*, 2022) and globally (Morga-Jacquot *et al.*, 2021), consistent with oyster farming practices. Furthermore, they have highlighted differentiation between genotypes infecting *M. gigas* and *O. edulis*, with a greater viral diversity observed in *M. gigas* (Pelletier *et al.*, 2023). However, despite significant advancements in understanding OsHV-1 epidemiological and evolutionary processes, several knowledge gaps persist regarding its evolutionary mechanisms and how it adapts to host genetic backgrounds. Indeed, substantial genetic variation for resistance to OsHV-1 has been reported worldwide allowing selective breeding program to sustain oyster production (Camara *et al.*, 2017; Dégremont *et al.*, 2015b; Divilov *et al.*, 2019; Gutierrez *et al.*, 2020).

The evolution of viruses is driven by a range of processes, including mutation, genetic recombination, genetic drift and natural selection. RNA viruses provide vivid examples of high mutation rates and rapid evolution. These viruses are known for their ability to accumulate genetic changes swiftly, resulting in

substantial levels of viral polymorphism (Sanjuán, 2012; Sanjuán & Domingo-Calap, 2016, 2016). This genetic diversity empowers them to adapt to newly infected cellular environments and develop strategies to evade vaccines and antiviral drugs (Lauring, 2020). Although large dsDNA viruses exhibit lower mutation rates (*i.e.* between 10E-03 and 10E-07 nucleotide substitution/site/year) compared to RNA viruses (*i.e.* between 10E-02 and 10E-04 nucleotide substitution/site/year) due to their use of high-fidelity proofreading polymerases (Sanjuán *et al.*, 2010; Sanjuán, 2012; Sanjuán & Domingo-Calap, 2016), their genome stability is increasingly controversial because they sometimes display noteworthy genetic diversity (Sanjuán *et al.*, 2016).

Viral evolutionary processes can be investigated through various approaches including sequence analyses, metagenomics, phylogenetics, phylodynamics, and even experimental evolution (EE). EE involves subjecting viruses to controlled laboratory conditions, enabling the real-time observation and manipulation of evolutionary processes (Kawecki *et al.*, 2012). The emergence and dynamics of mutations, immune escape variants, and evolutionary trade-offs could be assessed by monitoring viral populations under different selective pressures, such as antiviral drug treatments or interactions with immune systems (Kawecki *et al.*, 2012).

While most studies using EE focus on easily replicable model organisms such as Escherichia coli, Drosophila spp, various bacteriophages, or Saccharomyces cerevisiae (Kawecki *et al.*, 2012), recent research has expanded EE to diverse viruses, including Herpesviruses (Fuandila *et al.*, 2022; Kuny *et al.*, 2020). For instance, in vitro EE experiments conducted with populations of Herpes Simplex Virus type 1 (HSV-1) have revealed the accumulation of minor genetic variants and changes in population diversity. These changes have been observed to lead to relatively rapid modifications, at the consensus level, of the viral population. (Kuny *et al.*, 2020). Another EE study, conducted on Koi herpesvirus (KHV), which infects the common carp *Cyprinus carpio*, has highlighted the key role of structural variations in the virus evolution in vitro (Fuandila *et al.*, 2022). Lastly, a recent study investigating the in vivo evolution of the Dengue virus (DENV) in two mosquito species has demonstrated that positive selection driven by the mosquito species acting as vectors enables the virus to adapt to different mosquito species (Bellone *et al.*, 2020).

Due to the unavailability of oyster cell cultures for replicating the OsHV-1 virus, we are compelled to resort to in vivo techniques. In the context of OsHV-1, our understanding of the evolutionary processes contributing to the diversity of the viral population remains limited. The primary objective of the present study is to gain insights into the evolutionary mechanisms driving the adaptation and diversification of the OsHV-1 genotypes based on oyster infection susceptibility. To achieve this goal,

we employed an experimental evolution approach spanning from 13 to 28 generations under controlled condition, combined with an extensive sequencing analysis of ninety OsHV-1 genomes.

Material and methods

Oysters production

To explore the impact of oyster genetic backgrounds on the evolutionary dynamics of Ostreid herpesvirus-1 (OsHV-1), three oyster populations with varying levels of resistance against OsHV-1 infection were generated. In December 2021, wild oysters were collected from two sites in Charente-Maritime (France) based on their proximity to oyster farms.

The first site, La Floride (Lat: 45.803; Long.: -1.153) is located in an area surrounded by oyster farms and is regularly exposed to OsHV-1 infections (Dégremont *et al.*, 2019). Due to the strong genetic basis for resistance to OsHV-1 in this population, these oysters are potentially more resistant to the viral infection (de Lorgeril *et al.*, 2020).

The second site, Chaucre (Lat: 45.982; Long: -1.396), is situated on the West coast of Oléron island in a non-farming area. Oysters collected from Chaucre are less likely to be affected by OsHV-1 infections compared to oysters from farming area, and thus, they may be more susceptible to OsHV-1 infection.

In addition to the wild oyster populations, an oyster line developed at Ifremer, was included in our study. This oyster line is highly susceptible to OsHV-1, with mortality exceeding 80%, and has been described as the control of a mass selection experiments (Dégremont *et al.*, 2020). These oysters were initially sampled in 2008 from a farming area, and since 2009, they have been reproduced over six successive generations using oysters from the second generation that have never been exposed to OsHV-1. Therefore, it can be assumed that this oyster line is naive to OsHV-1 since 2009, as animals have always been kept in a secured facility using UV-treated seawater (40 mj/cm2).

The three populations (La Floride, Chaucre, and the naive line) were placed in separate tanks in the conditioning room at the Ifremer hatchery in La Tremblade in December 2021. This separation was necessary because wild stocks may have contained asymptomatic adults infected with OsHV-1, *V. aestuarianus* or other pathogens. Keeping the populations separate helped prevent potential horizontal transmission of pathogens. The seawater temperature was increased and maintained at 20 ◦C, and a cultured phytoplankton diet (*Isochrysis galbana*, *Tetraselmis suecica* and *Skeletonema costatum*) was provided to favor gametogenesis (Dégremont *et al.* 2005).

In March 2022, 30 oysters per population were opened and sexed. Gametes were collected through gonad stripping. For females, all gametes were combined and then sieved using 20-µm and 100-µm screens to remove small and large debris, respectively. Eggs were then divided by the number of males, from which the sperm was individually collected. Eggs were fertilized with the sperm and, 10 min after fertilization, all eggs were combined and then transferred into 30 L tanks. This method reduces sperm competition and thus maximizes the effective population size (Boudry *et al.*, 2002). The methodology used for the larval and spat cultures is described in two previous studies (Dégremont *et al.*, 2005, 2007). The progenies were maintained in our controlled facilities using UV-treated seawater until their evaluations under experimental infections by OsHV-1. The progenies of three populations, La Floride, Chaucre and the naïve oyster line, were hereafter renamed F, NF and C, respectively, corresponding to farming, non-farming and control oysters.

Before the experiment, spat were acclimated in 120 L tanks with a continuous flow of UV-filtered seawater heated at 19°C. The seawater was enriched with a phytoplankton mixture containing *Skeletonema costatum*, *Isochrysis galbana*, and *Tetraselmis suecica*. The acclimation process spanned at least 2 weeks to ensure optimal growth and feeding conditions, essential for observing effective virus replication leading to oyster mortalities (Azéma, Maurouard, *et al.*, 2017).

Viral suspension preparation

To generate a heterogeneous viral inoculum for infecting all oyster lines, nine viral suspensions were produced between the years 2020 and 2021. These viral suspensions were generated using spat oysters collected during monitoring of OsHV-1 infections from various areas in France. Briefly, approximately 1000 specific-pathogen-free (SPF) spat oysters, reared in the Ifremer experimental facilities located in Argenton (Brittany, France), following a standardized methodology (Petton *et al.*, 2015), were deployed into four oyster farming areas (Figure 1) during periods of disease outbreaks. These areas included i) Brest harbor (BR) (2 viral suspensions) (Logonna-Daoulas, lat.: 48.335 long.: -4.318), ii) Marennes-Oléron bay (MO) (4 viral suspensions) (La Floride, lat.: 45.803 and long.: -1.153), iii) Arcachon basin (ARC) (2 viral suspensions) (Grahude, lat.: 44.653 and long.: -1.073), iv) Leucate lagoon (LEU) (1 viral suspensions) (Leucate, lat.: 43.379 and long.: 3.571). Seven days after the oysters were transplanted into the field, and as mortality began to occur, oysters were brought back to laboratory and maintained at a temperature of 20°C for 7 days. During this period, moribund oysters were systematically sampled daily and then stored at -80°C.

In total, nine viral suspensions were prepared according to an established protocol (Schikorski *et al.*, 2011) (Table 1, Figure 1).

Figure 1: Origin of samples collected to produce viral suspensions. *The samples used to produce the nine viral suspensions were collected between 2020 and 2021 from four locations: Brest harbor (red point), Marennes-Oléron basin (orange point), Arcachon bay (blue point) and Leucate lagoon (green point).*

Table 1: Viral suspension produced and used to infect oysters, sampling dates and viral load quantification.

Viral suspension	Origin	Sampling date	Viral load (copy/µL seawater)	Volume (ml)
1	Arcachon bay	08/07/2020	2,54E+04	11.37
2	Marennes-Oléron basin	08/07/2020	2,30E+04	8.97
3	Brest harbor	24/07/2020	9,24E+04	3.38
4	Marennes-Oléron basin	05/08/2020	1,82E+05	1.71
5	Leucate lagoon	21/05/2021	4,51E+04	5.54
6	Arcachon bay	28/05/2021	7,61E+04	3.29
7	Marennes-Oléron basin	31/05/2021	3,33E+05	0.75
8	Marennes-Oléron basin	07/07/2021	3,67E+05	0.68
9	Brest harbor	07/07/2021	3,33E+04	7.50

Viral load quantification for each inoculum and, preparation of OsHV-1 viral suspension mix

To prepare an equimolar viral inoculum for injection into oysters, viral load of each of the nine viral suspension was assessed. To do so, DNA was extracted from 100 µl of each viral suspension using the MagAttract® HMW DNA kit according to manufacturer's protocol. DNA purity and concentration were assessed by Nano-Drop ND-1000 spectrometer (Thermo Scientific) and Qubit® dsDNA BR assay kits (Molecular Probes Life Technologies), respectively. Then quantification of viral copies was carried out by quantitative PCR using a Mx3005 P Thermocycler (Agilent) (Pepin *et al.* 2008). The outcomes were expressed as the log-transformed copy number of viral DNA per microliter of sea water (cp/µL).

These viral suspensions were pooled to produce the initial viral inoculum (Table 1). To account for variability in the viral copy numbers of each suspension, the pool was generated to ensure equimolarity of the viral copy numbers (Table 1).

Experimental evolution design

The experimental evolution was conducted under controlled laboratory conditions in a Type 2 laboratory at Ifremer's La Tremblade facilities. To initiate the experimental infections, 50 oysters from each oyster line underwent prior myorelaxation using hexahydrate MgCl2 (ACROS, 50 g/L) (Gagnaire *et al.*, 2007). For each oyster line, oysters were individually injected with a 26-gauge needle attached to a multi-dispensing hand pipette. They received 100 µL of pooled inoculum into the abductor muscle (Figure 2, step 1). Concurrently, for each oyster line, a group of 10 oysters underwent myorelaxation and were injected with artificial sea water, serving as a control. The experiment was conducted in duplicate, (*i.e.* 2 batches of oysters in 2 tanks) and placed in 5 L plastic tanks filled with seawater at a temperature of 20°C and with a phytoplankton mixture. Daily mortality monitoring was carried out for 7 days.

After the initial infection, nine tanks were set up: two with infected oysters and one with ASW-injected oysters for each of the three oyster lines. These nine tanks served as the starting point, referred to as "Generation 0" (Figure 2, step 2). To create the first generation of infection (G1) through cohabitation, 25 healthy oysters from each oyster line were introduced 24 hours after infecting the donor oysters (Figure 2, step 3). These healthy oysters were placed in their respective oyster line tank (*i.e.*, 25 individuals from the F oyster line added to each F oyster line's infection replicate, and so on). After

introducing the healthy oysters to the contaminated tank, they were moved to fresh, pathogen-free seawater 24 hours later to release viral particles (Figure 2, step 4), initiating a second generation of infection. The process of adding healthy oysters and moving them for excretion was repeated for the second generation, following the same procedure as outlined above. During this period, the donor oysters from the cohabitation step were monitored for 7 days to estimate mortality rates (Figure 2, step 5), and all moribund oysters were collected for DNA extraction (see below). In total, we performed 13, 18, and 28 generations for the F, C, and NF oyster lines, respectively.

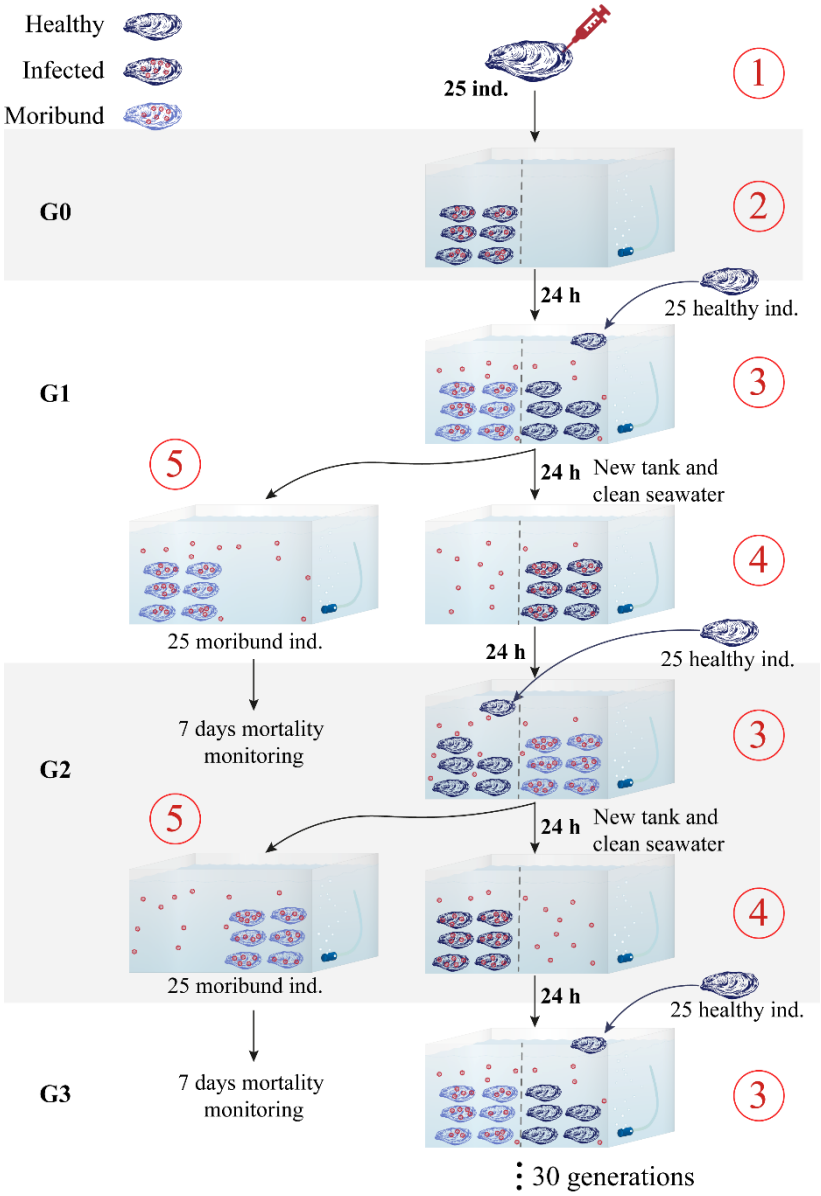


Figure 2: Experimental evolution design for one oyster line oyster. *The experimental evolution was conducted using three oyster lines: a resistant oyster line originated from a farming area (F), a moderately resistant oyster line originated from Ifremer facilities (C) and a sensitive oyster line originated from a non-farming area (NF). Healthy oysters are represented in dark blue, infected oysters in dark blue and red points and moribund oysters in light blue with red points. All experiments were conducted in duplicates. The steps of experimentation consist of 1) infection of 25 healthy oysters by intramuscular injection, 2) placement of infected oysters in a tank of clean seawater to initiate viral particles production, 3) 24 hours post infection, addition of 25 new healthy oysters to the tank to be infected, 4) 24 hours after the cohabitation of healthy oysters with infected oysters, removal of newly infected oysters from the tank and placement in a new tank with clean sea water to initiate viral particles release and 5) counting of moribund oysters for seven days to estimate mortality rates.*

DNA extraction and sequencing

DNA extraction was performed on oysters collected during mortality monitoring (Figure 2, step 5) at generations G0, G5, G10, and G13 for the F oyster line, at generations G0, G5, G10, G15, and G18 for the C oyster line, and at generations G0, G5, G10, G15, G20, G24, and G28 for the NF oyster line. For each oyster line, infection generation, and replicate, seven samples of 35 mg of mantle tissue were collected for DNA extractions. DNA extractions were performed using the MagAttract® HMW DNA kit according to manufacturer's protocols. DNA purity and concentration were assessed employing a Nano-Drop ND-1000 spectrometer (Thermo Scientific) and Qubit® dsDNA BR assay kits (Molecular Probes Life Technologies), respectively. Then, quantification of viral copies was carried out by quantitative PCR using a Mx3005 P Thermocycler (Agilent), as detailed earlier. From each oyster line and replicates, we selected three individuals oyster line collected at different generations: G0, G5, G10, G13 for oysters from the F oyster line, G0, G5, G10, G15, G18 for oysters from the C oyster line and G0, G5, G10, G15 G20, G24 and G28 for oysters from the NF oyster line. The selection was based on viral load and DNA concentration. Overall, 96 samples were selected for sequencing. Illumina DNA-seq libraries were prepared using the Twist library Preparation Enzymatic Fragmentation Kit 1.0 kit (Twist Bioscience) and the sequencing was conducted by IntegraGen SA (Evry, France).

Creation of a global consensus

Following the previously outlined procedure, the viral inoculum employed for oyster infections was produced using an equimolar pool of 9 viral suspensions from oysters collected in LEU (1), ARC (2), MO (4) and BR (2) across the years 2020 and 2021. In a prior study that focused on the global diversity of OsHV-1, genomic sequencing was conducted on 126 samples from the same batch that was utilized for generating the initial nine viral suspensions (Pelletier *et al.* 2023). Specifically, the genomes were obtained from a range of origins: 10 individuals were collected from ARC in 2020, 20 from MO in 2020, 10 from BR in 2020, 44 from MO in 2021, 12 from LEU in 2021, and 15 each from ARC and BR in 2021.

To establish a consensus representative of the global viral suspension used to infect oysters, we performed a multiple alignment of these 126 previously assembled genomes (Figure 3) using MAFFT v1.4.0 (Katoh *et al.*, 2002). The consensus sequence was derived from the alignment and is composed by all the major alleles of OsHV-1 genotypes within the viral suspension. In the forthcoming sections of this manuscript, this consensus will be referred as the “artificial OsHV-1 major NR-consensus”.

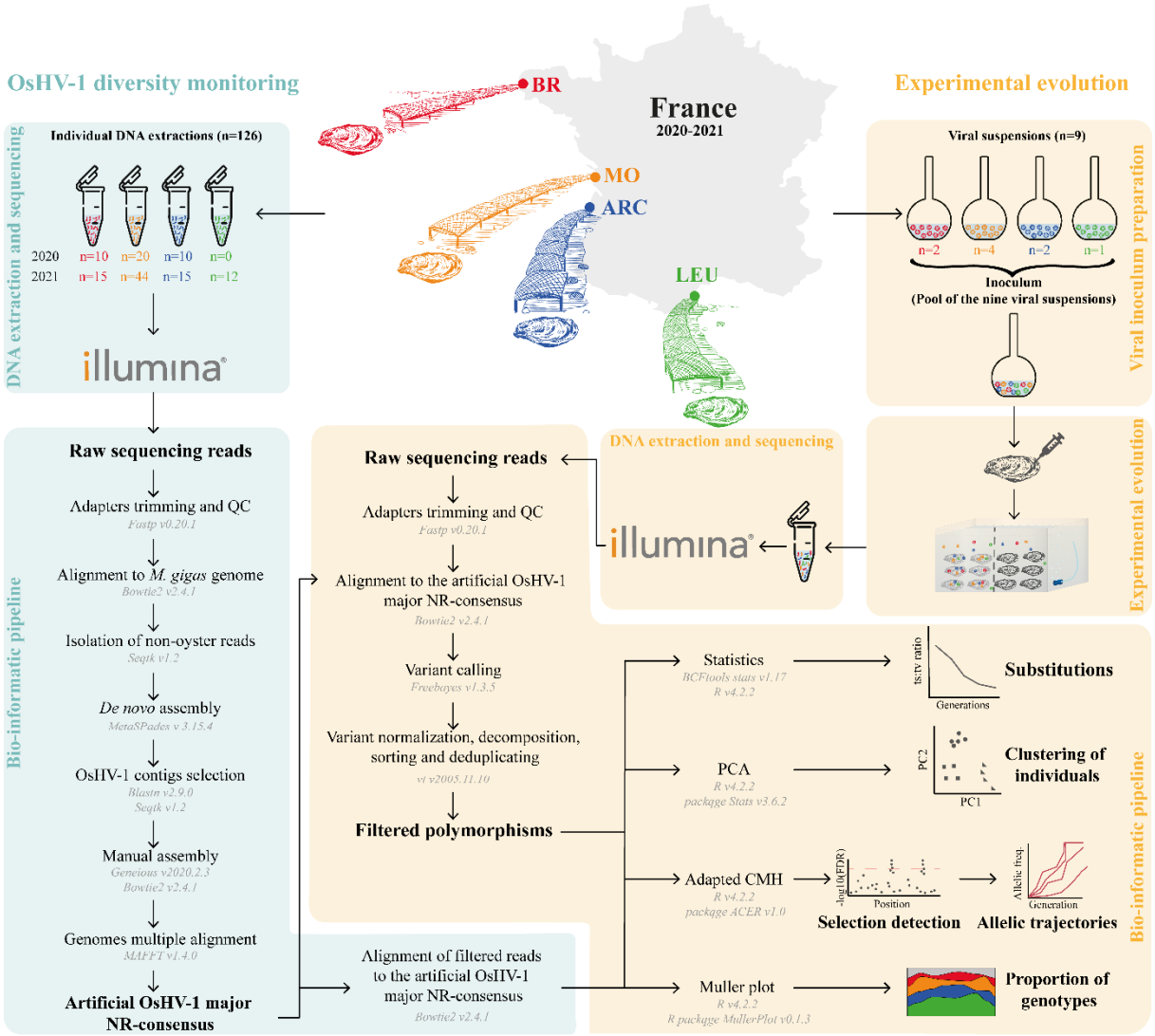


Figure 3: Bioinformatic pipeline used to analyze the dataset and detect selection patterns within OsHV-1 genomes. *In blue are represented the steps performed during monitoring of OsHV-1 diversity in France between 2020 and 2021. Pieces of mantle collected from moribund oysters has been used to extract DNA. DNA has been sequenced via the Illumina technology and raw reads has been used to assemble OsHV-1 genomes de novo through a specific bioinformatic pipeline previously developed for genomic analysis of OsHV-1 (Dotto-Maurel-Pelletier et al. 2019). In orange are depicted the steps performed specifically for the analysis of experimental evolution samples and the detection of selection patterns. Moribund oysters collected between 2020 and 2021 has been used to produce nine viral suspensions which have been pooled according to equimolarity of viral copy numbers and named "Inoculum". Experimental evolution has been performed using this inoculum has the primary source of infection and as described in Figure 2. Pieces of mantle collected from moribund oysters has been used to extract DNA which has been sequenced via the Illumina technology. A specific bioinformatic pipeline has been developed to detect and characterize selection*

Variant calling

The variant calling step was performed using an adapted approach previously employed to detect polymorphisms (Figure 3, see Delmotte-Pelletier *et al.*, 2022 for the detailed bioinformatic pipeline).

The raw data have been deposited in the EBI database under accession numbers xx for future reference and accessibility (Table S1).

Analysis of substitutions types and clustering of individuals based on SNPs

To obtain detailed information about transitions (*i.e.* substitution of a purine by a purine or of a pyrimidine by a pyrimidine) and transversions (*i.e.* substitution of a purine by a pyrimidine or of a pyrimidine by a purine) number within OsHV-1 genome across generations and among host oyster lines, polymorphisms detected in all libraries were used to compute statistics with BCFtools stats v1.17 (Figure 3, Danecek *et al.*, 2021). Substitution statistics were averaged by generation and by oyster line and were then plotted using R v4.2.2 and the ggplot package v3.4.0 (Wickham, 2009).

Exploring sample clustering and SNP data quality assessment through Principal Component Analysis

In the context of EE studies, Principal Component Analysis (PCA) is employed to assess sample clustering and evaluate the quality of Single Nucleotide Polymorphisms (SNPs) data, aiding in uncovering underlying population structure and identifying potential data anomalies (Patterson *et al.*, 2006).

Utilizing SNP frequencies, the individuals have been projected onto a bidimensional space to check the clustering of samples (Figure 3). The PCA analysis was performed with the function pcrcomp of the factoextra v1.0.7 package (Kassambara & Mundt, 2020) in R v4.2.2. This information was linked to those of oyster lines, infection generation and replicates.

Selection detection

Selection patterns were detected using a SNP-based method. As the experiment was performed in duplicates for three host oyster lines, selected SNPs could be detected within each oyster line, in all replicates or some replicates (Figure 3). Selection of particular OsHV-1 genotypes within the different oyster lines is independent, and detected SNPs were not comparable among oyster lines. Consequently, analyses were carried out for each oyster line independently. Initially, to identify SNPs demonstrating significant shifts in allele frequency (AF), a comparison was made between the initial (G0) and evolved populations (reaching G13 for F, G18 for C, and G28 for NF oyster lines, respectively). To achieve this, an adapted Cochran-Mantel-Haenszel (CMH) test was employed, which calculated p-values for each SNP detected within the datasets using the ACER v1.0 package (Barghi *et al.*, 2020; Spitzer *et al.*, 2020) in R. These p-values reflect the variance in SNP frequencies across generations while accounting for the influences of genetic drift and sampling (Spitzer *et al.*, 2020). Subsequent to this, the p-values underwent adjustment using the false discovery rate (FDR) method, as implemented in the p.adjust function from the stats v4.2.2 package (R Development Core Team, 2005). The resultant adjusted p-values were then plotted onto the OsHV-1 genome using a Manhattan plot, employing the ggplot2 v3.4.0 package (Wickham, 2009). Alleles were considered as selected when p-values were higher than the CMH cutoff determined as 5 % of false discovery rate. Finally, using the adjusted p-values and the allelic frequencies of SNPs across generations, allelic frequency trajectories were generated for each replicate. This visualization was achieved using the ggplot v3.4.0 package (Wickham, 2009).

Genotypes proportions

A Muller plot was used to visualize the appearance or disappearance of genotypes specific to the origins of the initial viral suspension used to infect the oysters (Figure 3). As described previously, the initial viral suspension injected into oysters at G0 was produced by pooling of nine viral suspensions collected from LEU (1), BR (2), MO (4) and ARC (2). Libraries associated with the sequencing of the individuals collected during the same mortality events within each basin were used to determined SNPs specific to each origin. Raw reads were then filtered and trimmed as described previously, and aligned to the artificial OsHV-1 major NR-consensus. Variant were called using Freebayes.

Since each sample was taken from a specific location, each SNP detected in the libraries was assigned to the corresponding location. SNPs detected during EE of OsHV-1 was classified by origin and then used to plot the relative abundance of SNPs detected within specific origins using the MullerPlot package v0.13 (Farahpour *et al.*, 2022) in R.

Results

Evolution of survival rates and viral load in Pacific Oyster spat infected with OsHV-1 across generations

To assess the impact of the infection by the pool of OsHV-1 genotypes on Pacific oyster spat, the survival rate was monitored for 7 days post-infection (dpi) in both infected replicates and control tanks for each of the three oyster lines. Throughout the entire experiment, no oyster mortality was observed in the control conditions.

Across all three oyster lines, the overall oyster mortality rates exhibited fluctuations between successive infection generations. However, a predominant trend emerged, revealing a decline in oyster mortalities across all three oyster lines over the course of the experiment (Figure 4A). Specifically, at G0 mortality rates were 88 % for F oyster line, 96 % for C oyster line and 84 for NF oyster line.

Different variations in mortality rates were observed for the oyster lines C compared to the lines F and NF oyster lines. In the oyster line C, the mortality rate continuously decreased until it reached 42% at G15. Subsequently, the mortality rate increased, reaching 56% at G18, leading to the termination of the experiment because no mortalities were observed at G19. In contrast, mortality rates of oysters in lines F and NF varied from one generation to the next. At G5 oysters from lines F experienced a significant decrease in mortality rate, reaching 58%. At G10, there was an increase in mortality rates reaching 87%. After G10, mortality rates fell drastically reaching 2% for the oyster line F. This observation led to the end of the experiment for the oyster lines F. At G5 oysters from lines F

experienced a significant decrease in mortality rate reaching 50 %, followed by an increase in mortality rates that reach 99 % at G10. After G10, mortality rates decrease to 51%. Finally, the mortality rate increased between G15 and G20 to reach 89%, decreased between G20 and G24 to reach 28%, and then increased again until G28 to 48%.

Cohabitation-induced infection involves the maintenance of the infection solely through replication within the host. The challenges associated with inducing reinfection through successive generations *via* cohabitation led to the interruption of the experimental evolution at the 13th generation for the oyster line F, the 18th generation for oyster line C, and the 28th generation for oyster line NF.

Variations in viral loads was observed both among individuals and among oyster lines (Figure 4B). The mean viral load within individuals from the oyster lines F exhibited a slight increase across generations, rising from 8.38x105 ± 5.49x105 viral DNA copies/ng of DNA at G0 to 1.13x106 ± 4.81x105 viral DNA copies/ng of DNA (Figure 4B). Conversely, the average viral load within individuals from the oyster lines C and NF demonstrated a slight decrease across generations. For C oysters, the viral load varied from 6.62x105 ± 4.49x105 viral DNA copies/ng of DNA at G0 to 3.34x105 ± 2.54x105 viral DNA copies/ng of DNA at G18. In the case of NF oysters, the viral load ranged from 2.09x106 ± 1.20x106 viral DNA copies/ng of DNA at G0 to 9.35x105 ± 1.11x106 viral DNA copies/ng of DNA at G28 (Figure 4B).

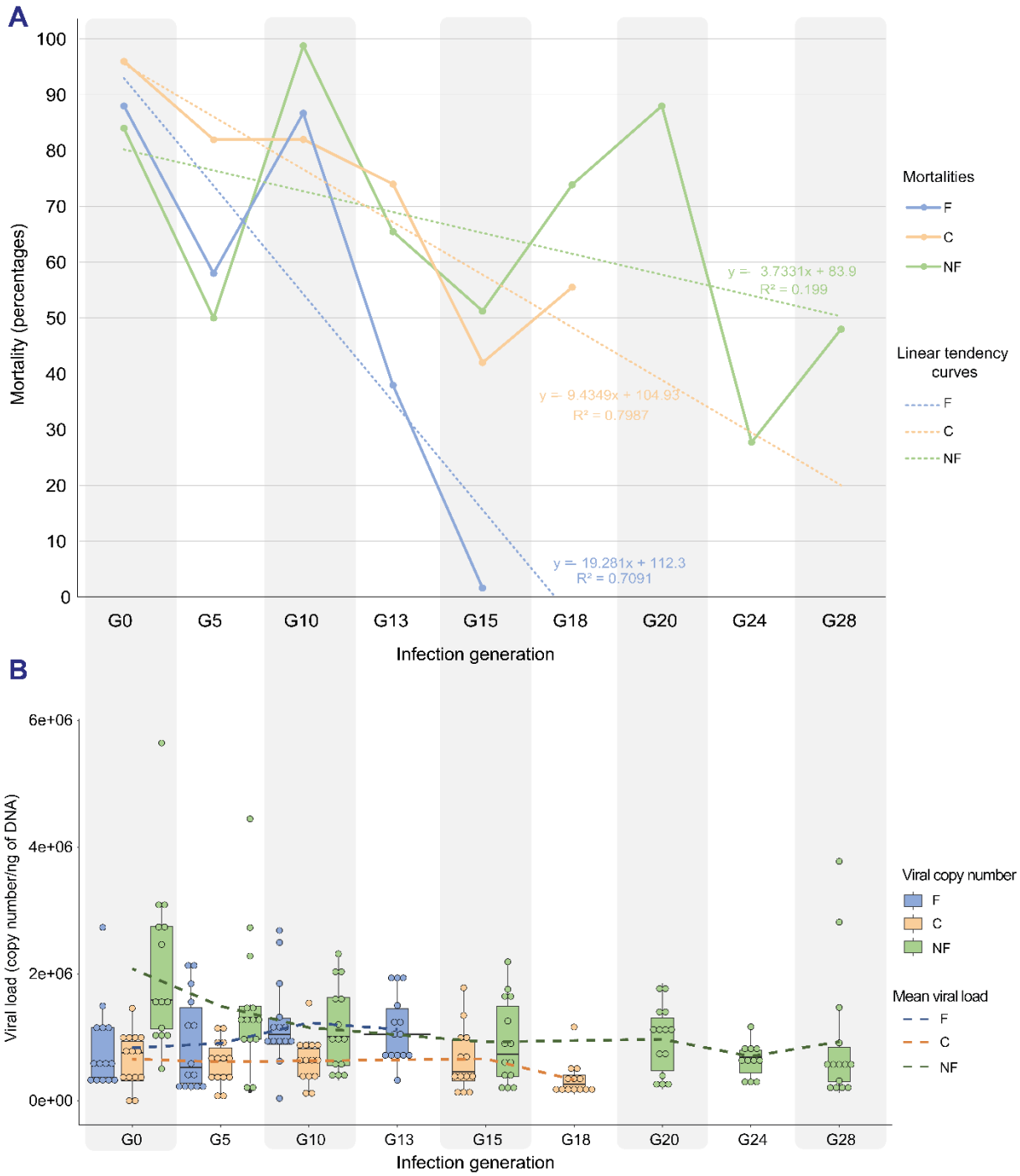


Figure 4: Oyster mortalities and viral load from the three oyster lines at generations 0, 5, 10, 13, 15, 18, 20, 24 and 28. *A) Oyster mortality was monitored daily to estimate mortality rates within each oyster line for selected generations. The solid curves and dotted lines correspond to mortality rates and the linear tendency of mortality rates respectively, for oysters from the line F oyster line (in blue), from the C oyster line (in orange), and from the NF oyster line (in green). B) Viral loads were assessed from seven individuals for each generation, oyster line, and replicate to monitor the evolution of the infection over generations. Viral loads of individuals are represented by points (in blue, orange, and green for*

oysters from lines F, C, and NF, respectively). The box plots represent median values, the first and third quartiles, and the standard deviation of viral loads within an oyster line for a specific generation. The dotted lines correspond to the viral load for oysters from the F oyster line (in blue), from the C oyster line (in orange), and from the NF oyster line (in green).

Dynamics of substitutions and transition:transversion ratio

Studying the variation in both the number and types of substitutions across generations of viral infection during experimental evolution is of paramount importance for gaining insights into the dynamic evolutionary processes that shape genetic diversity and adaptation within populations (Bromham, 2020).

A prominent trend emerges for all oyster lines, indicating a rapid reduction in the count of SNPs in the OsHV-1 genome across all three lines throughout the generations of infection (Figure 5). The trend is more pronounced for the NF oyster line extending from 372 ± 55 SNPs at G0 to 294 ± 43 SNPs at G5 (Figure 5A).

A statistical analysis, based on the identified SNPs, was conducted to assess the transition:transversion (ts:tv) bias (Figure 5B). The ts:tv ratio of the virus remained consistently below 0 across all generations and lines, except for the samples collected from the F oyster lines at G0, for which the virus displayed a ts:tv ratio of 1.26, indicating an excess of transitions (Figure 5B). Additionally, the viral ts:tv ratio demonstrates a decreasing trend over generations for both the F and C oyster lines. Regarding the NF oyster lines, the overarching pattern leans towards a decline in the ts:tv ratio; however, an upward shift in the ratio is noticeable in G10.

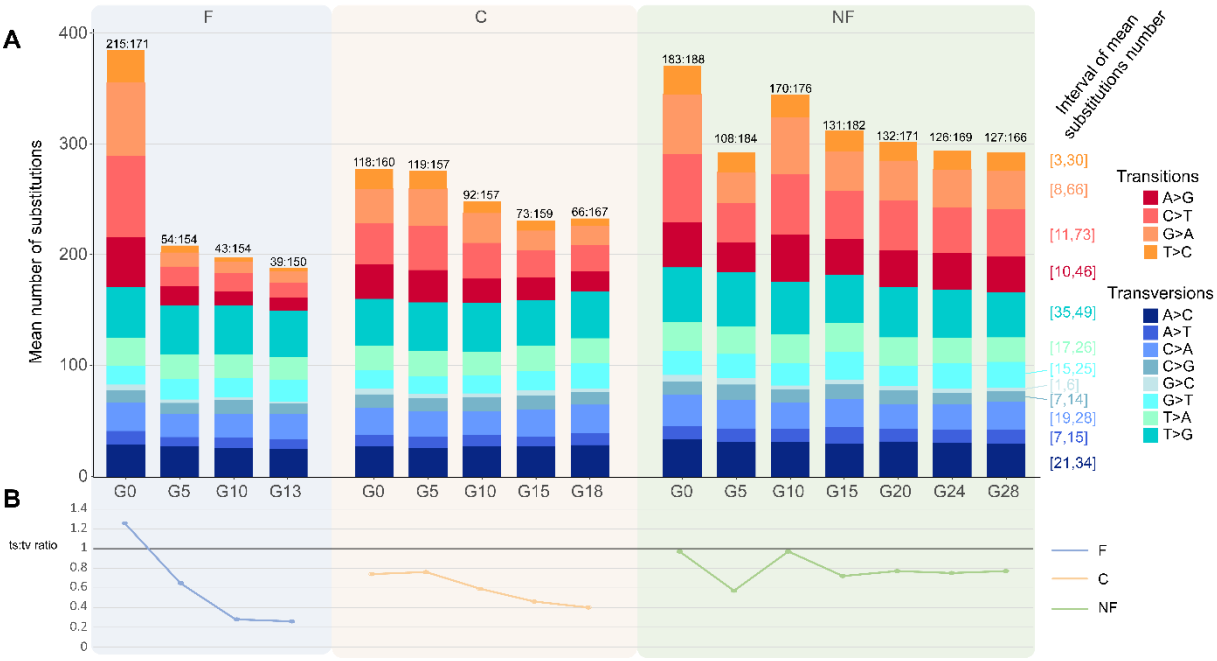


Figure 5: Substitution types and transition:transversion ratio observed within OsHV-1 populations. *A) The mean number of viral substitution types per generation and per oyster line is depicted. Transitions (substitution of a pyrimidine by a pyrimidine or a purine by a purine) are represented with a gradient from red to orange, while transversions (substitution of a purine by a pyrimidine or the reverse) are represented with a gradient from blue to green. The numbers above the stacked bars indicate the transition and transversion numbers. To the right of the bars, numbers in square brackets indicate the average number of each substitution type in the samples. B) Curves displaying variations in the ts:tv ratio among generations for each oyster line. Lines are colored according to oyster line.*

Unveiling OsHV-1 population structure and anomalies in experimental evolution through SNP Data analysis and principal component projection

Using SNP frequencies, we projected the OsHV-1 genomes of individuals onto a bidimensional space (Figure 6). The initial principal component axis accounted for 12.61% of the variance in the data, while the subsequent principal component axis explains 9.89% of the variance (Figure 6). Clearly, the first principal component aligns with the progression of infection generations, while the second component is indicative of oyster line associations or susceptibility traits within oyster lineages. The graphical representation reveals the presence of four distinct clusters corresponding to the interplay between oyster lines and infection generations.

