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

En Biologie des Interactions

École doctorale GAIA

Unité de recherche UMR 5244, Interactions Hôtes Pathogènes Environnements (IHPE)

La diversité génétique virale : quel est son rôle dans l'interaction du virus Ostreid Herpesvirus 1 avec son hôte l'huître *Crassostrea gigas* ?

Présentée par Jean DELMOTTE

Le jeudi 01 avril 2021

Sous la direction de Jean-Michel ESCOUBAS

Devant le jury composé de

Maryline HOUSSIN, Chargé de recherche LABEO, Caen	Rapporteur
Elisabeth HERNIOU, Directrice de Recherche CNRS, Tours	Rapporteur
Stéphane BLANC, Directeur de recherche INRAE, Montpellier	Examinateur
Germain CHEVIGNON, Chargé de recherche Ifremer, La Tremblade	Examinateur
Mathieu SICARD, Professeur à l'Université de Montpellier, Montpellier	Président du jury
Jean-Michel ESCOUBAS, Chargé de recherche CNRS, Montpellier	Directeur

UNIVERSITÉ DE MONTPELLIER

‘I think we live in an unscientific age in which almost all the buffeting of communications and television - words, books, and so on - are unscientific. As a result, there is a considerable amount of intellectual tyranny in the name of science.’

Richard Phillips Feynman *(1918-1988) What is Science?*, presented at the 15th annual meeting of the National Science Teachers Association (1966)

La diversité génétique virale : quel est son rôle dans l'interaction du virus Ostreid herpesvirus 1 avec son hôte l'huître *Crassostrea gigas* ?



Avant-propos

Ce manuscrit de thèse présente les résultats scientifiques que j'ai obtenu pendant 3 ans, d'octobre 2017 à décembre 2020, au sein de l'unité mixte de recherche (UMR5244) *Interactions Hôtes Pathogènes Environnements* (IHPE, directeur : Guillaume MITTA) - affiliée au Centre national de la recherche scientifique (CNRS), à l'Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) et à l'Université de Montpellier (UM) -, dans l'équipe *Mécanisme d'Interaction et d'Adaptation en Milieu Marin* (MIMM, responsable: Delphine DESTOUMIEUX-GARZON), sous la direction de Jean-Michel ESCOUBAS (CNRS) et le co-encadrement de Caroline Montagni (IFREMER). Ce travail a été financé suite à l'attribution d'une bourse de l'école doctorale GAIA en 2017, ainsi qu'une partie de deux financements de projet attribués au laboratoire. Le projet HemoVir soutenu par l'IFREMER et le projet HaploFit soutenu par le labEx CeMEB. Enfin, ces recherches ont bénéficié de l'appui logistique de l'UM et de l'IFREMER à travers l'accès aux outils de calcul datarmor.

Les résultats présentés dans ce manuscrit sont séparés en 5 chapitres dont deux articles scientifique. Le chapitre I correspond à l'introduction générale du sujet en développant le cadre théorique dans lequel s'inscrit ce travail en détaillant les modèles utilisés et enfin en introduisant les questions scientifiques liées à ces recherches, à savoir la caractérisation de la diversité génétique et génomique du virus Ostreid Herpèsvirus 1 (OsHV-1) et l'évaluation de son rôle potentiel dans le Syndrôme de mortalité des huîtres du Pacifique (Pacific Oyster Mortality Syndrome, POMS). Le chapitre II correspond à un premier article paru en 2020 (Delmotte et al., 2020). Le chapitre III correspond à un second article en fin d'écriture qui devrait être soumis en février 2021. Des sections d'introductions, de discussions et de conclusions entre les articles sont rédigées en français et apportent les connaissances utiles à la bonne compréhension du chapitre ainsi que le fil conducteur de la thèse. Le chapitre IV apporte une discussion sur l'évolution des connaissances globales grâce à mes recherches dans le rôle de la diversité génomique virale dans les interactions entre l'huître et le virus OsHV-1. Enfin le chapitre V présente ou illustre les trois grandes perspectives ou approches méthodologiques pour mettre en évidence le rôle joué par la diversité génétique du virus OsHV-1 dans les POMS.

Enfin, l'essentiel de l'innovation présente dans mes travaux porte davantage sur l'amélioration des outils bioinformatiques et des méthodes d'analyses que j'ai pu mettre en place plutôt que sur les expériences que j'ai réalisées. Ainsi, j'ai fait le choix de ne pas mettre en avant les outils d'analyses dans ce manuscrit mais il m'est apparu raisonnable de les adosser dans l'annexe en fin de thèse ou de renvoyer le lecteur curieux directement vers le dépôt de code afin de préserver l'unité de la démarche plutôt que compiler l'ensemble du matériel additionnel. Pour aider à la lecture de ce manuscrit les termes importants pour la compréhension sont en gras et leurs traductions anglaise est entre parenthèses en italique. Enfin, les encarts en gris aideront à préciser certains détails tandis qu'un résumé à la fin de chaque partie permettra de récapituler les informations importantes.

Table de matière

Introduction générale 1

L'Interaction hôte-pathogène 3

La coévolution hôte-pathogène 4

Les dynamiques des interactions hôtes et pathogènes 5

Les modes, processus, et mécanismes de sélection 6

La signature génomique des interactions hôte pathogène 8

Les différentes modifications du génome 9

La caractérisation de la diversité génétique 11

Les mécanismes de diversification génétique 12

La diversité génétique chez les organismes eucaryotes pluricellulaire à reproduction sexuée 13

Chez les vertébrés 13

Chez les invertébrés 14

L'impact de la diversité génétique chez les organismes eucaryotes pluricellulaire à reproduction sexuée 15

Les bases de la diversité génétique chez les virus 15

Les virus à ARN 15

Les virus à ADN 17

Le virus Ostreid Herpèsvirus 1 19

La découverte de l'OsHV-1 19

La structure de l'OsHV-1 20

Le génome de l'OsHV-1 21

L'Hôte et le tropisme de l'OsHV-1 22

La diversité génétique de l'OsHV-1 22

La biologie et écologie de l'huître 25

Le mode de vie de l'huître 25

La reproduction et le cycle de vie de l'huître 25

L'anatomie de l'huître 27

L'ostréiculture 28

L'élevage des huîtres 28

L'impact économique 29

Le syndrôme de mortalité des huîtres du Pacifique (Pacific Oyster Mortality Syndrome-POMS) 31

La cause 31

L'histoire des episodes de mortalité des huîtres 33

La manifestation de la maladie 34

Les traitements 35

L'objectifs de la thèse *37*

Chapitre II Contribution de la diversité génomique virale à la sensibilité des huîtres dans le syndrôme de mortalité des huîtres du Pacifique *37*