Interestingly, the first group encompasses samples drawn from all oyster lines at G0, centralizing within the plot and thereby exhibiting similarity across individuals OsHV-1 genomes (Figure 6A, group 1). These samples groups with those collected from the replicate 1 of the C line at G0, G5, G10, G15, and G18.

The second cluster (Figure 6A, group 2) encompasses samples primarily collected from F at G5, G10, and G13, as well as the replicate 2 from C line at G5, G10, G15, and G18.

Finally, the third cluster comprises specimens sourced from the NF oyster lines, further bifurcated into two subsets. The initial subset (Figure 6A, group 3.1) exclusively represents samples obtained during G5, while the second subset (Figure 6A, group 3.2) corresponds to samples collected between G15 and G28.

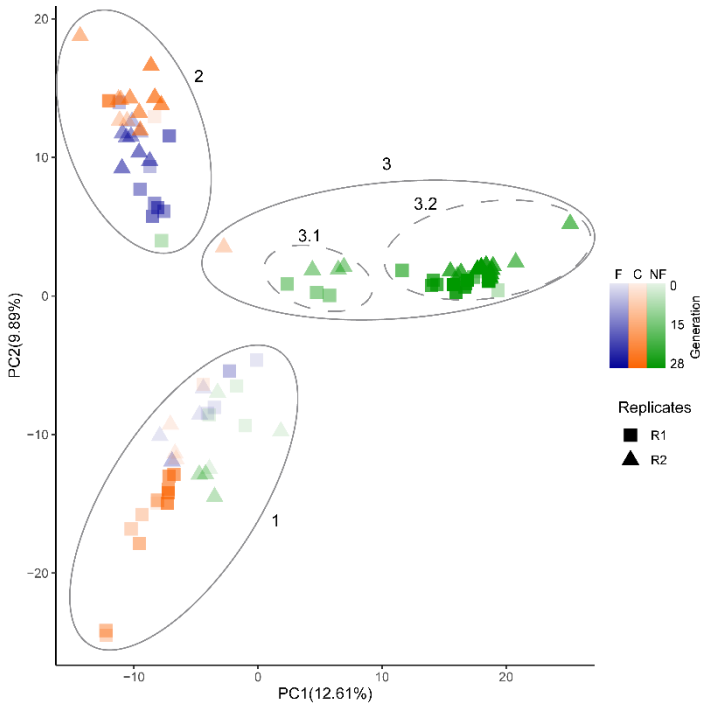


Figure 6: Principal component analysis reveals OsHV-1 population structure and generation-oyster lines associations. *Projection of individual similarities onto the first and second principal components. The colors of the points correspond to the infected oyster lines: individuals from the F oyster line are colored in blue, individuals from the C oyster line are colored in orange, and individuals from the NF oyster line are colored in green. Infection generations are represented by a gradient of colors, and replicates are indicated by different point types (square for replicate 1 and triangle for replicate 2).*

Detection of regions under selection pressure and allelic frequency trajectories

The Cochran-Mantel-Haenszel test (CMH) is a valuable analytical tool widely used for discerning selection pressures within experimental evolution and time series datasets. It offers a comprehensive approach to assess shifts in allele frequencies over successive generations (Spitzer *et al.*, 2020).

The CMH test, employed on SNPs detected within each viral population collected from the three lines independently (Figure 7), revealed that in oyster lines F (Figure 7A) and C (Figure 7B), no regions were significantly impacted by selection. However, within the NF oyster lines, a total of 117 positions demonstrated candidate selection signatures (Figure 7C). These positions are distributed across the entire OsHV-1 genome, with a distinct emphasis on specific regions, such as repetitive segments or the UL region spanning 50 kb to 80 kb. More precisely, 28 potentially selected SNPs were located within intergenic regions, including one SNP in the stem loop, and 89 SNPs were located within ORFs (Figure 8).

Specifically, among the 89 SNPs situated within ORFs, 58 correspond to ORFs with undefined functions, while 12 are found within ORFs encoding transmembrane proteins. Additionally, 8 SNPs are situated within ORFs encoding secreted proteins, four within ORF 100 coding for a DNA polymerase, three within ORFs encoding Zinc-finger, Ring-type proteins, two within ORFs encoding BIR repeat motifs, one within ORF115 associated with an Origin-binding protein, and the final SNP within ORF20 encoding a ribonucleotide reductase.

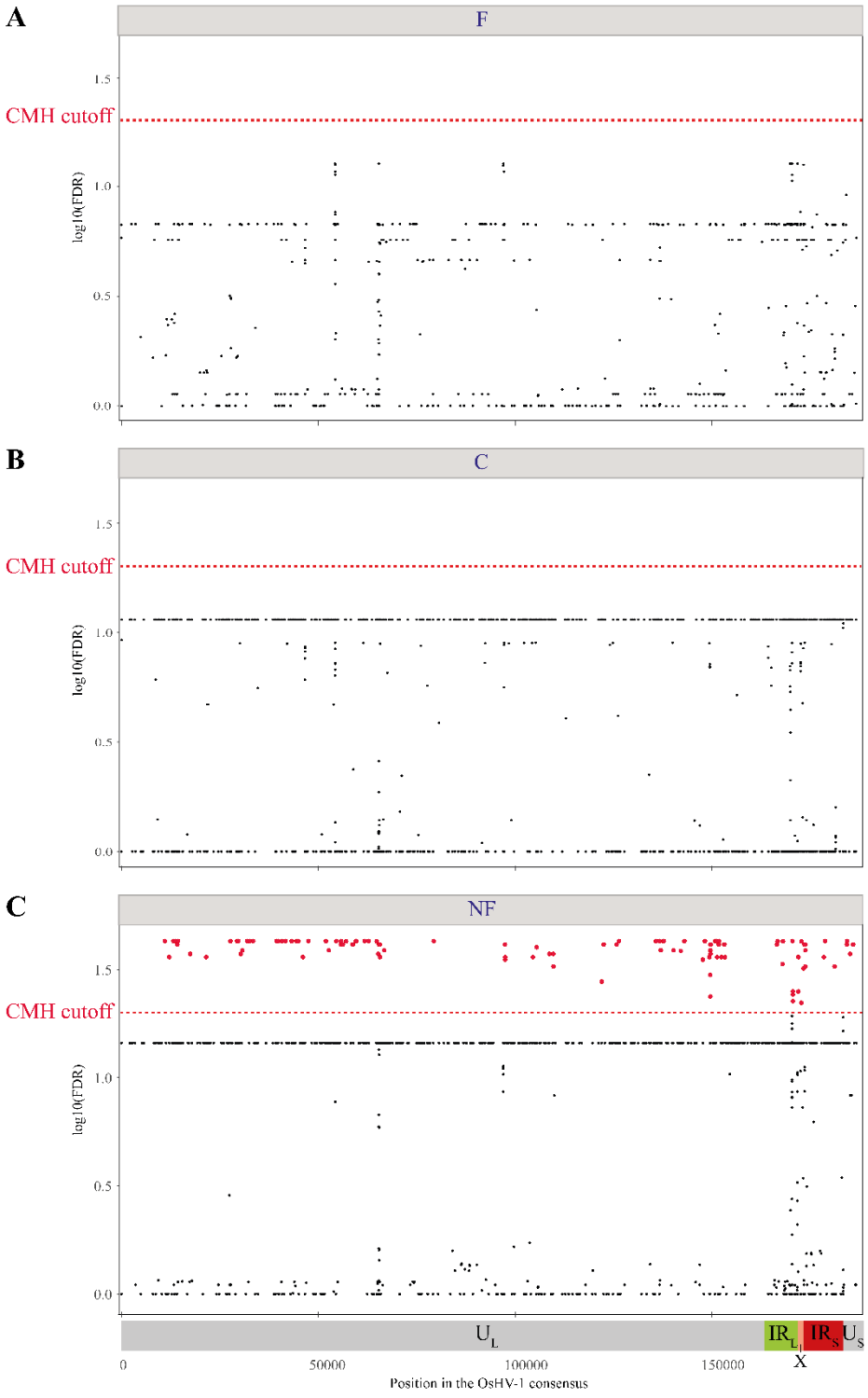


Figure 7: Genomic distribution of candidate SNPs in oyster lines. *The scatter plots show the negative log10-transformed p-values of SNPs corresponding to genomic positions for A) Farming, B) Control and C) Non-Farming oyster lines. These p-values were calculated using the CMH test, which involved comparing the viral populations of the founder and evolved states. Significance thresholds are indicated by red dotted lines, representing the CMH cutoff at a 5% false positive rate. Significant p-values are*

highlighted by red points. The genomic architecture of the OsHV-1 genome is represented below the graphs, depicting unique and repeated regions as follows: U_L/U_S (unique long/short) and IR_L/IR_S (inverted repeat long/short).

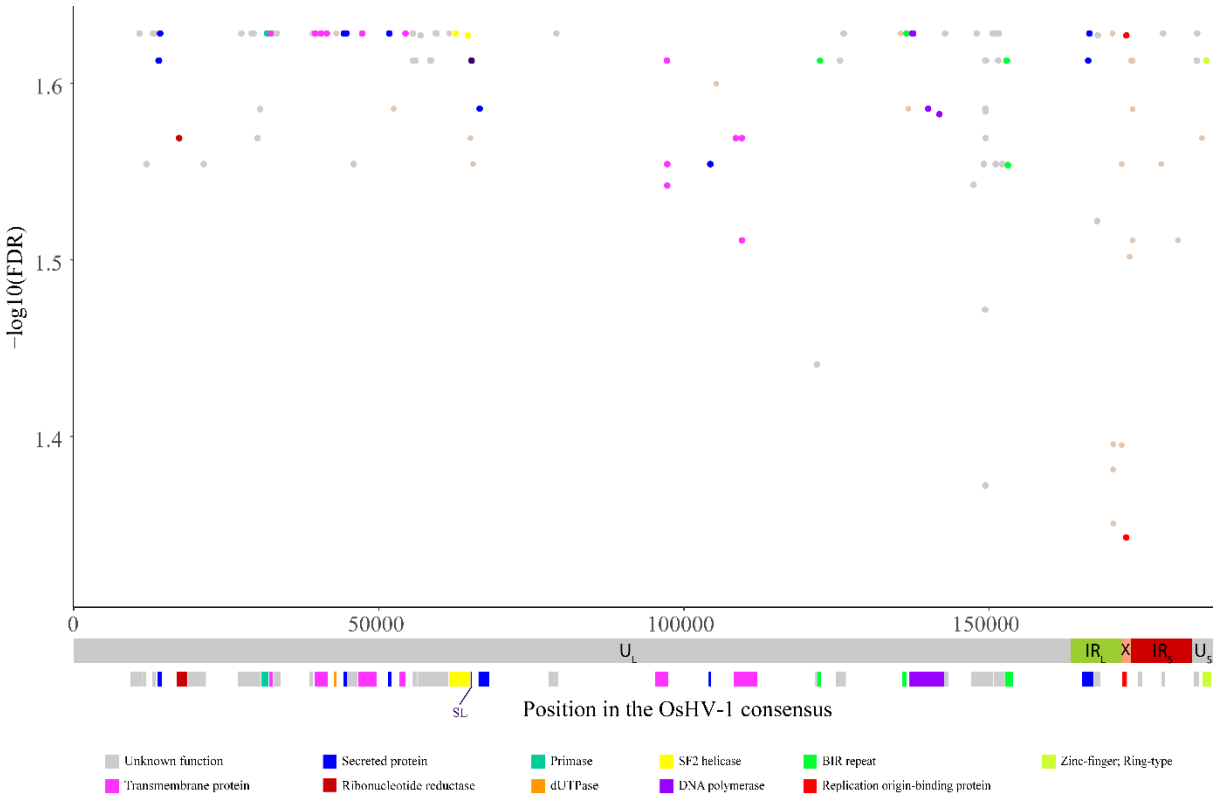


Figure 8: Zoom on the genomic distribution of selected SNPs in Non-Farming oyster line. *The scatter plot shows the negative log10-transformed p-values of selected SNPs corresponding to genomic positions. These p-values were calculated using the CMH test, which involved comparing the viral populations of the founder and evolved states. Significance thresholds representing the CMH cutoff at a 5% false positive rate are displayed. Below the plots, the genomic structure of the OsHV-1 genome is presented, highlighting distinct segments such as unique and repeated regions denoted as U_L/U_S (unique long/short) and IR_L/IR_S (inverted repeat long/short). Moreover, the genomic architecture is supplemented by the depiction of ORFs (Open Reading Frames) housing the selected SNPs, color-coded based on their encoded functions or domains. SL: Stem-loop. The SNPs themselves are also color-coded according to the specific impact they have on the function of the corresponding ORF.*

An allelic frequency trajectories (AFT) plot was employed for candidate SNPs detected within the NF oyster line to track changes in allelic frequencies over generations and get insights into the evolutionary dynamics within populations, especially for positive selection (Barghi *et al.*, 2020). The analysis of allelic frequency trajectories has been performed separately for both replicates (Figure 9).

The AFT plot of candidate SNPs reveals a general trend of increasing allelic frequencies, ultimately resulting in the fixation of selected alleles in both replicates. However, in the first replicate, the allelic frequencies of candidate SNPs increase between G0 and G5, then decrease between G5 and G10, and finally increase again between G10 and G15 to reach fixation (Figure 9A). In the second replicate, a similar trend is observed, but only for a 77 SNPs (Figure 9B). Moreover, alleles of 66 SNPS for replicate 1 and 26 SNPs for replicate 2, represented at the bottom of the plot, show an increase in frequencies, albeit to a lesser extent (Figure 9A and B).

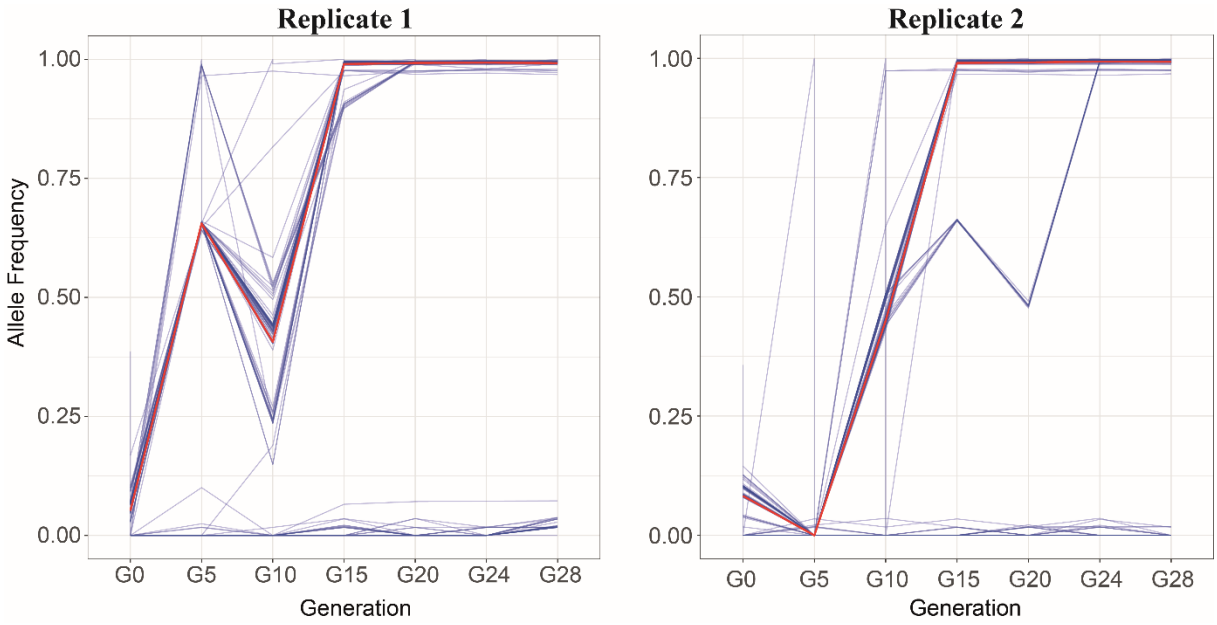


Figure 9: Allelic frequency trajectories plot for the top 100 candidate SNPs detected within the OsHV-1 samples of replicate 1 and 2 collected from oysters of the NF oyster line. *Each blue line represents the allelic frequency trajectory of a specific SNP across generations. Red line represents the median allelic frequency trajectory.*

Relative abundance of SNP from specific origin

The initial inoculum used to infect donor oysters was produced by mixing nine viral suspensions collected from BR, MO, ARC, and LEU. To track the presence of viral genotypes specific to one of the geographic origins and their fate across generations of infection, a Muller plot was conducted (Figure 10).

The results revealed that multiple genotypes from different origins coexist from the beginning to the end of the experiment within each oyster lines. The initial viral inoculum had been produced by pooling nine viral suspensions in equimolar quantities and had been injected into donor oysters at G0. At G0, the relative abundance (RA) of genotypes specific to MO and LEU is higher than that of genotypes specific to ARC and BR.

This trend is generally maintained across generations of infections for all oyster lines. However, depending on the oyster lines, the RA varies over the generations. Specifically, for the F oyster line, OsHV-1 SNPs specific to MO are present at a RA of 0.2 in G0, then their RA increases to reach 40% in G10 before slightly decreasing. The trend is generally similar for OsHV-1 SNPs specific to ARC and BR, but with much lower RA (between 0.02 and 2.0). As for OsHV-1 SNPs specifically from LEU, the trend is reversed. Indeed, the initial RA decreased from 0.4, to 0,1 at G5.

Moreover, over the course of the experiment, SNPs that were not initially present in any of the viral suspensions began to emerge. At G0, the RA of these SNPs was relatively low within farming and non-farming oyster lines (RA = 0.1). However, as the infection generations progressed, their RA increased. In contrast, the RA of “other” SNPs within the control oyster lines at the beginning of the experiment was 0.3, and then it experienced fluctuations across generations. First, it decreased to RA=0.2 until G5, then increased to RA=0.5 until G10 and decreased to RA=0.2 until G15.

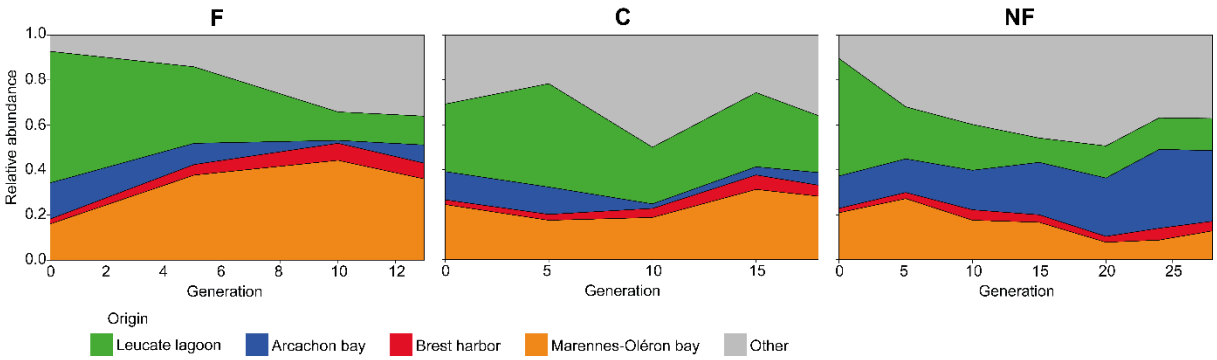


Figure 10: Muller plot representing the relative abundance of viral SNPs collected from specific origins for each oyster line during the experimental evolution. *The relative abundance of viral SNPs is color-coded according to their geographical origin of the viral suspension: the orange area corresponds to genotypes collected specifically in MO, the red area corresponds to genotypes collected specifically in BR, the blue area corresponds to genotypes collected specifically in ARC, the green area corresponds to genotypes collected specifically in LEU, and the grey area corresponds to genotypes that appeared during the experimental evolution.*

Discussion

To mitigate the impact of infectious diseases, it is essential to unravel the mechanisms that shape the evolution and dissemination of disease-causing microorganisms. Variability in host resistance levels is known to play a significant role in driving these processes (Gandon & Michalakis, 2000; Kubinak & Potts, 2013). This study provides valuable insights into the intricate evolutionary dynamics of OsHV-1 and its interactions with Pacific oysters, focusing on how the virus adapts and diversifies in response to variations in host genetic backgrounds. These findings shed light into mechanisms driving viral evolution, a crucial step in understanding and addressing the devastating effects of this virus on the oyster farming industry.

During this experimental evolution, we observed a decrease in both oyster mortalities and viral load across successive generations of OsHV-1 infection. This decline in mortality rates can be attributed to several factors. First and foremost, it can be explained by the selection of less virulent viral variants. As is well-established, the interactions between a pathogen and its host are characterized by a delicate balance (Anderson & May, 1982). In conditions where the oyster population is limited, such as in our experimental setup with 25 individuals per tank, infected hosts succumb rapidly. One plausible strategy employed by the virus may involve reducing its virulence to maintain a stable host population

while ensuring efficient replication without lethality. It is a well-documented phenomenon that there exists a trade-off between virulence and transmission, where increased transmission is associated with reduced virulence, and conversely, heightened virulence leads to diminished transmission (Anderson & May, 1982; Gandon & Michalakis, 2000; Kun *et al.*, 2023). Similar results to ours were obtained in an experimental evolution study involving tobacco and pepper plants infected with the tobacco etch potyvirus (TEV). In this study, viral load and virulence significantly decreased across generations, a phenomenon linked to the host lineage, which plays a pivotal role in shaping the fate of viruses (Cuevas *et al.*, 2015).

Secondly, the reduction in mortality rates can be attributed to a lack of complementation between genotypes (Froissart *et al.*, 2004; Montville *et al.*, 2005). In fact, host infection rarely occurs with a single genotype; instead, it often involves co-infection where multiple viral genotypes contribute to the infection and associated symptoms. It has been previously demonstrated that certain genotypes, which have developed disadvantageous mutations in important genes, can be complemented by other genotypes that do not carry these mutations and whose corresponding proteins remain functional (Froissart *et al.*, 2004; Montville *et al.*, 2005). This mechanism allows low-fitness genotypes to gain a phenotypic advantage by utilizing intracellular proteins produced by co-infecting high-fitness strains (Froissart *et al.*, 2004; Montville *et al.*, 2005). In the context of our study, where genetic diversity has likely been drastically reduced, possibly due to a bottleneck effect, it is possible that some genotypes have been purged from the viral population, leading to limited complementation events and consequently a reduction in virulence over successive generations.

A third hypothesis to explain the decline in mortality rates and viral loads is the temporal shift in the peak of viral particle release into seawater as infection generations progress. Consequently, the viral load would be high when cohabitation begins and progressively decreasing over time. In the case of OsHV-1, the excretion peak by oysters occurs between 24 and 48 hours post-infection (hpi) (Delmotte *et al.*, 2020), but this timing can vary depending on several factors, including oyster susceptibility, the initial viral load in infected oysters, the number of donors, food availability, oyster age, water temperature, and salinity (Dégremont, 2011; Delmotte *et al.*, 2020; Evans *et al.*, 2017; Pernet *et al.*, 2018, 2018; Petton *et al.*, 2015; Schikorski *et al.*, 2011). Furthermore, previous research has demonstrated that oyster infections through cohabitation result in fewer fatalities (Cain *et al.*, 2021). While it may not be the optimal method for experimental evolution, successive infection by cohabitation was chosen to mimics the natural field infection of oysters by the OsHV-1 virus.

Genetic mutations play a pivotal role in shaping the evolutionary landscape of organisms, with transitions and transversions representing two distinct classes of point mutations that contribute to

genetic diversity and adaptation. Molecular evolutionary hypothesis suggests that natural selection promotes amino acid substitutions mostly through transitions (Stoltzfus & Norris, 2016). The substantial decline in the count of single nucleotide polymorphisms (SNPs) and the decrease in the transition:transversion (ts:tv) ratio across all three oyster lines over successive infection generations suggest a trend toward genetic homogenization within the viral populations. This trend could arise from selection favoring specific viral genotypes or from a reduction in genetic diversity due to population bottlenecks (McCrone & Lauring, 2018). Interestingly, similar patterns of reduction in viral variants have been observed in two distinct experimental evolution studies. Firstly, in the case of the Cucumber mosaic virus injected into tobacco plants during an experimental infection, viral populations underwent a population bottleneck due to genetic drift, resulting in a significant reduction in viral diversity (Li & Roossinck, 2004). Similarly, when Potato virus Y passed through three different potato plant genotypes with varying levels of susceptibility to the virus, a combination of strong selective sweeps and bottlenecks led to a reduced viral diversity by the fifth generation (Kutnjak *et al.*, 2017). Given the context of OsHV-1 infecting different oyster genetic backgrounds and considering host population size, it is plausible that the viral population underwent a bottleneck, leading to a reduction in the number of viral variants.

The PCA analysis of SNP frequencies unveiled distinct clusters that corresponded to the different oyster lines and infection generations, signifying genetic differentiation. Through the process of experimental evolution, viral populations exhibited adaptability across successive generations of infection, with individuals organizing into clusters according to their infection histories. Particularly noteworthy, the clustering of individuals from the NF and F oyster lines is clearly delineated and distinct, indicating a contrasting evolution within these two lines. In contrast, C oysters exhibited mixed clustering, likely owing to their intermediate genetic status. These findings strongly suggest ongoing genetic differentiation, which could be attributed to varying levels of oysters susceptibility to OsHV-1 infection (Kubinak & Potts, 2013). Within the NF oyster line, distinct clusters emerged, indicating evolutionary changes across infection generations in response to selective pressures.

The identification of 117 candidate selection signatures in the OsHV-1 genomes, in conjunction with the allelic frequency trajectories (AFT) plots collected from the S oyster line, provides compelling evidence of selective pressures acting on OsHV-1 during the course of the experimental evolution. The trend toward increasing allelic frequencies, ultimately leading to allele fixation in both replicates, strongly indicates positive selection (Barghi *et al.*, 2020). Notably, selective pressure was exclusively detected within the OsHV-1 genome during successive infection in the NF oyster line.

One plausible hypothesis to explain this observation is that in more susceptible oysters, viral replication is favored (Dégremont, 2011), leading to an increase in mutation rates and enabling the fixation of advantageous alleles within populations (Elena & Sanjuán, 2005; Peck & Lauring, 2018). Given that viral diversity and natural selection processes in viruses are primarily influenced by environmental changes or by the host's immune system acting as a filter to limit viral diversity and adaptation, one might have expected to detect selection signals in viral genomes infecting the F oyster line (Kubinak & Potts, 2013). However, since no selection signals were detected for the F oyster line, despite the observed reduction in viral diversity, this strengthens the hypothesis that the viral populations infecting the F oyster line may have been influenced by a bottleneck effect due to genetic drift.

These selection signatures are distributed across the viral genome, with a specific emphasis on repetitive segments and the U_L region. The presence of selection signatures within ORFs encoding various proteins suggests that both regulatory and functional elements of the virus are under selective pressure. Specifically, regions responsible for vital biological functions or domains related to the attachment of viral particles to host cells, DNA replication, synthesis, packaging, as well as transmembrane and membrane proteins, could be crucial for the adaptation to resistant or susceptible hosts. It is worth noting that viral entry proteins play a pivotal role in enhancing viral fitness, and even a single amino acid mutation within these proteins can modify the virus's ability to infect hosts with varying levels of susceptibility to infection (van Sluijs *et al.*, 2017). Moreover, studies have revealed that various domains, including envelope domains (such as membrane glycoproteins and transmembrane receptors), auxiliary domains (like Zinc-finger, RING type, dinucleoside kinase), and modulation and control domains (such as Interleukin, Interferon-regulatory factor, or Zinc-finger), have been acquired, duplicated, or lost during the evolution of Herpesviridae (Brito & Pinney, 2020). Many of these acquired domains enabled viruses to specifically bind to host cells and to evade or manipulate the host's immune system. In the present study, the majority of the signals of selection were identified within these domains, indicating a potential adaptation of essential proteins that leads to a specialization of viral genotypes to their respective host genetic background. Additionally, two previous studies conducted on OsHV-1 (Delmotte *et al.*, 2020; Pelletier *et al.*, 2023) have demonstrated that selective pressure plays a key role in the adaptation of the virus to different host species and environments through mutations, particularly in ORFs encoding viral membrane-related and metabolic-related proteins, further supporting the hypothesis of immune selection of OsHV-1 within the S oyster line. Based on the results obtained in this study, it is evident that adaptation can also occur at a finer scale within the host's genetic background, which is influenced in this case by the geographic origin of the parental individuals (*i.e.* farmed and not farmed areas).

Conclusion

This study provides valuable insights into the evolutionary forcing mechanisms of OsHV-1 and its interactions with Pacific oysters, with a particular emphasis on viral adaptation and diversification in response to host infection susceptibility. Employing an experimental evolution approach, it sheds light on the dynamic nature of host-virus interactions, the potential for viral adaptation, and the potential role of genetic diversity in shaping the outcome of these interactions. The evolution of OsHV-1 appears to be driven by both genetic drift and positive selection linked to the oyster genetic background, resulting in a reduction in viral diversity and the fixation of specific alleles within ORFs dedicated to host-virus interactions and virus functional maintenance. Further research into the functional consequences of these changes is warranted. These findings contribute to our understanding of the mechanisms governing viral evolution and hold significant implications for mitigating the devastating impacts of OsHV-1 on oyster farming industries. Future studies should continue to explore the intricate dynamics of host-virus interactions and their implications for viral evolution and disease management.

References

Abbadi, M., Zamperin, G., Gastaldelli, M., Pascoli, F., Rosani, U., Milani, A., Schivo, A., Rossetti, E., Turolla, E., Gennari, L., Toffan, A., Arcangeli, G., & Venier, P. (2018). Identification of a newly described OsHV-1 µvar from the North Adriatic Sea (Italy). Journal of General Virology, 99(5), Article 5. https://doi.org/10.1099/jgv.0.001042

Anderson, R. M., & May, R. M. (1982). Coevolution of hosts and parasites. Parasitology, 85(2), 411–426. https://doi.org/10.1017/S0031182000055360

Arzul, I., Nicolas, J.-L., Davison, A. J., & Renault, T. (2001). French Scallops: A New Host for Ostreid Herpesvirus-1. Virology, 290(2), Article 2. https://doi.org/10.1006/viro.2001.1186

Arzul, I., Renault, T., Lipart, C., & Davison, A. J. (2001). Evidence for interspecies transmission of oyster herpesvirus in marine bivalves. Journal of General Virology, 82(4), Article 4. https://doi.org/10.1099/0022-1317-82-4-865

Azéma, P., Lamy, J.-B., Boudry, P., Renault, T., Travers, M.-A., & Dégremont, L. (2017). Genetic parameters of resistance to Vibrio aestuarianus, and OsHV-1 infections in the Pacific oyster, Crassostrea gigas, at three different life stages. Genetics, Selection, Evolution : GSE, 49, 23. https://doi.org/10.1186/s12711-017-0297-2

Azéma, P., Maurouard, E., Lamy, J.-B., & Dégremont, L. (2017). The use of size and growing height to improve Crassostrea gigas farming and breeding techniques against OsHV-1. Aquaculture, 471, 121–129. https://doi.org/10.1016/j.aquaculture.2017.01.011

Bai, C.-M., Morga, B., Rosani, U., Shi, J., Li, C., Xin, L.-S., & Wang, C.-M. (2019). Long-range PCR and high-throughput sequencing of Ostreid herpesvirus 1 indicate high genetic diversity and complex evolution process. Virology, 526, 81–90. https://doi.org/10.1016/j.virol.2018.09.026

Barghi, N., Hermisson, J., & Schlötterer, C. (2020). Polygenic adaptation: A unifying framework to understand positive selection. Nature Reviews Genetics, 21(12), 769–781. https://doi.org/10.1038/s41576-020-0250-z

Bellone, R., Lequime, S., Jupille, H., Göertz, G. P., Aubry, F., Mousson, L., Piorkowski, G., Yen, P.-S., Gabiane, G., Vazeille, M., Sakuntabhai, A., Pijlman, G. P., de Lamballerie, X., Lambrechts, L., & Failloux, A.-B. (2020). Experimental adaptation of dengue virus 1 to Aedes albopictus mosquitoes by in vivo selection. Scientific Reports, 10, 18404. https://doi.org/10.1038/s41598-020-75042-4

Boudry, P., Collet, B., Cornette, F., Hervouet, V., & Bonhomme, F. (2002). High variance in reproductive success of the Pacific oyster žCrassostrea gigas, Thunberg/ revealed by microsatellite-based parentage analysis of multifactorial crosses.

Brito, A. F., & Pinney, J. W. (2020). The evolution of protein domain repertoires: Shedding light on the origins of the Herpesviridae family. Virus Evolution, 6(1), veaa001. https://doi.org/10.1093/ve/veaa001

Bromham, L. (2020). Substitution Rate Analysis and Molecular Evolution.

Burioli, E. A. V., Prearo, M., & Houssin, M. (2017). Complete genome sequence of Ostreid herpesvirus type 1 µVar isolated during mortality events in the Pacific oyster Crassostrea gigas in France and Ireland. Virology, 509, 239–251. https://doi.org/10.1016/j.virol.2017.06.027

Cain, G., Liu, O., Whittington, R. J., & Hick, P. M. (2021). Reduction in Virulence over Time in Ostreid herpesvirus 1 (OsHV-1) Microvariants between 2011 and 2015 in Australia. Viruses, 13(5), 946. https://doi.org/10.3390/v13050946

Comps, M. (1988). Epizootic diseases of oysters associated with viral infections. American Fisheries Society Special Publication, 18, Article 18.

Cuevas, J. M., Willemsen, A., Hillung, J., Zwart, M. P., & Elena, S. F. (2015). Temporal Dynamics of Intrahost Molecular Evolution for a Plant RNA Virus. Molecular Biology and Evolution, 32(5), 1132–1147. https://doi.org/10.1093/molbev/msv028

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience, 10(2), giab008. https://doi.org/10.1093/gigascience/giab008

Davison, A. J., Trus, B. L., Cheng, N., Steven, A. C., Watson, M. S., Cunningham, C., Deuff, R.-M. L., & Renault, T. (2005). A novel class of herpesvirus with bivalve hosts. Journal of General Virology, 86(1), Article 1. https://doi.org/10.1099/vir.0.80382-0

de Lorgeril, J., Petton, B., Lucasson, A., Perez, V., Stenger, P.-L., Dégremont, L., Montagnani, C., Escoubas, J.-M., Haffner, P., Allienne, J.-F., Leroy, M., Lagarde, F., Vidal-Dupiol, J., Gueguen, Y., & Mitta, G. (2020). Differential basal expression of immune genes confers Crassostrea gigas resistance to Pacific oyster mortality syndrome. BMC Genomics, 21(1), 63. https://doi.org/10.1186/s12864-020-6471-x

Dégremont, L. (2011). Evidence of herpesvirus (OsHV-1) resistance in juvenile Crassostrea gigas selected for high resistance to the summer mortality phenomenon. Aquaculture, 317(1–4), 94–98. https://doi.org/10.1016/j.aquaculture.2011.04.029

Dégremont, L., Azéma, P., Maurouard, E., & Travers, M.-A. (2020). Enhancing resistance to Vibrio aestuarianus in Crassostrea gigas by selection. Aquaculture, 526, 735429. https://doi.org/10.1016/j.aquaculture.2020.735429

Dégremont, L., Bédier, E., Soletchnik, P., Ropert, M., Huvet, A., Moal, J., Samain, J.-F., & Boudry, P. (2005). Relative importance of family, site, and field placement timing on survival, growth, and yield of hatchery-produced Pacific oyster spat (Crassostrea gigas). Aquaculture, 249(1), 213–229. https://doi.org/10.1016/j.aquaculture.2005.03.046

Dégremont, L., Ernande, B., Bédier, E., & Boudry, P. (2007). Summer mortality of hatchery-produced Pacific oyster spat (Crassostrea gigas). I. Estimation of genetic parameters for survival and growth. Aquaculture, 262(1), 41–53. https://doi.org/10.1016/j.aquaculture.2006.10.025

Dégremont, L., Maurouard, E., Ledu, C., & Benabdelmouna, A. (2019). Synthesis of the “PLAN DE SAUVEGARDE” using selected all-triploid oysters to reduce the shortage of spat in France due to OsHV-1–associated mortality in Crassostrea gigas. Aquaculture, 505, 462–472. https://doi.org/10.1016/j.aquaculture.2019.03.014

Delmotte, J., Chaparro, C., Galinier, R., de Lorgeril, J., Petton, B., Stenger, P.-L., Vidal-Dupiol, J., Destoumieux-Garzon, D., Gueguen, Y., Montagnani, C., Escoubas, J.-M., & Mitta, G. (2020). Contribution of Viral Genomic Diversity to Oyster Susceptibility in the Pacific Oyster Mortality Syndrome. Frontiers in Microbiology, 11. https://doi.org/10.3389/fmicb.2020.01579

Delmotte, J., Pelletier, C., Morga, B., Galinier, R., Petton, B., Lamy, J.-B., Kaltz, O., Avarre, J.-C., Jacquot, M., Montagnani, C., & Escoubas, J.-M. (2022). Genetic diversity and connectivity of the Ostreid herpesvirus 1 populations in France: A first attempt to phylogeographic inference for a marine mollusc disease. Virus Evolution, 8(1), veac039. https://doi.org/10.1093/ve/veac039

Elena, S. F., & Sanjuán, R. (2005). Adaptive Value of High Mutation Rates of RNA Viruses: Separating Causes from Consequences. Journal of Virology, 79(18), 11555–11558. https://doi.org/10.1128/JVI.79.18.11555-11558.2005

Evans, O., Hick, P., & Whittington, R. J. (2017). Detection of Ostreid herpesvirus -1 microvariants in healthy Crassostrea gigas following disease events and their possible role as reservoirs of infection. Journal of Invertebrate Pathology, 148, 20–33. https://doi.org/10.1016/j.jip.2017.05.004

Farahpour, F., Saeedghalati, M., & Hoffmann, D. (2022). MullerPlot: Generates Muller Plot from Population/Abundance/Frequency Dynamics Data. R package version 0.1.3. (0.1.3) [Computer software].