Le contexte des recherches *37*

L'objectif de l'étude *38*

L'illustration du matériel et méthode *38*

Article 1 : Contribution of Viral Genomic Diversity to Oyster Susceptibility in the Pacific Oyster Mortality Syndrome *39*

Discussion *57*

Une stratégie d'analyse non optimale pour accéder à la diversité génétique des variants viraux *57*

Une analyse impactée par la sensibilité de l'hôte *57*

Les variations génétiques des transcrits viraux ne semblent pas façonner par le phénotype des huîtres *59*

L'origine des géniteurs des huîtres pourrait-il expliquer la différence de phénotype observé ? *60*

Conclusion *62*

Chapitre III Aperçu de la diversité génomique de l'OsHV-1 en France lors d'événements de mortalité dans différentes zones de production *62*

Le contexte des recherches *62*

L'objectif de l'étude *63*

L'illustration du matériel et méthode *63*

Article 2 : Aperçu de la diversité génomique de l'OsHV-1 en France lors d'événements de mortalité dans différentes zones de production *64*

Discussion *106*

Une approche non biaisée dans l'analyse de la diversité virale *106*

Les virus appartiennent au même génotype *106*

Une propagation du virus OsHV via des transferts d'huîtres *107*

Le maintien des populations virales dans le bassin de Marennes-Oléron *108*

Des populations virales locales supplantaient par l'apport des variants viraux issus de Marennes-Oléron *108*

Conclusion *109*

Discussion générale *109*

La comparaison des approches entre les deux études *110*

La caractérisation de la diversité génétique du virus OsHV-1 ..... *111*

L'influence du système immunitaire de l'huître sur la diversité virale *114*

L'espèce hôte, un facteur primordial pour expliquer la diversité virale de l'OsHV-1 ..... *115*

Les facteurs de la triade épidémiologique expliquent davantage la variabilité phénotypique des huîtres lors des POMS *116*

Perspectives *117*

L'urgence d'harmoniser l'annotation des gènes d'OsHV-1 ..... *117*
L'approche pangénomique *121*
L'assemblage de longs fragments du génome d'OsHV-1 ..... *123*
Conclusion *123*
Référence *125*
Annexe *143*
Remerciements *158*
Index *160*
Résumé *162*

Table des illustrations

Figure 1. ***Représentation schématique de la coévolution.*** 4

Figure 2. ***Schéma simplifié des modèles d'interactions hôte-pathogène.*** 6

Figure 3. ***Représentation schématique de la coévolution sélective par sélection directionnelle avec un processus de balayage sélectif dans les pools de gènes d'une population hôte et parasitaire.*** 8

Figure 4. ***Différents types de mutations.*** 10

Figure 5. ***Illustration de la diversité génétique.*** 11

Figure 6. ***Diversité des populations virales peut être un déterminant de la virulence.*** 17

Figure 7. ***Répartition mondiale des lieux où l'OsHV-1 a été détecté.*** 20

Figure 8. ***Illustration de la structure de l'OsHV-1.*** 21

Figure 9. ***Schéma illustrant la structure du genome actuelle de l'OsHV-1.*** 21

Figure 10. ***Génomes complets de l'OsHV disponibles dans les bases de données.*** 233

Figure 11. ***Schéma représentant le cycle de vie de l'huître creuse.*** 26

Figure 12. ***Représentation schématique de l'anatomie de l'huître creuse.*** 27

Figure 13. ***Cycle de production de Crassostrea gigas.*** 29

Figure 14. ***Production aquacole ostréicole en termes de poids de production.*** 30

Figure 15. ***Triade épidémiologique adaptée pour l'huître.*** 32

Figure 16. ***Succession des événements entraînant la mortalité des huîtres lors des POMS.*** 33

Figure 17. ***Impact de mortalité cumulé sur l'élevage des huîtres.*** 34

Figure 18. ***Illustration du plan expérimental utilisé dans l'article I.*** 39

Figure 19. ***Courbe de raréfaction virale.*** 58

Figure 20. ***Analyse d'ensemble complète pour mettre en évidence l'effet famille.*** 60

Figure 21. ***Fonds génétiques des géniteurs utilisés pour la production des familles d'huîtres.*** 61

Figure 22. ***Illustration du plan expérimental utilisé dans l'article II.*** 64

Figure 23. ***Erreur d'annotation pour les génomes déjà disponible dans les bases de données.*** 119

Figure 24. ***Comparaison d'une même séquence avec l'annotation prokka avec celle existante.*** 120

Figure 25. ***Visualisation de l'erreur d'annotation de l'ORF PT4 et PT5.*** 121

Figure 26. ***Pangénome des virus connus d'OsHV-1.*** 122

Liste des abréviations et sigles

***ADN** = **A**cide **D**ésoxyribo**N**ucléique*
***AID** = **A**ctivation-**I**nduced cytidine **D**eaminase (cytidine désaminase induite par activation)*
***ARN** = **A**cide **R**ibo**N**ucléique*
***ARNm** = **A**cide **R**ibo**N**ucléique **m**essager*
***Atl** = **Atl**antique*
***BCR** = **B C**ells **R**eceptor (recepteur des cellules B)*
***Br** = **Br**est*
***CMH** = **C**omplexe **M**ajeur d'**H**istocompatibilité*
***DAMP** = **D**amage-**A**ssociated **M**olecular **P**atterns (motifs moléculaires associés aux dégâts)*
***FAO** = **F**ood and **A**griculture **O**rganization of the United Nations (Organisation des Nations unies pour l'alimentation et l'agriculture)*
***FREP** = **F**ibrinogen-**RE**lated **P**rotein (protéine liée au fibrinogène)*
***GWAS** = **G**enome-**W**ide **A**ssociation **S**tudy (études d'association à l'échelle du génome)*
***HaploFit** = **Haplo**type **Fit**ness (projet de recherche)*
***HemoVir** = **Hemo**cyte **Vir**us (projet de recherché)*
HSV**-1 = **H**erpes **S**implex **V**irus **1
***HTS** = **H**igh **T**hroughput **S**equencing (séquençage à haut débit)*
***IEM** = **I**nfections **E**xpérimentales en **M**ésocosme*
***IFREMER** = **I**nstitut **F**rançais de **R**echerche pour l'**E**xploitation de la **M**er*
***IgSF** = **I**mmuno**G**lobulin **S**uper**F**amily (superfamille d'immunoglobulines)*
***IGV** = **I**ntegrative **G**enomics **V**iewer (logiciel qui permet de visualiser des genomes)*
***IHPE** = **I**nteractions **H**ôtes **P**athogènes **E**nvironnements*
***INDEL** = **IN**sertion **DEL**étion*
***jpf** = **j**our **p**ost **f**econdation*
***Lt** = **L**a **t**remblade*
***Med** = **Méd**iterranée*
***µVar** = **microVar**iant*
***MO** = **M**arennes-**O**léron*
***NSI** = **N**aissains **S**tandardisés **I**FREMER*
***ORF** = **O**pen **R**eading **F**rames (cadre de lecture ouvert)*
OsHV-1** = **Os**treid **H**erpes**V**irus type **1
***pb** = **p**aire de **b**ase*
***PRR** = **P**attern **R**ecognition **R**eceptor (récepteurs de reconnaissance de motifs moléculaires)*
***POMS** = **P**acific **O**yster **M**ortality **S**yndrome (syndrome de mortalité des huîtres du Pacifique)*
***PCR** = **P**olymerase **C**hain **R**eaction (réaction de polymérisation en chaîne)*
***qPCR** = **q**uantitative **P**olymerase **C**hain **R**eaction*
***RNAseq** = **R**ibo**N**ucleic **A**cid **seq**uencing (séquençage à ARN)*
***SNP** = **S**ingle-**N**ucleotide **P**olymorphism (polymorphisme d'un seul nucléotide)*
***SV** = **S**tructural **V**ariations (Variation structurelle génomique)*
***TCR** = **T C**ells **R**eceptor (recepteur des cellules T)*
***Th** = Etang de **Th**au*
VHC** = **V**irus de l'**H**épatite **C
***VIH** = **V**irus de l'**I**mmunodéficience Humaine*
***ZC** = **Z**one **C**onchylicole*
***ZNC** = **Z**one **N**on **C**onchylicole*

Introduction générale

La théorie de l'évolution est à la biologie ce que la relativité générale est à la physique. C'est une théorie fondamentale difficile à appréhender mais c'est un outil puissant pour poser des questions sur la Nature. Les théories évolutionnistes ont mis plusieurs siècles à maturer et ce sont les travaux de Charles Darwin et Alfred Russel Wallace sur la sélection naturelle au XIXe siècle qui posèrent les fondements des cadres conceptuels contemporains (Darwin, 1859).

Cette théorie a constamment évolué avec les nouvelles découvertes en génétique et d'autres disciplines jusqu'à aujourd'hui. En 1866, le travail du moine autrichien Gregor Mendel a proposé trois principes d'hérédité qui décrivent la transmission des caractères génétiques (Mendel, 1865). En 1942, le biologiste britannique Julian Huxley a décrit ce cadre de travail émergent dans un livre intitulé *Evolution : The Modern Synthesis*. Cette synthèse arrive à combiner génétique des populations et génétique mendélienne, sélection naturelle et hérédité dans une seule et même théorie de l'évolution (Huxley, 1942). Elle est considérée comme l'une des réalisations majeures de la biologie et a conduit à l'adoption généralisée de plusieurs hypothèses de base : (i) Les mutations géniques qui se produisent à basse fréquence indépendamment de la force et de la direction de la sélection naturelle créent des variations phénotypiques sur le plan de l'évolution; (ii) Les mutations qui sont favorables conduisent généralement à des changements phénotypiques faible se traduisant par de lents changements phénotypiques progressifs; (iii) l'hérédité est génétique; (iv) la sélection naturelle est la seule explication de l'adaptation; (v) la macroévolution est le résultat des différences accumulées dans le processus de microévolution (Laland et al., 2015).

Les données expérimentales ont remis en question plusieurs hypothèses de cette synthèse, lui permettant d'évoluer et de couvrir un plus large éventail de phénomènes. Par exemple, suite au progrès de la biologie moléculaire et l'observation que la majorité des variations génomiques entre deux individus ne sont ni bénéfiques, ni nuisibles à la survie et à la reproduction de l'organisme qui l'exprime, mais aussi, que les variations observées dans les régions fonctionnelles sont bien souvent défavorables à l'organisme. Le cadre conceptuel global a évolué en incluant la théorie neutraliste de l'évolution (*Neutral theory of molecular evolution*) (Kimura, 1983). Celle-ci soutient que la plupart des changements évolutifs au niveau moléculaire, et la plupart des variations au sein et entre les espèces, sont dus à une **dérive génétique** aléatoire d'allèles mutants qui sont sélectivement neutres (Kimura, 1983; King and Jukes, 1969).

Aujourd'hui, le cadre conceptuel dominant utilisé est la **théorie synthétique de l'évolution** ou théorie moderne de l'évolution (*The synthetic theory of evolution*) qui incorpore les découvertes théoriques et empiriques les plus récentes (par exemple, la théorie neutraliste de l'évolution, théorie de la valeur sélective inclusive). Cette théorie considère l'évolution et l'adaptation comme un processus lent, causé par des mutations génétiques aléatoires. La sélection n'est plus la seule pression évolutive et surtout elle ne se réduit pas à un effet adaptatif simpliste, où la survie du plus apte conduit nécessairement à l'adaptation. Ainsi la sélection

naturelle (ou artificielle) peut prendre de nombreuses formes et varie en fonction de l'environnement conjointement avec des phénomènes de migration et de dérive génétique.

Du fait de tous ces mécanismes, il n'est pas simple de déterminer la proportion de chaque force évolutive en jeu au sein d'une population et/ou entre des populations d'un organisme. Or, c'est uniquement au travers de la compréhension de ces forces que nous pouvons donner du sens à la diversité génétique d'une population en faisant la différence entre l'action sous-jacente du hasard ou d'un processus évolutif complexe ayant un rôle important. Dans le domaine des maladies infectieuses, l'émergence continue de nouveaux agents pathogènes et l'augmentation de la propagation de la résistance nous rappellent que les micro-organismes sont fondamentalement soumis à ces forces évolutives (Alizon and Méthot, 2018; Bordier, 1888). Ainsi, la compréhension des **interactions hôte-pathogène** (*host-pathogen interactions*), c'est-à-dire les méchanismes d'un organisme pathogène pour se maintenir dans un organism hôte lequel utilise des méchanismes pour s'en défaire, sont au cœur des luttes modernes contre les organismes pathogènes.