Farley, C. A., Banfield, W. G., Kasnic, G., & Foster, W. S. (1972). Oyster Herpes-Type Virus. Science, 178(4062), Article 4062. https://doi.org/10.1126/science.178.4062.759

Froissart, R., Wilke, C. O., Montville, R., Remold, S. K., Chao, L., & Turner, P. E. (2004). Co-infection Weakens Selection Against Epistatic Mutations in RNA Viruses. Genetics, 168(1), 9–19. https://doi.org/10.1534/genetics.104.030205

Fuandila, N. N., Gosselin-Grenet, A.-S., Tilak, M.-K., Bergmann, S. M., Escoubas, J.-M., Klafack, S., Lusiastuti, A. M., Yuhana, M., Fiston-Lavier, A.-S., Avarre, J.-C., & Cherif, E. (2022). Structural variation turnovers and defective genomes: Key drivers for the in vitro evolution of the large double-stranded DNA koi herpesvirus (KHV). Peer Community Journal, 2, e44. https://doi.org/10.24072/pcjournal.154

Gagnaire, B., Gay, M., Huvet, A., Daniel, J.-Y., Saulnier, D., & Renault, T. (2007). Combination of a pesticide exposure and a bacterial challenge: In vivo effects on immune response of Pacific oyster, Crassostrea gigas (Thunberg). Aquatic Toxicology, 84(1), 92–102. https://doi.org/10.1016/j.aquatox.2007.06.002

Gandon, S., & Michalakis, Y. (2000). Evolution of parasite virulence against qualitative or quantitative host resistance. Proceedings. Biological Sciences, 267(1447), 985–990. https://doi.org/10.1098/rspb.2000.1100

Kassambara, A., & Mundt, F. (2020). Extract and Visualize the Results of Multivariate Data Analyses [R package factoextra v1.0.7]. https://rpkgs.datano*via*.com/factoextra/index.html

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research, 30(14), Article 14. https://doi.org/10.1093/nar/gkf436

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., & Whitlock, M. C. (2012). Experimental evolution. Trends in Ecology & Evolution, 27(10), 547–560. https://doi.org/10.1016/j.tree.2012.06.001

Kubinak, J. L., & Potts, W. K. (2013). Host resistance influences patterns of experimental viral adaptation and virulence evolution. Virulence, 4(5), 410–418. https://doi.org/10.4161/viru.24724

Kun, Á., Hubai, A. G., Král, A., Mokos, J., Mikulecz, B. Á., & Radványi, Á. (2023). Do pathogens always evolve to be less virulent? The virulence–transmission trade-off in light of the COVID-19 pandemic. Biologia Futura, 1–12. https://doi.org/10.1007/s42977-023-00159-2

Kuny, C. V., Bowen, C. D., Renner, D. W., Johnston, C. M., & Szpara, M. L. (2020). In vitro evolution of herpes simplex virus 1 (HSV-1) reveals selection for syncytia and other minor variants in cell culture. Virus Evolution, 6(1), veaa013. https://doi.org/10.1093/ve/veaa013

Kutnjak, D., Elena, S. F., & Ravnikar, M. (2017). Time-Sampled Population Sequencing Reveals the Interplay of Selection and Genetic Drift in Experimental Evolution of Potato Virus Y. Journal of Virology, 91(16), e00690-17. https://doi.org/10.1128/JVI.00690-17

Lauring, A. S. (2020). Within-host Viral Diversity, a Window into Viral Evolution. Annual Review of Virology, 7(1), 63–81. https://doi.org/10.1146/annurev-virology-010320-061642

Li, H., & Roossinck, M. J. (2004). Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. Journal of Virology, 78(19), 10582–10587. https://doi.org/10.1128/JVI.78.19.10582-10587.2004

McCrone, J. T., & Lauring, A. S. (2018). Genetic bottlenecks in intraspecies virus transmission. Current Opinion in Virology, 28, 20–25. https://doi.org/10.1016/j.coviro.2017.10.008

Montville, R., Froissart, R., Remold, S. K., Tenaillon, O., & Turner, P. E. (2005). Evolution of Mutational Robustness in an RNA Virus. PLOS Biology, 3(11), e381. https://doi.org/10.1371/journal.pbio.0030381

Morga, B., Jacquot, M., Pelletier, C., Chevignon, G., Dégremont, L., Biétry, A., Pepin, J.-F., Heurtebise, S., Escoubas, J.-M., Bean, T. P., Rosani, U., Bai, C.-M., Renault, T., & Lamy, J.-B. (2021). Genomic Diversity of the Ostreid Herpesvirus Type 1 Across Time and Location and Among Host Species. Frontiers in Microbiology, 12, 711377. https://doi.org/10.3389/fmicb.2021.711377

Nicolas, J.-L., Comps, M., & Cochennec, N. (1992). Herpes-like virus infecting Pacific oyster larvae. (1). 12(1), Article 1.

Patterson, N., Price, A. L., & Reich, D. (2006). Population Structure and Eigenanalysis. PLOS Genetics, 2(12), e190. https://doi.org/10.1371/journal.pgen.0020190

Peck, K. M., & Lauring, A. S. (2018). Complexities of Viral Mutation Rates. Journal of Virology, 92(14), 10.1128/jvi.01031-17. https://doi.org/10.1128/jvi.01031-17

Pelletier, C., Chevignon, G., Faury, N., Arzul, I., Garcia, C., Chollet, B., Renault, T., Morga, B., & Jacquot, M. (2023). Genetic differentiation and host specialization among OsHV-1 infecting two oyster species in France [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2023.08.23.554398

Pepin, J. F., Riou, A., & Renault, T. (2008). Rapid and sensitive detection of ostreid herpesvirus 1 in oyster samples by real-time PCR. Journal of Virological Methods, 149(2), Article 2. https://doi.org/10.1016/j.jviromet.2008.01.022

Pernet, F., Fuhrmann, M., Petton, B., Mazurié, J., Bouget, J.-F., Fleury, E., Daigle, G., & Gernez, P. (2018). Determination of risk factors for herpesvirus outbreak in oysters using a broad-scale spatial epidemiology framework. Scientific Reports, 8(1), Article 1. https://doi.org/10.1038/s41598-018-29238-4

Petton, B., Boudry, P., Alunno-Bruscia, M., & Pernet, F. (2015). Factors influencing disease-induced mortality of Pacific oysters Crassostrea gigas. Aquaculture Environment Interactions, 6(3), Article 3. https://doi.org/10.3354/aei00125

R Development Core Team. (2005). R: A language and environment for statistical computing. [Computer software]. http://www.R-project.org

Renault & Arzul. (2001). Herpes-like virus infections in hatchery-reared bivalve larvae in Europe: Specific viral DNA detection by PCR. Journal of Fish Diseases, 24(3), Article 3. https://doi.org/10.1046/j.1365-2761.2001.00282.x

Renault, T. (1998). Infection herpétiques chez les invertébrés: Détection de virus de type herpès chez les mollusques bivalves marins. Annales de l'Institut Pasteur., 2, Article 2. https://archimer.ifremer.fr/doc/00000/2903/

Renault, T., Cochennec, R., Le Deuff, R. M., & Chollet, B. (1994). HERPES-LIKE VIRUS INFECTING JAPANESE OYSTER (CRASSOSTREA GIGAS) SPAT. 3.

Renault, T., Lipart, C., & Arzul, I. (2001). A herpes-like virus infects a non-ostreid bivalve species: Virus replication in Ruditapes philippinarum larvae. Diseases of Aquatic Organisms, 45, 1–7. https://doi.org/10.3354/dao045001

Sanjuán, R. (2012). From Molecular Genetics to Phylodynamics: Evolutionary Relevance of Mutation Rates Across Viruses. PLOS Pathogens, 8(5), e1002685. https://doi.org/10.1371/journal.ppat.1002685

Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. Cellular and Molecular Life Sciences, 73(23), 4433–4448. https://doi.org/10.1007/s00018-016-2299-6

Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., & Belshaw, R. (2010). Viral Mutation Rates. Journal of Virology, 84(19), 9733–9748. https://doi.org/10.1128/JVI.00694-10

Sanjuán, R., Pereira-Gómez, M., & Risso, J. (2016). Chapter 3—Genome Instability in DNA Viruses. In I. Kovalchuk & O. Kovalchuk (Eds.), Genome Stability (pp. 37–47). Academic Press. https://doi.org/10.1016/B978-0-12-803309-8.00003-3

Schikorski, D., Renault, T., Saulnier, D., Faury, N., Moreau, P., & Pépin, J.-F. (2011). Experimental infection of Pacific oyster Crassostrea gigas spat by ostreid herpesvirus 1: Demonstration of oyster spat susceptibility. Veterinary Research, 42(1), Article 1. https://doi.org/10.1186/1297-9716-42-27

Spitzer, K., Pelizzola, M., & Futschik, A. (2020). Modifying the Chi-square and the CMH test for population genetic inference: Adapting to overdispersion. The Annals of Applied Statistics, 14(1). https://doi.org/10.1214/19-AOAS1301

Stoltzfus, A., & Norris, R. W. (2016). On the Causes of Evolutionary Transition:Transversion Bias. Molecular Biology and Evolution, 33(3), 595–602. https://doi.org/10.1093/molbev/msv274

van Sluijs, L., Pijlman, G. P., & Kammenga, J. E. (2017). Why do Individuals Differ in Viral Susceptibility? A Story Told by Model Organisms. Viruses, 9(10), 284. https://doi.org/10.3390/v9100284

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer New York. https://doi.org/10.1007/978-0-387-98141-3

Xia, J., Bai, C., Wang, C., Song, X., & Huang, J. (2015). Complete genome sequence of Ostreid herpesvirus-1 associated with mortalities of Scapharca broughtonii broodstocks. Virology Journal, 12(1), Article 1. https://doi.org/10.1186/s12985-015-0334-0

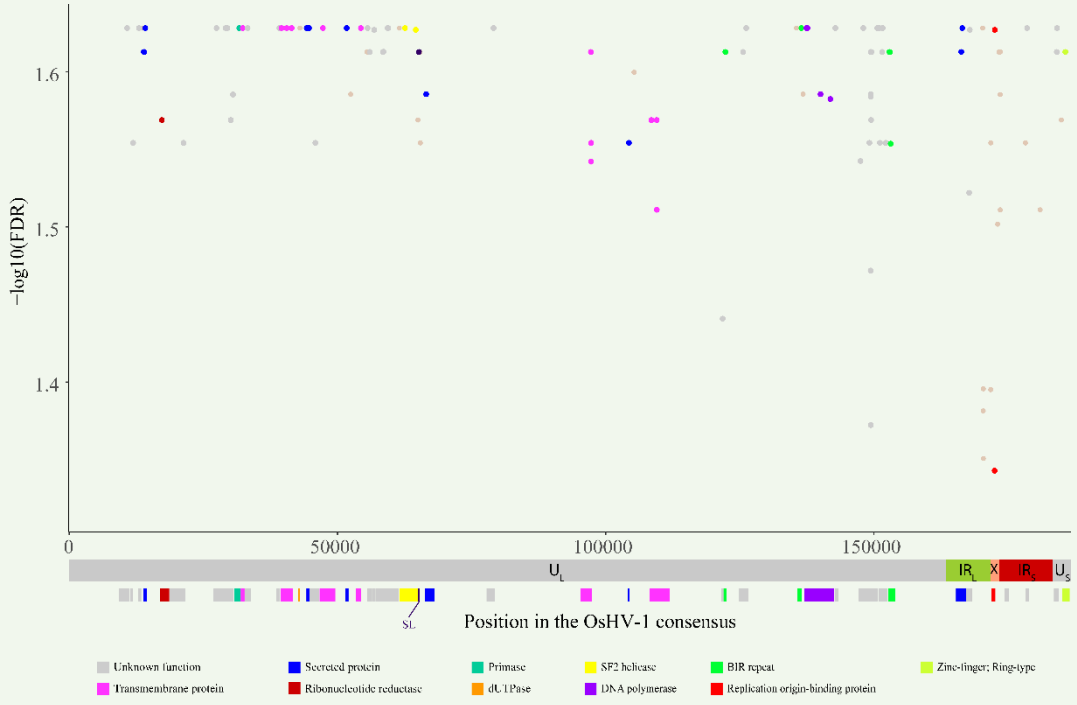
Messages clés :

- Une diminution significative de la mortalité des huîtres ainsi que de la charge virale a été observée sur plusieurs générations d'infection. Cette réduction de la mortalité peut être attribuée à la **sélection de variants viraux moins virulents**, ce qui suggère l'existence d'un compromis entre la virulence du virus et sa capacité de transmission.

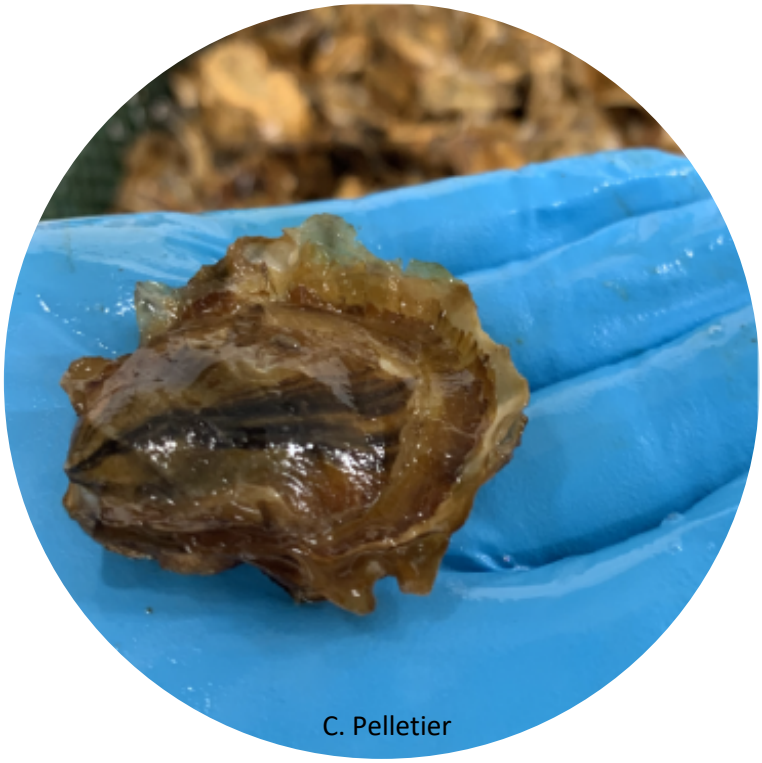
- Une tendance vers **l'homogénéisation génétique des populations virales** a été observée au fil des générations d'infection. Cela pourrait résulter de la sélection favorisant certains génotypes viraux ou d'une réduction de la diversité génétique due à des **goulots d'étranglement** de population.

- L'analyse des fréquences des SNPs a révélé une **différenciation génétique** distincte entre les différentes familles d'huîtres et les générations d'infection. **Les populations virales ont évolué au cours du temps**.

- Des signatures de **sélection positive** ont été identifiées dans le génome d'OsHV-1, en particulier dans les régions répétées et la région U_L. Ces signatures suggèrent que des éléments régulateurs et fonctionnels du virus sont soumis à une pression de sélection.



Discussion et Perspectives



C. Pelletier

Au début de ma thèse, il était impératif de développer des méthodologies permettant le séquençage massif à partir d'échantillons de tissus d'huîtres infectées par OsHV-1, puis de développer un pipeline d'analyse génétique et génomique de ses divers génotypes. Suite à ce développement méthodologique, j'ai pu caractériser la diversité des populations mondiales du virus OsHV-1, en prenant en considération différentes espèces hôtes et en suivant leur évolution temporelle. De plus, j'ai initié l'étude des mécanismes qui sous-tendent l'évolution génomique du virus OsHV-1 (Figure 37). En résumé, les questions abordées pendant ma thèse étaient les suivantes :

- ***i)*** Comment séquencer le virus OsHV-1 et analyser sa diversité ? (Figure 37A)
- ***ii)*** Est-ce que les lignées virales qui circulent chez deux espèces d'huîtres, *M. gigas* et *O. edulis,* sont les mêmes et est-ce que les événements de transmission inter-espèces sont fréquents ? (Figure 37B)
- ***iii)*** Quelle est la diversité spatio-temporelle du virus OsHV-1 chez *M. gigas* ? (Figure 37C)
- ***iv)*** Comment le virus OsHV-1 évolue en fonction du fond génétique de son hôte, *M.gigas* ? (Figure 37D)

L'ensemble des travaux que j'ai réalisé pendant cette période de doctorat ont permis de répondre à ces quatre questions et d'ouvrir de nouvelles perspectives de recherche sur OsHV-1.

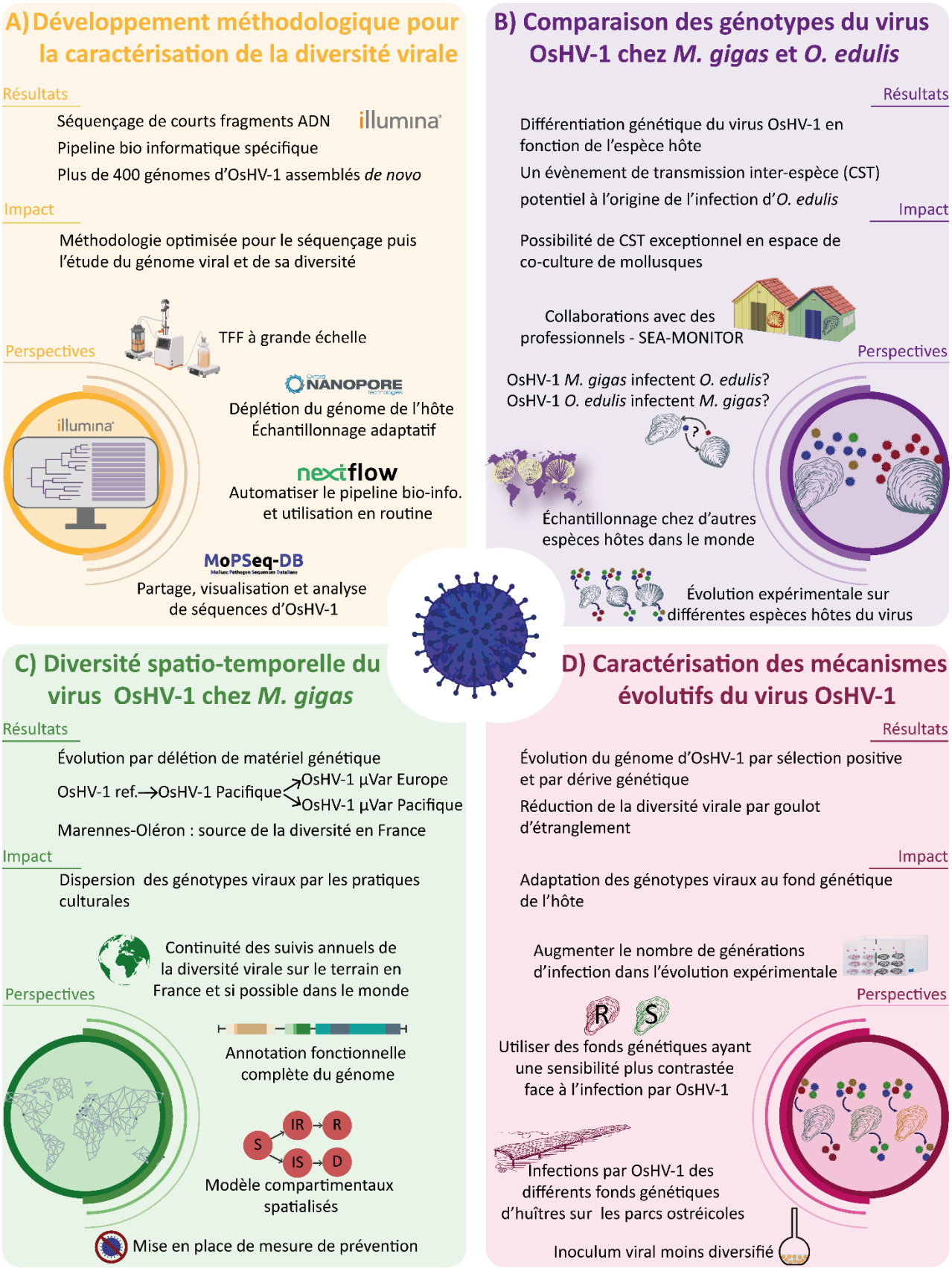


Figure 37 : Représentation graphique des principales problématiques abordées dans cette thèse, principaux résultats, impacts pour les professionnels collaborateurs et perspectives.

Levée des verrous technologiques associés à l'étude génomique du virus OsHV-1

Le séquençage du génome d'OsHV-1

Lorsque que j'ai débuté mes travaux de recherche sur OsHV-1, en 2019, l'acquisition de génomes pour ce type de virus utilisait des approches relativement complexes, non compatibles avec le besoin d'obtenir plusieurs centaines de génomes par an. En effet et jusqu'à récemment, le séquençage du génome d'OsHV-1 nécessitait des étapes coûteuses et fastidieuses de purification des particules virales (Burioli et al., 2017; Davison et al., 2005; Ren et al., 2013; Xia et al., 2015). Ensuite, grâce à l'émergence des nouvelles technologies de séquençage (NGS) et plus particulièrement de l'essor de la technologie de séquençage Illumina, qui permet d'atteindre des profondeurs de séquençage sans précédent, le virus OsHV-1 a pu être séquencé directement à partir de tissus d'huîtres infectées (Abbadi et al., 2018). Mais cette dernière approche n'était pas encore optimale car la préparation de la librairie de séquençage reposait sur une méthode d'amplification de l'ADN et les « gaps » identifiés à la suite de l'assemblage étaient comblés par le séquençage d'amplicons *via* la méthode Sanger (Abbadi et al., 2018). Avant le début de ma thèse, notre laboratoire ASIM en collaboration avec le laboratoire IHPE, a commencé à obtenir des génomes d'OsHV-1 séquencés directement à partir de tissus d'huîtres infectées *via* la technologie Illumina et sans amplification de librairie (article 3 et article 4). Ces études pionnières nous ont permis de valider l'approche pour quelques dizaines de génomes d'OsHV-1, approche qui a ensuite été utilisée pour l'ensemble des génomes acquis durant mon doctorat.

En parallèle au déroulé de mon projet de thèse, les laboratoires ASIM et IHPE ont débuté l'évaluation de la technologie de séquençage ONT pour la caractérisation des génomes d'OsHV-1. Dans ce but, des particules virales d'OsHV-1 ont été purifiées par une approche de filtration à flux tangentiel (TFF) et leurs ADN ont été extrait puis séquencés *via* les technologies Illumina et Nanopore (article 1). Cette étude a permis de confirmer le potentiel incroyable du séquenceur portable ONT pour réaliser du séquençage au plus près des échantillons. De plus, cette approche long-fragment permet de faciliter l'assemblage du génome d'OsHV-1 et d'identifier les différents isomères viraux. Cependant le taux d'erreurs au sein des séquences ne permet pas d'atteindre le niveau de confiance nécessaire à des analyses de type intra-individuelles qui m'ont permis d'étudier les co-infections virales.

C'est lors de cette évaluation de la technologie de séquençage ONT que j'ai eu l'opportunité de développer le pipeline bio-informatique (Figure 38) nécessaire au traitement des données Illumina pour l'assemblage des génomes d'OsHV-1 et qui a ensuite été utilisé tout au long de ma thèse.

L'assemblage du génome d'OsHV-1

Le pipeline bioinformatique que j'ai développé m'a permis d'assembler plus de 400 génomes du virus OsHV-1 (Figure 38). Spécifiquement, 40 nouveaux génomes d'OsHV-1 ont été assemblés pour l'étude de la différentiation génétique (article 2) du virus chez O. *edulis* (n=23 génomes) et chez *M. gigas* (n=17 génomes), 279 nouveaux génomes d'OsHV-1 ont été assemblés pour l'étude spatio-temporelle du virus OsHV-1 dans le monde entre 1993 et 2022 (article 3, article 4 et article 5). Une centaine d'autres génomes du virus OsHV-1 ont été assemblés, mais ne sont pas encore valorisés. Ce pipeline m'a également permis d'assembler 90 génomes d'OsHV-1 produit durant l'évolution expérimentale (article 6). C'est une avancée considérable pour l'étude du génome viral d'OsHV-1, puisqu'au début de ma thèse, seulement six génomes viraux étaient disponibles dans les bases de données. L'ensemble de ces génomes a été intégré à une base de données, développée et gérée par l'unité ASIM, pour être partagés avec la communauté scientifique. Cette base, nommée « Mollusc Pathogen Sequences DataBase » (MoPSeq-DB) sera bientôt disponible en ligne (Battistel, 2023) et référence les génomes de différents agents pathogènes, et notamment ceux du virus OsHV-1 (Figure 39, Batistel *et al.,* in prep).

Ce pipeline est donc fonctionnel et efficace pour assembler des centaines de génomes d'OsHV-1, mais il devra être optimisé et automatisé afin de le rendre accessible et facilement utilisable par la communauté scientifique. En effet, la redondance des régions répétées dans le génome viral et la présence des différents isomères génomiques rendent impossible la résolution de ces jonctions par les approches de séquençage de courts fragments d'ADN de type Illumina. De plus, quelques bases sont manquantes entre la jonction de certains contigs et doivent donc être ajoutées manuellement après alignement des séquences sur les contigs. Ceci s'explique par l'utilisation de l'outil d'assemblage *de novo* Spades qui produit un assemblage de contigs en testant plusieurs k-mers (*i.e.* sous-chaînes de longueur définie, multiple de trois, dans une séquence biologique) et sélectionne celui qui est optimal pour le résultat final. La base manquante est donc expliquée par le fait qu'elle n'est pas comprise dans les tri-nucléotides testés par les k-mer. Une étape manuelle et fastidieuse est donc actuellement nécessaire pour joindre les différents contigs produits et pour ajouter les bases manquantes lors de l'étape d'assemblage *de novo*. Plusieurs outils permettant l'extension des contigs ont été testés mais n'ont pas donnés de résultats satisfaisants. Pour résoudre ce problème, il serait nécessaire de développer un script qui prendrait en compte l'alignement des séquences virales brutes sur les contigs générés accompagné d'un assemblage guidé utilisant un outil d'alignement sur un génome de référence.

Concernant l'automatisation de ce pipeline, il pourrait être réécrit en NextFlow, un ordonnanceur de taches qui permet : *i)* de mieux gérer les dépendances des logiciels utilisés, *ii)* de paralléliser les différentes tâches pour accélérer l'exécution du script, *iii)* d'assurer une reproductibilité des analyses et *iv)* de permettre la portabilité du script sur différents clusters de calcul (Di Tommaso et al., 2017).

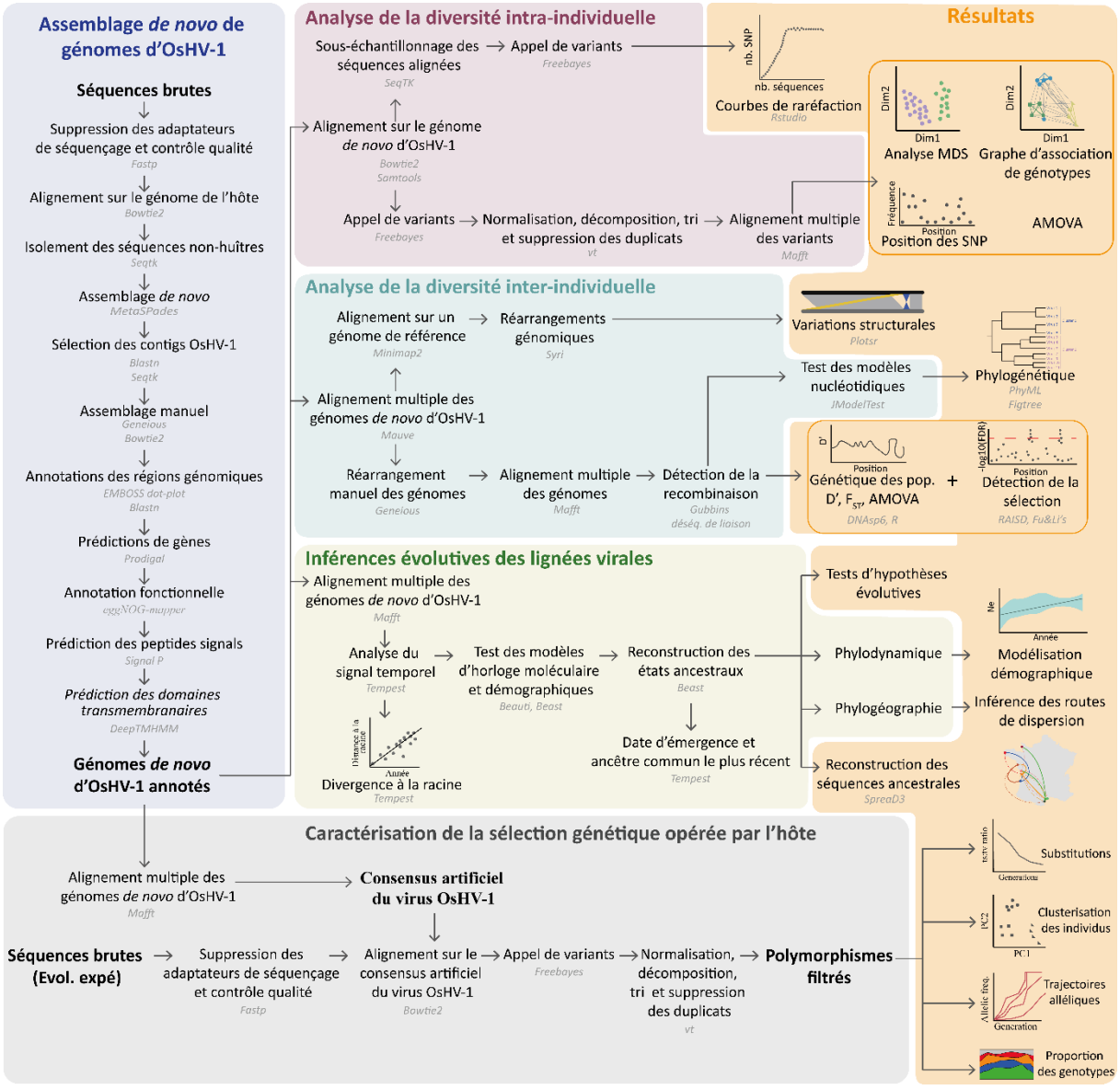


Figure 38 : Pipeline bionformatique utilisé pour l'analyse génomique du virus OsHV-1.

Ce pipeline permet d'assembler *de novo* des génomes du virus OsHV-1 (partie bleue), puis d'analyser sa diversité à l'échelle intra- (partir violette) et inter-individuelle (partie bleu turquoise) puis d'effectuer des inférences de phylodynamiques pour suivre l'évolution des lignées virales (partie verte). Enfin, il permet également d'étudier les pressions de sélection et la dérive génétique opérée par l'espèce hôte au cours d'une évolution expérimentale (partie grise). A droite sont indiqués les résultats qui peuvent être obtenus grâce à ce pipeline (partie orange).

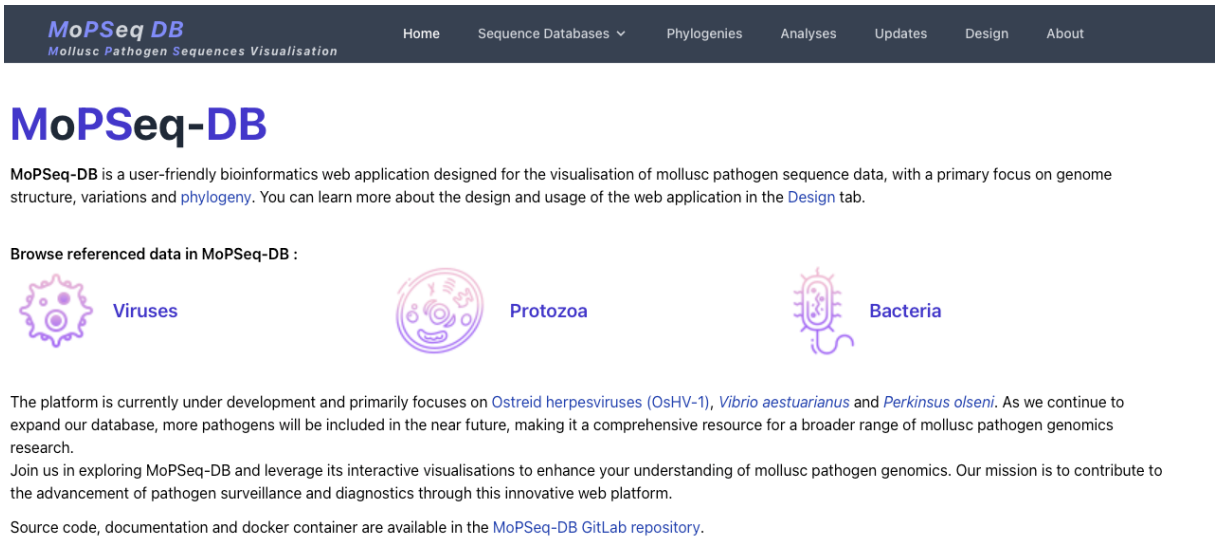


Figure 39 : Capture d'écran de la page d'accueil de la plateforme web MoPSeq-DB développée par ASIM pour la gestion et la visualisation des données de séquences des mollusques marins.

OsHV-1 évolue de façon mesurable

Les variations structurales : moteur de l'évolution du virus OsHV-1

Le génome du virus OsHV-1 a évolué entre sa première caractérisation effectuée sur des échantillons collectés en 1992 chez *M. gigas* (Davison *et al.*, 2005) et aujourd'hui. A l'échelle du génome complet, cette évolution se fait principalement au travers de réarrangement et de délétion de matériel génétique.

Premièrement, la comparaison des génotypes viraux collectés à l'échelle mondiale chez *M. gigas*, a mis en évidence une continuité dans l'évolution de l'architecture génomique du virus OsHV-1 au cours du temps (Figure 40, article 5). Le premier génotype à avoir émergé et qui a été détecté exclusivement entre 1991 et 2008, est le génotype de référence caractérisé par un génome de 207 kb organisé en régions uniques et répétées. Cette organisation peut être résumée sous la forme TR_L-U_L-IR_L-X-IR_S-U_S-TR_S. Certains des génomes viraux contiennent une inversion d'une portion de 4 kb dans la région U_L. Ces génotypes avaient déjà été caractérisés lors de la publication du premier génome d'OsHV-1 (Davison *et al.*, 2005). Par la suite, deux génotypes ont émergé dans l'océan Pacifique (OsHV-1 microvariants, C.-M. Bai et al., 2019; Burge et al., 2021; Jenkins et al., 2013; Keeling et al., 2014b; Shimahara et al., 2012) et en Europe (OsHV-1 µVar, (Burioli et al., 2017; Renault et al., 2012; Segarra et al., 2010). Dans l'océan Pacifique (USA, Corée, Australie et Nouvelle-Zélande), le génotype présente

d'importantes insertions et délétions dans la région U_L et dans les régions répétées induisant une variation de 5,5 kb dans le génome. De plus, les génomes présentent une inversion accompagnée d'une translocation de la fin de la région IRS. Dans notre jeu de données, ce génotype a été détecté à partir de 2006 aux USA et à partir de 2011 en Australie et Nouvelle-Zélande. En Corée et Nouvelle-Zélande entre 2011 et 2018, OsHV-1 a finalement encore évolué par l'intermédiaire de délétions supplémentaires, notamment dans la région U_L. Ces résultats sont en adéquation avec ceux observés dans de précédentes études qui utilisaient des approches de PCR pour caractériser les microvariants du virus (C.-M. Bai et al., 2019; Burge et al., 2021; Jenkins et al., 2013; S. Keeling et al., 2014a; Shimahara et al., 2012). En Europe (France, Allemagne, Espagne, Irlande, Italie), le génotype OsHV-1 de référence a évolué vers le génotype OsHV-1 µVar qui a déjà été très bien décrit (Burioli *et al.*, 2017; Renault *et al.*, 2014; Segarra *et al.*, 2010). Ce génotype est caractérisé par une délétion spécifique de 12 pb dans le microsatellite localisé en amont de l'ORF4. De plus, tout comme pour les génotypes détectés dans l'océan Pacifique, ce génotype présente de nombreuses délétions dans la région U_L et dans les régions répétées. Au total, environ 10 kb ont été impactées par des insertions ou délétions. Aucune trace de la présence d'OsHV-1 µVar avant son émergence en 2008 n'a pu être trouvé. De la même manière, le génotype de référence n'a plus jamais été retrouvé après cette date. OsHV-1 µVar a donc très rapidement totalement remplacé le génotype de référence en Europe. Sachant que, malgré les mortalités massives observées en 2008, la production d'huître est restée importante (grande taille efficace de la population d'hôte), le remplacement et/ou l'extinction du génotype de référence est probablement la conséquence d'un avantage sélectif du génotype délété OsHV-1 µVar.

Deuxièmement, les génotypes infectant l'huître plate *O. edulis* sont structuralement très différents de ceux infectant *M. gigas* (Figure 40, article 2). L'analyse des variations structurales a permis d'identifier neufs architectures génomiques distinctes au sein des génomes d'OsHV-1 infectant *M. gigas* (n=6, article 2 et article 5) et *O. edulis* (n=3, article 2). Ces génotypes sont différenciés par de grandes insertions-délétions (Indels), des translocations génomiques et des inversions, principalement dans les régions U_L et IR_L, entraînant ainsi des variations de taille totale d'environ 6 kpb entre les génomes représentatifs de ces génotypes. Cependant, les génomes d'OsHV-1 prélevés chez *M. gigas* et chez *O. edulis* ne présentent pas de délétion ou d'insertion dans les mêmes portions génomiques. Toutefois, en raison de différences notables dans les variations structurales du génome viral détectées chez ces deux espèces d'huîtres et de la limitation de la taille de l'échantillon, il n'est pas possible de tirer des conclusions définitives sur l'ordre d'apparition des génotypes viraux chez *M. gigas* et *O. edulis*.