C'est au travers de ces théories que la biologie évolutive offre des réponses pertinentes pour mieux comprendre le rôle de la diversité génétique dans les interactions hôte-pathogène au niveau moléculaire, cellulaire, de l'individu ou de la population. Celles-ci sont donc incontournables pour mieux comprendre le rôle de la diversité génétique entre un organisme pathogène et son hôte. Durant les travaux qui vont être présenté par la suite, l'organisme pathogène sera le virus Ostreid Herpesvirus 1 (OsHV-1) et l'organisme hôte sera l'huître *Crassostrea gigas*. C'est pourquoi l'introduction générale va permettre de délimiter (i) le cadre conceptuel des interactions hôte-pathogène à travers des concepts fondamentaux de biologie évolutive. (ii) Décrire quelles sont les stratégies de diversifications chez les organismes eucaryotes pluricellulaires et les virus à l'aide d'exemples d'immunologie et de microbiologie. (iii) Nous décrirons le principal protagoniste de ce sujet de thèse à savoir le virus Ostreid Herpèsvirus 1 (OsHV-1). (vi) Puis nous décrirons son hôte biologique le mieux étudié : l'huître *Crassostrea gigas*. (vi) Enfin, nous présenterons le syndrôme de mortalité des huîtres du Pacifique dans lequel l'OsHV-1 et l'huître interagissent. Ensembles ces descriptions permettront de mieux comprendre comment les questions scientifiques de ce projet ont émergé et quels sont les objectifs scientifiques et les enjeux des recherches menées durant ma thèse.

L'Interaction hôte-pathogène

Les micro-organismes responsables de maladies sont appelés des agents **pathogènes** (*pathogens*), la pathogenèse étant le processus par lequel ils déclenchent ces maladies. Ebert et Fields définissent les **parasites** (*parasites*) et leurs hôtes comme des organismes (souvent des pathogènes) qui tirent avantages d'un autre organisme (leurs hôtes) suscitant ainsi un processus de sélection pour lequel un hôte se défend contre son parasite. Ainsi lorsqu'un agent pathogène rencontre un organisme hôte, il s'engage alors une lutte constante. Dans ce combat, la sélection naturelle favorise, en général, les pathogènes susceptibles de développer des mécanismes de virulence efficaces tandis qu'à l'inverse, elle favorise l'hôte qui est capable de mettre en place des stratégies de résistance adéquates (**Fig.1. p4**). La sélection va agir au niveau du support de l'information ADN ou ARN des individus au travers de la diversité génétique et génomique. La **diversité génétique**[1] (*genetic diversity*) est le substrat de cette sélection, elle peut composer la variation d'une séquence d'ADN entre des individus distincts d'une espèce (ou population) donnée (Ellegren and Galtier, 2016). Les agents pathogènes et leurs hôtes ont des effets évolutifs les uns sur les autres et ces effets sont réciproques, cela suggère qu'ils coévoluent. Ainsi la **coévolution** (*coevolution*) se définit comme les changements génétiques adaptatifs et réciproques entre les espèces en interaction. La coévolution peut être étudiée en termes de traits phénotypiques appariés, tels que la résistance et l'infectiosité, en termes d'interaction entre les molécules hôtes et pathogènes ou en termes de gènes ou de séquences nucléotidiques directement (Janzen, 1980; Woolhouse et al., 2002). Ce processus est peut-être l'un des plus importants moteurs de diversité génétique et biologique au cours des 3,5 derniers milliards d'années (Majerus et al., 1996).

Dans cette première partie nous allons décrire (i) le phénomène de coévolution entre un hôte et l'un de ses pathogènes, (ii) passer en revu les processus évolutifs et écologiques importants susceptibles d'affecter la dynamique et (iii) les processus de sélection qui vont agir sur la diversité génétique. (iv) Nous allons voir que la coévolution laisse des signatures génomiques chez les deux protagonistes. (v) Nous décrirons ces modifications génomiques qui peuvent (vi) caractériser la diversité génétique.

[1] Cette définition est communément acceptée, toutefois, elle va être redéfinie ultérieurement

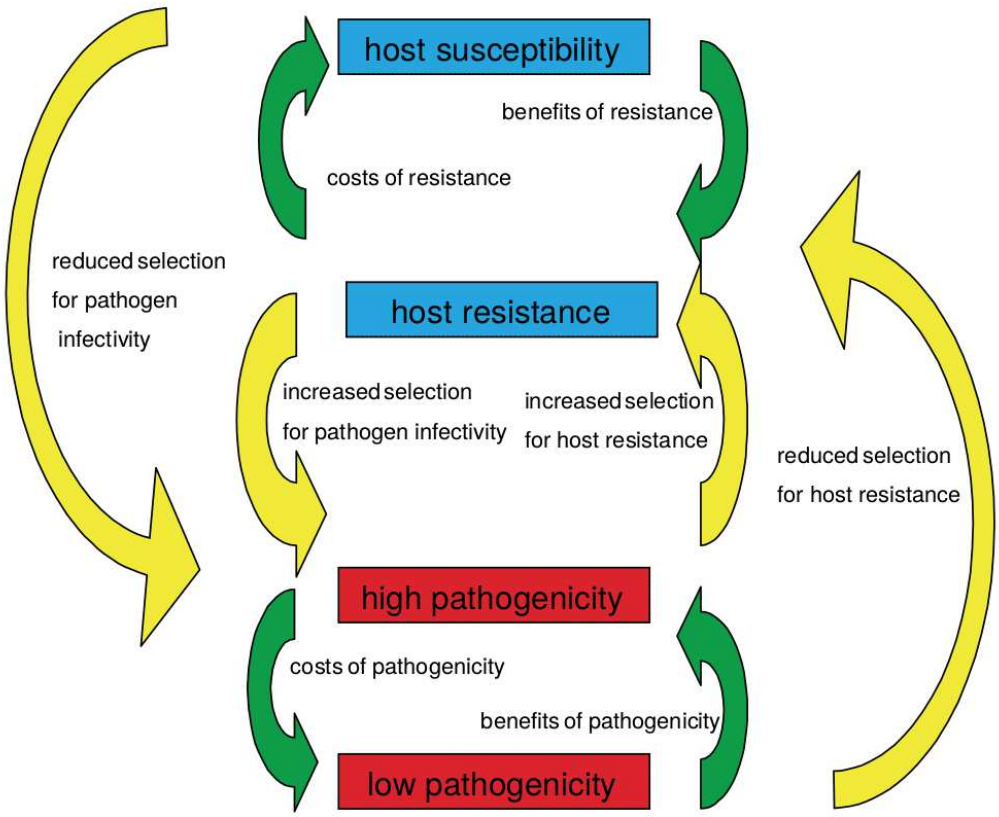


Figure 1. **Représentation schématique de la coévolution**.
Le diagramme schématique de la coévolution met l'accent sur la réciprocité, car un changement de fréquence allélique causé par la sélection dans une espèce imposera la sélection d'un changement de fréquence allélique chez une autre espèce. Les flèches vertes indiquent la sélection intraspécifique; les flèches jaunes indiquent la sélection interspécifique. Illustration tirée de la figure 1 de (Woolhouse et al., 2002).