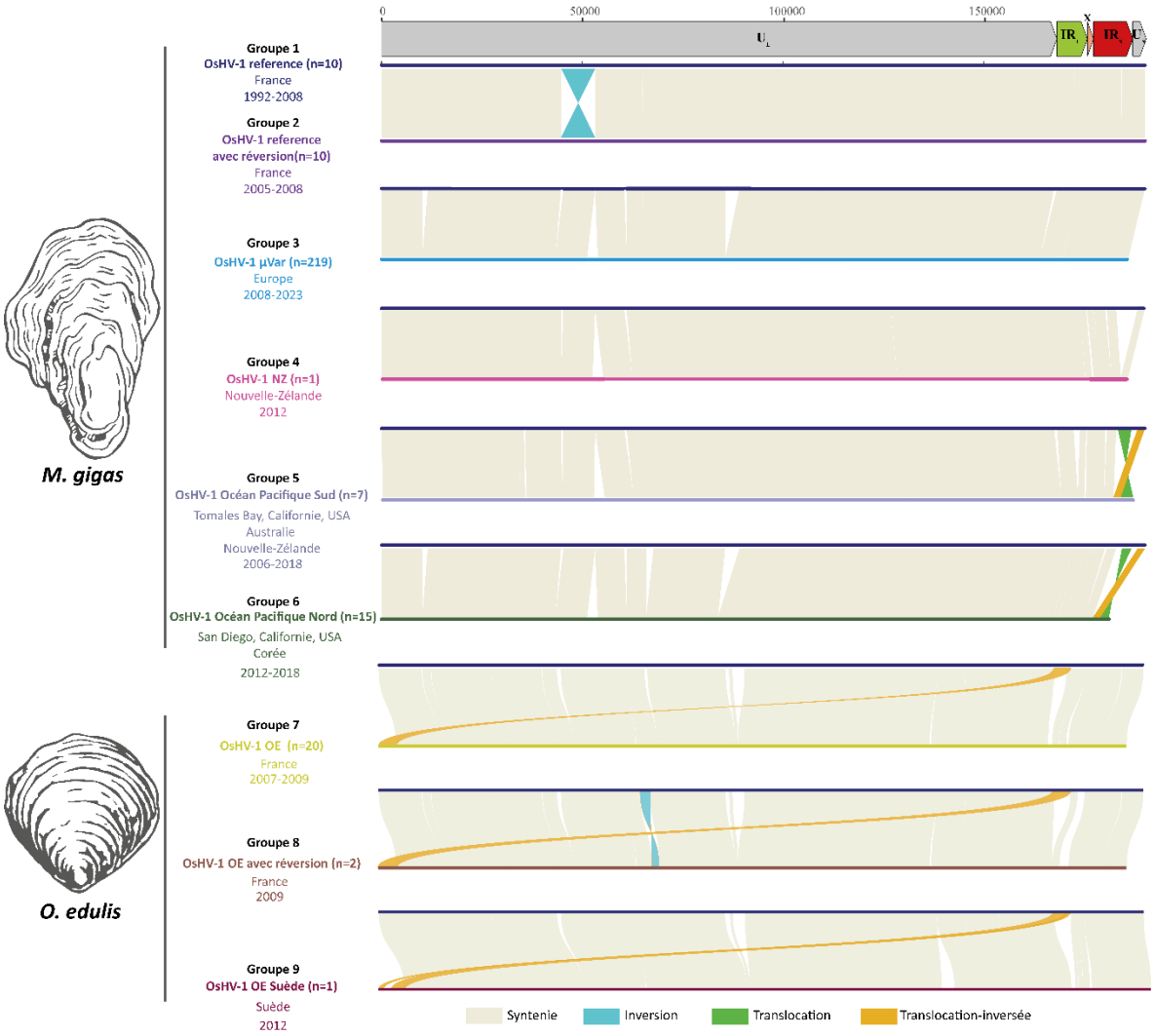


Figure 40 : Génomique comparative des 9 génomes représentatifs des génotypes du virus OsHV-1 mis en évidence durant ma thèse.

Six génotypes ont été détectés chez *M. gigas* et trois génotypes ont été détectés chez *O. edulis*. La description des variations structurales du génome était basée sur une comparaison par paires avec le génome de référence d'OsHV-1 (NC_005881.2, Davison *et al*. 2005), effectuée à l'aide de l'outil Mummer. Dans chaque comparaison par paires, le génome de référence est représenté par une barre bleu foncé. Les régions synténiques sont représentées en beige, les inversions en bleu clair, les translocations en vert et les translocations inversées en orange. Au-dessus du graphique des variations structurales se trouve l'architecture et la position dans le génome de référence d'OsHV-1.

Ces génomes évoluent par réarrangements génomiques, notamment des délétions, des translocations ou inversions. Le mécanisme d'apparition de ces réarrangements est pour l'instant inconnu, mais il est probable qu'ils soient associés à la réplication ou à des événements de recombinaison. Les taux de recombinaison n'ont pour l'instant pas été caractérisé, mais il serait nécessaire de les étudier, notamment par les approches de ratio r/m (section 1.2.2.2) ou de coalescent structuré (section 1.3.2.6) pour étblir la fréquence d'apparition de ces évènements.

Les isomères du génome viral qui ont été mis en évidence dans l'article 1 sont également considérés comme des réarrangements génomiques. Chez d'autres *Herpesvirus*, notamment le virus *herpès simplex 1*, ces réarrangements sont étroitement liés aux mécanismes de réplication et de recombinaison au sein de ces virus, se produisant tôt dans leur cycle de vie. Lors de la réplication de l'ADN, l'apparition de nombreuses cassures aléatoires des brins double est un catalyseur efficace d'événements de recombinaison, conduisant finalement à la création de nouveaux réarrangements génomiques (Pérez-Losada *et al.*, 2015). Ces réarrangements génomiques peuvent avoir des conséquences importantes sur la fonctionnalité des gènes présents dans le génome viral, mais le niveau d'annotation actuel du génome viral ne permet pas de tirer de réelles conclusions sur l'impact de ces événements sur les populations virales. Il est donc nécessaire et urgent de procéder à l'annotation fonctionnelle précise des génomes d'OsHV-1 par des approches de transcriptomiques et de protéomiques, afin de mieux mesurer l'impact des réarrangements fréquents observés.

OsHV-1 évolue à des taux de substitution plus élevés qu'attendu

Les taux moyen d'évolution d'OsHV-1 inférés au cours de ma thèse, présentent d'importantes différences suivant les échantillonnages utilisés (6,787E-05 subs/site/an (article 3) et 3,6093E-03 subs/site/an (article 5)). Ceux-ci sont plus élevés que ceux attendus selon la théorie de la coévolution du virus avec les espèces hôtes. Par exemple, pour divers *Herpesvirus*, notamment le virus de l'herpès simplex 1 (HSV1), le virus de la varicelle-zona (VZV) et le virus de la maladie de Marek (MDV), les taux d'évolution ont été estimés à 8,21E-05 subs/site/an, 6,56E-06 subs/site/an et 8,25E-06 subs/site/an, respectivement (Sanjuán, 2012).

Des biais potentiels inhérents aux méthodes d'estimations de ces taux ont été démontrés (Debruyne & Poinar, 2009; Emerson, 2007; S. Y. W. Ho, 2005). En particulier, il a été montré que les taux d'évolution moléculaire sont accélérés lorsque les études sont menées à de courtes échelles de temps (Ho, 2005). L'explication la plus probable de ce phénomène est la persistance de mutations légèrement délétères, qui, grâce à la sélection purificatrice, auraient été éliminées sur des échelles de temps plus longues (S. Ho & Larson, 2006; S. Y. W. Ho et al., 2005; S. Y. W. Ho, Kolokotronis, et al., 2007; S. Y. W. Ho, Shapiro, et al., 2007). Dans le cadre de l'étude la diversité globale du virus OsHV-1 (article 5), la couverture temporelle de l'échantillonnage est principalement limitée aux cinq dernières années. Une dépendance temporelle de l'estimation de l'horloge moléculaire est donc possible et les taux de substitutions sur-estimés.

Cependant, ces résultats sont cohérents avec des études soutenant que les taux évolutifs de certains virus à ADN double brin sont comparables à ceux des virus à ARN et à simple brin d'ADN (Hughes et al., 2010; Patrício et al., 2012). Pour OsHV-1, des taux d'évolution rapides peuvent s'expliquer par des densités d'huîtres élevées (grande taille effective) (Garcia et al., 2011). De plus, la cohabitation des

huîtres creuses avec diverses espèces de mollusques bivalves cultivées, telles que les moules, d'autres espèces d'huîtres, les ormeaux, les palourdes ou les coques, favorise potentiellement l'occurrence d'évènements de recombinaison entre les différents génotypes co-infectant ces espèces hôtes (Casto *et al.*, 2020; Crespo-Bellido *et al.*, 2021) et conduit à l'apparition de nouvelles combinaisons alléliques. En effet, il est important de noter que chaque individu peut être infecté par de multiples génotypes viraux (article 2, article 3 et article 4). Deux hypothèses, non exclusives, peuvent être formulées pour expliquer cette observation. Premièrement, la diversité observée au sein d'un individu hôte peut être liée à une co-infection par plusieurs lignées virales en circulation. La seconde hypothèse concerne l'apparition fréquente de mutations *de novo* au sein des individus hôte au cours de l'infection (Figure 6) pouvant donner naissance à de nouvelles lignées virales (Sanjuán & Domingo-Calap, 2016). Tandis que l'estimation du nombre de lignées circulantes au cours d'un épisode de mortalités pourrait s'effectuer *via* l'utilisation de la filtration tangentielle sur de grands volumes d'eau de mer prélevée à proximité des parcs à huîtres, l'apparition possible de nouvelles lignées virales, sur de courtes échelles de temps (20 générations en moyenne) a déjà été montré au cours de l'évolution expérimentale (article 6), concomitante avec l'exercice de pressions de sélection et de la dérive génétique.

Effets de la sélection et de la dérive génétique

Pour étudier les mécanismes évolutifs du virus OsHV-1 en fonction la lignée d'huîtres *M. gigas*, une évolution expérimentale a été mise en place pour la première fois sur le modèle *M. gigas* et OsHV-1 (article 6). Durant cette expérimentation, trois fonds génétiques présentant un niveau de sensibilité différent à l'infection par le virus OsHV-1 ont été infectés avec un inoculum viral hétérogène, génétiquement très diversifié. De nombreux échantillons d'huîtres infectées par OsHV-1 ont été séquencés pour suivre l'évolution des fréquences alléliques au cours des générations et déterminer les pressions évolutives qui s'exercent sur les populations (Figure 38C).

Durant cette expérimentation, une diminution à la fois de la mortalité des huîtres et de la charge virale a été observée au fil des générations successives d'infection (article 6). Cette diminution de la mortalité peut s'expliquer par la sélection de variants viraux moins virulents. Il est bien établi qu'il existe un équilibre délicat dans les interactions entre un pathogène et son hôte, où une virulence réduite peut favoriser une transmission plus efficace (Anderson & May, 1982). Il est donc envisageable, que la diminution des taux de mortalités observés au cours des générations, s'explique par la sélection des lignées virales les mois virulentes.

De plus, au cours des générations d'infections, la diversité virale tend vers une homogénéisation pour l'ensemble des familles. Cette homogénéisation se traduit par une diminution du nombre de polymorphismes détectés et du ratio transition:transversion. Des résultats similaires ont été observés chez différents virus infectant le tabac et les pommes de terre (Kutnjak et al., 2017; H. Li & Roossinck, 2004). Ces résultats indiquent un probable impact de la dérive génétique dans l'épuration des populations virales au sein des trois familles. Considérant la taille des populations virales dans les bacs d'infection il est plus que plausible qu'un goulot d'étranglement des populations ait eu lieu à cause des phénomènes de dérive génétique. Cependant, l'effet et la dynamique de ce vraissemblable goulot d'étranglement n'a pu être observé que partiellement, à la cinquième génération d'infection (séquençage à G0 puis G5). Séquencer les échantillons des générations intermédiaires permettrait de gagner en précision dans la description du phénomène.

Des signaux de sélection positive ont pu être détectés dans les régions répétés et la région U_L du génome viral uniquement pour la famille sensible. Il est possible que la réplication virale soit favorisée dans les huîtres les plus sensibles (Dégremont, 2011), conduisant à une augmentation des taux de mutation et à la fixation d'allèles avantageux au sein des populations virales (Elena & Sanjuán, 2005; Peck & Lauring, 2018). Ces résultats suggèrent donc un role clé de la selection dans l'évolution du virus.

En effet, l'analyse des pressions de sélection effectuée dans le cadre de l'évolution expérimentale (article 6), et la comparaison des génotypes viraux détectés chez *M. gigas* et *O. edulis* (article 2) a permis de détecter des signaux de sélection naturelle dans les régions codantes pour des protéines associées à des fonctions biologiques essentielles ou des domaines impliqués dans la liaison des particules virales aux cellules hôtes, la synthèse d'ADN, la réplication et l'emballage de l'ADN, ainsi que les protéines membranaires et transmembranaires. Ces résultats suggèrent que l'adaptation des protéines essentielles a conduit à la spécialisation des génotypes viraux vis-à-vis de leurs espèces hôtes respectives et laissent même imaginer l'existence potentielle de plusieurs « espèces virales » qui seraient responsables de l'infection respective de chaque espèce hôte (Peck & Lauring, 2018).

Isolement génétique et spécialisation d'hôte

L'analyse phylogénétique du virus OsHV-1 chez les deux espèces d'huîtres a révélé l'existence de deux clades distincts, l'un correspondant aux virus infectant *M. gigas* et l'autre aux virus infectant *O. edulis* (article 2). La comparaison interspécifique des génomes assemblés a mis en évidence des pourcentages d'identité variant de 91 % à 94 % entre les génomes de *M. gigas* et ceux de *O. edulis*. Des résultats similaires avaient été préalablement obtenus pour six génomes d'OsHV-1 prélevés chez différentes espèces de mollusques (C.-M. Bai et al., 2019). L'ancêtre commun le plus récent des lignées de ces

deux clades a été attribué à *M. gigas*, ce qui suggère que cette espèce hôte a joué un rôle essentiel dans l'évolution et l'émergence du génotype de référence. L'émergence du génotype d'OsHV-1 détecté chez *O. edulis* a été attribué à un évènement de transmission inter-espèce qui a été facilité par un évènement de recombinaison qui aurait eu lieu entre 1998 et 2004.

Pour qu'une transmission inter-espèce du virus donne lieu à un saut d'espèce, il est nécessaire que le génotype viral en question soit adapté à sa nouvelle espèce hôte et puisse se répliquer puis se transmettre efficacement aux autres individus (Longdon et al., 2014). Au vu de l'adaptation spécifique du virus OsHV-1 à *O. edulis* mis en évidence dans l'article 2, il semblerait que ce génotype ait été transmis à *O. edulis* par un saut d'espèce. Cependant, à l'heure actuelle, aucune donnée n'est disponible sur les taux de réplication du virus OsHV-1 chez *O. edulis*. Pour étudier les taux de réplication et d'excrétion des génotypes viraux majoritaires chez ces deux espèces hôtes, une infection expérimentale pourrait être imaginée accompagnée d'un suivi temporel de la charge virale dans les tissus de l'hôte et dans l'eau de mer. Cette étude a été amorcée durant ma thèse, mais aucune mortalité d'huîtres plates n'a été observée après injection intramusculaire de suspension virale générée à partir d'huîtres creuses infectées. Cependant, des quantifications de la charge virale dans les huîtres plates ont été réalisées toutes les 24h par qPCR, indiquant la présence de virus dans les tissus jusqu'à 3 jours post-injection. L'infection a donc bien eu lieu, mais le génotype utilisé semble non pathogène pour l'huître plate et laisse imaginer que le génotype circulant actuellement chez les huîtres creuses ne peut être transmis à *O. edulis* que par débordement d'espèce. L'infection des huîtres saines a été effectuée uniquement avec des génotypes viraux contemporains collectés chez l'huître du Pacifique *M. gigas* car aucune suspension de virus anciens ni infectant l'huître plate *O. edulis* ne sont disponibles pour réaliser une infection expérimentale. Ces résultats viennent tout de même confirmer l'hypothèse de spécialisation des génotypes du virus OsHV-1 en fonction de l'espèce hôte.

En Charente-Maritime et dans le bassin d'Arcachon par exemple, ces deux espèces d'huîtres vivent en sympatrie et ont des modes de vie similaires (Danic-Tchaleu et al., 2011). De même, les écloseries commerciales produisent du naissain d'huîtres creuses, mais également d'autres espèces de mollusques telles que les palourdes *R. decussatus* et *R. philipinarum* ou l'huître plate *O. edulis* (Rodgers et al., 2019). Malgré cette proximité entre les différentes espèces de mollusques bivalves, en milieu naturel ou en espace réduit de co-culture, les événements de transmission inter-espèces semblent très peu fréquents pour le virus OsHV-1 (article 2), puisque dans cette étude un seul événement de saut d'espèce a pu être détecté et aucun débordement d'espèce entre 1993 et 2022. Il ne semble donc pas y avoir de risque majeur liée à la co-culture des espèces de bivalves.

Les contributions relatives des différentes espèces hôtes à la diversité du virus OsHV-1 seront cependant importantes à estimer, pour mieux comprendre l'impact des hôtes sur la variabilité génétique de ce virus. Des modèles fréquentistes peuvent être utilisés pour effectuer ces estimations (Jackson et al., 2021). Ces modèles fournissent un cadre analytique solide pour évaluer de manière quantitative et statistiquement robuste les variations de la diversité virale en relation avec les différents hôtes. En combinant des données épidémiologiques, génomiques et environnementales, l'approche fréquentiste offre la possibilité de clarifier les relations complexes entre les espèces hôtes et le virus.

La figure 42 présente les relations évolutives entre les génotypes d'OsHV-1 collectés chez différentes espèces de bivalves, actuellement disponibles dans les bases de données. Très peu de génomes viraux assemblés chez *A. farreri* (n=2) ou *S. broughtonii* (n=1) sont disponibles, ce qui limite fortement les conclusions à propos de l'évolution du virus chez ces espèces hôtes. Bien évidemment pour avoir une idée encore plus globale de la diversité du virus et de son évolution au sein de ses différents hôtes alternatifs, il sera nécessaire de collecter des génotypes viraux chez d'autres espèces hôtes, en milieu naturel, dans le monde et si possible avec des dates de prélèvements répartis entre 1990 et aujourd'hui. En revanche, les travaux présentés dans le chapitre 4 ont permis d'assembler 279 nouveaux génomes du virus OsHV-1 prélevés entre 1992 et 2022 chez *M. gigas*. Grâce à ces génomes, nous avons pu retracer l'historique de diffusion du virus au sein des plus gros bassins ostréicoles mondiaux depuis son émergence.

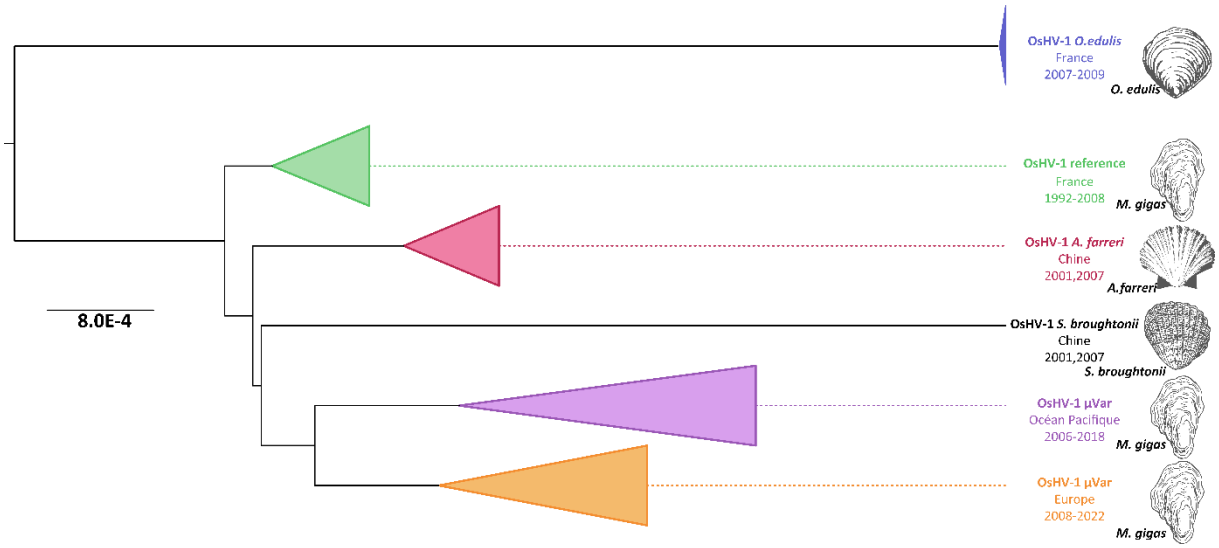


Figure 41 : Phylogénie de maximum de vraisemblance montrant les relations évolutives entre les 308 génotypes viraux collectés chez différentes espèces de bivalves dans le monde.

Historique de diffusion mondiale du virus OsHV-1 chez *M. gigas*

L'analyse phylogéographique suggère l'émergence du génotype d'OsHV-1 présent dans l'océan Pacifique à partir du génotype OsHV-1 de référence (Davison et al., 2005). Le génotype OsHV-1 µVar aurait ensuite émergé en Europe en 2008 à partir du génotype d'OsHV-1 présent dans l'océan Pacifique. A la lumière des résultats obtenus sur les mécanismes d'évolution du virus, ces émergences sont la conséquence plausible de l'occurrence successive de différents évènements de réarrangement architecturaux, de recombinaison et de substitutions, sous l'effet de pressions de sélection et de la dérive génétique.

Cette dernière hypothèse a pu être corroborée par la reconstruction des routes de dispersion du virus en France qui montre l'implication du bassin de Marennes-Oléron comme source de diversité. Ceci pourrait s'expliquer par le fait que le bassin de Marennes-Oléron est devenu l'une des plus grandes zones ostréicoles d'Europe, caractérisé par une aquaculture intensive et une surcharge chronique de la population d'huîtres (Buestel *et al.*, 2009). De nombreux génotypes issus du bassin de Marennes-Oléron semblent être régulièrement dispersés vers les autres bassins ostréicoles Français. Cependant, les distances géographiques importantes séparant les bassins de production étudiés, ne permettent pas la dispersion virale par le milieu (courants hydrologiques). Ceci suggère donc que ces dispersions sont favorisées par un autre facteur (M. J. Keeling *et al.*, 2001). Une étude des mouvements des huîtres a révélé que celles-ci sont déplacées jusqu'à neuf fois au cours de leur cycle de production, en particulier pendant les saisons de printemps et d'automne, périodes où les mortalités liées à OsHV-1 sont les plus fréquentes (Lupo *et al.*, 2016). Ces pratiques culturales exercent une influence significative sur la diversité génétique d'OsHV-1 le long des côtes françaises, favorisant ainsi sa dispersion longue-distance du virus (Lupo *et al.*, 2016) pendant les épizooties. De surcroît, des études antérieures ont déjà confirmé que le virus peut être transmis à des huîtres saines par le biais de la cohabitation avec des huîtres ayant survécus à une infection par OsHV-1 (Degremont & Benabdelmouna, 2014), potentiellement porteuses du virus sous forme persistante (*e.g.* latence). Ceci met en lumière l'impact du transfert fréquent d'animaux entre les bassins et l'importance d'évaluer l'apport potentiel de programmes de surveillance active de la maladie et de l'établissement de mesures de traçabilité de mouvements des animaux tout au long de l'année.

La dispersion locale du virus OsHV-1 à une échelle spatiale fine et sur de courtes distances est un aspect que je n'ai pas abordé dans ma thèse, mais qui peut être crucial dans la dynamique de cette maladie au sein des bassins ostréicoles. Cette dispersion peut être influencée par divers facteurs tels que les courants marins, les densités d'huîtres, la distance entre les parcs ou les facteurs environnementaux et climatiques. Comprendre ces mécanismes de dispersion locale est essentiel pour le contrôle efficace

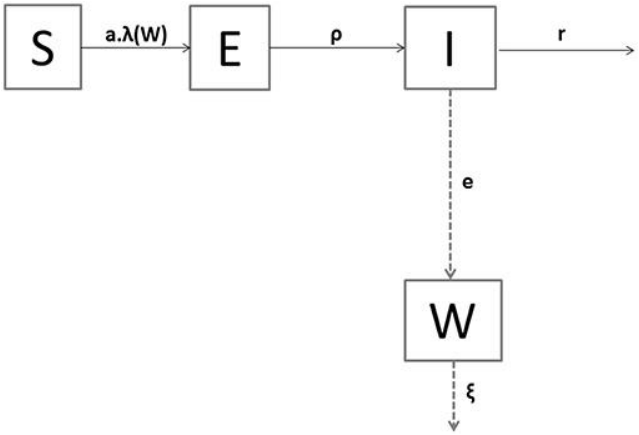
de la maladie, car ils peuvent permettre de cibler les zones à haut risque et d'adopter des stratégies de gestion adaptées à chaque zone. Des approches de phylogéographie du paysage pourraient être appliquées pour déterminer les effets relatifs des facteurs anthropiques et environnementaux sur la diffusion du virus dans les différentes zones de productions ostréicoles (Dellicour et al., 2016, 2020).

Vers une meilleure gestion de la maladie

La prévention des maladies virales dans les élevages de mollusques bivalves est d'une importance cruciale pour maintenir la santé des populations et garantir une production stable. Pour limiter la dispersion des génotypes viraux et éviter l'émergence de nouvelles lignées virales problématiques, il est nécessaire de mettre en place des mesures de prévention. Tout d'abord, à chaque transfert, une mise en quarantaine des stocks d'huîtres pourrait être mise en place pour limiter le nombre d'introduction de nouveaux génotypes du virus dans les élevages. Le naissain d'huître issus des écloseries ou du captage naturel pourrait être maintenu à l'écart des populations déjà au champ, pendant une période de surveillance afin de s'assurer qu'elles ne portent pas le virus (quarantaine). Bien que déjà existante dans la plupart des grandes structures ostréicoles, des mesures d'hygiène strictes pourraient également être mises en place dans les installations aquacoles pour minimiser la transmission du virus. Cela inclut le nettoyage régulier des équipements, la désinfection des installations et le contrôle de la qualité de l'eau pour maintenir un environnement non infecté. De plus, la surveillance continue de la santé des huîtres est essentielle. Des tests réguliers devraient être effectués pour détecter la présence du virus permettant une réponse rapide en cas d'épidémie pour limiter les pertes économiques. En cas d'épizootie, des mesures de gestion pourraient être mises en place pour réduire les pertes. Cela peut inclure l'isolement des lots infectés ou la suppression des individus malades.

L'impact de ces potentielles mesures mises en place pour contrôler la maladie pourrait être évalué par l'utilisation de modèles à compartiments spatialisés (Ezanno et al., 2012). Les modèles à compartiments spatialisés sont des outils mathématiques qui permettent de représenter la propagation d'une maladie ou d'un phénomène dans un espace géographique en le divisant en différentes zones ou compartiments. Chaque compartiment représente un état ou une catégorie spécifique, comme les individus susceptibles d'être infectés, les individus infectés, ou les individus guéris. Ces modèles intègrent également des informations sur les taux de transmission, les déplacements spatiaux, les contacts entre individus, et d'autres facteurs environnementaux influençant la diffusion de la maladie. En utilisant des données empiriques et des hypothèses appropriées, les modèles compartimentaux spatialisés permettent de simuler et de prédire la manière

dont la maladie se propage dans un environnement donné, ce qui est essentiel pour évaluer l'efficacité des mesures de contrôle mises en place, optimiser leur gestion, et prendre des décisions pour minimiser l'impact de la maladie. Un modèle a déjà été développé pour l'étude de la transmission de *Vibrio aestuarianus*, une bactérie pathogène de l'huître creuse (Figure 42, Lupo et al., 2019). Un second modèle est en cours de développement pour étudier la transmission du virus OsHV-1 chez les huîtres creuses dans le cadre du projet IDEAL, financé par l'Agence Nationale de la Recherche.



a : taux d'exposition pour contaminer l'eau (jour^{-1})
K : Concentration en bactéries dans l'eau de mer qui induit 50% de chance d'être infecté en un jour ou par la moitié de la dose infectieuse (bactéries/mL)
1/ξ : temps de survie des bactéries libres dans l'eau de mer (jours)
1/ρ : Période de latence (jours)
e : taux d'excrétion des bactéries (bactéries/mL par jour et par huître)
1/r : période d'infectiosité (jours)
Λ(W) : probabilité d'être infecté par *V. aestuarianus*, dépendant de la concentration dans l'eau et du paramètre K

Figure 42 : Modèle à compartiments de transmission de *Vibrio aestuarianus* chez les huîtres et dans leur environnement.

Les états de santé des huîtres sont S pour sensible, E pour exposée et I pour infectée. W représente la concentration de la bactérie das l'eau de mer. Les flèches pleines indiquent les transitions des huîtres entre les états de santé. Les flèches en pointillés représentent la dynamique de la population du pathogène (Lupo *et al.*, 2019).

De plus, il sera intéressant que, dans le cadre d'un programme de surveillance, une routine d'échantillonnages pour suivre les mortalités d'huîtres liées au virus OsHV-1 soit mise en place. Dans ce contexte, des échantillonnages pourront être effectués dans les principaux bassins de productions ostréicoles, à savoir en Normandie, dans la rade de Brest, le bassin de Marennes-Oléron, le bassin d'Arcachon et l'étang de Thau. Ces prélèvements permettraient de faire des études quantitatives et qualitatives sur les variants détectés au sein de ces bassins et de continuer l'étude de l'évolution des lignées virales en France. De plus, ces suivis permettront de détecter de manière précoce l'émergence de nouveaux variants dans les différents bassins. Le développement à plus grande échelle du projet SEA-MONITOR (Surveillance et Épidémiologie moléculaire des Agents pathogènes de MOllusques mariNs: Impact de la Technologie nanopORe, https://asim.ifremer.fr/Projets/SEA-MONITOR-Depuis-2021) en cours depuis 2021 et financé par Direction de l'Innovation de l'Ifremer, permettrait notamment d'effectuer ces suivis en temps réel. L'objectif du projet est de fournir une solution concrète pour la surveillance en temps réel des agents pathogènes viraux des mollusques marins pour en améliorer significativement la prévention et le contrôle. Ce projet implique des ostréiculteurs et

des laboratoires de recherches intéressés. Il s'appuie sur le développement d'une valise portable qui contient le matériel de laboratoire miniaturisé nécessaire à la réalisation d'extractions et de quantification de l'ADN puis de son séquençage (Figure 43). Dans ce projet, le séquençage est réalisé *via* la technologie ONT.

Figure 43 : Image de la valise développée dans le cadre du projet SEA-MONITOR financé par la Direction de l'Innovation de l'Ifremer pour améliorer la surveillance en temps réel des agents pathogènes viraux des mollusques marins.

Cette valise est composée du matériel nécessaire au prélèvement de tissus de mollusques, à l'extraction d'ADN, à la réalisation de gel d'électrophorèse pour la quantification de l'ADN extrait et au séquençage de l'ADN via la technologie ONT (ordinateur et séquenceur).

Finalement, en perspective globale de ce travail de thèse, on pourrait imaginer la création d'un groupe de recherche international (GDRI), pour continuer l'étude de la diversité et l'évolution de ce virus et élaborer des stratégies de gestion efficaces. Le GDRI sur l'OsHV-1 permettra de favoriser la collaboration entre les chercheurs du monde entier et de centraliser les données de séquençage du virus OsHV-1. Ses principaux objectifs seraient :

i) De favoriser des collaborations entre chercheurs des différents pays

Le groupe de recherche rassemblera des experts de renommée internationale en virologie marine, en génomique, en écologie des écosystèmes marins et en aquaculture. Il pourrait être ouvert à la participation de chercheurs, d'instituts de recherche, d'organisations gouvernementales et d'entreprises privées du monde entier.

ii) De faciliter la collecte d'échantillons d'OsHV-1 infectant différentes espèces de mollusques bivalves, à la fois en France et à l'étranger, dans les plus gros bassins de productions mondiaux.

Les membres du GDRI coordonneront des expéditions pour collecter des échantillons d'OsHV-1 dans divers environnements marins. Les sites d'échantillonnage incluront des zones d'aquaculture, des habitats sauvages et des zones touchées par des épidémies passées.

iii) D'analyser en temps réel les données génomiques du virus OsHV-1.

Les échantillons seront utilisés pour continuer la caractérisation de la diversité génétique et de l'évolution d'OsHV-1 à l'aide de techniques avancées de séquençage, de génomique et de phylodynamique.

iv) De mettre en place des mesures de contrôle

Ces mesures de contrôle permettront de limiter la dispersion des variants et de minimiser l'impact de la maladie sur les populations d'huîtres et de mollusques bivalves afin de réduire les pertes économiques.

v) Développement à plus large échelle de la plateforme MoPSeq-DB

La plateforme MoPSeq-DB, actuellement en développement pour centraliser les données de séquençage des différents agents pathogènes des mollusques bivalves, les génomes et les informations sur la diversité génomique pourra être améliorée par l'ajout de données épidémiologiques et des protocoles de gestion. Cette base de données sera accessible à tous les membres du GDRI et plus largement diffusée si besoin.

La création d'un GDRI sur l'OsHV-1 marquerait une étape cruciale dans la compréhension et la gestion de la maladie causée par ce virus. En favorisant la collaboration internationale, la récupération d'échantillons et la centralisation des données, ce groupe viserait à améliorer la résilience des populations d'huîtres et de mollusques bivalves face à la menace du virus OsHV-1, tout en contribuant à la protection l'industrie ostréicole à l'échelle mondiale.

Conclusion

L'étude du virus OsHV-1 et de ses interactions avec les différentes espèces hôtes constitue un domaine de recherche complexe mais essentiel pour l'industrie ostréicole. Ce travail de thèse a permis d'assembler *de novo* plus de 500 nouveaux génomes du virus OsHV-1. C'est une avancée considérable sachant qu'en 2019, seulement six génomes du virus étaient disponibles dans les bases de données. Les avancées méthodologiques présentées dans ce manuscrit, telles que les techniques de séquençage de fragments d'ADN court et long et le développement d'un pipeline bio-informatique dédié à l'étude du génome d'OsHV-1, ont considérablement amélioré notre capacité à étudier ce virus. Les résultats montrent que la diversité génétique du virus est structurée spatialement, temporellement et en fonction de l'espèce hôte, qui sont les facteurs clés de son évolution. La diversité virale observée dans les différents bassins de production ostréicole est étroitement liée aux pratiques culturales et notamment aux densités de populations et aux transferts d'huîtres entre bassins. De plus, la première évolution expérimentale menée pour étudier les mécanismes évolutifs du virus OsHV-1 a mis en évidence l'importance des pressions de sélection sur l'évolution du virus, en particulier de la sélection positive sur les portions du génome codant pour les protéines essentielles du virus, qui permettent son entrée dans le noyau de l'hôte, sa maintenance et sa réplication.

Il est nécessaire de continuer les recherches concernant la diversité et l'évolution du virus OsHV-1, d'une part dans une optique de surveillance, pour détecter précocement l'émergence de nouveaux variants et d'autre part pour prendre en compte d'autres facteurs, comme les autres micro-organismes présents dans les huîtres ou les phénomènes de persistance, qui pourraient influencer la maladie et sa résurgence.

Finalement, ces travaux de thèse contribuent à renforcer notre capacité à comprendre et éventuellement lutter contre les épidémies d'OsHV-1 chez les huîtres. La gestion de la santé des huîtres et la surveillance continue de la diversité virale sont essentielles pour assurer la durabilité de cette industrie et la préservation de ces espèces marines emblématiques.

Références bibliographiques

Abbadi, M., Zamperin, G., Gastaldelli, M., Pascoli, F., Rosani, U., Milani, A., Schivo, A., Rossetti, E., Turolla, E., Gennari, L., Toffan, A., Arcangeli, G., & Venier, P. (2018). Identification of a newly described OsHV-1 µvar from the North Adriatic Sea (Italy). *Journal of General Virology*, *99*(5), Article 5. https://doi.org/10.1099/jgv.0.001042

Adler, B., Sattler, C., & Adler, H. (2017). Herpesviruses and Their Host Cells: A Successful Liaison. *Trends in Microbiology*, *25*(3), 229–241. https://doi.org/10.1016/j.tim.2016.11.009

Agaba, B. B., Rugera, S. P., Mpirirwe, R., Atekat, M., Okubal, S., Masereka, K., Erionu, M., Adranya, B., Nabirwa, G., Odong, P. B., Mukiibi, Y., Ssewanyana, I., Nabadda, S., & Muwanguzi, E. (2022). Asymptomatic malaria infection, associated factors and accuracy of diagnostic tests in a historically high transmission setting in Northern Uganda. *Malaria Journal*, *21*(1), 392. https://doi.org/10.1186/s12936-022-04421-1

Alveshere, A. J. (2019). *Forces of Evolution*. https://pressbooks-dev.oer.hawaii.edu/explorationsbioanth/chapter/__unknown__-3/

Amemiya, I. (1929). *On the Sex-change of the Japanese Common Oyster, Ostrea gigas Thunberg. 5*, 284–286.

Anderson, R. M., & May, R. M. (1982). Coevolution of hosts and parasites. *Parasitology*, *85*(2), 411–426. https://doi.org/10.1017/S0031182000055360

Artika, I. M., Wiyatno, A., & Ma'roef, C. N. (2020). Pathogenic viruses: Molecular detection and characterization. *Infection, Genetics and Evolution*, *81*, 104215. https://doi.org/10.1016/j.meegid.2020.104215

Arzul, I., Corbeil, S., Morga, B., & Renault, T. (2017). Viruses infecting marine molluscs. *Journal of Invertebrate Pathology*, *147*, 118–135. https://doi.org/10.1016/j.jip.2017.01.009

Arzul, I., Nicolas, J.-L., Davison, A. J., & Renault, T. (2001). French Scallops: A New Host for Ostreid Herpesvirus-1. *Virology*, *290*(2), Article 2. https://doi.org/10.1006/viro.2001.1186

Arzul, I., Renault, T., & Lipart, C. (2001). Experimental herpes-like viral infections in marine bivalves: Demonstration of interspecies transmission. *Diseases of Aquatic Organisms*, *46*, 1–6. https://doi.org/10.3354/dao046001

Arzul, I., Renault, T., Lipart, C., & Davison, A. J. (2001). Evidence for interspecies transmission of oyster herpesvirus in marine bivalves. *Journal of General Virology*, *82*(4), Article 4. https://doi.org/10.1099/0022-1317-82-4-865

Arzul, I., Renault, T., Thébault, A., & Gérard, A. (2002). Detection of oyster herpesvirus DNA and proteins in asymptomatic Crassostrea gigas adults. *Virus Research*, *84*(1–2), Article 1–2. https://doi.org/10.1016/s0168-1702(02)00007-2

Attwood, S. W., Hill, S. C., Aanensen, D. M., Connor, T. R., & Pybus, O. G. (2022). Phylogenetic and phylodynamic approaches to understanding and combating the early SARS-CoV-2 pandemic. *Nature Reviews Genetics*, *23*(9), Article 9. https://doi.org/10.1038/s41576-022-00483-8

Azab, W., Dayaram, A., Greenwood, A. D., & Osterrieder, N. (2018). How Host Specific Are Herpesviruses? Lessons from Herpesviruses Infecting Wild and Endangered Mammals. *Annual Review of Virology*, *5*(1), 53–68. https://doi.org/10.1146/annurev-virology-092917-043227

Bai, C., Wang, C., Xia, J., Sun, H., Zhang, S., & Huang, J. (2015). Emerging and endemic types of Ostreid herpesvirus 1 were detected in bivalves in China. *Journal of Invertebrate Pathology*, *124*, 98–106. https://doi.org/10.1016/j.jip.2014.11.007

Bai, C.-M., Morga, B., Rosani, U., Shi, J., Li, C., Xin, L.-S., & Wang, C.-M. (2019). Long-range PCR and high-throughput sequencing of Ostreid herpesvirus 1 indicate high genetic diversity and complex evolution process. *Virology*, *526*, 81–90. https://doi.org/10.1016/j.virol.2018.09.026

Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriological Reviews*, *35*(3), 235–241. https://doi.org/10.1128/br.35.3.235-241.1971

Bandín, I., & Dopazo, C. P. (2011). Host range, host specificity and hypothesized host shift events among viruses of lower vertebrates. *Veterinary Research*, *42*(1), 67. https://doi.org/10.1186/1297-9716-42-67

Banerjee, N., & Mukhopadhyay, S. (2016). Viral glycoproteins: Biological role and application in diagnosis. *VirusDisease*, *27*(1), 1–11. https://doi.org/10.1007/s13337-015-0293-5

Barbieri, E. S., Medina, C. D., Vázquez, N., Fiorito, C., Martelli, A., Wigdorovitz, A., Schwindt, E., Morga, B., Renault, T., Parreño, V., & Barón, P. J. (2019). First detection of Ostreid herpesvirus 1 in wild Crassostrea gigas in Argentina. *Journal of Invertebrate Pathology*, *166*, 107222. https://doi.org/10.1016/j.jip.2019.107222

Batista, F. M., López-Sanmartín, M., Grade, A., Morgado, I., Valente, M., Navas, J. I., Power, D. M., & Ruano, F. (2015). Sequence variation in ostreid herpesvirus 1 microvar isolates detected in dying and asymptomatic Crassostrea angulata adults in the Iberian Peninsula: Insights into viral origin and spread. *Aquaculture*, *435*, 43–51. https://doi.org/10.1016/j.aquaculture.2014.09.016

Battistel, C. (2023, June 27). *MoPSeq-DB. Molluscs Pathogen Sequences DataBase.* [Poster]. Journée Ouvertes en Biologie, Informatique et Mathématiques - JOBIM, Nancy.