La coévolution hôte-pathogène

Un phénomène de coévolution peut se produire entre toutes les populations qui sont en interactions. Dans le cadres des processus hôte-parasite l'effet sur l'adaptabilité de l'hôte et du pathogène est en général oppose : ce qui est bon pour le pathogène est nocif pour l'hôte, et vice versa, on parle alors de **coévolution antagoniste** (Ebert and Fields, 2020; Janzen, 1980; Woolhouse et al., 2002). Le processus de coévolution antagoniste peut être décrit par deux types de modèles : des modèles de **coévolution spécifique** ou l'hôte interagit avec un parasite (parfois appelé coévolution par paires) ; et des modèles de coévolution non spécifique, mieux connus sous le nom de **coévolution diffuse** ou plusieurs hôtes et / ou espèces de parasites sont impliqués dans ce processus. Nous allons maintenant nous intéresser uniquement à la coévolution antagoniste spécifique car celle-ci est moins complexe et constitue une étape incontournable vers la seconde. Cette confrontation détermine la coévolution, dans laquelle les fréquences alléliques d'une espèce déterminent l'adéquation des génotypes de l'autre espèce (**Fig.1.p4**), entraînant des changements dans la diversité génétique pour la résistance et la tolérance chez les hôtes, et pour l'infectivité et la virulence chez les parasites.

Les dynamiques des interactions hôtes et pathogènes

On s'attend à ce que la coévolution entre les hôtes et leurs parasites suit une gamme de dynamiques possibles, les deux cas extrêmes sont appelés guerre des tranchées (ou reine rouge) et course aux armements (Tellier et al., 2014) (**Fig.2. p6**). On les retrouve sous de nombreux termes dans la littérature, parfois confondu avec le mode de sélection de ces dynamiques.

> « - Ma foi, dans mon pays à moi, répondit Alice, encore un peu essoufflée, on arriverait généralement à un autre endroit si on courait très vite pendant longtemps, comme nous venons de le faire.
> - On va bien lentement dans ton pays ! Ici, vois-tu, on est obligé de courir tant qu'on peut pour rester au même endroit. Si on veut aller ailleurs, il faut courir au moins deux fois plus vite que ça ! »
>
> *De l'autre côté du miroir, Lewis Carroll, 1871*

La guerre des tranchées (*trench warfare*) (Stahl et al., 1999) est également appelée «reine rouge» dans la littérature sur les hôtes et les parasites. Cette dynamique se produit lorsque l'adéquation de plusieurs génotypes fonctionnels chez les deux espèces dépend négativement de leur fréquence dans la population (Retel et al., 2019). L'illustration classique de la Reine rouge est utile pour comprendre cette dynamique, en effet cette hypothèse de la biologie évolutive proposée par (Van Valen, 1973), un paléontologue évolutionniste américain, qui est une métaphore issue d'un épisode du roman de Lewis Carroll : *« De l'autre côté du miroir »* dont le passage est cité juste avant. Le paysage, c'est l'environnement qui subit des variations incessantes, la Reine rouge c'est l'espèce, la course c'est l'évolution (Lefevre et al., 2016). L'environnement est constitué par les autres espèces en interactions avec la Reine Rouge. Pour le parasite, l'hôte constitue le paysage adaptatif du parasite et inversement (Lefevre et al., 2016). Van Valen suppose que l'adaptabilité des organismes à l'environnement est principalement attribuée à d'autres espèces, qui par leurs changements évolutifs exercent une nouvelle pression sélective sur les espèces avec lesquelles ils interagissent. De nombreux exemples de cette dynamique ont été observés comme pour les interactions chez *Arabidopsis thaliana* avec un gène conférant la capacité de reconnaître les pathogènes *Pseudomonas* (Stahl et al., 1999), le lin et son pathogène fongique (Thrall et al., 2012) ou encore *Daphnia magna* et son endoparasite bactérien *Pasteuria ramosa* (Decaestecker et al., 2007).

Dans la dynamique de **course à l'armement** (*arms race*) la coévolution est motivée par les augmentations réciproques consécutives de la fréquence des nouveaux génotypes qui fournissent un avantage évolutif (Ebert, 2008; Retel et al., 2019; Woolhouse et al., 2002) Ce concept s'illustre facilement avec le développement des armes atomiques entre l'Union soviétique et les États-Unis pendant la guerre froide où chaque protagoniste va augmenter son stock d'armes en réponse au développement d'armes de son antagoniste. Cette dynamique a été observée dans les interactions gène pour gène (*gene for gene*) des gènes R végétaux avec les agents pathogènes (Bergelson et al., 2001) où l'on observe des «améliorations» accumulées dans les deux populations (Woolhouse et al., 2002). D'autres exemples connus de courses aux armements ont pu être observés dans la coévolution des phages de Flavobacterium dans les

fermes piscicoles (Laanto et al., 2017; Retel et al., 2019) et la résistance de la drosophile au virus sigma (Retel et al., 2019; Wilfert and Jiggins, 2012).

Pour comprendre quel est le type de dynamique évolutive en jeu, et donc mettre en place des moyens de luttes efficaces contre les agents infectieux sur le long terme, il est nécessaire de comprendre quel est le mode et le processus de sélection à l'œuvre. Ces deux dynamiques évolutives font intervenir des mécanismes de sélection différents, les recherches sur les interactions hôte-pathogène se concentrent principalement sur deux de sélection connue, la sélection d'équilibrage et la sélection directionnelle.

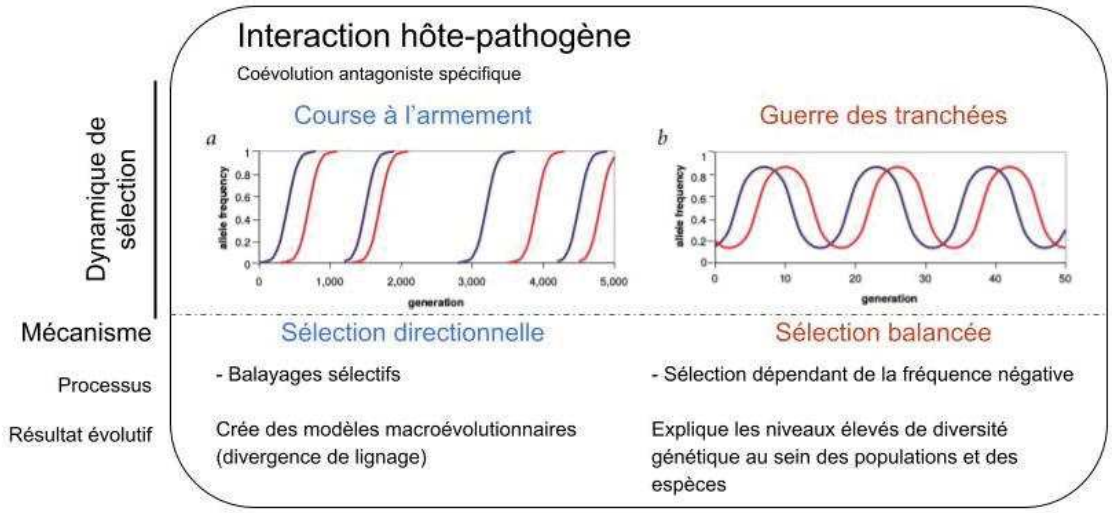


Figure 2. **Schéma simplifié des modèles d'interactions hôte-pathogène**.
Les graphes (*a*) et (*b*) sont issus de (Woolhouse et al., 2002) et représentent les changements de fréquence d'allèle entraînés par la coévolution au cours des générations (***a***) Une série de balayages sélectifs des allèles hôte (ligne bleue) et pathogène (rouge) dérivés par mutation. La sélection est directionnelle, c'est-à-dire que le changement génétique s'accumule dans les deux populations. À n'importe quel stade donné du processus, il peut y avoir des polymorphismes dans l'une, les deux ou aucune des deux populations. (***b***) Polymorphismes dynamiques à la fois chez l'hôte (bleu) et le pathogène (rouge) agissant sur la variation génétique existante. L'évolution n'est pas directionnelle. À toutes les étapes du processus, les deux populations sont polymorphes.

Les modes, processus, et mécanismes de sélection

Dans le cas de la dynamique de guerre de tranchées, on observe des polymorphismes stables à long terme au niveau des loci co évolutifs de l'hôte et du parasite, le mode de sélection sous-jacent est la sélection équilibrante (Tellier et al., 2014). On parle de **sélection balancée** (*balancing selection*), lorsque les allèles impliqués ont, en moyenne, un avantage sélectif qui est en corrélation négative avec leur fréquence au sein d'une population ou d'une espèce. Ce terme n'inclut pas les loci qui sont à un équilibre entre les variants gagnants et perdants, telle que la dérive de mutation et l'équilibre de sélection de mutation (Ebert and Fields, 2020). Ce mode de sélection fait notamment intervenir des mécanismes de sélection antagonistes dépendante de la fréquence. La **sélection fréquence dépendante** (*Negative frequency-dependent selection*) se produit lorsque la valeur sélective d'un variant (par rapport à d'autres variants) est fonction de

son abondance dans la population (par rapport à d'autres variants) de telle sorte que sa valeur relative augmente avec l'abondance relative, ou la fréquence, des variants (Brisson, 2018; Ebert and Fields, 2020). La diversité génétique est maintenue et les allèles adaptatifs oscillent en fréquence à la suite d'une sélection contre le type d'hôte ou d'agent pathogène le plus courant (Papkou et al., 2019).

Dans le cas de la dynamique de course à l'armement, les deux antagonistes accumulent séquentiellement des mutations adaptatives, chacune allant jusqu'à la fixation à travers la population hôte ou pathogène à travers la sélection directionnelle (**Fig.3. p8**). La **sélection directionnelle** (*Directional selection*) est un mode de sélection naturelle par lequel un variant génétique va se fixer et se propager dans la population (également connue sous le nom de sélection positive) (Ebert and Fields, 2020). Ce mode de sélection fait intervenir plusieurs processus de sélection dont des **balayages sélectifs** (*selective sweep*) récurrents pour lesquels un allèle avantageux va augmenter en fréquence et se fixer dans une population (**Fig.3. p8**). Les allèles liés peuvent également être fixés par «auto-stop» (*Genetic hitch-hiking*) (Woolhouse et al., 2002) et des mutants bénéfiques peuvent apparaître *de novo* ou se sont séparés dans la population avant le balayage et devenir bénéfiques après un changement de conditions (Ebert and Fields, 2020).

Les variants génétiques sous sélection balancée ne fournissent un avantage que par rapport à un variant génétique spécifique correspondant dans l'antagoniste, contrairement au modèle de sélection directionnelle par balayage sélectif, dans lequel les allèles mutants ont un avantage général sur les allèles de type sauvage. Ainsi, par rapport au reste du génome, les régions génomiques contenant des gènes sous sélection d'équilibrage devraient présenter une grande diversité génétique et un spectre de fréquences plus uniforme des variants. Pour mettre en évidence les différents modes évolutifs, il est nécessaire d'observer les différents processus dont sont sujets les antagonistes, dans notre cas un hôte et un pathogène, afin d'évaluer les résultats évolutifs de ces interactions ; soit en mesurant empiriquement l'adaptabilité réciproque de l'hôte et du pathogène au fil du temps dans des populations co évolutionnaires (Gandon et al., 2008; Tellier et al., 2014; Thrall et al., 2012) ; soit en étudiant les signatures moléculaires dans les séquences génétiques de l'hôte et du parasite. En effet, d'après la théorie synthétique de l'évolution il est défini que les phénomènes d'adaptation sont des processus lents, causés par des mutations génétiques aléatoires, on s'attend donc à ce que les deux dynamiques laissent des traces observables dans les génomes (Woolhouse et al., 2002).