Bayry, J. (2013). Emerging viral diseases of livestock in the developing world. *Indian Journal of Virology*, *24*(3), 291–294. https://doi.org/10.1007/s13337-013-0164-x

Bean, B., Moore, B. M., Sterner, B., Peterson, L. R., Gerding, D. N., & Balfour, H. H. (1982). Survival of influenza viruses on environmental surfaces. *The Journal of Infectious Diseases*, *146*(1), 47–51. https://doi.org/10.1093/infdis/146.1.47

Behringer, D. C., Butler, M. J., & Shields, J. D. (2006). Avoidance of disease by social lobsters. *Nature*, *441*(7092), Article 7092. https://doi.org/10.1038/441421a

Belding, D. L. (1965). Textbook of Parasitology. *Textbook of Parasitology.*, *3rd Edition*. https://www.cabdirect.org/cabdirect/abstract/19662901591

Biek, R., Pybus, O. G., Lloyd-Smith, J. O., & Didelot, X. (2015). Measurably evolving pathogens in the genomic era. *Trends in Ecology & Evolution*, *30*(6), 306–313. https://doi.org/10.1016/j.tree.2015.03.009

Bloomquist, E. W., Lemey, P., & Suchard, M. A. (2010). Three roads diverged? Routes to phylogeographic inference. *Trends in Ecology & Evolution*, *25*(11), 626–632. https://doi.org/10.1016/j.tree.2010.08.010

Boldogh, I., Albrecht, T., & Porter, D. D. (1996). Persistent Viral Infections. In S. Baron (Ed.), *Medical Microbiology* (4th ed.). University of Texas Medical Branch at Galveston. http://www.ncbi.nlm.nih.gov/books/NBK8538/

Bonneaud, C., & Longdon, B. (2020). Emerging pathogen evolution: Using evolutionary theory to understand the fate of novel infectious pathogens. *EMBO Reports*, *21*(9), e51374. https://doi.org/10.15252/embr.202051374

Boudry, P., Heurtebise, S., Collet, B., Cornette, F., & Gerard, A. (1998). *Differentiation between populations of the Portuguese oyster, Crassostrea angulata (Lamark) and the Pacific oyster, Crassostrea gigas (Thunberg), revealed by mtDNA RFLP analysis* (2). *226*(2), Article 2.

Brito, A. F., & Pinney, J. W. (2020). The evolution of protein domain repertoires: Shedding light on the origins of the Herpesviridae family. *Virus Evolution*, *6*(1), veaa001. https://doi.org/10.1093/ve/veaa001

Brooks, D. R., Hoberg, E. P., Boeger, W. A., & Trivellone, V. (2022). Emerging infectious disease: An underappreciated area of strategic concern for food security. *Transboundary and Emerging Diseases*, *69*(2), 254–267. https://doi.org/10.1111/tbed.14009

Broquard, C. (2019). *Le déterminisme sexuel de l'huître Crassostrea gigas: Du phénotype aux facteurs moléculaires sous-jacents.* https://tel.archives-ouvertes.fr/tel-03150111

Buck, J. C., & Weinstein, S. B. (2020). The ecological consequences of a pandemic. *Biology Letters*, *16*(11), 20200641. https://doi.org/10.1098/rsbl.2020.0641

Buestel, D., Ropert, M., Prou, J., & Goulletquer, P. (2009). History, Status, and Future of Oyster Culture in France. *Journal of Shellfish Research*, *28*(4), Article 4. https://doi.org/10.2983/035.028.0410

Burge, C. A., Friedman, C. S., Kachmar, M. L., Humphrey, K. L., Moore, J. D., & Elston, R. A. (2021). The first detection of a novel OsHV-1 microvariant in San Diego, California, USA. *Journal of Invertebrate Pathology*, *184*, 107636. https://doi.org/10.1016/j.jip.2021.107636

Burge, C. A., Griffin, F. J., & Friedman, C. S. (2006). Mortality and herpesvirus infections of the Pacific oyster Crassostrea gigas in Tomales Bay, California, USA. *Diseases of Aquatic Organisms*, *72*(1), 31–43. https://doi.org/10.3354/dao072031

Burge, C. A., Mark Eakin, C., Friedman, C. S., Froelich, B., Hershberger, P. K., Hofmann, E. E., Petes, L. E., Prager, K. C., Weil, E., Willis, B. L., Ford, S. E., & Harvell, C. D. (2014). Climate Change Influences on Marine Infectious Diseases: Implications for Management and Society. *Annual Review of Marine Science*, *6*(1), 249–277. https://doi.org/10.1146/annurev-marine-010213-135029

Burioli, E. A. V., Prearo, M., & Houssin, M. (2017). Complete genome sequence of Ostreid herpesvirus type 1 µVar isolated during mortality events in the Pacific oyster Crassostrea gigas in France and Ireland. *Virology*, *509*, 239–251. https://doi.org/10.1016/j.virol.2017.06.027

Canier, L., Combette, A., Lupo, C., & Garcia, C. (2019). *Bilan 2019 du dispositif national de surveillance de la santé des mollusques marins REPAMO*.

Casadevall, A., & Pirofski, L. (2014). Microbiology: Ditch the term pathogen. *Nature*, *516*(7530), Article 7530. https://doi.org/10.1038/516165a

Casto, A. M., Roychoudhury, P., Xie, H., Selke, S., Perchetti, G. A., Wofford, H., Huang, M.-L., Verjans, G. M. G. M., Gottlieb, G. S., Wald, A., Jerome, K. R., Koelle, D. M., Johnston, C., & Greninger, A. L. (2020). Large, Stable, Contemporary Interspecies Recombination Events in Circulating Human Herpes Simplex Viruses. *The Journal of Infectious Diseases*, *221*(8), 1271–1279. https://doi.org/10.1093/infdis/jiz199

Chang, P., Kuo, S., Lai, S., Yang, H., Ting, Y., Hsu, C., & Chen, H. (2005). Herpes-like virus infection causing mortality of cultured abalone Haliotis diversicolor supertexta in Taiwan. *Diseases of Aquatic Organisms*, *65*, 23–27. https://doi.org/10.3354/dao065023

Chollier, M., Tomkinson, C., & Philibert, P. (2016). Stigmatisation et impact sur le soin: L'exemple des IST/VIH. *Sexologies*, *25*(4), 179–183. https://doi.org/10.1016/j.sexol.2016.03.006

Churchyard, G., Kim, P., Shah, N. S., Rustomjee, R., Gandhi, N., Mathema, B., Dowdy, D., Kasmar, A., & Cardenas, V. (2017). What We Know About Tuberculosis Transmission: An Overview. *The Journal of Infectious Diseases*, *216*(Suppl 6), S629–S635. https://doi.org/10.1093/infdis/jix362

Clegg, T. A., Morrissey, T., Geoghegan, F., Martin, S. W., Lyons, K., Ashe, S., & More, S. J. (2014). Risk factors associated with increased mortality of farmed Pacific oysters in Ireland during 2011. *Preventive Veterinary Medicine*, *113*(2), Article 2. https://doi.org/10.1016/j.prevetmed.2013.10.023

Cochennec, N., Baud, J.-P., Pépin, J.-F., & et al. (2011). *Les surmortalités des naissains d'huîtres creuses, Crassostrea gigas: Acquis des recherches en 2010*. https://archimer.ifremer.fr/doc/00033/14423/11719.pdf

Cohen, J. I. (2020). Herpesvirus latency. *Journal of Clinical Investigation*, *130*(7), 3361–3369. https://doi.org/10.1172/JCI136225

Combe, M., Reverter, M., Caruso, D., Pepey, E., & Gozlan, R. E. (2023). Impact of Global Warming on the Severity of Viral Diseases: A Potentially Alarming Threat to Sustainable Aquaculture Worldwide. *Microorganisms*, *11*(4), Article 4. https://doi.org/10.3390/microorganisms11041049

Comps, M. (1988). Epizootic diseases of oysters associated with viral infections. *American Fisheries Society Special Publication*, *18*, Article 18.

Comps, M., & Cochennec, N. (1993). A Herpes-like Virus from the European Oyster Ostrea edulis L. *Journal of Invertebrate Pathology*, *62*(2), 201–203. https://doi.org/10.1006/jipa.1993.1098

Cox, P. A. (1983). Extinction of the Hawaiian Avifauna Resulted in a Change of Pollinators for the ieie, Freycinetia arborea. *Oikos*, *41*(2), 195. https://doi.org/10.2307/3544263

Crespo-Bellido, A., Hoyer, J. S., Dubey, D., Jeannot, R. B., & Duffy, S. (2021). Interspecies Recombination Has Driven the Macroevolution of Cassava Mosaic Begomoviruses. *Journal of Virology*, *95*(17), e00541-21. https://doi.org/10.1128/JVI.00541-21

Cunliffe, J. (2008). A Proliferation of Pathogens through the 20th Century. *Scandinavian Journal of Immunology*, *68*(2), 120–128. https://doi.org/10.1111/j.1365-3083.2008.02130.x

Danic-Tchaleu, G., Heurtebise, S., Morga, B., & Lapègue, S. (2011). Complete mitochondrial DNA sequence of the European flat oyster Ostrea edulis confirms Ostreidae classification. *BMC Research Notes*, *4*(1), 400. https://doi.org/10.1186/1756-0500-4-400

Darwin, C. (1889). *Journal of Researches Into the Natural History and Geology of the Countries Visited During the Voyage of H.M.S. "Beagle" Round the World: Under the Command of Capt. Fitz Roy, R.N.* Ward, Lock and Company.

Davison, A. J. (1998). The Genome of Salmonid Herpesvirus 1. *Journal of Virology*, *72*(3), 1974–1982. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC109490/

Davison, A. J. (2010). Herpesvirus systematics. *Veterinary Microbiology*, *143*(1), Article 1. https://doi.org/10.1016/j.vetmic.2010.02.014

Davison, A. J., Eberle, R., Ehlers, B., Hayward, G. S., McGeoch, D. J., Minson, A. C., Pellett, P. E., Roizman, B., Studdert, M. J., & Thiry, E. (2009). The order Herpesvirales. *Archives of Virology*, *154*(1), Article 1. https://doi.org/10.1007/s00705-008-0278-4

Davison, A. J., Trus, B. L., Cheng, N., Steven, A. C., Watson, M. S., Cunningham, C., Deuff, R.-M. L., & Renault, T. (2005). A novel class of herpesvirus with bivalve hosts. *Journal of General Virology*, *86*(1), Article 1. https://doi.org/10.1099/vir.0.80382-0

de Lorgeril, J., Lucasson, A., Petton, B., Toulza, E., Montagnani, C., Clerissi, C., Vidal-Dupiol, J., Chaparro, C., Galinier, R., Escoubas, J.-M., Haffner, P., Dégremont, L., Charrière, G. M., Lafont, M., Delort, A., Vergnes, A., Chiarello, M., Faury, N., Rubio, T., … Mitta, G. (2018). Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-06659-3

De Pelsmaeker, S., Romero, N., Vitale, M., & Favoreel, H. W. (2018). Herpesvirus Evasion of Natural Killer Cells. *Journal of Virology*, *92*(11), e02105-17. https://doi.org/10.1128/JVI.02105-17

Debré, P. (2020). Épidémies: Leçons d'Histoire. *médecine/sciences*, *36*(6–7), 642–646. https://doi.org/10.1051/medsci/2020111

Debruyne, R., & Poinar, H. N. (2009). Time Dependency of Molecular Rates in Ancient DNA Data Sets, A Sampling Artifact? *Systematic Biology*, *58*(3), 348–360. https://doi.org/10.1093/sysbio/syp028

Dégremont, L. (2011). Evidence of herpesvirus (OsHV-1) resistance in juvenile Crassostrea gigas selected for high resistance to the summer mortality phenomenon. *Aquaculture*, *317*(1–4), 94–98. https://doi.org/10.1016/j.aquaculture.2011.04.029

Dégremont, L. (2013). Size and genotype affect resistance to mortality caused by OsHV-1 in Crassostrea gigas. *Aquaculture*, *416–417*, 129–134. https://doi.org/10.1016/j.aquaculture.2013.09.011

Degremont, L., & Benabdelmouna, A. (2014). Mortality associated with OsHV-1 in spat Crassostrea gigas: Role of wild-caught spat in the horizontal transmission of the disease. *Aquaculture International*, *22*(6), 1767–1781. https://doi.org/10.1007/s10499-014-9781-7

Dellicour, S., Lemey, P., Suchard, M. A., Gilbert, M., & Baele, G. (2022). Accommodating sampling location uncertainty in continuous phylogeography. *Virus Evolution*, *8*(1), veac041. https://doi.org/10.1093/ve/veac041

Dellicour, S., Lequime, S., Vrancken, B., Gill, M. S., Bastide, P., Gangavarapu, K., Matteson, N. L., Tan, Y., du Plessis, L., Fisher, A. A., Nelson, M. I., Gilbert, M., Suchard, M. A., Andersen, K. G., Grubaugh, N. D., Pybus, O. G., & Lemey, P. (2020). Epidemiological hypothesis testing using a phylogeographic and phylodynamic framework. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19122-z

Dellicour, S., Rose, R., Faria, N. R., Lemey, P., & Pybus, O. G. (2016). SERAPHIM: Studying environmental rasters and phylogenetically informed movements. *Bioinformatics*, *32*(20), 3204–3206. https://doi.org/10.1093/bioinformatics/btw384

Dhama, K., Nainu, F., Frediansyah, A., Yatoo, Mohd. I., Mohapatra, R. K., Chakraborty, S., Zhou, H., Islam, Md. R., Mamada, S. S., Kusuma, H. I., Rabaan, A. A., Alhumaid, S., Mutair, A. A., Iqhrammullah, M., Al-Tawfiq, J. A., Mohaini, M. A., Alsalman, A. J., Tuli, H. S., Chakraborty, C., & Harapan, H. (2023). Global emerging Omicron variant of SARS-CoV-2: Impacts, challenges and strategies. *Journal of Infection and Public Health*, *16*(1), 4–14. https://doi.org/10.1016/j.jiph.2022.11.024

Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature Biotechnology*, *35*(4), 316–319. https://doi.org/10.1038/nbt.3820

Djaharuddin, I., Munawwarah, S., Nurulita, A., Ilyas, M., Tabri, N. A., & Lihawa, N. (2021). Comorbidities and mortality in COVID-19 patients. *Gaceta Sanitaria*, *35 Suppl 2*, S530–S532. https://doi.org/10.1016/j.gaceta.2021.10.085

Dollfus, R. PH. (1921). *Résumé de nos principales connaissances pratiques sur les Maladies et les Ennemis de l'Huître*. https://archimer.ifremer.fr/doc/00016/12766/9703.pdf

Donohoe, O., Zhang, H., Delrez, N., Gao, Y., Suárez, N. M., Davison, A. J., & Vanderplasschen, A. (2021). Genomes of Anguillid Herpesvirus 1 Strains Reveal Evolutionary Disparities and Low Genetic Diversity in the Genus Cyprinivirus. *Microorganisms*, *9*(5), 998. https://doi.org/10.3390/microorganisms9050998

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed Phylogenetics and Dating with Confidence. *PLOS Biology*, *4*(5), e88. https://doi.org/10.1371/journal.pbio.0040088

Drummond, A. J., Nicholls, G. K., Rodrigo, A. G., & Solomon, W. (2002). Estimating Mutation Parameters, Population History and Genealogy Simultaneously From Temporally Spaced Sequence Data. *Genetics*, *161*(3), 1307–1320. https://doi.org/10.1093/genetics/161.3.1307

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*(1), 214. https://doi.org/10.1186/1471-2148-7-214

Duffy, S. (2018). Why are RNA virus mutation rates so damn high? *PLoS Biology*, *16*(8), e3000003. https://doi.org/10.1371/journal.pbio.3000003

Dundon, W. G., Arzul, I., Omnes, E., Robert, M., Magnabosco, C., Zambon, M., Gennari, L., Toffan, A., Terregino, C., Capua, I., & Arcangeli, G. (2011). Detection of Type 1 Ostreid Herpes variant (OsHV-1 µvar) with no associated mortality in French-origin Pacific cupped oyster Crassostrea gigas farmed in Italy. *Aquaculture*, *314*(1–4), Article 1–4. https://doi.org/10.1016/j.aquaculture.2011.02.005

Elena, S. F., & Sanjuán, R. (2005). Adaptive Value of High Mutation Rates of RNA Viruses: Separating Causes from Consequences. *Journal of Virology*, *79*(18), 11555–11558. https://doi.org/10.1128/jvi.79.18.11555-11558.2005

Ellwanger, J. H., & Chies, J. A. B. (2021). Zoonotic spillover: Understanding basic aspects for better prevention. *Genetics and Molecular Biology*, *44*(1 Suppl 1), e20200355. https://doi.org/10.1590/1678-4685-GMB-2020-0355

Emerson, B. C. (2007). Alarm Bells for the Molecular Clock? No Support for Ho et al.'s Model of Time-Dependent Molecular Rate Estimates. *Systematic Biology*, *56*(2), 337–345. https://doi.org/10.1080/10635150701258795

Eybpoosh, S., Haghdoost, A. A., Mostafavi, E., Bahrampour, A., Azadmanesh, K., & Zolala, F. (2017). Molecular epidemiology of infectious diseases. *Electronic Physician*, *9*(8), 5149–5158. https://doi.org/10.19082/5149

Ezanno, P., Vergu, E., Langlais, M., & Gilot-Fromont, E. (2012). Modelling the Dynamics of Host-Parasite Interactions: Basic Principles. In S. Morand, F. Beaudeau, & J. Cabaret (Eds.), *New Frontiers of Molecular Epidemiology of Infectious Diseases* (pp. 79–101). Springer Netherlands. https://doi.org/10.1007/978-94-007-2114-2_5

Fabioux, C., Huvet, A., Lapegue, S., Heurtebise, S., & Boudry, P. (2002). Past and present geographical distribution of populations of Portuguese (Crassostrea angulata) and Pacific (C. gigas) oysters along the European and north African Atlantic coats. *Haliotis*, *31*, 33–44. https://archimer.ifremer.fr/doc/00000/2785/

Farley, C. A., Banfield, W. G., Kasnic, G., & Foster, W. S. (1972). Oyster Herpes-Type Virus. *Science*, *178*(4062), Article 4062. https://doi.org/10.1126/science.178.4062.759

Field, H. J. (2001). Herpes simplex virus antiviral drug resistance—Current trends and future prospects. *Journal of Clinical Virology*, *21*(3), 261–269. https://doi.org/10.1016/S1386-6532(00)00169-4

Fleury, E., Barbier, P., Petton, B., Normand, J., Thomas, Y., Pouvreau, S., Daigle, G., & Pernet, F. (2020). Latitudinal drivers of oyster mortality: Deciphering host, pathogen and environmental risk factors. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-64086-1

Fortier, G., van Erck, E., Fortier, C., Richard, E., Pottier, D., Pronost, S., Miszczak, F., Thiry, E., & Lekeux, P. (2009). Herpesviruses in respiratory liquids of horses: Putative implication in airway inflammation and association with cytological features. *Veterinary Microbiology*, *139*(1), 34–41. https://doi.org/10.1016/j.vetmic.2009.04.021

François, C., Joly, J.-P., Garcia, C., Lupo, C., Travers, M.-A., Pepin, J.-F., Hatt, P.-J., Arzul, I., Omnes, E., Tourbiez, D., Faury, N., Haffner, P., Huchet, E., & Dubreuil, C. (2013). *Bilan 2012 du réseau Repamo—Réseau national de surveillance de la santé des mollusques marins*.

Friedman, C., Estes, R., Stokes, N., Burge, C., Hargove, J., Barber, B., Elston, R., Burreson, E., & Reece, K. (2005). Herpes virus in juvenile Pacific oysters Crassostrea gigas from Tomales Bay, California, coincides with summer mortality episodes. *Diseases of Aquatic Organisms*, *63*, 33–41. https://doi.org/10.3354/dao063033

Fuhrmann, L., Jablonski, K. P., & Beerenwinkel, N. (2021). Quantitative measures of within-host viral genetic diversity. *Current Opinion in Virology*, *49*, 157–163. https://doi.org/10.1016/j.coviro.2021.06.002

Garcia, C., Thébault, A., Dégremont, L., Arzul, I., Miossec, L., Robert, M., Chollet, B., François, C., Joly, J.-P., Ferrand, S., Kerdudou, N., & Renault, T. (2011). Ostreid herpesvirus 1 detection and relationship with Crassostrea gigas spat mortality in France between 1998 and 2006. *Veterinary Research*, *42*(1), Article 1. https://doi.org/10.1186/1297-9716-42-73

Gay, L., Melenotte, C., Lakbar, I., Mezouar, S., Devaux, C., Raoult, D., Bendiane, M.-K., Leone, M., & Mège, J.-L. (2021). Sexual Dimorphism and Gender in Infectious Diseases. *Frontiers in Immunology*, *12*. https://www.frontiersin.org/articles/10.3389/fimmu.2021.698121

Gerritzen, M. A., Lambooij, E., Stegeman, J. A., & Spruijt, B. M. (2006). Slaughter of poultry during the epidemic of avian influenza in the Netherlands in 2003. *Veterinary Record*, *159*(2), 39–42. https://doi.org/10.1136/vr.159.2.39

Gethmann, J., Probst, C., & Conraths, F. J. (2020). Economic Impact of a Bluetongue Serotype 8 Epidemic in Germany. *Frontiers in Veterinary Science*, *7*, 65. https://doi.org/10.3389/fvets.2020.00065

Gittenberger, A., Voorbergen-Laarman, M. A., & Engelsma, M. Y. (2016). Ostreid herpesvirus OsHV-1 µVar in Pacific oysters *Crassostrea gigas* (Thunberg 1793) of the Wadden Sea, a UNESCO world heritage site. *Journal of Fish Diseases*, *39*(1), Article 1. https://doi.org/10.1111/jfd.12332

Glynn, J. R., & Moss, P. A. H. (2020). Systematic analysis of infectious disease outcomes by age shows lowest severity in school-age children. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-00668-y

Gómez-López, V. M., Jubinville, E., Rodríguez-López, M. I., Trudel-Ferland, M., Bouchard, S., & Jean, J. (2021). Inactivation of Foodborne Viruses by UV Light: A Review. *Foods*, *10*(12), Article 12. https://doi.org/10.3390/foods10123141

González-Candelas, F., Patiño-Galindo, J. Á., & Valiente-Mullor, C. (2018). Population Genomics of Human Viruses. *Population Genomics: Microorganisms*, 267–296. https://doi.org/10.1007/13836_2018_31

González-Torres, P., Rodríguez-Mateos, F., Antón, J., & Gabaldón, T. (2019). Impact of Homologous Recombination on the Evolution of Prokaryotic Core Genomes. *mBio*, *10*(1), 10.1128/mbio.02494-18. https://doi.org/10.1128/mbio.02494-18

Gorbalenya, A. E., & Lauber, C. (2017). Phylogeny of Viruses. *Reference Module in Biomedical Sciences*, B978-0-12-801238-3.95723-4. https://doi.org/10.1016/B978-0-12-801238-3.95723-4

Goudenège, D., Travers, M. A., Lemire, A., Petton, B., Haffner, P., Labreuche, Y., Tourbiez, D., Mangenot, S., Calteau, A., Mazel, D., Nicolas, J. L., Jacq, A., & Le roux, F. (2015). A single regulatory gene is sufficient to alter Vibrio aestuarianus pathogenicity in oysters. *Environmental Microbiology*, *17*(11), 4189–4199. https://doi.org/10.1111/1462-2920.12699

Gowthaman, V., Kumar, S., Koul, M., Dave, U., Murthy, T. R. G. K., Munuswamy, P., Tiwari, R., Karthik, K., Dhama, K., Michalak, I., & Joshi, S. K. (2020). Infectious laryngotracheitis: Etiology, epidemiology, pathobiology, and advances in diagnosis and control – a comprehensive review. *The Veterinary Quarterly*, *40*(1), 140–161. https://doi.org/10.1080/01652176.2020.1759845

Grad, Y. H., & Lipsitch, M. (2014). Epidemiologic data and pathogen genome sequences: A powerful synergy for public health. *Genome Biology*, *15*(11), 538. https://doi.org/10.1186/s13059-014-0538-4

Gray, A., & Sharara, F. (2022). Global and regional sepsis and infectious syndrome mortality in 2019: A systematic analysis. *The Lancet Global Health*, *10*, S2. https://doi.org/10.1016/S2214-109X(22)00131-0

Griffin, B. D., Verweij, M. C., & Wiertz, E. J. H. J. (2010). Herpesviruses and immunity: The art of evasion. *Veterinary Microbiology*, *143*(1), 89–100. https://doi.org/10.1016/j.vetmic.2010.02.017

Grijalva-Chon, J. M., Castro-Longoria, R., Ramos-Paredes, J., Enríquez-Espinoza, T. L., & Mendoza-Cano, F. (2013). Detection of a new OsHV-1 DNA strain in the healthy Pacific oyster, *Crassostrea gigas* Thunberg, from the Gulf of California. *Journal of Fish Diseases*, n/a-n/a. https://doi.org/10.1111/jfd.12028

Grizel, H., & Heral, M. (1991). Introduction into France of the Japanese oyster (Crassostrea gigas). *Journal Du Conseil - Conseil International Pour l'Exploration de La Mer*, *47*, Article 47. https://archimer.ifremer.fr/doc/00000/2760/

Hanley, K. A. (2011). The double-edged sword: How evolution can make or break a live-attenuated virus vaccine. *Evolution*, *4*(4), 635–643. https://doi.org/10.1007/s12052-011-0365-y

Hanson, L., Dishon, A., & Kotler, M. (2011). Herpesviruses that Infect Fish. *Viruses*, *3*(11), 2160–2191. https://doi.org/10.3390/v3112160

Hayek, M. N. (2022). The infectious disease trap of animal agriculture. *Science Advances*, *8*(44), eadd6681. https://doi.org/10.1126/sciadv.add6681

Hessling, M., Gierke, A.-M., Sicks, B., Fehler, N., & Vatter, P. (2022). Sensitivity of influenza virus to ultraviolet irradiation. *GMS Hygiene and Infection Control*, *17*, Doc20. https://doi.org/10.3205/dgkh000423

Hine, P., & Thorne, T. (1997). Replication of herpes-like viruses in haemocytes of adult flat oysters Ostrea angasi:an ultrastructural study. *Diseases of Aquatic Organisms*, *29*, 189–196. https://doi.org/10.3354/dao029189

Ho, S., & Larson, G. (2006). Molecular clocks: When timesare a-changin'. *Trends in Genetics*, *22*(2), 79–83. https://doi.org/10.1016/j.tig.2005.11.006

Ho, S. Y. W. (2005). Accuracy of Rate Estimation Using Relaxed-Clock Models with a Critical Focus on the Early Metazoan Radiation. *Molecular Biology and Evolution*, *22*(5), 1355–1363. https://doi.org/10.1093/molbev/msi125

Ho, S. Y. W., Kolokotronis, S.-O., & Allaby, R. G. (2007). Elevated substitution rates estimated from ancient DNA sequences. *Biology Letters*, *3*(6), 702–705. https://doi.org/10.1098/rsbl.2007.0377

Ho, S. Y. W., Phillips, M. J., Cooper, A., & Drummond, A. J. (2005). Time Dependency of Molecular Rate Estimates and Systematic Overestimation of Recent Divergence Times. *Molecular Biology and Evolution*, *22*(7), 1561–1568. https://doi.org/10.1093/molbev/msi145

Ho, S. Y. W., Shapiro, B., Phillips, M. J., Cooper, A., & Drummond, A. J. (2007). Evidence for Time Dependency of Molecular Rate Estimates. *Systematic Biology*, *56*(3), 515–522. https://doi.org/10.1080/10635150701435401

Honardoost, M., Rajabpour, A., & Vakili, L. (2018). Molecular epidemiology; New but impressive. *Medical Journal of the Islamic Republic of Iran*, *32*, 53. https://doi.org/10.14196/mjiri.32.53

Hughes, A. L., Irausquin, S., & Friedman, R. (2010). The evolutionary biology of poxviruses. *Infection, Genetics and Evolution*, *10*(1), 50–59. https://doi.org/10.1016/j.meegid.2009.10.001

Huson, D. H., & Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, *23*(2), 254–267. https://doi.org/10.1093/molbev/msj030

Hwang, J. Y., Park, J. J., Yu, H. J., Hur, Y. B., Arzul, I., Couraleau, Y., & Park, M. A. (2013). Ostreid herpesvirus 1 infection in farmed Pacific oyster larvae *Crassostrea gigas* (Thunberg) in Korea. *Journal of Fish Diseases*, n/a-n/a. https://doi.org/10.1111/jfd.12093

Issimov, A., Kutumbetov, L., Orynbayev, M. B., Khairullin, B., Myrzakhmetova, B., Sultankulova, K., & White, P. J. (2020). Mechanical Transmission of Lumpy Skin Disease Virus by Stomoxys spp. (Stomoxys calsitrans, Stomoxys sitiens, Stomoxys indica), Diptera: Muscidae. *Animals : An Open Access Journal from MDPI*, *10*(3), 477. https://doi.org/10.3390/ani10030477

Jackson, M. L., Ferdinands, J., Nowalk, M. P., Zimmerman, R. K., Kieke, B., Gaglani, M., Murthy, K., Petrie, J. G., Martin, E. T., Chung, J. R., Flannery, B., & Jackson, L. A. (2021). Differences between Frequentist and Bayesian inference in routine surveillance for influenza vaccine effectiveness: A test-negative case-control study. *BMC Public Health*, *21*(1), 516. https://doi.org/10.1186/s12889-021-10543-z

Jacob, S. T., Crozier, I., Fischer, W. A., Hewlett, A., Kraft, C. S., Vega, M.-A. de L., Soka, M. J., Wahl, V., Griffiths, A., Bollinger, L., & Kuhn, J. H. (2020). Ebola virus disease. *Nature Reviews Disease Primers*, *6*(1), Article 1. https://doi.org/10.1038/s41572-020-0147-3

James, P. B., Wardle, J., Steel, A., & Adams, J. (2020). An assessment of Ebola-related stigma and its association with informal healthcare utilisation among Ebola survivors in Sierra Leone: A cross-sectional study. *BMC Public Health*, *20*(1), 182. https://doi.org/10.1186/s12889-020-8279-7

Jee, Bo Young, Lee, Su Jin, Cho, Mi Young, 이순정, Kim, Jin-Woo, Choi, Seung Hyuk, JEONG, Hyun-Do, & Kim, Ki Hong. (2013). Detection of Ostreid Herpesvirus 1 from adult Pacific Oysters Crassostrea gigas Cultured in Korea. *Fisheries and Aquatic Sciences*, *16*(2), Article 2. https://doi.org/10.5657/FAS.2013.0131

Jenkins, C., Hick, P., Gabor, M., Spiers, Z., Fell, S., Gu, X., Read, A., Go, J., Dove, M., O'Connor, W., Kirkland, P., & Frances, J. (2013). Identification and characterisation of an ostreid herpesvirus-1 microvariant (OsHV-1 µ-var) in Crassostrea gigas (Pacific oysters) in Australia. *Diseases of Aquatic Organisms*, *105*(2), Article 2. https://doi.org/10.3354/dao02623

Jouaux, A., Lafont, M., Blin, J.-L., Houssin, M., Mathieu, M., & Lelong, C. (2013). Physiological change under OsHV-1 contamination in Pacific oyster Crassostrea gigas through massive mortality events on fields. *BMC Genomics*, *14*(1), 590. https://doi.org/10.1186/1471-2164-14-590

Kaján, G. L., Doszpoly, A., Tarján, Z. L., Vidovszky, M. Z., & Papp, T. (2020). Virus–Host Coevolution with a Focus on Animal and Human DNA Viruses. *Journal of Molecular Evolution*, *88*(1), 41–56. https://doi.org/10.1007/s00239-019-09913-4

Kalkauskas, A., Perron, U., Sun, Y., Goldman, N., Baele, G., Guindon, S., & Maio, N. D. (2021). Sampling bias and model choice in continuous phylogeography: Getting lost on a random walk. *PLOS Computational Biology*, *17*(1), e1008561. https://doi.org/10.1371/journal.pcbi.1008561

Karlsson, E. K., Baranowska, I., Wade, C. M., Salmon Hillbertz, N. H. C., Zody, M. C., Anderson, N., Biagi, T. M., Patterson, N., Pielberg, G. R., Kulbokas, E. J., Comstock, K. E., Keller, E. T., Mesirov, J. P., Von Euler, H., Kämpe, O., Hedhammar, Å., Lander, E. S., Andersson, G., Andersson, L., & Lindblad-Toh, K. (2007). Efficient mapping of mendelian traits in dogs through genome-wide association. *Nature Genetics*, *39*(11), 1321–1328. https://doi.org/10.1038/ng.2007.10

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., & Whitlock, M. C. (2012). Experimental evolution. *Trends in Ecology & Evolution*, *27*(10), 547–560. https://doi.org/10.1016/j.tree.2012.06.001

Keeling, M. J., Woolhouse, M. E. J., Shaw, D. J., Matthews, L., Chase-Topping, M., Haydon, D. T., Cornell, S. J., Kappey, J., Wilesmith, J., & Grenfell, B. T. (2001). Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. *Science*, *294*(5543), 813–817. https://doi.org/10.1126/science.1065973

Keeling, S., Brosnahan, C., Williams, R., Gias, E., Hannah, M., Bueno, R., McDonald, W., & Johnston, C. (2014a). New Zealand juvenile oyster mortality associated with ostreid herpesvirus 1—An opportunistic longitudinal study. *Diseases of Aquatic Organisms*, *109*(3), Article 3. https://doi.org/10.3354/dao02735

Keeling, S., Brosnahan, C., Williams, R., Gias, E., Hannah, M., Bueno, R., McDonald, W., & Johnston, C. (2014b). New Zealand juvenile oyster mortality associated with ostreid herpesvirus 1—An opportunistic longitudinal study. *Diseases of Aquatic Organisms*, *109*(3), 231–239. https://doi.org/10.3354/dao02735

Kibenge, F. S. (2019). Emerging viruses in aquaculture. *Current Opinion in Virology*, *34*, 97–103. https://doi.org/10.1016/j.coviro.2018.12.008

Knobler, S., Mahmoud, A., Lemon, S., & Pray, L. (2006). Summary and Assessment. In *The Impact of Globalization on Infectious Disease Emergence and Control: Exploring the Consequences and Opportunities: Workshop Summary*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK56579/

Kong, J., Lu, Y., Ren, Y., Chen, Z., & Chen, M. (2021). The virus removal in UV irradiation, ozonation and chlorination. *Water Cycle*, *2*, 23–31. https://doi.org/10.1016/j.watcyc.2021.05.001

Koonin, E. V. (2016). Horizontal gene transfer: Essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. *F1000Research*, *5*, 1805. https://doi.org/10.12688/f1000research.8737.1

Koonin, E. V., Krupovic, M., & Agol, V. I. (2021). The Baltimore Classification of Viruses 50 Years Later: How Does It Stand in the Light of Virus Evolution? *Microbiology and Molecular Biology Reviews*, *85*(3), e00053-21. https://doi.org/10.1128/MMBR.00053-21

Krkosek, M. (2010). Host density thresholds and disease control for fisheries and aquaculture. *Aquaculture Environment Interactions*, *1*(1), 21–32. https://doi.org/10.3354/aei0004

Kutnjak, D., Elena, S. F., & Ravnikar, M. (2017). Time-Sampled Population Sequencing Reveals the Interplay of Selection and Genetic Drift in Experimental Evolution of *Potato Virus Y*. *Journal of Virology*, *91*(16), e00690-17. https://doi.org/10.1128/JVI.00690-17

Lai, M. M. C. (1992). RNA Recombination in Animal and Plant Viruses. *MICROBIOL. REV.*, *56*.

Laketa, V. (2018). Microscopy in Infectious Disease Research—Imaging Across Scales. *Journal of Molecular Biology*, *430*(17), 2612–2625. https://doi.org/10.1016/j.jmb.2018.06.018

Lamers, S. L., Newman, R. M., Laeyendecker, O., Tobian, A. A. R., Colgrove, R. C., Ray, S. C., Koelle, D. M., Cohen, J., Knipe, D. M., & Quinn, T. C. (2015). Global Diversity within and between Human Herpesvirus 1 and 2 Glycoproteins. *Journal of Virology*, *89*(16), 8206–8218. https://doi.org/10.1128/JVI.01302-15

Lauring, A. S. (2020). Within-host Viral Diversity, a Window into Viral Evolution. *Annual Review of Virology*, *7*(1), 63–81. https://doi.org/10.1146/annurev-virology-010320-061642

Le Deuff, R. M., & Renault, T. (1999). Purification and partial genome characterization of a herpes-like virus infecting the Japanese oyster, Crassostrea gigas. *Journal of General Virology*, *80*(5), Article 5. https://doi.org/10.1099/0022-1317-80-5-1317

Lecoq, H. (2001). Découverte du premier virus, le virus de la mosaïque du tabac: 1892 ou 1898 ? *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie*, *324*(10), 929–933. https://doi.org/10.1016/S0764-4469(01)01368-3

Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G., & Smith, D. B. (2018). Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research*, *46*(D1), Article D1. https://doi.org/10.1093/nar/gkx932

Lemey, P., Rambaut, A., Bedford, T., Faria, N. R., Bielejec, F., Baele, G., Russell, C. A., Smith, D. J., Pybus, O. G., Brockmann, D., & Suchard, M. A. (2012). *The seasonal flight of influenza: A unified framework for spatiotemporal hypothesis testing* (arXiv:1210.5877). arXiv. http://arxiv.org/abs/1210.5877

Lemey, P., Rambaut, A., Welch, J. J., & Suchard, M. A. (2010). Phylogeography Takes a Relaxed Random Walk in Continuous Space and Time. *Molecular Biology and Evolution*, *27*(8), 1877–1885. https://doi.org/10.1093/molbev/msq067

Leroy, É. M. (2015). L'Émergence du virus EBOLA chez l'homme: Un long processus pas totalement élucidé. *Bulletin De L'Academie Nationale De Medecine*, *199*(4), 651–671. https://doi.org/10.1016/S0001-4079(19)30940-9

Levine, O. S., & Levine, M. M. (1991). Houseflies (Musca domestica) as Mechanical Vectors of Shigellosis. *Reviews of Infectious Diseases*, *13*(4), 688–696. https://doi.org/10.1093/clinids/13.4.688

Lewontin, R. C. (1964). THE INTERACTION OF SELECTION AND LINKAGE. I. GENERAL CONSIDERATIONS; HETEROTIC MODELS. *Genetics*, *49*(1), 49–67. https://doi.org/10.1093/genetics/49.1.49

Li, H., & Roossinck, M. J. (2004). Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. *Journal of Virology*, *78*(19), 10582–10587. https://doi.org/10.1128/JVI.78.19.10582-10587.2004

Li, X., Han, M., Zhang, H., Liu, F., Pan, Y., Zhu, J., Liao, Z., Chen, X., & Zhang, B. (2022). Structures and biological functions of zinc finger proteins and their roles in hepatocellular carcinoma. *Biomarker Research*, *10*(1), 2. https://doi.org/10.1186/s40364-021-00345-1

Lin, J. T., Saunders, D. L., & Meshnick, S. R. (2014). The role of submicroscopic malaria in malaria transmission: What is the evidence? *Trends in Parasitology*, *30*(4), 183–190. https://doi.org/10.1016/j.pt.2014.02.004

Lipart, C., & Renault, T. (2002). Herpes-like virus detection in infected Crassostrea gigas spat using DIG-labelled probes. *Journal of Virological Methods*, *101*(1–2), Article 1–2. https://doi.org/10.1016/s0166-0934(01)00413-x

Liu, K.-K., & Shan, C.-X. (2023). Viral inactivation by irradiation rays. *Light: Science & Applications*, *12*(1), Article 1. https://doi.org/10.1038/s41377-023-01108-3

Loehle, C. (1995). Social barriers to pathogen transmission in wild animal populations. *Ecology*, *76*(2), 326–335.

Longdon, B., Brockhurst, M. A., Russell, C. A., Welch, J. J., & Jiggins, F. M. (2014). The Evolution and Genetics of Virus Host Shifts. *PLOS Pathogens*, *10*(11), e1004395. https://doi.org/10.1371/journal.ppat.1004395

Louten, J. (2016). Virus Replication. *Essential Human Virology*, 49–70. https://doi.org/10.1016/B978-0-12-800947-5.00004-1

Lupo, C., Ezanno, P., Arzul, I., Garcia, C., Jadot, C., Joly, J.-P., Renault, T., & Bareille, N. (2016). How network analysis of oyster movements can improve surveillance and control programs of infectious diseases? *Frontiers in Veterinary Science*, *3*. https://doi.org/10.3389/conf.FVETS.2016.02.00044

Lupo, C., Travers, M.-A., Tourbiez, D., Barthélémy, C. F., Beaunée, G., & Ezanno, P. (2019). Modeling the Transmission of Vibrio aestuarianus in Pacific Oysters Using Experimental Infection Data. *Frontiers in Veterinary Science*, *6*. https://www.frontiersin.org/articles/10.3389/fvets.2019.00142

Lynch, M. (2010). Evolution of the mutation rate. *Trends in Genetics*, *26*(8), 345–352. https://doi.org/10.1016/j.tig.2010.05.003

Lynch, M., Ackerman, M. S., Gout, J.-F., Long, H., Sung, W., Thomas, W. K., & Foster, P. L. (2016). Genetic drift, selection and the evolution of the mutation rate. *Nature Reviews Genetics*, *17*(11), Article 11. https://doi.org/10.1038/nrg.2016.104

Lynch, S. A., Carlsson, J., Reilly, A. O., Cotter, E., & Culloty, S. C. (2012). A previously undescribed ostreid herpes virus 1 (OsHV-1) genotype detected in the pacific oyster, Crassostrea gigas, in Ireland. *Parasitology*, *139*(12), Article 12. https://doi.org/10.1017/S0031182012000881

Maclachlan, N. J., Mayo, C. E., Daniels, P. W., Savini, G., Zientara, S., & Gibbs, E. P. J. (2015). Bluetongue: Fièvre catarrhale ovine. *Revue Scientifique et Technique de l'OIE*, *34*(2), 329–340. https://doi.org/10.20506/rst.34.2.2360

Magouras, I., Brookes, V. J., Jori, F., Martin, A., Pfeiffer, D. U., & Dürr, S. (2020). Emerging Zoonotic Diseases: Should We Rethink the Animal–Human Interface? *Frontiers in Veterinary Science*, *7*. https://www.frontiersin.org/articles/10.3389/fvets.2020.582743

Maiden, M. C. J., Jansen van Rensburg, M. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A., & McCarthy, N. D. (2013). MLST revisited: The gene-by-gene approach to bacterial genomics. *Nature Reviews. Microbiology*, *11*(10), 728–736. https://doi.org/10.1038/nrmicro3093

Marshall, N., Priyamvada, L., Ende, Z., Steel, J., & Lowen, A. C. (2013). Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathogens*, *9*(6), e1003421. https://doi.org/10.1371/journal.ppat.1003421

Martenot, C., Fourour, S., Oden, E., Jouaux, A., Travaillé, E., Malas, J. P., & Houssin, M. (2012). Detection of the OsHV-1 µVar in the Pacific oyster Crassostrea gigas before 2008 in France and description of two new microvariants of the Ostreid Herpesvirus 1 (OsHV-1). *Aquaculture*, *338–341*, 293–296. https://doi.org/10.1016/j.aquaculture.2011.12.030

Martenot, C., Oden, E., Travaillé, E., Malas, J.-P., & Houssin, M. (2011). Detection of different variants of Ostreid Herpesvirus 1 in the Pacific oyster, Crassostrea gigas between 2008 and 2010. *Virus Research*, *160*(1–2), Article 1–2. https://doi.org/10.1016/j.virusres.2011.04.012

Martinez, M., Grijalva, J., Castro, R., Chavez, J., Enriquez, T., Maeda, A., & Pena, E. (2017). OsHV 1 and notifiable protozoa in healthy Crassostrea corteziensis cultured in two distant areas of the Gulf of California. *Latin American Journal of Aquatic Research*, *45*(4), 699–707. https://doi.org/10.3856/vol45-issue4-fulltext-6

Mausner, J. S., & Bahn, A. K. (1974). *Epidemiology: An introductory text*. Saunders.

Mayer, A. (1886). Über die Mosaikkrankheit des Tabaks. In *Die Landwirtschaftlichen Versuchs-Stationen* (Vol. 32, pp. 451–467).

Maynard, J., & Haigh, J. (2007). The hitch-hiking effect of a favourable gene. *Genetics Research*, *89*(5–6), 391–403. https://doi.org/10.1017/S0016672308009579

McCallum, H. (2012). Disease and the dynamics of extinction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1604), 2828–2839. https://doi.org/10.1098/rstb.2012.0224

McGeoch, D. J., Dolan, A., & Ralph, A. C. (2000). Toward a Comprehensive Phylogeny for Mammalian and Avian Herpesviruses. *Journal of Virology*, *74*(22), 10401–10406. https://doi.org/10.1128/JVI.74.22.10401-10406.2000

McGeoch, D. J., Rixon, F. J., & Davison, A. J. (2006). Topics in herpesvirus genomics and evolution. *Virus Research*, *117*(1), 90–104. https://doi.org/10.1016/j.virusres.2006.01.002

McHardy, A. C., & Adams, B. (2009). The Role of Genomics in Tracking the Evolution of Influenza A Virus. *PLOS Pathogens*, *5*(10), e1000566. https://doi.org/10.1371/journal.ppat.1000566

McLaughlin, R. N., Jr, & Malik, H. S. (2017). Genetic conflicts: The usual suspects and beyond. *Journal of Experimental Biology*, *220*(1), 6–17. https://doi.org/10.1242/jeb.148148

Mello, D. F., Danielli, N. M., Curbani, F., Pontinha, V. A., Suhnel, S., Castro, M. A. M., Medeiros, S. C., Wendt, N. C., Trevisan, R., Magalhães, A. R. M., & Dafre, A. L. (2018). First evidence of viral and bacterial oyster pathogens in the Brazilian coast. *Journal of Fish Diseases*, *41*(3), Article 3. https://doi.org/10.1111/jfd.12755

Meyers, T., Burton, T., Evans, W., & Starkey, N. (2009). Detection of viruses and virus-like particles in four species of wild and farmed bivalve molluscs in Alaska, USA, from 1987 to 2009. *Diseases of Aquatic Organisms*, *88*, 1–12. https://doi.org/10.3354/dao02154

Miller, A. W., Reynolds, A. C., Sobrino, C., & Riedel, G. F. (2009). Shellfish Face Uncertain Future in High CO2 World: Influence of Acidification on Oyster Larvae Calcification and Growth in Estuaries. *PLOS ONE*, *4*(5), e5661. https://doi.org/10.1371/journal.pone.0005661

Moin, A. T., Chowdhury, M. A., Riana, S. H., Ullah, M. A., Araf, Y., Sarkar, B., & Shohael, A. M. (2021). An Updated Overview of Herpes Simplex Virus-1 Infection: Insights from Origin to Mitigation Measures. *Electronic Journal of General Medicine*, *18*(4), em299. https://doi.org/10.29333/ejgm/10869

Moreau, P., Faury, N., Burgeot, T., & Renault, T. (2015). Pesticides and Ostreid Herpesvirus 1 Infection in the Pacific Oyster, Crassostrea gigas. *PLOS ONE*, *10*(6), e0130628. https://doi.org/10.1371/journal.pone.0130628

Morens, D. M., Folkers, G. K., & Fauci, A. S. (2004). The challenge of emerging and re-emerging infectious diseases. *Nature*, *430*(6996), 242–249. https://doi.org/10.1038/nature02759

Morissette, G., & Flamand, L. (2010). Herpesviruses and Chromosomal Integration. *Journal of Virology*, *84*(23), Article 23. https://doi.org/10.1128/JVI.01169-10

Morrissey, T., McCleary, S., Collins, E., Henshilwood, K., & Cheslett, D. (2015). An investigation of ostreid herpes virus microvariants found in Crassostrea gigas oyster producing bays in Ireland. *Aquaculture*, *442*, 86–92. https://doi.org/10.1016/j.aquaculture.2015.02.026

Morse, S. S. (1995). Factors in the Emergence of Infectious Diseases. *Emerging Infectious Diseases*, *1*(1), 7–15. https://doi.org/10.3201/eid0101.950102

Mortensen, S., Strand, Å., Bodvin, T., Alfjorden, A., Skår, C., Jelmert, A., Aspán, A., Sælemyr, L., Naustvoll, L., & Albretsen, J. (2016). Summer mortalities and detection of ostreid herpesvirus microvariant in Pacific oyster Crassostrea gigas in Sweden and Norway. *Diseases of Aquatic Organisms*, *117*(3), Article 3. https://doi.org/10.3354/dao02944

Moss, J., Burreson, E., Cordes, J., Dungan, C., Brown, G., Wang, A., Wu, X., & Reece, K. (2007). Pathogens in Crassostrea ariakensis and other Asian oyster species: Implications for non-native oyster introduction to Chesapeake Bay. *Diseases of Aquatic Organisms*, *77*, 207–223. https://doi.org/10.3354/dao01829

Muylaert, I., Tang, K.-W., & Elias, P. (2011). Replication and Recombination of Herpes Simplex Virus DNA. *Journal of Biological Chemistry*, *286*(18), 15619–15624. https://doi.org/10.1074/jbc.R111.233981

Nahata, K. D., Bollen, N., Gill, M. S., Layan, M., Bourhy, H., Dellicour, S., & Baele, G. (2021). On the Use of Phylogeographic Inference to Infer the Dispersal History of Rabies Virus: A Review Study. *Viruses*, *13*(8), Article 8. https://doi.org/10.3390/v13081628

Neely, A. N., & Orloff, M. M. (2001). Survival of Some Medically Important Fungi on Hospital Fabrics and Plastics. *Journal of Clinical Microbiology*, *39*(9), 3360–3361. https://doi.org/10.1128/JCM.39.9.3360-3361.2001

Nelson, C. W., & Hughes, A. L. (2015). Within-Host Nucleotide Diversity of Virus Populations: Insights from Next-Generation Sequencing. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *0*, 1–7. https://doi.org/10.1016/j.meegid.2014.11.026

Netea, M. G., Schlitzer, A., Placek, K., Joosten, L. A. B., & Schultze, J. L. (2019). Innate and Adaptive Immune Memory: An Evolutionary Continuum in the Host's Response to Pathogens. *Cell Host & Microbe*, *25*(1), 13–26. https://doi.org/10.1016/j.chom.2018.12.006

Nicolas, J.-L., Comps, M., & Cochennec, N. (1992). *Herpes-like virus infecting Pacific oyster larvae.* (1). *12*(1), Article 1.

Northfield, T. D., & Ives, A. R. (2013). Coevolution and the Effects of Climate Change on Interacting Species. *PLoS Biology*, *11*(10), e1001685. https://doi.org/10.1371/journal.pbio.1001685

Nova, N., Athni, T. S., Childs, M. L., Mandle, L., & Mordecai, E. A. (2022). Global Change and Emerging Infectious Diseases. *Annual Review of Resource Economics*, *14*(1), 333–354. https://doi.org/10.1146/annurev-resource-111820-024214

Obbard, D. J., & Dudas, G. (2014). The genetics of host–virus coevolution in invertebrates. *Current Opinion in Virology*, *8*, 73–78. https://doi.org/10.1016/j.coviro.2014.07.002

Onihary, A. M., Razanajatovo, I. M., Rabetafika, L., Bastaraud, A., Heraud, J.-M., & Rasolofo, V. (2021). Genotype Diversity and Spread of White Spot Syndrome Virus (WSSV) in Madagascar (2012–2016). *Viruses*, *13*(9), 1713. https://doi.org/10.3390/v13091713

Ottmann, M. (2022). Ces virus qui nous habitent et qui nous visitent: Le virome humain. *Médecine/Sciences*, *38*(12), 1028–1038. https://doi.org/10.1051/medsci/2022161

Otto, S. P., & Whitlock, M. C. (1997). The Probability of Fixation in Populations of Changing Size. *Genetics*, *146*(2), 723–733. https://doi.org/10.1093/genetics/146.2.723

Pace, N. R., Sapp, J., & Goldenfeld, N. (2012). Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life. *Proceedings of the National Academy of Sciences*, *109*(4), 1011–1018. https://doi.org/10.1073/pnas.1109716109

Pappas, G. (2022). The Lanzhou Brucella Leak: The Largest Laboratory Accident in the History of Infectious Diseases? *Clinical Infectious Diseases*, *75*(10), 1845–1847. https://doi.org/10.1093/cid/ciac463

Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E.-C., Burke, D. S., Calisher, C. H., Laughlin, C. A., Saif, L. J., & Daszak, P. (2008). Cross-Species Virus Transmission and the Emergence of New Epidemic Diseases. *Microbiology and Molecular Biology Reviews*, *72*(3), 457–470. https://doi.org/10.1128/MMBR.00004-08

Pasteur, L. (1822-1895) A. du texte. (1922). *Oeuvres de Pasteur. Tome 4 / réunies par Pasteur Vallery-Radot,...* https://gallica.bnf.fr/ark:/12148/bpt6k73599

Patrício, A. R., Herbst, L. H., Duarte, A., Vélez-Zuazo, X., Santos Loureiro, N., Pereira, N., Tavares, L., & Toranzos, G. A. (2012). Global phylogeography and evolution of chelonid fibropapilloma-associated herpesvirus. *Journal of General Virology*, *93*(5), 1035–1045. https://doi.org/10.1099/vir.0.038950-0

Paul-Pont, I., Evans, O., Dhand, N. K., Rubio, A., Coad, P., & Whittington, R. J. (2014). Descriptive epidemiology of mass mortality due to Ostreid herpesvirus-1 (OsHV-1) in commercially farmed Pacific oysters (Crassostrea gigas) in the Hawkesbury River estuary, Australia. *Aquaculture*, *422–423*, 146–159. https://doi.org/10.1016/j.aquaculture.2013.12.009

Peck, K. M., & Lauring, A. S. (2018). Complexities of Viral Mutation Rates. *Journal of Virology*, *92*(14), 10.1128/jvi.01031-17. https://doi.org/10.1128/jvi.01031-17

Peeler, E. J., Allan Reese, R., Cheslett, D. L., Geoghegan, F., Power, A., & Thrush, M. A. (2012). Investigation of mortality in Pacific oysters associated with Ostreid herpesvirus-1µVar in the Republic of Ireland in 2009. *Preventive Veterinary Medicine*, *105*(1–2), Article 1–2. https://doi.org/10.1016/j.prevetmed.2012.02.001

Perdoncini Carvalho, C., Ren, R., Han, J., & Qu, F. (2022). Natural Selection, Intracellular Bottlenecks of Virus Populations, and Viral Superinfection Exclusion. *Annual Review of Virology*, *9*(1), 121–137. https://doi.org/10.1146/annurev-virology-100520-114758

Pérez-Losada, M., Arenas, M., Galán, J. C., Palero, F., & González-Candelas, F. (2015). Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. *Infection, Genetics and Evolution*, *30*, 296–307. https://doi.org/10.1016/j.meegid.2014.12.022

Pérez-Losada, M., Cabezas, P., Castro-Nallar, E., & Crandall, K. A. (2013). Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infection, Genetics and Evolution*, *16*, 38–53. https://doi.org/10.1016/j.meegid.2013.01.009

Pernet, F., Gachelin, S., Stanisière, J.-Y., Petton, B., Fleury, E., & Mazurié, J. (2019). Farmer monitoring reveals the effect of tidal height on mortality risk of oysters during a herpesvirus outbreak. *ICES Journal of Marine Science*, *76*(6), Article 6. https://doi.org/10.1093/icesjms/fsz074

Pernet, F., Lagarde, F., Jeannée, N., Daigle, G., Barret, J., Le Gall, P., Quere, C., & D'orbcastel, E. R. (2014). Spatial and Temporal Dynamics of Mass Mortalities in Oysters Is Influenced by Energetic Reserves and Food Quality. *PLoS ONE*, *9*(2), Article 2. https://doi.org/10.1371/journal.pone.0088469

Petton, B., Destoumieux-Garzón, D., Pernet, F., Toulza, E., de Lorgeril, J., Degremont, L., & Mitta, G. (2021). The Pacific Oyster Mortality Syndrome, a Polymicrobial and Multifactorial Disease: State of Knowledge and Future Directions. *Frontiers in Immunology*, *12*, 630343. https://doi.org/10.3389/fimmu.2021.630343

Petton, B., Pernet, F., Robert, R., & Boudry, P. (2013). Temperature influence on pathogen transmission and subsequent mortalities in juvenile Pacific oysters Crassostrea gigas. *Aquaculture Environment Interactions*, *3*(3), Article 3. https://doi.org/10.3354/aei00070

Poirotte, C., Massol, F., Herbert, A., Willaume, E., Bomo, P. M., Kappeler, P. M., & Charpentier, M. J. E. (2017). Mandrills use olfaction to socially avoid parasitized conspecifics. *Science Advances*, *3*(4), e1601721. https://doi.org/10.1126/sciadv.1601721

Polack, B., Boulouis, H.-J., Guillot, J., & Chermette, R. (2015). Les zoonoses (tableaux synthétiques: Animaux réservoirs de pathogènes et modes de transmission). *Revue Francophone Des Laboratoires*, *2015*(477), 67–79. https://doi.org/10.1016/S1773-035X(15)30318-X

Pybus, O. G., & Rambaut, A. (2009). Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics*, *10*(8), Article 8. https://doi.org/10.1038/nrg2583

Pybus, O. G., Rambaut, A., & Harvey, P. H. (2000). An Integrated Framework for the Inference of Viral Population History From Reconstructed Genealogies. *Genetics*, *155*(3), 1429–1437. https://doi.org/10.1093/genetics/155.3.1429

Rao, M., & Larsen, T. (2010). Ecological consequences of extinction. *Network of Conservation Educators and Practitioners*, *3*, 25–53. https://www.amnh.org/content/download/141367/2285419/file/ncep.amnh.org/linc/

Ren, W., Chen, H., Renault, T., Cai, Y., Bai, C., Wang, C., & Huang, J. (2013). Complete genome sequence of acute viral necrosis virus associated with massive mortality outbreaks in the Chinese scallop, Chlamys farreri. *Virology Journal*, *10*(1), 110. https://doi.org/10.1186/1743-422X-10-110

Renault & Arzul. (2001). Herpes-like virus infections in hatchery-reared bivalve larvae in Europe: Specific viral DNA detection by PCR. *Journal of Fish Diseases*, *24*(3), Article 3. https://doi.org/10.1046/j.1365-2761.2001.00282.x

Renault, T. (1998). Infection herpétiques chez les invertébrés: Détection de virus de type herpès chez les mollusques bivalves marins. *Annales de l'Institut Pasteur.*, *2*, Article 2. https://archimer.ifremer.fr/doc/00000/2903/

Renault, T., Cochennec, R., Le Deuff, R. M., & Chollet, B. (1994). *HERPES-LIKE VIRUS INFECTING JAPANESE OYSTER (CRASSOSTREA GIGAS) SPAT*. 3.

Renault, T., Deuff, R. M. L., Cochennec, N., & Maffart, P. (1994). *Herpesviruses associated with mortalities among Pacifie oyster, Crassostrea gigas, in France · Comparative study*. 8.

Renault, T., Faury, N., Barbosa-Solomieu, V., & Moreau, K. (2011). Suppression substractive hybridisation (SSH) and real time PCR reveal differential gene expression in the Pacific cupped oyster, Crassostrea gigas, challenged with Ostreid herpesvirus 1. *Developmental & Comparative Immunology*, *35*(7), Article 7. https://doi.org/10.1016/j.dci.2011.02.004

Renault, T., Le Deuff, R., Chollet, B., Cochennec, N., & Gerard, A. (2000). Concomitant herpes-like virus infections in hatchery-reared larvae and nursery-cultured spat Crassostrea gigas and Ostrea edulis. *Diseases of Aquatic Organisms*, *42*, 173–183. https://doi.org/10.3354/dao042173

Renault, T., Moreau, P., Faury, N., Pepin, J.-F., Segarra, A., & Webb, S. (2012). Analysis of Clinical Ostreid Herpesvirus 1 (Malacoherpesviridae) Specimens by Sequencing Amplified Fragments from Three Virus Genome Areas. *Journal of Virology*, *86*(10), 5942–5947. https://doi.org/10.1128/JVI.06534-11

Renault, T., Tchaleu, G., Faury, N., Moreau, P., Segarra, A., Barbosa-Solomieu, V., & Lapègue, S. (2014). Genotyping of a microsatellite locus to differentiate clinical Ostreid herpesvirus 1 specimens. *Veterinary Research*, *45*(1), Article 1. https://doi.org/10.1186/1297-9716-45-3

Rennekamp, A. J., & Lieberman, P. M. (2010). Initiation of lytic DNA replication in Epstein–Barr virus: Search for a common family mechanism. *Future Virology*, *5*(1), 65–83. https://doi.org/10.2217/fvl.09.69

Riley, L. W., & Blanton, R. E. (2018). Advances in Molecular Epidemiology of Infectious Diseases: Definitions, Approaches, and Scope of the Field. *Microbiology Spectrum*, *6*(6), 6.6.01. https://doi.org/10.1128/microbiolspec.AME-0001-2018

Roberts, J. D., Bebenek, K., & Kunkel, T. A. (1988). The Accuracy of Reverse Transcriptase from HIV-1. *Science*, *242*(4882), 1171–1173. https://doi.org/10.1126/science.2460925

Roque, A., Carrasco, N., Andree, K. B., Lacuesta, B., Elandaloussi, L., Gairin, I., Rodgers, C. J., & Furones, M. D. (2012). First report of OsHV-1 microvar in Pacific oyster (Crassostrea gigas) cultured in Spain. *Aquaculture*, *324–325*, 303–306. https://doi.org/10.1016/j.aquaculture.2011.10.018

Rosenberg, N. A., & Nordborg, M. (2002). Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics*, *3*(5), Article 5. https://doi.org/10.1038/nrg795

Ryu, W.-S. (2017). Virus Life Cycle. *Molecular Virology of Human Pathogenic Viruses*, 31–45. https://doi.org/10.1016/B978-0-12-800838-6.00003-5

Rzeżutka, A., & Cook, N. (2004). Survival of human enteric viruses in the environment and food. *FEMS Microbiology Reviews*, *28*(4), 441–453. https://doi.org/10.1016/j.femsre.2004.02.001

Saatkamp, H. W., Mourits, M. C. M., & Howe, K. S. (2016). A Framework for Categorization of the Economic Impacts of Outbreaks of Highly Contagious Livestock Diseases. *Transboundary and Emerging Diseases*, *63*(4), 422–434. https://doi.org/10.1111/tbed.12286

Said, M. S., Tirthani, E., & Lesho, E. (2023). Animal Zoonotic Related Diseases. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK570559/

Samuel, M. D., Woodworth, B. L., Atkinson, C. T., Hart, P. J., & LaPointe, D. A. (2015). Avian malaria in Hawaiian forest birds: Infection and population impacts across species and elevations. *Ecosphere*, *6*(6), art104. https://doi.org/10.1890/ES14-00393.1

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., Slocombe, P. M., & Smith, M. (1977). Nucleotide sequence of bacteriophage ϕX174 DNA. *Nature*, *265*(5596), Article 5596. https://doi.org/10.1038/265687a0

Sanjuán, R. (2012). From Molecular Genetics to Phylodynamics: Evolutionary Relevance of Mutation Rates Across Viruses. *PLOS Pathogens*, *8*(5), e1002685. https://doi.org/10.1371/journal.ppat.1002685

Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. *Cellular and Molecular Life Sciences*, *73*(23), 4433–4448. https://doi.org/10.1007/s00018-016-2299-6

Sanjuán, R., & Domingo-Calap, P. (2021). *Genetic Diversity and Evolution of Viral Populations.* 53–61. https://doi.org/10.1016/B978-0-12-809633-8.20958-8

Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., & Belshaw, R. (2010). Viral Mutation Rates. *Journal of Virology*, *84*(19), 9733–9748. https://doi.org/10.1128/JVI.00694-10

Sano, M., Fukuda, H., & Sano, T. (1990). Isolation and characterization of a new herpesvirus from eel. *Pathology in Marine Science*, 15–31.

Schrenzel, M. D., Tucker, T. A., Donovan, T. A., Busch, M. D. M., Wise, A. G., Maes, R. K., & Kiupel, M. (2008). New Hosts for Equine Herpesvirus 9. *Emerging Infectious Diseases*, *14*(10), 1616–1619. https://doi.org/10.3201/eid1410.080703

Segarra, A. (2015). *Étude des interactions hôte/virus chez l'huître creuse, Crassostrea gigas, et son virus Ostreid herpesvirus 1*. 240.

Segarra, A., Baillon, L., Tourbiez, D., Benabdelmouna, A., Faury, N., Bourgougnon, N., & Renault, T. (2014). Ostreid herpesvirus type 1 replication and host response in adult Pacific oysters,

Crassostrea gigas. *Veterinary Research*, *45*(1), 103. https://doi.org/10.1186/s13567-014-0103-x

Segarra, A., Mauduit, F., Faury, N., Trancart, S., Dégremont, L., Tourbiez, D., Haffner, P., Barbosa-Solomieu, V., Pépin, J.-F., Travers, M.-A., & Renault, T. (2014). Dual transcriptomics of virus-host interactions: Comparing two Pacific oyster families presenting contrasted susceptibility to ostreid herpesvirus 1. *BMC Genomics*, *15*(1), Article 1. https://doi.org/10.1186/1471-2164-15-580

Segarra, A., Pépin, J. F., Arzul, I., Morga, B., Faury, N., & Renault, T. (2010). Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, Crassostrea gigas, in France in 2008. *Virus Research*, *153*(1), Article 1. https://doi.org/10.1016/j.virusres.2010.07.011

Serquiña, A. K., & Ziegelbauer, J. M. (2017). How herpesviruses pass on their genomes. *The Journal of Cell Biology*, *216*(9), Article 9. https://doi.org/10.1083/jcb.201708077

Shimahara, Y., Kurita, J., Kiryu, I., Nishioka, T., Yuasa, K., Kawana, M., Kamaishi, T., & Oseko, N. (2012). Surveillance of Type 1 Ostreid Herpesvirus (OsHV-1) Variants in Japan. *Fish Pathology*, *47*(4), Article 4. https://doi.org/10.3147/jsfp.47.129

Sigwart, J. (2009). Coalescent Theory: An Introduction. *Systematic Biology*, *58*(1), 162–165. https://doi.org/10.1093/schbul/syp004

Simmonds, P., Aiewsakun, P., & Katzourakis, A. (2019). Prisoners of war—Host adaptation and its constraints on virus evolution. *Nature Reviews Microbiology*, *17*(5), Article 5. https://doi.org/10.1038/s41579-018-0120-2

Slatkin, M. (2008). Linkage disequilibrium—Understanding the evolutionary past and mapping the medical future. *Nature Reviews. Genetics*, *9*(6), 477–485. https://doi.org/10.1038/nrg2361

Small, H. J., & Pagenkopp, K. M. (2011). Reservoirs and alternate hosts for pathogens of commercially important crustaceans: A review. *Journal of Invertebrate Pathology*, *106*(1), 153–164. https://doi.org/10.1016/j.jip.2010.09.016

Smith, L. M., McWhorter, A. R., Shellam, G. R., & Redwood, A. J. (2013). The genome of murine cytomegalovirus is shaped by purifying selection and extensive recombination. *Virology*, *435*(2), 258–268. https://doi.org/10.1016/j.virol.2012.08.041

Steere, A. C., Coburn, J., & Glickstein, L. (2004). The emergence of Lyme disease. *Journal of Clinical Investigation*, *113*(8), 1093–1101. https://doi.org/10.1172/JCI200421681

Steinhauer, D. A., Domingo, E., & Holland, J. J. (1992). Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene*, *122*(2), 281–288. https://doi.org/10.1016/0378-1119(92)90216-C

Stephan, W. (2019). Selective Sweeps. *Genetics*, *211*(1), 5–13. https://doi.org/10.1534/genetics.118.301319

Sumption, K. J., & Flowerdew, J. R. (1985). The ecological effects of the decline in Rabbits (Oryctolagus cuniculus L.) due to myxomatosis. *Mammal Review*, *15*(4), 151–186. https://doi.org/10.1111/j.1365-2907.1985.tb00396.x

Szpara, M. L., Gatherer, D., Ochoa, A., Greenbaum, B., Dolan, A., Bowden, R. J., Enquist, L. W., Legendre, M., & Davison, A. J. (2014). Evolution and diversity in human herpes simplex virus genomes. *Journal of Virology*, *88*(2), Article 2. https://doi.org/10.1128/JVI.01987-13

Tabassum, A., Ramesh, S. V., Zhai, Y., Iftikhar, R., Olaya, C., & Pappu, H. R. (2021). Viruses Without Borders: Global Analysis of the Population Structure, Haplotype Distribution, and Evolutionary Pattern of Iris Yellow Spot Orthotospovirus (Family Tospoviridae, Genus Orthotospovirus). *Frontiers in Microbiology*, *12*, 633710. https://doi.org/10.3389/fmicb.2021.633710

Tadesse, F. G., Slater, H. C., Chali, W., Teelen, K., Lanke, K., Belachew, M., Menberu, T., Shumie, G., Shitaye, G., Okell, L. C., Graumans, W., van Gemert, G.-J., Kedir, S., Tesfaye, A., Belachew, F., Abebe, W., Mamo, H., Sauerwein, R., Balcha, T., … Bousema, T. (2018). The Relative Contribution of Symptomatic and Asymptomatic Plasmodium vivax and Plasmodium

falciparum Infections to the Infectious Reservoir in a Low-Endemic Setting in Ethiopia. *Clinical Infectious Diseases*, *66*(12), 1883–1891. https://doi.org/10.1093/cid/cix1123

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, *123*(3), 585–595. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1203831/

Tills, D. (2008). The Use of the FST; Statistic of Wright for Estimating the Effects of Genetic Drift, Selection and Migration in Populations, with Special Reference to Ireland. *Human Heredity*, *27*(2), 153–159. https://doi.org/10.1159/000152865

Trovato, M., Sartorius, R., D'Apice, L., Manco, R., & De Berardinis, P. (2020). Viral Emerging Diseases: Challenges in Developing Vaccination Strategies. *Frontiers in Immunology*, *11*. https://www.frontiersin.org/articles/10.3389/fimmu.2020.02130

Van Borm, S., Belák, S., Freimanis, G., Fusaro, A., Granberg, F., Höper, D., King, D. P., Monne, I., Orton, R., & Rosseel, T. (2014). Next-Generation Sequencing in Veterinary Medicine: How Can the Massive Amount of Information Arising from High-Throughput Technologies Improve Diagnosis, Control, and Management of Infectious Diseases? *Veterinary Infection Biology: Molecular Diagnostics and High-Throughput Strategies*, *1247*, 415–436. https://doi.org/10.1007/978-1-4939-2004-4_30

van Doorn, H. R. (2014). Emerging infectious diseases. *Medicine*, *42*(1), 60–63. https://doi.org/10.1016/j.mpmed.2013.10.014

van Seventer, J. M., & Hochberg, N. S. (2017). Principles of Infectious Diseases: Transmission, Diagnosis, Prevention, and Control. *International Encyclopedia of Public Health*, 22–39. https://doi.org/10.1016/B978-0-12-803678-5.00516-6

Vasquez Yeomans, R., Cáceres-Martínez, J., & Huerta, A. F. (2004). Herpes-like virus associated with eroded gills of the Pacific oyster Crassostrea gigas in Mexico. *Journal of Shellfish Research*, *23*, 417–419.

Vijaykrishna, D., Mukerji, R., & Smith, G. J. D. (2015). RNA Virus Reassortment: An Evolutionary Mechanism for Host Jumps and Immune Evasion. *PLoS Pathogens*, *11*(7), e1004902. https://doi.org/10.1371/journal.ppat.1004902

Vogeler, S., Bean, T. P., Lyons, B. P., & Galloway, T. S. (2016). Dynamics of nuclear receptor gene expression during Pacific oyster development. *BMC Developmental Biology*, *16*(1), 33. https://doi.org/10.1186/s12861-016-0129-6

Volz, E. M., Koelle, K., & Bedford, T. (2013). Viral Phylodynamics. *PLoS Computational Biology*, *9*(3), e1002947. https://doi.org/10.1371/journal.pcbi.1002947

Webb, S. C., Fidler, A., & Renault, T. (2007). Primers for PCR-based detection of ostreid herpes virus-1 (OsHV-1): Application in a survey of New Zealand molluscs. *Aquaculture*, *272*(1–4), Article 1–4. https://doi.org/10.1016/j.aquaculture.2007.07.224

Weidner-Glunde, M., Kruminis-Kaszkiel, E., & Savanagouder, M. (2020). Herpesviral Latency—Common Themes. *Pathogens*, *9*(2), 125. https://doi.org/10.3390/pathogens9020125

Weyer, J., Grobbelaar, A., & Blumberg, L. (2015). Ebola Virus Disease: History, Epidemiology and Outbreaks. *Current Infectious Disease Reports*, *17*(5), 21. https://doi.org/10.1007/s11908-015-0480-y

WHO. (2008). Anthrax in Humans and animals—Etiology and ecology. In *Anthrax in Humans and Animals. 4th edition*. World Health Organization. https://www.ncbi.nlm.nih.gov/books/NBK310478/

WHO. (2020). *Transmission of SARS-CoV-2: Implications for infection prevention precautions*. https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions

WOAH. (2007). *Economic analysis—Prevention versus outbreak costs*.

WOAH. (2023a). *Avian Influenza*. WOAH - World Organisation for Animal Health. https://www.woah.org/en/disease/avian-influenza/

WOAH. (2023b). *Bluetongue*. WOAH - World Organisation for Animal Health. https://www.woah.org/en/disease/bluetongue/

Woolhouse, M. E. J., Taylor, L. H., & Haydon, D. T. (2001). Population Biology of Multihost Pathogens. *Science*, *292*(5519), 1109–1112. https://doi.org/10.1126/science.1059026

Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics*, *32*(4), Article 4. https://doi.org/10.1038/ng1202-569

Woźniakowski, G., & Samorek-Salamonowicz, E. (2015). Animal herpesviruses and their zoonotic potential for cross-species infection. *Annals of Agricultural and Environmental Medicine*, *22*(2), 191–194. https://doi.org/10.5604/12321966.1152063

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*(2), 97.

Xia, J., Bai, C., Wang, C., Song, X., & Huang, J. (2015). Complete genome sequence of Ostreid herpesvirus-1 associated with mortalities of Scapharca broughtonii broodstocks. *Virology Journal*, *12*(1), Article 1. https://doi.org/10.1186/s12985-015-0334-0

Young, A. D., & Gillung, J. P. (2020). Phylogenomics—Principles, opportunities and pitfalls of big-data phylogenetics. *Systematic Entomology*, *45*(2), 225–247. https://doi.org/10.1111/syen.12406

Zarrouk, K., Piret, J., & Boivin, G. (2017). Herpesvirus DNA polymerases: Structures, functions and inhibitors. *Virus Research*, *234*, 177–192. https://doi.org/10.1016/j.virusres.2017.01.019

Zhang, Q.-Y., Ke, F., Gui, L., & Zhao, Z. (2022). Recent insights into aquatic viruses: Emerging and reemerging pathogens, molecular features, biological effects, and novel investigative approaches. *Water Biology and Security*, *1*(4), 100062. https://doi.org/10.1016/j.watbs.2022.100062

Zhao, M., dos Santos Tavares, C. P., & Schott, E. J. (2021). Diversity and classification of reoviruses in crustaceans: A proposal. *Journal of Invertebrate Pathology*, *182*, 107568. https://doi.org/10.1016/j.jip.2021.107568

Annexes

Annexe 1

Article 3

Genomic diversity of the Ostreid Herpesvirus Type 1 across time and location and among host species

Benjamin Morga, Maude Jacquot, Camille Pelletier, Germain Chevignon, Lionel Dégremont, Antoine Biétry, Jean-François Pepin, Serge Heurtebise, Jean-Michel Escoubas, Tim Bean, *et al.*

frontiers in Microbiology

ORIGINAL RESEARCH
published: 13 July 2021
doi: 10.3389/fmicb.2021.711377

Genomic Diversity of the Ostreid Herpesvirus Type 1 Across Time and Location and Among Host Species

Benjamin Morga[1]†, Maude Jacquot[1†], Camille Pelletier[1], Germain Chevignon[1], Lionel Dégremont[1], Antoine Biétry[1], Jean-François Pepin[2], Serge Heurtebise[1], Jean-Michel Escoubas[3], Tim P. Bean[4,5], Umberto Rosani[6], Chang-Ming Bai[7], Tristan Renault[8] and Jean-Baptiste Lamy[1]*

[1] Ifremer, RBE-SGMM-LGPMM, La Tremblade, France, [2] Ifremer, ODE-Littoral-Laboratoire Environnement Ressources des Pertuis Charentais (LER-PC), La Tremblade, France, [3] IHPE, CNRS, Ifremer, Université de Montpellier – Université de Perpignan Via Domitia, Montpellier, France, [4] The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian, United Kingdom, [5] Centre for Environment, Fisheries and Aquaculture Science, Weymouth, United Kingdom, [6] Department of Biology, University of Padua, Padua, Italy, [7] Yellow Sea Fisheries Research Institute, CAFS, Qingdao, China, [8] Ifremer, RBE, Nantes, France

OPEN ACCESS

Edited by:
Andrew S. Lang, Memorial University of Newfoundland, Canada

Reviewed by:
Giuseppe Arcangeli, Experimental Zooprophylactic Institute of Venice (IZSVe), Italy
Arshan Nasir, Los Alamos National Laboratory (DOE), United States

***Correspondence:**
Benjamin Morga
Benjamin.Morga@ifremer.fr

†*These authors share first authorship*

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

Received: 18 May 2021
Accepted: 21 June 2021
Published: 13 July 2021

Citation:
Morga B, Jacquot M, Pelletier C, Chevignon G, Dégremont L, Biétry A, Pepin J-F, Heurtebise S, Escoubas J-M, Bean TP, Rosani U, Bai C-M, Renault T and Lamy J-B (2021) Genomic Diversity of the Ostreid Herpesvirus Type 1 Across Time and Location and Among Host Species. Front. Microbiol. 12:711377. doi: 10.3389/fmicb.2021.711377

The mechanisms underlying virus emergence are rarely well understood, making the appearance of outbreaks largely unpredictable. This is particularly true for pathogens with low per-site mutation rates, such as DNA viruses, that do not exhibit a large amount of evolutionary change among genetic sequences sampled at different time points. However, whole-genome sequencing can reveal the accumulation of novel genetic variation between samples, promising to render most, if not all, microbial pathogens measurably evolving and suitable for analytical techniques derived from population genetic theory. Here, we aim to assess the measurability of evolution on epidemiological time scales of the Ostreid herpesvirus 1 (OsHV-1), a double stranded DNA virus of which a new variant, OsHV-1 µVar, emerged in France in 2008, spreading across Europe and causing dramatic economic and ecological damage. We performed phylogenetic analyses of heterochronous (n = 21) OsHV-1 genomes sampled worldwide. Results show sufficient temporal signal in the viral sequences to proceed with phylogenetic molecular clock analyses and they indicate that the genetic diversity seen in these OsHV-1 isolates has arisen within the past three decades. OsHV-1 samples from France and New Zealand did not cluster together suggesting a spatial structuration of the viral populations. The genome-wide study of simple and complex polymorphisms shows that specific genomic regions are deleted in several isolates or accumulate a high number of substitutions. These contrasting and non-random patterns of polymorphism suggest that some genomic regions are affected by strong selective pressures. Interestingly, we also found variant genotypes within all infected individuals. Altogether, these results provide baseline evidence that whole genome sequencing could be used to study population dynamic processes of OsHV-1, and more broadly herpesviruses.