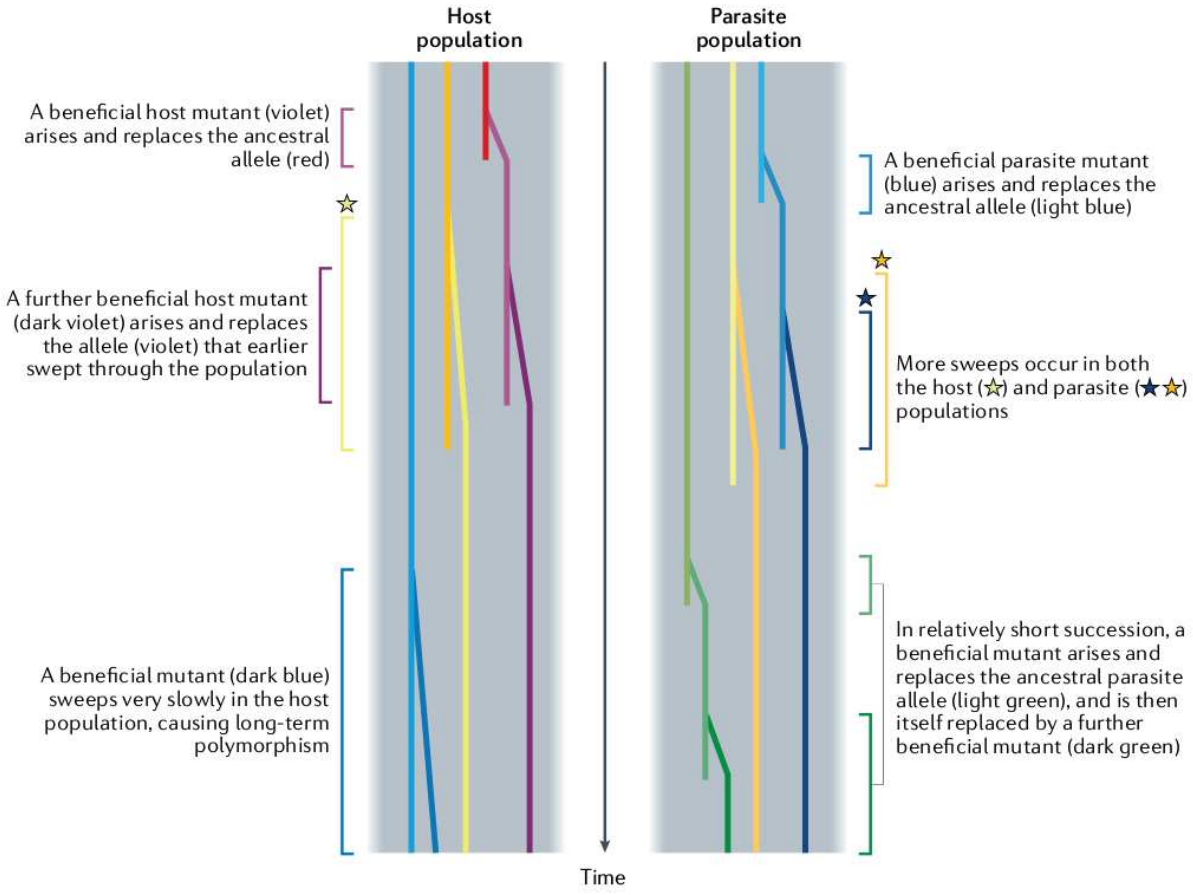


Figure 3. **Représentation schématique de la coévolution sélective par sélection directionnelle avec un processus de balayage sélectif dans les pools de gènes d'une population hôte et parasitaire**. Pour chaque antagoniste, les chronologies des allèles à différents loci sont affichées dans différentes couleurs. Au cours de la période de temps indiquée, chacune des allèles est remplacée au moins une fois par un mutant, indiquant des balayages sélectifs. La variation génétique au niveau des loci - c'est-à-dire la présence de plusieurs allèles dans le pool génétique à un moment donné - n'est visible que pendant le balayage. Les balayages hôte et parasite ne sont pas liés. Les locus peuvent être présents n'importe où dans le génome. D'autres loci peuvent subir des balayages en même temps pour des raisons sans rapport avec les interactions co évolutives. Illustration tirée de la figure 1 de Ebert et Fields (2020).

La signature génomique des interactions hôte pathogène

La coévolution hôte-parasite transforme par définition les génomes des deux protagonistes. On peut donc caractériser des **signatures génomiques** c'est-à-dire des motifs caractéristiques de variation génétique, observés dans une région génomique dans un échantillon de génomes (Ebert and Fields, 2020). L'empreinte de la coévolution et les signatures génomiques des deux antagonistes partages des contraintes inhérentes dont les deux protagonistes ne peuvent pas se soustraire pour arriver à se maintenir (Hsu and Du Pasquier, 2015). (i) Investir dans des ressources pour des mécanismes de virulence ou de résistance signifie que ces ressources ne sont pas redistribuées à la production de progéniture, ce qui conduit inévitablement à une diminution de la valeur sélective. (ii) Le génome est contraint par sa taille et son nombre de gènes, il est donc nécessaire de coder les informations pour la

virulence et la défense avec le moins d'informations génétiques possibles sous peine d'avoir un coût supplémentaire inapproprié. Malgré ces contraintes communes, on ne peut pas s'attendre à ce que les signatures génomiques de la coévolution chez les hôtes et les parasites se ressemblent, car les deux antagonistes vivent la sélection de manières très différentes. Premièrement, les hôtes ont généralement un temps de génération et un temps de vie bien plus important que les pathogènes (*Generation Gap*) qui peuvent donc fixer de nouveaux mécanismes de résistance beaucoup plus rapidement que l'hôte ne peut fixer des stratégies de résistance complémentaires. Deuxièmement, lors de sa réponse immunitaire, l'hôte doit faire en sorte d'éviter de causer des dégâts irréversibles à ses tissus et organes lorsqu'il élimine l'agent pathogène. Troisièmement, les agents pathogènes et en particulier certains parasites se doivent de ne pas être trop virulents, au risque de tuer leurs hôtes et de ne pas pouvoir se maintenir.

Ce n'est qu'en suivant ces traces laissées par la diversité génétique dans les génomes des organismes en interaction que l'on peut juger de l'importance de cette dernière dans un processus biologique comme la pathogenèse. Et donc par corollaire, c'est au travers des modèles de biologie évolutive qu'il sera possible de faire la différence d'une diversité génétique effective ou plus simplement d'une dérive génétique. Des travaux sur les analyses des génomes d'organismes hôte ont révélés de nombreuses signatures de régions sous sélection d'équilibrage, les plus connues étant les gènes majeurs d'histocompatibilité (CMH) (Leffler et al., 2013) et le système de groupe sanguin ABO des primates supérieurs (Bubb et al., 2006). Moins d'attention est accordée à la sélection chez les parasites que chez les hôtes. Toutefois, suite au développement des outils bioinformatiques, l'exemple des loci où une sélection ayant lieu a été observée récemment grâce à des études d'association à l'échelle du génome (GWAS, G*enome-Wide Association Study*) pour une série de pathogènes comme des bactéries (Andras et al., 2020; Thomas et al., 2012), ou des virus (Bartha et al., 2013; Brackney et al., 2009; Ryabov et al., 2019). Ces études sur les variations génétiques chez de nombreux individus, déterminent leurs corrélations avec des traits phénotypiques et mettent en évidence quelles sont les variations génétiques impliquées dans une pathologie.

Les différentes modifications du génome

Le maintien de signatures génomiques dans les populations est directement influencé par la sélection naturelle telles que les interactions hôte-pathogène décrites précédemment mais il existe également trois autres forces évolutives qui font varier les fréquences alléliques et génotypiques dans des populations naturelles. Les phénomènes de **flux de gènes** (ou migrations), c'est-à-dire le mouvement des allèles entre les populations génétiquement structurées entraînant un changement de la fréquence des alleles (Wolf and Ellegren, 2017). La **dérive génétique**, c'est-à-dire des changements aléatoires de la fréquence des allèles entre les générations, conséquence d'un échantillonnage stochastique dans une population finie (Wolf and Ellegren, 2017). Les fréquences alléliques fluctuent d'autant plus que les populations sont de petite taille. Enfin la **mutation**, un changement irréversible du matériel génétique (Lefevre et al., 2016). Cette dernière est la source fondamentale de variation, mais le devenir des mutations va dépendre des autres forces évolutives qui agissent en interactions. Les modifications spontanées d'un état allélique par mutation entraîne des formes variantes d'une séquence génétique sont appelées **allèles**, et la position à laquelle elles se produisent est appelée **locus**.