Keywords: marine virus, OsHV-1, diversity, shellfish farming, measurably evolution

INTRODUCTION

A common paradigm is that double-stranded DNA (dsDNA) viruses, and particularly herpesviruses, are inherently stable (Sanjuán et al., 2010; Sanjuán and Domingo-Calap, 2016) making them unsuitable for analytical techniques derived from population genetic theory. At the molecular level, this view stems from low per-site mutation rates due to the high-fidelity polymerases and error correction systems typical of these viruses. However, recent progress has been made in understanding the genomic diversity and evolution of mainly human herpesviruses, driven by the rapid expansion and application of high throughput sequencing (HTS), bioinformatics, and comparative genomics in virology. The ability of dsDNA viruses to accrue standing variation, and to undergo recombination with neighboring genomes, creates many opportunities for selective pressures to induce rapid genetic shifts (for a review of factors see Sanjuán and Domingo-Calap, 2016). In the meantime, estimates of the substitution rates of human herpesviruses and cytomegaloviruses are substantially higher than previously though (Kitchen et al., 2010; Jaramillo et al., 2013; Parsons et al., 2015; Renner and Szpara, 2017) and evidences that several viral haplotypes can be found within one infected host (at the individual scale) have been provided (Renzette et al., 2011; Cudini et al., 2019). Together, these data have reshaped our sense of the stability of herpesvirus genomes. Because contemporary viral genomic diversity is a result of the dynamic interaction of past ecological and evolutionary processes (Retel et al., 2019), opportunities to assess their evolution, at a genome-wide scale, to infer their population dynamics from genomic data (Biek et al., 2015) offer new perspectives for better understanding and preventing herpesviruses outbreaks.

Ostreid herpesvirus 1 (OsHV-1) is a dsDNA herpesvirus and the unique member of the genus Ostreavirus (family Malacoherpesviridae, order Herpesvirales). It was the first herpesvirus isolated from invertebrates in the early 1990s (Nicolas et al., 1992) and likely is the causative agent of the last decades increased mortality events in Pacific oysters, *Crassostrea gigas* and therefore responsible for dramatic economic damages (EFSA, 2010; Segarra et al., 2010). Since the appearance of the OsHV-1 μVar variant in 2008, we have observed high mortality following the detection of this variant, the means to fight against the infection of this virus in oyster are limited.

Ostreid herpesvirus 1 infects mainly spat and juvenile *C. gigas* and is currently detected in most oyster producing area worldwide (Nicolas et al., 1992; Garcia et al., 2011; Paul-Pont et al., 2013). Furthermore, this virus is able to infect at least four more distantly related marine bivalve species (*Ostrea edulis*, *Chlamys farreri*, *Pecten maximus Cerastoderma edule*, *and Anadara broughtonii*; Arzul et al., 2001, 2017; Davison et al., 2005; Bai et al., 2015; Bookelaar et al., 2020). However, whether viral populations infecting the different host species are the same remains unknown.

The first variant of OsHV-1 was described on *P. maximus* larvae and was named OsHV-1 Var but the associated disease was not well documented (Arzul et al., 2001). Since 2008, microvariants of OsHV-1 have been associated with increasing mortality of *C. gigas* in France, and in a number of countries across Europe, as well as in Australia and New Zealand (Segarra et al., 2010; Jenkins et al., 2013). OsHV-1 μVar is a single variant which differs from the reference genome published in 2005 by Davidson (accession number NC_005881.2) mainly because of sequence variations in a microsatellite locus upstream of the Open Reading Frame (ORF) 4, in ORF4 and in ORF42/43. The term "microvariants" is now used to refer to all haplotypes with mutations in and upstream of ORF 4 and in ORF 42/43. Other variants have also been described, for example OsHV-1-SB which has been associated with mass mortalities of *S. broughtonii* in China (Bai et al., 2015, 2016).

Following the emergence of OsHV-1 μVar, a number of studies have been undertaken to investigate the diversity of OsHV-1 using single- (Renault et al., 2014) or multi-genomic regions (Renault et al., 2012). Three OsHV-1 genome sequences were obtained with bacteriophage lambda libraries or genome walking procedures (Davison et al., 2005; Ren et al., 2013; Xia et al., 2015). More recently, three additional OsHV-1 μVar genomes were sequenced using Illumina shotgun DNA-seq technology (Burioli et al., 2017; Abbadi et al., 2018).

While sequencing allows us to characterize polymorphism and genomic variation at the genome scale with unprecedented power, in the case of OsHV-1, diversity has not yet been explored at the genome scale, or across different dimensions such as time, location and host species. Recently a study of viral diversity was performed at the viral mRNA level, which demonstrated viral populations differ depending on the genetic background of their hosts. This result suggests a co-evolution process between OsHV-1 μVar and oyster populations (Delmotte et al., 2020).

Therefore, using a reference mapping approach from historical collection of deep sequenced infected animals, we aim to assess the measurability of evolution on epidemiological time scales of OsHV-1. We also aim at characterizing the potential of the approach to test (i) if infected individuals belonging to different host species harbor different viral populations with large genomic differences possibly due to viral speciation and, (ii) if infected individual from the same host species but originated from different world region harbor viral populations with the same genomic structure and simple polymorphisms.

MATERIALS AND METHODS

Sample Collection

Samples analyzed in this study come from a combination of sources, including an historical collection and the European and National Reference Laboratory for Bivalve Mollusc Diseases. For the rest of this paper, we refer to "*isolates*" in the sense of viral populations isolated from infected individual host. Most of the data derives from whole-genome sequencing of infected individual host or pools of infected individual hosts (**Table 1**). The 22 isolates in this study were collected from

TABLE 1 | Description of the 22 sequenced libraries in this study: country of origin, host species, year of collection, sample stage, sequencing technology, and number of reads aligned against Ostreid herpesvirus 1 (*OsHV-1*) reference genome (Accession number: NC_005881.2).

Sample's ID	Country of origin	Species	Year	Stage	Sequencer, sequencing protocol, platform	Aligned reads
VIV46-2-m	France	*Crassostrea gigas*	2017	Juvenile	HiSeq 4000, Paired-end 150 bp, LIGAN-PM Genomics	281,336
VIV56-10-m	France	*Crassostrea gigas*	2017	Juvenile	HiSeq 4000, Paired-end 150 bp, LIGAN-PM Genomics	346,491
VIV48-4-m	France	*Crassostrea gigas*	2017	Juvenile	HiSeq 4000, Paired-end 150 bp, LIGAN-PM Genomics	354836
VIV49-5-m88	France	*Crassostrea gigas*	2017	Juvenile	HiSeq 4000, Paired-end 150 bp, LIGAN-PM Genomics	183,665
VIV58-12-m	France	*Crassostrea gigas*	2017	Juvenile	HiSeq 4000, Paired-end 150 bp, LIGAN-PM Genomics	155,976
VIV47-3-m	France	*Crassostrea gigas*	2017	Juvenile	HiSeq 4000, Paired-end 150 bp, LIGAN-PM Genomics	370,770
VIV57-11-m99	France	*Crassostrea gigas*	2017	Juvenile	HiSeq 4000, Paired-end 150 bp, LIGAN-PM Genomics	49,985
Poole-Harbour United Kingdom	United-Kingdom	*Crassostrea gigas*	2015	na	MiSeq, Paired-end 150 bp	55,056
LI	France	*Crassostrea gigas*	2010	Pool of larvae	HiSeq 2500, Paired-end 100 bp, GenoToul	400,021
PR	France	*Crassostrea gigas*	2008	Pool of juveniles	HiSeq 2500, Single-end 100 bp, GATC Biotech	344,852
MV	France	*Crassostrea gigas*	2010	Ripe Adult	HiSeq 2500, Single-end 100 bp, GATC Biotech	331,278
NZ	New-Zealand	*Crassostrea gigas*	2010	Pool of larvae	HiSeq 2500, Single-end 100 bp, GATC Biotech	341256
NZ16	New-Zealand	*Crassostrea gigas*	2011	Adult	HiSeq 2500, Paired-end 100 bp, GATC Biotech	105,610
NZ17	New-Zealand	*Crassostrea* gigas	2011	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	202,246
NZ18	New-Zealand	*Crassostrea gigas*	2011	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	150,315
IRL15*	Ireland	*Crassostrea gigas*	2011	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	1,802
JP2*	Japan	*Crassostrea gigas*	na	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	2,115
JP6*	Japan	*Crassostrea gigas*	na	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	3,280
NL4*	Netherlands	*Crassostrea gigas*	na	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	4,143
SP16*	Spain	*Crassostrea gigas*	na	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	1,259
SW3*	Sweden	*Ostrea edulis*	2012	Pool of larvae	HiSeq 2500, Paired-end 100 bp, GenoToul	1,839
SW6	Sweden	*Ostrea edulis*	2012	Adult	HiSeq 2500, Paired-end 100 bp, GenoToul	402,082

Samples for which the ID starts by "VIV" are from research project Vivaldi and the remaining samples are from historical collections.
**Those samples did not pass the quality check and were discarded.*

several countries around the world: France (eight infected juveniles *C. gigas*, two pool of infected *C. gigas* oysters, and one *C. gigas* adult), New-Zealand (one infected pool of *C. gigas* larvae and three infected *C. gigas* adults), Ireland (one *C. gigas* infected adult), Japan (two *C. gigas* infected adults), Sweden (two pooled infected *O. edulis* larvae and one infected adult), United-Kingdom (one *C. gigas* individual), Netherlands (one *Crassostrea gigas* infected adult), and Spain (one *C. gigas* infected adult; **Table 1**).

The historical samples were from moribund oysters collected during mortality outbreaks. For spat, juvenile and adults, a piece of mantle or gill was collected and stored in 90% ethanol until DNA extraction. For larvae samples, a pool of individuals was frozen (−20°C) and stored until DNA extraction.

All "VIV" oysters were sampled during a mortality outbreak in July 2017 (**Table 1**). Briefly, spat were placed into the field (*17390 La Tremblade Ronce-les-Bains France, a sandbank called "La Floride»: LAT 45.80°* and *LONG -1.15°*). As soon as the

onset of the mortality, spat were transferred in the laboratory and checked twice a day (from 18/07/2017 to 21/07/2017) to sample fresh moribund individuals.

Genomic Library Preparation and Sequencing Runs

A range of 60–200 mg of pooled larval samples were mechanically crushed using unique piston pellet before DNA extraction. About 50 mg of fresh tissue were collected from spat-juveniles-adults. Nucleic acid extraction was performed using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's handbook. All extracted DNA samples were stored at −25°C to avoid any degradation and DNA concentrations were measured with the Nanodrop™ (Labtech, France). After storage, DNA quality was assessed with Nanodrop™ and checked again by the sequencing platform using fluorometric measurement of nucleic acids.

All the libraries were sequenced using Illumina technology, with HiSeq 2000 sequencers on two platforms, GATC (Mulhouse, France) and GenoToul (Toulouse, France), with the exception of United Kingdom samples that were sequenced with MiSeq on in-house sequencers, and the HiSeq 4000 for the most recent samples on LIGAN (CNRS, Lille, France). Regarding the sequencing protocols, three libraries were sequenced on single-end protocol (1 × 100 bp) and remaining libraries were sequenced with paired-end protocol (2 × 150 bp) (**Table 1**), with the exception of the Miseq data, which was generated using V3-600 Miseq specific cartridge. Raw sequence data are available on the NCBI Sequence Read Archive (SRA) under the BioProject ID: PRJNA721248.

Bioinformatics and Genomic Framework

Read Mapping and Filtration

A reference mapping approach was used to explore the viral genomic diversity within and between samples. This approach has been extensively used for eukaryotic organisms and is best suited to call single nucleotide polymorphisms (SNPs) and genomic structural variation (SV) from a collection of individuals or pooled libraries in comparison to the more widely used *de novo* assembly strategies in the viral community (Baaijens et al., 2017). All sequenced libraries were aligned against the OsHV-1 reference genome (Accession number: NC_005881.2) (Davison et al., 2005) using BWA MEM v7.1.0 (Li and Durbin, 2009; Li, 2013). Only samples with a mean of coverage depth equal to or above 100X have been included for polymorphism calling and 10X for haplotype reconstruction (**Figure 1A**).

Local Haplotypes Reconstruction

We first aimed to reconstruct local haplotypes. As far as we are aware, no software dedicated to haplotype reconstruction is able to manage long indels, as it is observed in OsHV-1 genomes, with short reads (i.e., ∼100 bp reads long). To circumvent this limitation, we decided to make local haplotype reconstruction on each gene coding DNA sequence (CDS) that were present in every sample and outside of inverted and repeat regions. Each CDS coordinate were extracted from the GenBank file (Clark et al., 2016) and we use them to subset the bam file for each sample (Locus_tag, **Figure 1B**). We feed QuasiRecomb (Töpfer et al., 2013), haploclique (Töpfer et al., 2014), and shorah (Zagordi et al., 2010, 2011) with CDS-split bam files (**Figure 1E**). Only results from Shorah software are showed on the main text because it represents the most conservative short read local assembler, whereas the other software outputs are showed in supplementary figures. Briefly, Shorah employs a model-based probabilistic clustering algorithm to correct errors, infer haplotypes and their frequencies, and estimate, using a Bayesian approach, the quality of the reconstruction by computing the full joint posterior probability distribution of all parameters of interest (Zagordi et al., 2010).

Subsequent plotting were made in R Development Core Team (2008) with basics functions and the ggplot2 package (Wickham, 2016) (**Figure 1H**).

Short and Long Polymorphism Calling

The raw bam files were also processed with Freebayes software (Garrison and Marth, 2012; **Figure 1D**) to call short polymorphisms. Freebayes is one of the most versatile variants calling software, particularly when the libraries included pool of viral haplotypes as it is in the present study and has been extensively tested and validated on various organisms with different protocols (Olson et al., 2005; Wang et al., 2005). We discarded genomic regions with low complexity (entropy > 1) and regions with a high number of misalignments namely TRL, IRL, IRS, and TRS genomic regions that are duplicated.

Variation of coverage depth and, when available, the orientation of the paired-reads against the reference genome were jointly analyzed to detect SV, deletions, duplications, insertions, inversions, and translocations (Davison et al., 2005). We use Sniffles to perform the SV detection (**Figure 1C**; Sedlazeck et al., 2018), all the inferred structural variants were checked manually and filtered out as soon as we suspected some alignments artefacts. Most inferred deletions were visually checked with IGV (Thorvaldsdóttir et al., 2013), the remaining types of SV (insertions, duplications, inversions, and translocations) were rare, difficult to check manually and consequently we did not consider those in the rest of the paper. Information from both VCF files, short and long polymorphisms, were merged using vcflib (Garrison, 2019).

Consensus Sequence Reconstruction

For each sample, we reconstructed a "consensus sequence" based on the most frequent alleles for SNPs and short insertions/deletions. Into this consensus, we also integrated all long deletions found in the SV calling step, only all "large deletions" inferred by Sniffles and visually checked (**Figure 1C**, **Supplementary Table 1**, and **Supplementary Figure 1**). To create these consensus sequences, we use a combination of bcftools (Li et al., 2009; Li, 2011) and bedtools (Quinlan and Hall, 2010; **Figure 1F**).

Phylogenomic Analysis of Consensus Sequence

We choose to work on the full-length genome (**Figures 1C,D,F,G**) because among the Malacoherpesviridae family most of the transcripts and proteins have no functional annotation.

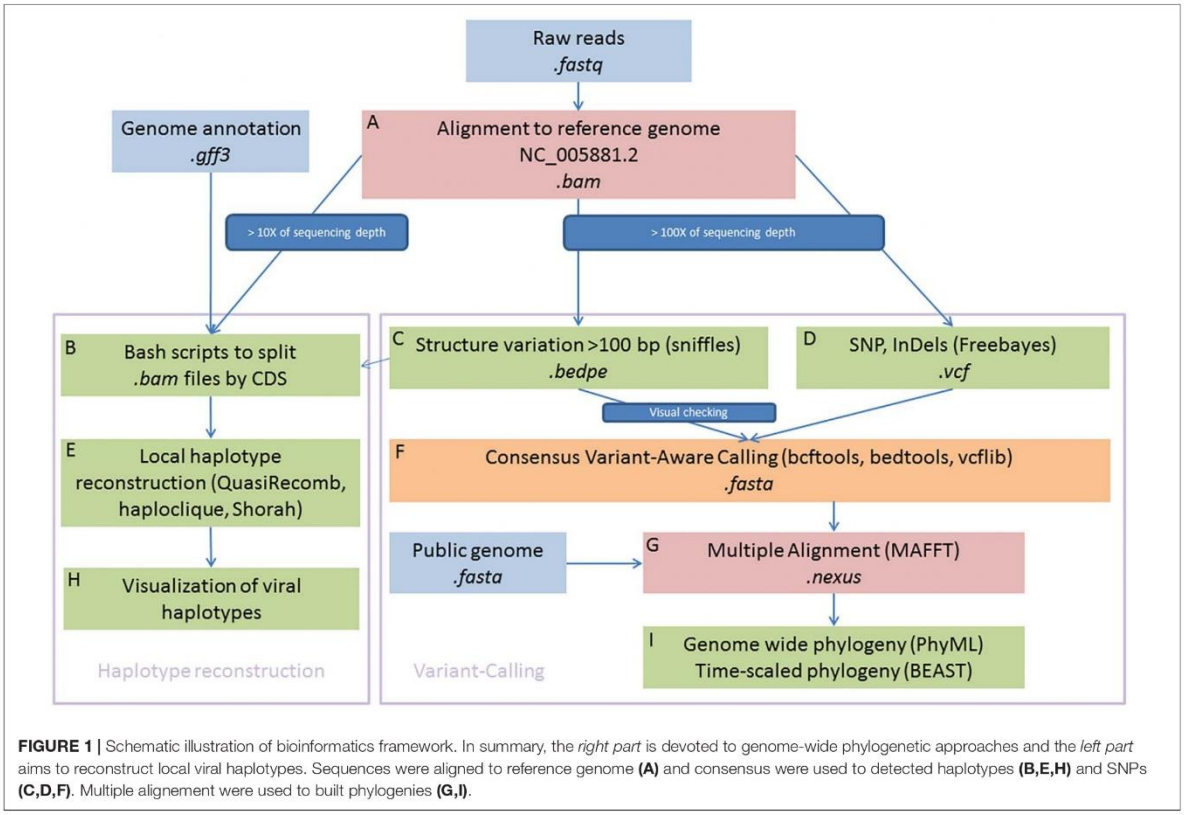


FIGURE 1 | Schematic illustration of bioinformatics framework. In summary, the *right part* is devoted to genome-wide phylogenetic approaches and the *left part* aims to reconstruct local viral haplotypes. Sequences were aligned to reference genome **(A)** and consensus were used to detected haplotypes **(B,E,H)** and SNPs **(C,D,F)**. Multiple alignement were used to built phylogenies **(G,I)**.

All consensus sequences for each samples, plus the publicly available full length OsHV-1 genomes as of March 2021 (MF509813.1, NC_005881.2, KY242785.1, and KY271630.1) and the genome of the acute viral necrotic virus (AVNV) GQ153938.1 were merged and aligned with MAFFT (Katoh and Standley, 2013; **Figure 1G**). For the multiple sequence alignment, we did not perform any trimming or masking steps as recommended in Tan et al. (2015).

We used a general time reversible evolutionary model with a gamma distribution of the rate of variation among sites and a proportion of invariable sites (GTR + I + G) considered most appropriate according to jModelTest v2.1.10 (Posada, 2008) and based on the Akaike Information Criterion corrected for small sample size (AICc).

We assessed whether there is sufficient temporal signal in our data to proceed with phylogenetic molecular clock analysis and checked there were no sequences whose genetic divergence and sampling date were incongruent with TempEst v1.5.3 (Rambaut et al., 2016) from a maximum likelihood tree build using PhyML v3.1 (Guindon et al., 2010). Time-scaled Bayesian phylogenetic trees were estimated in BEAST v1.10.4 (Suchard et al., 2018) using a GTR + I + G substitution model under a relaxed molecular clock with a lognormal distribution of rates (**Figure 1I**). A Gaussian Markov Random Field Bayesian skyride coalescent model was used as the tree prior (Minin et al., 2008). A Monte Carlo Markov chain (MCMC) was run for the number of generations needed to achieve stationary distributions (200,000,000 generations) and sampling frequency adjusted accordingly to yield 10,000 samples from the posterior. We used Tracer v1.6 (Rambaut et al., 2016) to visualize the posterior distribution of each parameter and to obtain an estimate of the effective sample size (ESS) after removal of the initial 10% burn-in. We assumed the run had reached sufficient mixing as the ESS of all the parameters was above 200. A maximum clade credibility (MCC) tree was produced after removal of the initial 10% burn-in based on common ancestor heights method (Heled and Bouckaert, 2013) using the auxiliary program TreeAnnotator included in the BEAST package.

Statistical Analysis

To test for the correlation between polymorphism and time, host species and stage, we calculated a mean frequency of the targeted polymorphism type (snp, deletion, and standardized on 1,000 bp) across all genes for each individual (to avoid pseudo-replication). We then used a generalized linear model with fixed factors with the following syntax (we add into brackets information about variable nature and modalities numbers).

Polymorphism (SNP or deletion) ~ sampling date [continuous variable] + host species [factor with two modalities *C. gigas*, *O. edulis*] + host stage [factor with three modalities:

juvenile, pool_larvae, and adult]. All statistics were conducted with R 3.4 (R Development Core Team, 2008) using the linear model, built in "lm" function.

RESULTS

Data Quality Observations

As expected, the sequencing depth and the viral genome coverage was correlated to the total number of reads for the different libraries (data not shown). However, the sequencing of fresh moribund individuals, or even better, a pool of infected larvae (as the cost of losing information about the within-host dynamic of viral haplotypes) improves coverage and depth of sequencing. A pragmatic rule of thumb in order to have meaningful and sufficient reads from the virus genome, when sequencing an infected individual host, is to choose a fresh moribund oyster (i.e., an adult, a juvenile or a pool of larvae) harboring more than 5.0×10^6 OsHV-1 copies/mg of fresh flesh and sequence it with a yield of at least 10×10^6 reads per sample (**Supplementary Figure 1**). This procedure enabled a mean coverage depth of 100X of the whole viral genome for the following samples: LI, NZW17, NZW18, SW6, VIV46-2-m, VIV47-3-m, VIV48-4-m, VIV49-5-m88, VIV56-10-m, and VIV58-12-m. However, it should be kept in mind that most of the reads from a library are from the host's genome (approximately 95–99%, although this could be as low as 70% in case of pooled infected larvae).

Viral Haplotype Diversity

The local reconstruction of haplotype diversity gives an estimate of number of viral haplotypes segregating within one isolate (**Figure 2**, and **Supplementary Figure 2**) between 1 and 4 haplotypes (on the CDS basis). As expected, libraries sequenced with short reads and single end, namely PR, MV, and NZ, did not bridge different mutations (SNP or indels) and jeopardize the local assembler heuristics to disentangle sequencing error from true mutations. The coverage depth has a deep impact of the number of haplotypes that were reconstructed. For example, all the VIV samples showed a higher number of reconstructed haplotypes mainly due to a good sequencing depth. Indeed, it is hard to compare the number of viral haplotypes between samples from pool of larvae or juvenile and from one individual host. In summary, the number of viral haplotypes that segregate into infected individual is higher than unit except for samples Poole-Harbor United Kingdom and VIV57-11-m, but both samples have a rather low sequencing depth, and the Poole harbor isolate had previously been passaged through an oyster in a controlled environment. The variation of the number of haplotypes across the viral genome is hard to interpret but some regions seem to harbor more haplotypes nearly inverted and repeated regions.

Deletions, Substitution Rate and Most Recent Common Ancestors

The linear regression analysis showed that the accumulation of substitutions is correlated with time and not driven by a hidden species effect (DF = 6/9, F = 98.68, p value = 1.07×10^{-7}, time effect estimate = 0.12, time p value = 0.023). The greatest modalities of the stage did not have any effect except for the juveniles (or pool of juveniles), however, the unbalanced nature of these factors precludes any serious inferences. In comparison, accumulation of short deletions did not increase with time (model1, DF = 6/9, F = 113.8, p value = 5.72×10^{-8}, time effect estimate = 0.02, time p value = 0.68). It should be noted, the accumulation of deletions is far above the accumulation of substitutions and seems to be the main process of sequence evolution over time (**Figure 3**). Our reference mapping approach gives a biased picture of the insertion events into OsHV-1 genome as only sequences which are present in the reference genome can be detected and hence any new sequence inserts larger than read length will be lost during alignment process. Despite this limitation, 89 short insertions could be detected across 22 samples independently of the oyster species and the sampling date. Another important feature of OsHV-1 genome evolution is the presence of large deletions (**Figure 3**, **Supplementary Table 1**, and **Supplementary Figure 4**). The closest sample (PR) in time (and region) compare to OsHV-1 reference genome (Davison et al., 2005) shows no deletion whereas the sample taken from *O. edulis* in 2012 showed four large deletions (>>100 bp). A careful visual examination of reads alignment (**Figure 3** and **Supplementary Figure 1**) reveals numerous abrupt drop sequencing depths associated with clipped reads in the CIGAR motif. Both local sequencing depth variation and read clipping were used to infer large deletions and other structural variations (**Supplementary Table 1**). These large deletions do not seem to be random across the genome, which means that between isolates, deletions did not occur exactly at the same position but are located in the same genomic regions (**Supplementary Figure 3**, vertical reddish bands).

Regression of root-to-tip genetic distance against sampling time (**Supplementary Figure 5**) exposed sufficient temporal signal to proceed with phylogenetic molecular clock analyses. BEAST (Drummond et al., 2012) estimate of the coefficient of variation (CoV) was of 0.701 (95% HPD interval [0.4078, 1.0581]), indicating that strict clock-like evolution could be rejected and validating our choice to use a relaxed molecular clock. OsHV-1 viruses exhibited a mean evolutionary rate of 6.787E-05 nucleotide substitutions per site per year (95% HPD interval [3.5172E-05, 1.0335E-04]). The results indicate that the present genetic diversity seen in OsHV-1 isolates has arisen within the past three decades (**Figure 4**). Both herpesviruses found in *C. farreri* (GQ153938.1) and *A. broughtonii* (MF509813.1) group together (Ren et al., 2013; Xia et al., 2015). Their most recent common ancestor (MRCA) is very recent (approximatively 2001) and both species live in sympatry. The virus detected in *O. edulis* (SW6) is close to the reference genome (NC_005881.2) and from the OsHV-1 isolate collected from France in 2008 (PR) (Davison et al., 2005). PR and SW6 have their MRCA dated from 1993. Within the *C. gigas* host cluster, viral isolates cluster by provenance and sampling year. For example, all isolates from New-Zealand collected in 2010–2011 group together and appear distinct from isolates collected in Europe in 2010, 2015, and 2017 (MV, LI, Poole-Harbour United Kingdom, and all VIV samples). OsHV-1 μVars cluster

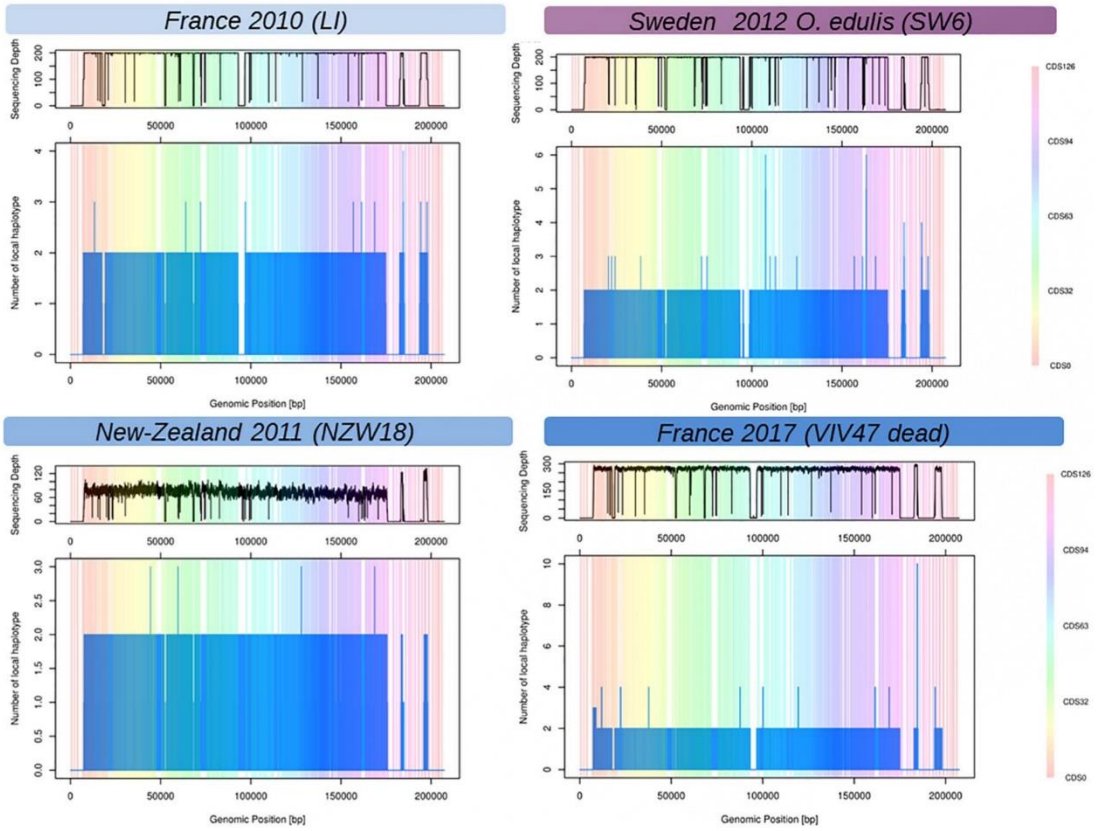


FIGURE 2 | Reconstructed local viral haplotypes for four samples using Shorah with amplicon mode (Zagordi et al., 2011). The X-axis represents the genomic position within the reference viral genome (Davison et al., 2005). The Y-axis is the number of reconstructed viral haplotypes on the focal genomic position. The vertical-colored bands represent the span of the coding DNA sequence (CDS) of the reference viral genome (Davison et al., 2005), one color represents one CDS numbered from 1 to 127. The top panels above the plots represent the coverage depth across the viral genome. Low or no sequencing depth on a genomic region could be interpreted as genomic deletion on the considered samples.

is represented by the LI and MV isolates from France in 2010 and by both μVar variant A and B (KY242785.1 and KY271630.1) that were isolated from France and Ireland between 2010 and 2014 (Burioli et al., 2017). As expected, all samples from the same epidemic event in 2017 clustered together.

DISCUSSION

Summary of the Main Findings

This study brings new insights about the diversity of OsHV-1 at the genomic scale. We show that the OsHV-1 genome has accumulated a large number of mutations since the publication of the reference genome in 2005 obtained from an infected oyster sampled in 1999. Deletions dominate the mutational landscape at the genome level, but our mapping reference approach is much powerful to detect deletions rather than insertions. Our phylogenetic analysis reveals that OsHV-1 diversity is structured by the host species (pectinids, clams, and oyster species), the geographical origin (Europe vs. Oceania) and most importantly the sampling date.

OsHV-1, a Measurably Evolving Pathogen

Examining the correlation between root-to-tip genetic distance and time suggests OsHV-1 is a measurably evolving pathogen when working at with whole-genome sequencing. The inferred mean evolutionary rate of 6.787E-05 nucleotide substitutions per site per year of OsHV-1 was higher than would be expected under the theory of co-evolution with host species. Potential biases inherent to the analysis conducted have been pointed out (Ho et al., 2005; Emerson, 2007; Debruyne and Poinar, 2009). In particular, it has been demonstrated that molecular rates are accelerated at short timescales on population level studies (Ho, 2005). The most likely explanation for this phenomenon is the persistence of slightly deleterious mutations, which, through

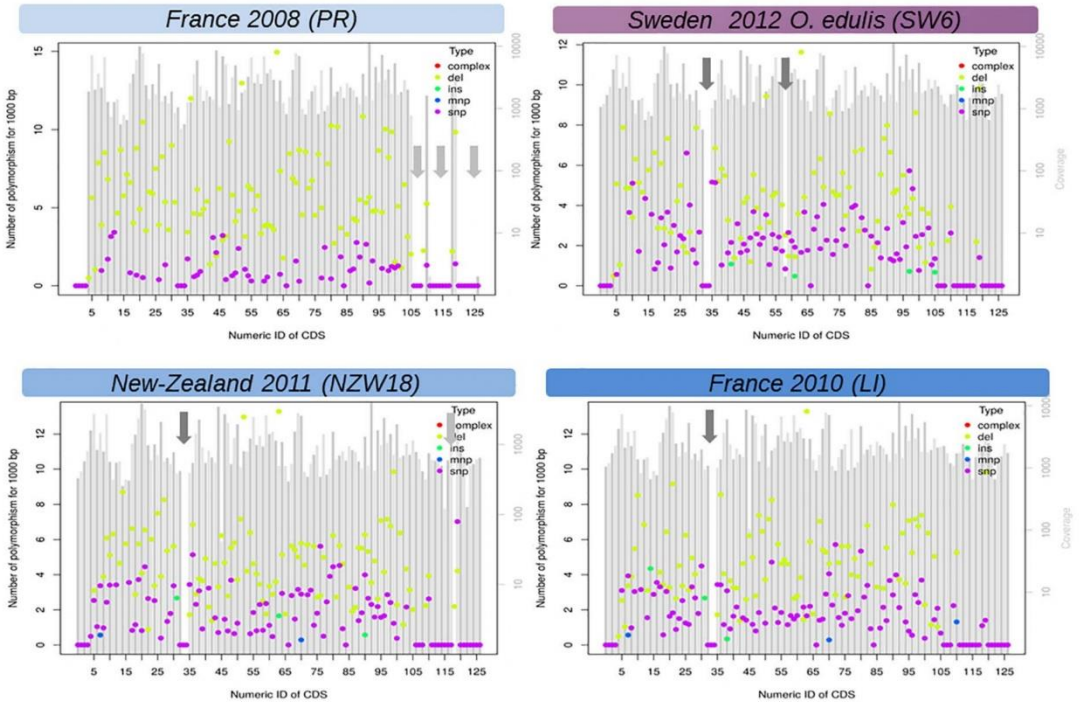


FIGURE 3 | Functional view of DNA variation (mutations) density along the genome of Ostreid herpesvirus 1 (OsHV-1). Here only LI, PR, NZ17, and SW6 are shown. The X-axis is the number of CDS found in the reference annotation genome last accessed (Davison et al., 2005) and the Y-axis is a standardized count of mutations found in 1,000 bp if the SNP density where the same in the targeted CDS. Genomic regions with a low complexity content (repeated sequences) were not considered in this analysis (light gray arrows). SNP, single nucleotide polymorphism; INS, short insertion (<50 bp); DEL, short insertion (<50 bp); MNP, short complex polymorphism, dark gray arrows: large deletion (>50 bp). Figures of all samples (detailed **Table 1**) are given in **Supplementary Files**.

purifying selection, would have been eliminated over long timescales (Ho et al., 2005, 2007a,b; Ho and Larson, 2006). While conducting selective pressure analyses on a bigger sampling would allow to test that our results were not a consequence of the time dependency of the molecular clock due to short time range of our dataset (i.e., >20 years), our findings are consistent with studies supporting that the evolutionary rates of some dsDNA viruses are comparable to those of RNA and single strand DNA viruses (Hughes et al., 2010; Patrício et al., 2012). Although we acknowledge current limitations of our study, mainly due to the small sample size available, our results demonstrate the feasibility of applying population genetic theory tools such as phylodynamics to a larger and well-designed sampling including virus from healthy animals (Burioli et al., 2016) of whole genome sequences of OsHV-1 and other herpesviruses.

OsHV-1 Within Host Diversity: Methodological Challenges and Biological Explanations

Viral diversity within infected individuals host is still hard to characterize because of two main biological viral features (Baaijens et al., 2017): (i) the abundance of variants is unknown and (ii) the great diversity and related high mutation rate could challenge classical tools used in bioinformatics (variant calling operations). This problem is even more acute when the abundance of variants is close to the sequencing error rate. Traditionally, in malacoherpesvirus literature, most of the studies have used *de novo* assembly methods to unravel viral diversity (Davison et al., 2005; Ren et al., 2013; Burioli et al., 2017; Bai et al., 2019a). However, at January 2018, most of the available short read assembly methods don't manage the mixture of haplotypes well, as they are not able to use the molecular phasing information from the reads (Baaijens et al., 2017). Indeed, we choose a pure mapping reference approach to describe viral diversity coupled with local *de novo* assembly. This process allows us to uncover the viral diversity within host. Our estimates about the true number of viral haplotypes within infected individuals were very conservative (lower bound) since we are not able to phase variants above 150 bp. To explain the source of the diversity within an infected individual, several non-exclusives hypotheses could be formulated: First, diversity within host is due to a multiple infection by several viral lineages. This means that several OsHV-1 haplotypes could be found in sea water during mortality outbreak. However, viral particles in sea water are hard to detect even when using passive sensor systems [lab detection

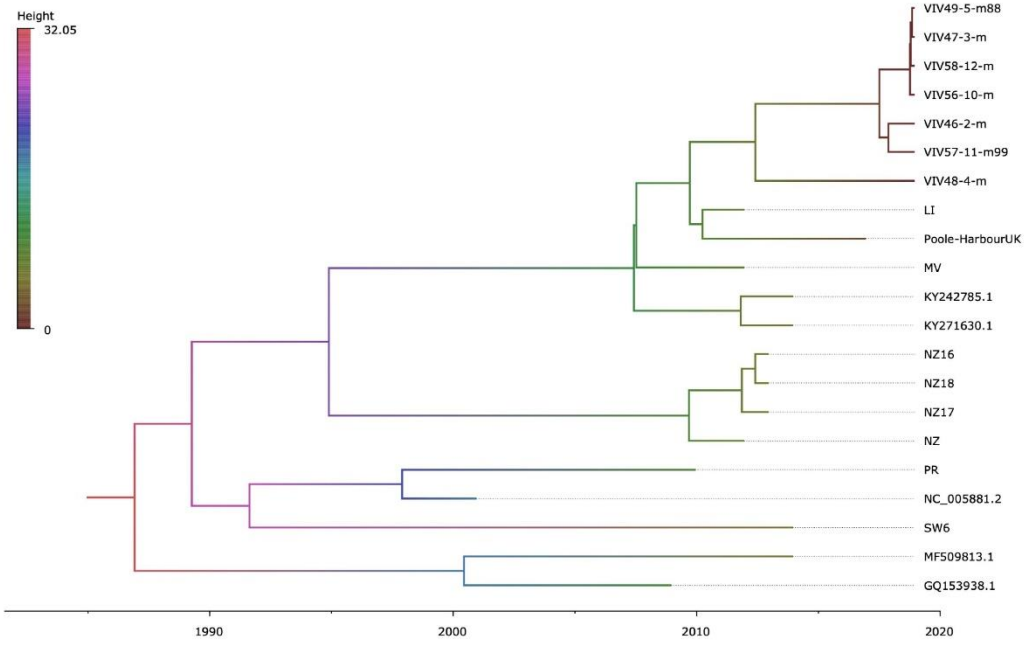


FIGURE 4 | Time-scaled phylogenetic tree of OsHV-1 and chlamys acute necrotic samples. Maximum clade credibility time-calibrated phylogenetic tree generated in BEAST. The tree is scaled in years, with the final sampling year being 2017. The branches are colored accordingly to their mean height across the posterior. This figure was produced with FigTree v1.4.3 (Rambaut, 2019).

of virus in seawater (Vincent-Hubert et al., 2017) and Hubert in prep (*in situ* detection of virus in sea water)]. Some alternatives approaches using ultra-filtration of large sea water volume are a promising way to check this hypothesis and we would help in finding if various haplotypes are present at the onset of the disease. Again, using ultra-filtration methods of sea water, we could expect an increase of the number of variants during the spread of the disease through a population. Such hypothesis could also be tested throughout the genomic relatedness among viral variants within individuals and variants from the previous years. The last hypothesis concerns *de novo* mutations during the infection process, viruses use host cellular machinery to replicate, there are several mechanisms which could give arise to *de novo* variants mediated by host antiviral defense systems (Sanjuán and Domingo-Calap, 2016).

Recently, extensive anti-viral mechanisms were characterized; Rosani et al. (2019) demonstrated that Adenosine deaminase enzymes of the ADAR family are found in the Pacific oyster and have the ability to post-transcriptionally modify OsHV-1 RNA in the form of A-to-I conversions. Indeed, investigators showed that viral genome have evolved to reduce the number of deamination targets along most of OsHV-1 CDS. The life cycle of herpes virus is also prone to mutation via recombination mechanisms which have been demonstrated in other herpesvirus or close-related as cytomegalovirus in human (Wilkinson and Weller, 2003; Norberg et al., 2007; Sijmons et al., 2015; Balloux et al., 2016; Kolb et al., 2017; Tomer et al., 2019). Although two DNA Recombination-Initiating Promoter Motifs have been reported as enriched among Malacoherpesviridae (Rosani and Venier, 2017), the evidence of such mechanisms is yet to be characterized in Malacoherpesviridae. The evolutionary advantage of having multiple viral haplotypes within one infected individual is of prime interest, since lethal or sub-lethal viral haplotypes could be maintained in the population thank to genetic complementation and be an additional reservoir of standing genetic variation (Segredo-Otero and Sanjuán, 2019).

Large Genomic Deletions Are a Good Markers of Virus Origin

In this study, we were able to compare the genomic structure of various viral isolates from several host species. All isolates were compared to the oldest reference genome published (Davison et al., 2005) which is likely to be the ancestor of most of our samples. Reads from all the samples were mapped to the Davison reference genome, however, visual inspection of sequencing depth and structural variant analysis reveals tremendous genetic variation between virus isolates. Notably virus isolates from *O. edulis* have two specific deletions in ORF101 (48 bp deleted) and ORF105 (49 bp deleted) that are not shared with the other isolates. We hypothesize that this particular haplotype who is attached to *O. edulis* (named OsHV-1 SW6) is a virus strain "adapted" to infect this species. If our expectation is true, this haplotype will be found in other infected *O. edulis*

population. In addition, isolates from New-Zealand (NZ16-NZ17-NZ18) share only two of the three observed larges deletions identified in France isolates (LI) probably because the New-Zealand samples are from a different virus reservoir. The exact molecular mechanism(s) of such deletion(s) are unknown, but they are not deleterious for the viral replication as they are maintained in the actual viral population. In addition, deletions in the OsHV-1 are not completely random across the samples they tend to occur in the regions 17,700–19,000 bp (ORF11), 52,000–52,800 bp (ORF35, ORF36, and ORF37), 67,972–68,573 bp (mainly ORF48), 93,000–98,500 bp (ORF62, ORF63, and ORF64). Burioli et al. (2017) observed the same deletions as we observe in the LI and VIV samples. The scarcities of annotation in these regions do not help us to formulate a functional hypothesis.

Refining Malacoherpiviradae Classification to Get a Better Understanding of Viral Micro Evolution

Ostreid herpesvirus 1 is usually described as a generalist herpesvirus that is able to infect various marine bivalves species across the world (Arzul et al., 2001, 2017; Davison et al., 2005; Gao et al., 2018; Kim et al., 2019). However, our phylogenetic analysis shows that such view is probably too simple and more elaborate hypothesis are needed. Recently, OsHV-1 found on two other bivalve species (*C. farreri* and *A. broughtonii* mainly in China) (Xia et al., 2015; Bai et al., 2017, 2018, 2019b) seems to have a recent common ancestor, the MRCA is dated from 1991 to 2006. In addition, there is a growing number of report of OsHV-1 DNA in novel host species *Scapharca subcrenata* and *Agropecten irridians* (Gao et al., 2018; Kim et al., 2019). In addition, we also observed a unique type of deletion in this isolate (**Figure 3**, SW6 sample) compare to other isolates from *C. gigas*. All these inferred host shifts could be explained by two hypotheses: (1) We are revealing an old viral diversity, i.e., various viral haplotypes are adapted to various host species, but this diversity was not accessible to previous studies because of technological limitations (MLST, gene concatenation, microsatellite, and Sanger sequencing). However, this hypothesis does explain why MRCA are so recent (less than 50 years). (2) We are observing an ongoing specialization (speciation) of a generalist virus [OsHV-1 census (Davison et al., 2005)] to a specialized viral populations OsHV-1-SB on *A. broughtonii*, AVNV on *C. farreri*, OsHV-1 Var on *P. maximus*, OsHV-1 μVars on *C. gigas*, and now OsHV-1 SW6 on *O. edulis*, a mechanism possibly supported by the different host antiviral responses. For the later haplotype, we hypothesize a host shift from *C. gigas* to *O. edulis* and the estimated date of this shift (1992) which is congruent with the history of the introduction of *C. gigas* in Scandinavia (Wrange et al., 2010; Mortensen et al., 2016; Anglès d'Auriac et al., 2017). These recent host shifts could be explained by two mechanisms well documented in other farming animal and plants (Jones et al., 2013; Bai et al., 2015; Mineur et al., 2015). The first mechanism is the introduction of *C. gigas* (and OsHV-1) world-wide for its extraordinary adaptive capacity and the second is an intensification of *C. gigas* farming (imported spat, spat from hatcheries, high field density, triploid's oyster, and extension of field surfaces). Indeed, OsHV-1 could reach quite high number of genome copies per infection events since the cultivated density of susceptible animals are much higher than natural beds, leading to an increase of contact with a larger number of bivalve's taxa than ever. Host shift is likely to be just a matter of time because of the abundant genomic viral diversity. We postulate that OsHV-1 taxon is under specialization in novel populations and/or bivalve species due to world-wide expansion of *C. gigas* shellfish farming in the 1970s.

In order to discriminate between the first and the second hypothesis regarding variants, we need more detailed study of isolates from various bivalve species, most notably including old and novel sympatric species of *C. gigas*. In fact, we need a better understanding of how OsHV-1 is constrained, or not, to a specific host species and how frequently it could jump from one species to another species (Brito et al., 2021). Only well-replicated phylogenetic studies encompassing the whole viral genome information will help-us to better understand OsHV-1 ecology and evolution.

Conclusion and Perspectives

We have characterized the viral genomic diversity within an infected individual; our estimates are probably lower bounds of the exact number of haplotypes. To refine and notably to understand complementation between haplotypes within individuals new techniques, such as long read sequencing, will be required (Bai et al., 2019a). The observed viral diversity is generally well structured according to host species, then geographical origins and finally sampling date. There is an urgent need to assess the effect of some mutations on viral phenotype namely pathogenicity (virulence and infectivity) for a given host species and populations, including basic information about which viral haplotype can infect which host species. These crucial questions about virus ecology should be explored to be able to understand the potential of OsHV-1 to jump from one species to the next.

This study thus lays the groundwork for future studies to further analyze the relationships between viral diversity and pathogenicity across bivalve species and strains.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the NCBI Sequence Read Archive (SRA) under repository, accession number of the BioProject ID: PRJNA721248.

AUTHOR CONTRIBUTIONS

This study is the result of a collective work. BM, J-BL, J-FP, and TR conceived this study, and participated in its design. SH and CP performed the sample preparation for Illumina sequencing. J-BL, AB, MJ, and BM performed genomic analyses. BM, MJ, and J-BL interpreted the results and drafted the manuscript. All authors read, corrected, and approved the final manuscript.

FUNDING

This work was financially supported by DGAL (French National Services of Food and Agriculture) through the National Reference Laboratory for Mollusc Diseases, Ifremer, La Tremblade and by the EU funded project VIVALDI (H2020 program, n°678589) within the WP 3.1.1 part about genetic of resistance to OsHV-1 infection.

ACKNOWLEDGMENTS

The authors thank the UMR 8199 LIGAN-PM Genomics platform (Lille, France) which belongs to the "*Fédération de Recherche*" 3508 Labex EGID (European Genomics Institute for Diabetes; ANR-10-LABX-46) and was supported by the ANR Equipex 2010 session (ANR-10-EQPX-07-01; "LIGAN-PM"). The LIGAN-PM Genomics platform (Lille, France) is also supported by the FEDER and the Region Nord-Pas-de-Calais-Picardie. The authors acknowledge the Pôle de Calcul et de Données Marines (PCDM) for providing DATARMOR storage, computational resources, and support services (URL: http://www.ifremer.fr/pcdm) with a special attention to Bioinformatic team: Laure Quintric and Patrick Durand. The authors thank Isabelle Arzul for the fruitful discussions about OsHV-1 biology.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.711377/full#supplementary-material

Supplementary Figure 1 | Correlation between the assessed viral copy numbers per µL per ng of gills from infected individuals for all the individuals from VIV batches and the sequencing depth of each sample. The labels of the sample should be interpreted as the family name (i.e., viv57) followed by the individual number (viv57-11) and the final status at the end of the experimental infection "v" stand for alive (viv57-11-v3) and "m" for dead (viv57-11-m99).

Supplementary Figure 2 | Reconstructed local viral haplotypes for four samples using Haploclique (Töpfer et al., 2013). The X-axis represents a genomic position within OsHV-1 reference genome in base pair; the Y-axis is the number of haplotypes reconstituted at the focal genomic position. Large deletions are visible in the sequencing top panels and in the number of reconstructed haplotypes.

Supplementary Figure 3 | Local viral haplotype reconstruction using QuasiRecomb software on genomic regions without any abrupt drop of sequencing depth (large deletions). Reconstructions have been done on CDS basis (using the reference genome annotation from Davison et al. (2005). The X-axis represents the genomic position within the reference viral genome. The Y-axis is the number of quasispecies. The black number corresponding to ORF. In red, we represented the larges deletion found in the consensus viral genome of each isolates, orange illustrate genomic regions where the alignment was not good enough to allow haplotype reconstruction. Large deletions were detected after visual inspection of bam file and sniffles detection.

Supplementary Figure 4 | Illustration of the variety of polymorphism found in OsHV-1 genome mapped against the reference genome from Davison et al. (2005). Some visual patterns observed in Bam files motivate the search of viral haplotypes. Here both Bam and Vcf files are viewed using IGV (Thorvaldsdóttir et al., 2013). The blue arrowed highlights some detected polymorphisms and red arrows point out undetected potential variants. Horizontal arrows showed the large deletions, i.e., structural variations, and vertical arrows pointed out short polymorphism, i.e., short indel and SNPs. Each picture (one per sample) is composed of a bandplot showing the sequencing depth (top panel) of the considered genomic regions followed by the alignment of individual reads (bottom panel). The sequencing depth is between 0 and 4,000 X for the LI and SW6 samples, and 0–190 for NZ17 samples. **(A)** The observed region is from 94,000 to 100,000 bp on LI, NC17, and SW6 isolates. **(B)** The observed region is from 71,400 to 73,000 bp. **(C)** Zoom in the 77,900–79,000 bp in the NZ17 samples, it seems that there is a physical linkage between a short deletion and a SNP leading to conclusion that there are at least two haplotypes segregating in this region. **(D)** Zoom in the 21,700–22,800 bp region of the SW6 samples. The extremely large sequence depth of this sample allows to infer that few haplotypes segregating in this population show duplication.

Supplementary Figure 5 | Clock-likeness of OsHV-1 genomes. Regression of root-to-tip genetic distance against sampling time were computed using TempEst v1.5.3 (Tan et al., 2015) using two rooting methods (best fit in TempEst and midpoint). Sufficient and robust temporal signal was observed to proceed with phylogenetic molecular clock analyses.

REFERENCES

Abbadi, M., Zamperin, G., Gastaldelli, M., Pascoli, F., Rosani, U., Milani, A., et al. (2018). Identification of a newly described OsHV-1 µvar from the North Adriatic Sea (Italy). *J. Gen. Virol.* 99, 693–703. doi: 10.1099/jgv.0.001042

Anglès d'Auriac, M. B., Rinde, E., Norling, P., Lapègue, S., Staalstrøm, A., Hjermann, D. Ø, et al. (2017). Rapid expansion of the invasive oyster *Crassostrea gigas* at its northern distribution limit in Europe: naturally dispersed or introduced? *PLoS One* 12:e0177481. doi: 10.1371/journal.pone.0177481

Arzul, I., Corbeil, S., Morga, B., and Renault, T. (2017). Viruses infecting marine molluscs. *J. Invertebr. Pathol.* 147, 118–35. doi: 10.1016/j.jip.2017.01.009

Arzul, I., Nicolas, J.-L., Davison, A. J., and Renault, T. (2001). French scallops: a new host for ostreid herpesvirus-1. *Virology* 290, 342–359. doi: 10.1006/viro.2001.1186

Baaijens, J. A., Aabidine, A. Z. E., Rivals, E., and Schönhuth, A. (2017). De novo assembly of viral quasispecies using overlap graphs. *Genome Res.* 27, 835–848. doi: 10.1101/gr.215038.116

Bai, C., Gao, W., Wang, C., Yu, T., Zhang, T., Qiu, Z., et al. (2016). Identification and characterization of ostreid herpesvirus 1 associated with massive mortalities of *Scapharca broughtonii* broodstocks in China. *Dis. Aquat. Organ.* 118, 65–75. doi: 10.3354/dao02958

Bai, C., Wang, C., Xia, J., Sun, H., Zhang, S., and Huang, J. (2015). Emerging and endemic types of ostreid herpesvirus 1 were detected in bivalves in China. *J. Invertebr. Pathol.* 124, 98–106. doi: 10.1016/j.jip.2014.11.007

Bai, C.-M., Morga, B., Rosani, U., Shi, J., Li, C., Xin, L.-S., et al. (2019a). Long-range PCR and high-throughput sequencing of ostreid herpesvirus 1 indicate high genetic diversity and complex evolution process. *Virology* 526, 81–90. doi: 10.1016/j.virol.2018.09.026

Bai, C.-M., Rosani, U., Xin, L.-S., Li, G.-Y., Li, C., Wang, Q.-C., et al. (2018). Dual transcriptomic analysis of ostreid herpesvirus 1 infected *Scapharca broughtonii* with an emphasis on viral anti-apoptosis activities and host oxidative bursts. *Fish Shellfish Immunol.* 82, 554–564. doi: 10.1016/j.fsi.2018.08.054

Bai, C.-M., Wang, Q.-C., Morga, B., Shi, J., and Wang, C.-M. (2017). Experimental infection of adult *Scapharca broughtonii* with ostreid herpesvirus SB strain. *J. Invertebr. Pathol.* 143, 79–82. doi: 10.1016/j.jip.2016.12.001

Bai, C.-M., Xin, L.-S., Rosani, U., Wu, B., Wang, Q.-C., Duan, X.-K., et al. (2019b). Chromosomal-level assembly of the blood clam, *Scapharca* (Anadara) *broughtonii*, using long sequence reads and Hi-C. *Gigascience* 8:giz067. doi: 10.1093/gigascience/giz067

Balloux, F., Lassalle, F., Tutill, H. J., Breuer, J., Reeves, M. B., Christiansen, M. T., et al. (2016). Islands of linkage in an ocean of pervasive recombination reveals two-speed evolution of human cytomegalovirus genomes. *Virus Evol.* 2:vew017. doi: 10.1093/ve/vew017

Biek, R., Pybus, O. G., Lloyd-Smith, J. O., and Didelot, X. (2015). Measurably evolving pathogens in the genomic era. *Trends Ecol. Evol.* 30, 306–313. doi: 10.1016/j.tree.2015.03.009

Bookelaar, B., Lynch, S. A., and Culloty, S. C. (2020). Host plasticity supports spread of an aquaculture introduced virus to an ecosystem engineer. *Parasites Vectors* 13:498.

Brito, A. F., Baele, G., Nahata, K. D., Grubaugh, N. D., and Pinney, J. W. (2021). Intrahost speciations and host switches played an important role in the evolution of herpesviruses. *Virus Evol.* 7:veab025.

Burioli, E. A. V., Prearo, M., and Houssin, M. (2017). Complete genome sequence of ostreid herpesvirus type 1 µVar isolated during mortality events in the Pacific oyster *Crassostrea gigas* in France and Ireland. *Virology* 509, 239–251. doi: 10.1016/j.virol.2017.06.027

Burioli, E. A. V., Prearo, M., Riina, M. V., Bona, M. C., Fioravanti, M. L., Arcangeli, G., et al. (2016). Ostreid herpesvirus type 1 genomic diversity in wild populations of Pacific oyster *Crassostrea gigas* from Italian coasts. *J. Invertebr. Pathol.* 137, 71–83. doi: 10.1016/j.jip.2016.05.004

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. (2016). GenBank. *Nucleic Acids Res.* 44, D67–D72.

Cudini, J., Roy, S., Houldcroft, C. J., Bryant, J. M., Depledge, D. P., Tutill, H., et al. (2019). Human cytomegalovirus haplotype reconstruction reveals high diversity due to superinfection and evidence of within-host recombination. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5693–5698. doi: 10.1073/pnas.1818130116

Davison, A. J., Trus, B. L., Cheng, N., Steven, A. C., Watson, M. S., Cunningham, C., et al. (2005). A novel class of herpesvirus with bivalve hosts. *J. Gen. Virol.* 86, 41–53. doi: 10.1099/vir.0.80382-0

Debruyne, R., and Poinar, H. N. (2009). Time dependency of molecular rates in ancient DNA data sets, a sampling artifact? *Syst. Biol.* 58, 348–360. doi: 10.1093/sysbio/syp028

Delmotte, J., Chaparro, C., Galinier, R., de Lorgeril, J., Petton, B., Stenger, P.-L., et al. (2020). Contribution of viral genomic diversity to oyster susceptibility in the Pacific oyster mortality syndrome. *Front. Microbiol.* 11:1579. doi: 10.3389/fmicb.2020.01579

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

EFSA (2010). Scientific opinion on the increased mortality events in Pacific oysters, *Crassostrea gigas*. *EFSA J.* 8:1894. doi: 10.2903/j.efsa.2010.1894

Emerson, B. C. (2007). Alarm bells for the molecular clock? No support for Ho et al.'s model of time-dependent molecular rate estimates. *Syst. Biol.* 56, 337–345. doi: 10.1080/10635150701258795

Gao, F., Jiang, J.-Z., Wang, J.-Y., and Wei, H.-Y. (2018). Real-time quantitative isothermal detection of ostreid herpesvirus-1 DNA in *Scapharca subcrenata* using recombinase polymerase amplification. *J. Virol. Methods* 255, 71–75. doi: 10.1016/j.jviromet.2018.02.007

Garcia, C., Thébault, A., Dégremont, L., Arzul, I., Miossec, L., Robert, M., et al. (2011). Ostreid herpesvirus 1 detection and relationship with *Crassostrea gigas* spat mortality in France between 1998 and 2006. *Vet. Res.* 42:73. doi: 10.1186/1297-9716-42-73

Garrison, E. (2019). A Simple C++ Library for Parsing and Manipulating VCF Files, + many Command-Line Utilities. GitHub Repository. Available online at: https://github.com/vcflib/vcflib (accessed October, 2020).

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv* [Preprint] 12073907. arXiv:1207.3907

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Heled, J., and Bouckaert, R. R. (2013). Looking for trees in the forest: summary tree from posterior samples. *BMC Evol. Biol.* 13:221. doi: 10.1186/1471-2148-13-221

Ho, S., and Larson, G. (2006). Molecular clocks: when timesare a-changin'. *Trends Genet.* 22, 79–83. doi: 10.1016/j.tig.2005.11.006

Ho, S. Y. W. (2005). Accuracy of rate estimation using relaxed-clock models with a critical focus on the early metazoan radiation. *Mol. Biol. Evol.* 22, 1355–1363. doi: 10.1093/molbev/msi125

Ho, S. Y. W., Kolokotronis, S.-O., and Allaby, R. G. (2007a). Elevated substitution rates estimated from ancient DNA sequences. *Biol. Lett.* 3, 702–705. doi: 10.1098/rsbl.2007.0377

Ho, S. Y. W., Phillips, M. J., Cooper, A., and Drummond, A. J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* 22, 1561–1568. doi: 10.1093/molbev/msi145

Ho, S. Y. W., Shapiro, B., Phillips, M. J., Cooper, A., and Drummond, A. J. (2007b). Evidence for time dependency of molecular rate estimates. *Syst. Biol.* 56, 515–522. doi: 10.1080/10635150701435401

Hughes, A. L., Irausquin, S., and Friedman, R. (2010). The evolutionary biology of poxviruses. *Infect. Genet. Evol.* 10, 50–59. doi: 10.1016/j.meegid.2009.10.001

Jaramillo, N., Domingo, E., Muñoz-Egea, M. C., Tabarés, E., and Gadea, I. (2013). Evidence of Muller's ratchet in herpes simplex virus type 1. *J. Gen. Virol.* 94, 366–375. doi: 10.1099/vir.0.044685-0

Jenkins, C., Hick, P., Gabor, M., Spiers, Z., Fell, S., Gu, X., et al. (2013). Identification and characterisation of an ostreid herpesvirus-1 microvariant (OsHV-1 µ-var) in *Crassostrea gigas* (Pacific oysters) in Australia. *Dis. Aquat. Organ.* 105, 109–126. doi: 10.3354/dao02623

Jones, B. A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M. Y., et al. (2013). Zoonosis emergence linked to agricultural intensification and environmental change. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8399–8404. doi: 10.1073/pnas.1208059110

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kim, H. J., Jun, J. W., Giri, S. S., Yun, S., Kim, S. G., Kim, S. W., et al. (2019). Mass mortality in Korean bay scallop (*Argopecten irradians*) associated with ostreid herpesvirus-1 µVar. *Transbound. Emerg. Dis.* 66, 1442–1448.

Kitchen, A., Shapiro, B., Firth, C., Holmes, E. C., Suchard, M. A., and Rambaut, A. (2010). Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol. Biol. Evol.* 27, 2038–2051. doi: 10.1093/molbev/msq088

Kolb, A. W., Lewin, A. C., Moeller Trane, R., McLellan, G. J., and Brandt, C. R. (2017). Phylogenetic and recombination analysis of the herpesvirus genus *Varicellovirus*. *BMC Genomics* 18:887. doi: 10.1186/s12864-017-4283-4

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* [Preprint] 13033997[q–bio]. arXiv:1303.3997 [q-bio.GN]

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078-2079. doi: 10.1093/bioinformatics/btp352

Mineur, F., Provan, J., and Arnott, G. (2015). Phylogeographical analyses of shellfish viruses: inferring a geographical origin for ostreid herpesviruses OsHV-1 (Malacoherpesviridae). *Mar. Biol.* 162, 181–192. doi: 10.1007/s00227-014-2566-8

Minin, V. N., Bloomquist, E. W., and Suchard, M. A. (2008). Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* 25, 1459–1471. doi: 10.1093/molbev/msn090

Mortensen, S., Strand, Å, Bodvin, T., Alfjorden, A., Skår, C., Jelmert, A., et al. (2016). Summer mortalities and detection of ostreid herpesvirus microvariant in Pacific oyster *Crassostrea gigas* in Sweden and Norway. *Dis. Aquat. Organ.* 117, 171–176. doi: 10.3354/dao02944

Nicolas, J. L., Comps, M., and Cochennec, N. (1992). Herpes-like virus infecting Pacific oyster larvae, *Crassostrea gigas*. *Bull. Eur. Assoc. Fish Pathol.* 12, 11–13.

Norberg, P., Kasubi, M. J., Haarr, L., Bergström, T., and Liljeqvist, J. -Å (2007). Divergence and recombination of clinical herpes simplex virus type 2 isolates. *J. Virol.* 81:13158. doi: 10.1128/JVI.01310-07

Olson, J. K., Ercolini, A. M., and Miller, S. D. (2005). A virus-induced molecular mimicry model of multiple sclerosis. *Curr. Top. Microbiol. Immunol.* 296, 39–53. doi: 10.1007/3-540-30791-5_3

Parsons, L. R., Tafuri, Y. R., Shreve, J. T., Bowen, C. D., Shipley, M. M., Enquist, L. W., et al. (2015). Rapid genome assembly and comparison decode intrastrain

variation in Human alphaherpesviruses. *mBio* 6:e02213-14. doi: 10.1128/mBio.02213-14

Patrício, A. R., Herbst, L. H., Duarte, A., Vélez-Zuazo, X., Santos Loureiro, N., Pereira, N., et al. (2012). Global phylogeography and evolution of chelonid fibropapilloma-associated herpesvirus. *J. Gen. Virol.* 93, 1035-45. doi: 10.1099/vir.0.038950-0

Paul-Pont, I., Dhand, N. K., and Whittington, R. J. (2013). Spatial distribution of mortality in Pacific oysters *Crassostrea gigas*: reflection on mechanisms of OsHV-1 transmission. *Dis. Aquat. Organ.* 105, 127–138. doi: 10.3354/dao02615

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256. doi: 10.1093/molbev/msn083

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rambaut, A. (2019). *FigTree*. Available online at: https://github.com/rambaut/figtree (accessed April, 2021).

Rambaut, A., Lam, T. T., Max Carvalho, L., and Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* 2:vew007. doi: 10.1093/ve/vew007

Ren, W., Chen, H., Renault, T., Cai, Y., Bai, C., Wang, C., et al. (2013). Complete genome sequence of acute viral necrosis virus associated with massive mortality outbreaks in the Chinese scallop, *Chlamys farreri*. *Virol. J.* 10:110. doi: 10.1186/1743-422X-10-110

Renault, T., Moreau, P., Faury, N., Pepin, J.-F., Segarra, A., and Webb, S. (2012). Analysis of clinical ostreid herpesvirus 1 (Malacoherpesviridae) specimens by sequencing amplified fragments from three virus genome areas. *J. Virol.* 86, 5942–5947. doi: 10.1128/JVI.06534-11

Renault, T., Tchaleu, G., Faury, N., Moreau, P., Segarra, A., Barbosa-Solomieu, V., et al. (2014). Genotyping of a microsatellite locus to differentiate clinical ostreid herpesvirus 1 specimens. *Vet. Res.* 45:3. doi: 10.1186/1297-9716-45-3

Renner, D. W., and Szpara, M. L. (2017). Impacts of genome-wide analyses on our understanding of human herpesvirus diversity and evolution. *J. Virol.* 92:e00908-17. doi: 10.1128/JVI.00908-17

Renzette, N., Bhattacharjee, B., Jensen, J. D., Gibson, L., and Kowalik, T. F. (2011). Extensive genome-wide variability of human cytomegalovirus in congenitally infected infants. *PLoS Pathog.* 7:e1001344. doi: 10.1371/journal.ppat.1001344

Retel, C., Märkle, H., Becks, L., and Feulner, P. (2019). Ecological and evolutionary processes shaping viral genetic diversity. *Viruses* 11:220. doi: 10.3390/v11030220

Rosani, U., Bai, C.-M., Maso, L., Shapiro, M., Abbadi, M., Domeneghetti, S., et al. (2019). A-to-I editing of Malacoherpesviridae RNAs supports the antiviral role of ADAR1 in mollusks. *BMC Evol. Biol.* 19:149. doi: 10.1186/s12862-019-1472-6

Rosani, U., and Venier, P. (2017). Oyster RNA-seq data support the development of Malacoherpesviridae genomics. *Front. Microbiol.* 8:1515. doi: 10.3389/fmicb.2017.01515

Sanjuán, R., and Domingo-Calap, P. (2016). Mechanisms of viral mutation. *Cell. Mol. Life Sci.* 73, 4433–4448. doi: 10.1007/s00018-016-2299-6

Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., and Belshaw, R. (2010). Viral mutation rates. *J. Virol.* 84, 9733–9748. doi: 10.1128/JVI.00694-10

Sedlazeck, F. J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., et al. (2018). Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* 15, 461–468. doi: 10.1038/s41592-018-0001-7

Segarra, A., Pépin, J. F., Arzul, I., Morga, B., Faury, N., and Renault, T. (2010). Detection and description of a particular ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. *Virus Res.* 153, 92–99. doi: 10.1016/j.virusres.2010.07.011

Segredo-Otero, E., and Sanjuán, R. (2019). The effect of genetic complementation on the fitness and diversity of viruses spreading as collective infectious units. *Virus Res.* 267, 41–48. doi: 10.1016/j.virusres.2019.05.005

Sijmons, S., Thys, K., Mbong Ngwese, M., Van Damme, E., Dvorak, J., Van Loock, M., et al. (2015). High-throughput analysis of human cytomegalovirus genome diversity highlights the widespread occurrence of gene-disrupting mutations and pervasive recombination. *J. Virol.* 89, 7673–7695. doi: 10.1128/JVI.00578-15

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 4:vey016. doi: 10.1093/ve/vey016/5035211

Tan, G., Muffato, M., Ledergerber, C., Herrero, J., Goldman, N., Gil, M., et al. (2015). Current methods for automated filtering of multiple sequence alignments frequently Worsen Single-gene phylogenetic inference. *Syst. Biol.* 64, 778–791. doi: 10.1093/sysbio/syv033

Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017

Tomer, E., Cohen, E. M., Drayman, N., Afriat, A., Weitzman, M. D., Zaritsky, A., et al. (2019). Coalescing replication compartments provide the opportunity for recombination between coinfecting herpesviruses. *FASEB J.* 33, 9388–9403. doi: 10.1096/fj.201900032R

Töpfer, A., Marschall, T., Bull, R. A., Luciani, F., Schönhuth, A., and Beerenwinkel, N. (2014). Viral quasispecies assembly via maximal clique enumeration. *PLoS Comput. Biol.* 10:e1003515. doi: 10.1371/journal.pcbi.1003515

Töpfer, A., Zagordi, O., Prabhakaran, S., Roth, V., Halperin, E., and Beerenwinkel, N. (2013). Probabilistic inference of viral quasispecies subject to recombination. *J. Comput. Biol.* 20, 113–123. doi: 10.1089/cmb.2012.0232

Vincent-Hubert, F., Morga, B., Renault, T., and Le Guyader, F. S. (2017). Adsorption of norovirus and ostreid herpesvirus type 1 to polymer membranes for the development of passive samplers. *J. Appl. Microbiol.* 122, 1039–1047. doi: 10.1111/jam.13394

Wang, E., Paessler, S., Aguilar, P. V., Smith, D. R., Coffey, L. L., Kang, W., et al. (2005). A novel, rapid assay for detection and differentiation of serotype-specific antibodies to Venezuelan equine encephalitis complex alphaviruses. *Am. J. Trop. Med. Hyg.* 72, 805–810. doi: 10.4269/ajtmh.2005.72.805

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Cham: Springer-Verlag New York. doi: 10.1007/978-3-319-24277-4

Wilkinson, D., and Weller, S. (2003). The role of DNA recombination in herpes simplex virus DNA replication. *IUBMB Life* 55, 451–458. doi: 10.1080/15216540310001612237

Wrange, A.-L., Valero, J., Harkestad, L. S., Strand, Ø, Lindegarth, S., Christensen, H. T., et al. (2010). Massive settlements of the Pacific oyster, *Crassostrea gigas*, in Scandinavia. *Biol. Invasions* 12, 1145–1152. doi: 10.1007/s10530-009-9535-z

Xia, J., Bai, C., Wang, C., Song, X., and Huang, J. (2015). Complete genome sequence of ostreid herpesvirus-1 associated with mortalities of *Scapharca broughtonii* broodstocks. *Virol. J.* 12:110. doi: 10.1186/s12985-015-0334-0

Zagordi, O., Bhattacharya, A., Eriksson, N., and Beerenwinkel, N. (2011). ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. *BMC Bioinformatics* 12:119. doi: 10.1186/1471-2105-12-119

Zagordi, O., Geyrhofer, L., Roth, V., and Beerenwinkel, N. (2010). Deep sequencing of a genetically heterogeneous sample: local haplotype reconstruction and read error correction. *J. Comput. Biol.* 17, 417–428. doi: 10.1089/cmb.2009.0164

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

Copyright © 2021 Morga, Jacquot, Pelletier, Chevignon, Dégremont, Biétry, Pepin, Heurtebise, Escoubas, Bean, Rosani, Bai, Renault and Lamy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Diversité et évolution de l'Ostreid Herpesvirus type 1

L'utilisation des nouvelles générations de séquençage (NGS) pour étudier et surveiller les agents pathogènes chez les mollusques bivalves est cruciale pour anticiper la détection de nouveaux agents pathogènes ou l'émergence de nouveaux génotypes. Améliorer nos connaissances et nos techniques de caractérisation génomique permet de mieux appréhender le cycle de vie de ces agents pathogènes, leurs interactions avec les hôtes et de caractériser leur capacité d'adaptation et d'évolution. L'Ostreid Herpevirus 1 (OsHV-1), responsable chaque année de fortes mortalités chez les larves et les juvéniles d'huîtres creuses *Magallana gigas*. Ce virus a également été détecté chez d'autres espèces de bivalves associés à des mortalités. L'étude du génome d'OsHV-1 est rendue complexe par *i)* l'impossibilité de le cultiver *in vitro* car aucune lignée cellulaire pour le propager n'est disponible *ii)* la présence de grandes régions répétées inversées. En 2019, seuls six génomes du virus OsHV-1 étaient disponibles dans les bases de données publiques, limitant ainsi considérablement notre compréhension de sa diversité spatio-temporelle et parmi différentes espèces hôtes. Dans ce contexte, mes travaux de thèse ont été consacrés à l'acquisition et à l'assemblage de plus de 400 génomes du virus OsHV-1, grâce au développement d'un pipeline bioinformatique spécifiquement conçu pour son analyse génomique. L'analyse de ces génomes a permis de comprendre la structuration spatio-temporelle des populations du virus OsHV-1, notamment en lien avec les mouvements d'huîtres entre différentes zones ostréicoles en France. En complément, les résultats obtenus ont permis de caractériser l'isolement génétique des génotypes viraux de l'huître creuse *M. gigas* vis-à-vis de ceux retrouvés chez l'huître plate *Ostrea edulis* et d'inférer des fréquences de transmissions du virus inter-espèces quasi-nulles. Enfin, dans une approche d'évolution expérimentale, l'impact du fond génétique de l'hôte sur la diversité virale a été étudié, révélant l'influence conjointe de la dérive génétique et de la sélection essentiel dans l'évolution du virus OsHV-1. L'ensemble de ces travaux a contribué de manière significative à l'acquisition d'un grand nombre de données génomiques, favorisant ainsi une compréhension plus approfondie de la diversité, de l'évolution et de l'adaptation du virus à son hôte.

Mots clés : Herpesvirus, mollusques bivalves, épidémiologie, diversité, évolution, NGS.

Diversity and evolution of the Ostreid Herpesvirus type 1

Using next generation sequencing technologies (NGS) to study and monitor pathogens in bivalve mollusks is crucial for anticipating the detection of new pathogens or the emergence of new genotypes. Improving our knowledge and genomic characterization techniques allows for a better understanding of the life cycle of these pathogens, their interactions with hosts, and characterizing their capacity for adaptation and evolution. The Ostreid Herpesvirus 1 (OsHV-1) is responsible for significant mortalities in Pacific oyster (*Magallana gigas)* larvae and juveniles every year. This virus has been detected in other bivalve species associated with mortalities. Studying the OsHV-1 genome is challenging due to *i)* the inability to cultivate it in vitro because no cell line is available for propagation, and *ii)* the presence of large inverted repeat regions. In 2019, only six OsHV-1 virus genomes were available in public databases, significantly limiting our understanding of its spatiotemporal diversity across different host species. In this context, my doctoral research was dedicated to acquiring and assembling over 400 genomes of OsHV-1 virus, thanks to the development of a bioinformatics pipeline specifically designed for its genomic analysis. The analysis of these genomes helped elucidate the spatiotemporal structuring of OsHV-1 virus populations, particularly in relation to oyster movements between different oyster farming areas in France. Furthermore, the results obtained characterized the genetic isolation of viral genotypes in Pacific oysters *M. gigas* compared to those found in flat oysters *Ostrea edulis* and inferred nearly negligible inter-species virus transmission frequencies. Finally, in an experimental evolution approach, the impact of the host's genetic background on viral diversity was studied, revealing the joint influence of genetic drift and essential selection in the evolution of OsHV-1 virus. This work has significantly contributed to the acquisition of a substantial amount of genomic data, thereby promoting a deeper understanding of the diversity, evolution, and adaptation of the virus to its host.

Key words: Herpesvirus, bivalve mollusks, epidemiology, diversity, evolution, NGS.