Les divergences observables sont des **polymorphismes** et représentent la coexistence de plusieurs allèles pour un gène ou locus donnés dans une population (Lefevre et al., 2016). Les mutations fréquemment rencontrées sont (**Fig.4. p10**), des changements ponctuels d'un nucléotide on parle de SNP (*single-nucleotide polymorphism*), des insertions / délétions avec l'ajout ou la suppression de nucléotide(s), on parle de IN/DEL. D'autres changements sont bien plus impactants comme les **variations structurelles** (*Structural variations, SV*) qui peuvent être définies sur le plan opérationnel comme des variantes de séquence d'une taille> 50 pb. Les formes les plus reconnues de SV incluent les suppressions, les duplications, les inversions, les insertions et les translocations (Ho et al., 2019). La plupart du temps, les petites mutations n'ont aucun effet. Mais, parfois, la combinaison de plusieurs mutations voire un seul acide aminé manquant ou modifié peut entraîner la production d'une protéine modifiée, ou pire aucune protéine du tout, avec de graves conséquences. La caractérisation de ces informations nous permet de nous renseigner et de trouver les signatures de coévolution d'une espèce au travers de l'analyse de la diversité génétique.

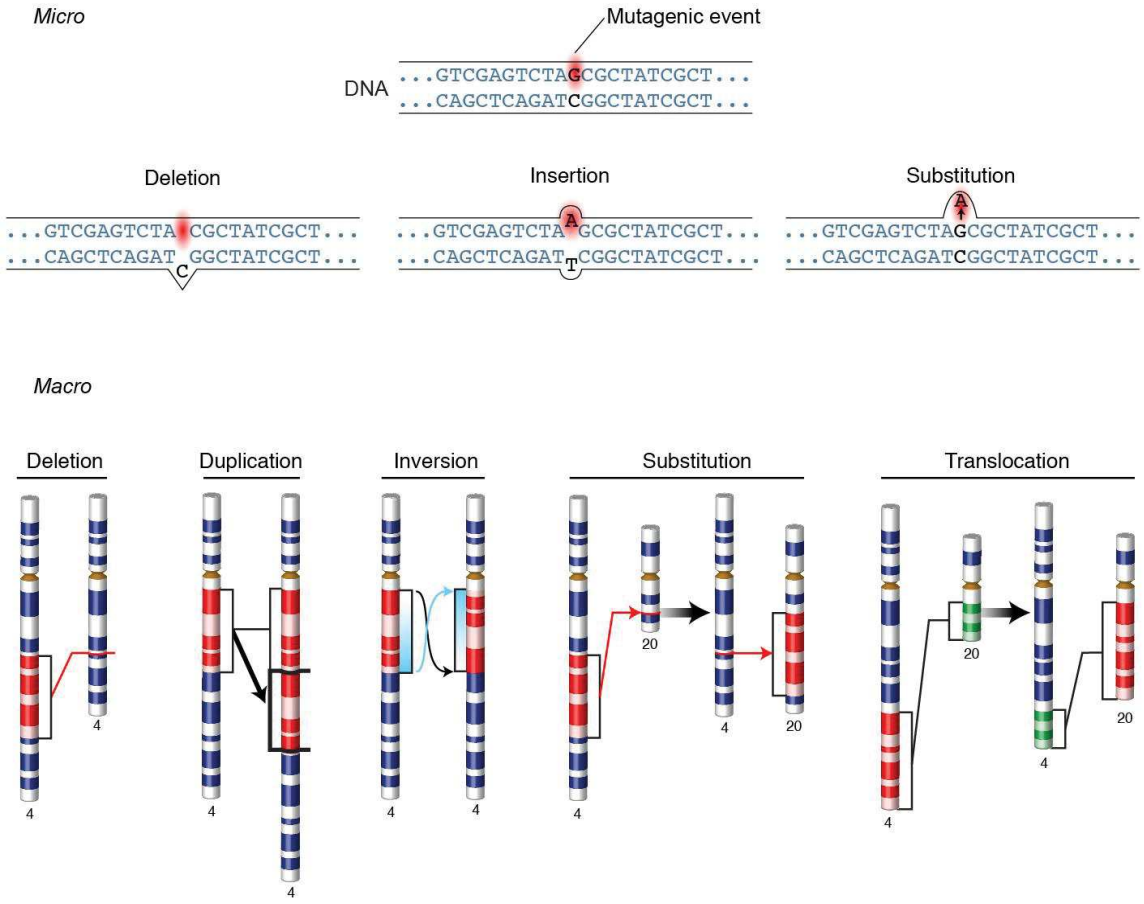


Figure 4. **Différents types de mutations**.
En haut les mutations sont de petits changements de séquences, en bas ceux sont des variations structurelles. Image issue du National Human Genome Research Institute

La caractérisation de la diversité génétique

Pour mesurer la diversité génétique il est nécessaire d'obtenir la séquence du matériel génétique (ADN ou ARN), c'est-à-dire de déterminer la succession des nucléotides d'une information génétique. Cela peut se faire grâce aux méthodes de **séquençage**, le processus de détermination de la séquence nucléotidique d'un fragment donné. Avec les progrès des technologies de séquençage à haut débit (HTS) il est maintenant possible d'évaluer la diversité génétique de populations à un niveau de détails sans précédent. Une fois l'information génétique séquencée, il est nécessaire de l'analyser pour déterminer la diversité génétique réelle. Il est important de considérer comment nous allons définir la diversité génétique, car il s'agit d'un terme complexe avec une gamme de définitions en fonction de la discipline et de la mesure employée (Avolio et al., 2012). Pour cette thèse nous la définirons comme "une mesure quantitative relative, basée sur le génome, qui quantifie l'ampleur des modifications du génome au niveau moléculaire entre individus d'une même espèce ou d'espèces différentes pour une séquence d'ADN apparentée". La variabilité génétique peut être observée avec une analyse bioinformatique qui utilise l'alignement de séquences. Ces alignements sont une façon d'organiser des séquences de l'ADN, ARN, ou protéine pour identifier des régions de similarité. Ainsi une comparaison de diversité génétique est illustrée (**Fig.5. p11**) ou les séquences en A montrent une plus grande diversité génétique que les séquences en B. Les séquences présentées ici sont simples, mais cet alignement de séquences pourrait être des génomes entiers, des portions de séquences codantes ou encore de courtes lectures (*reads*) d'ADN obtenues par HTS.

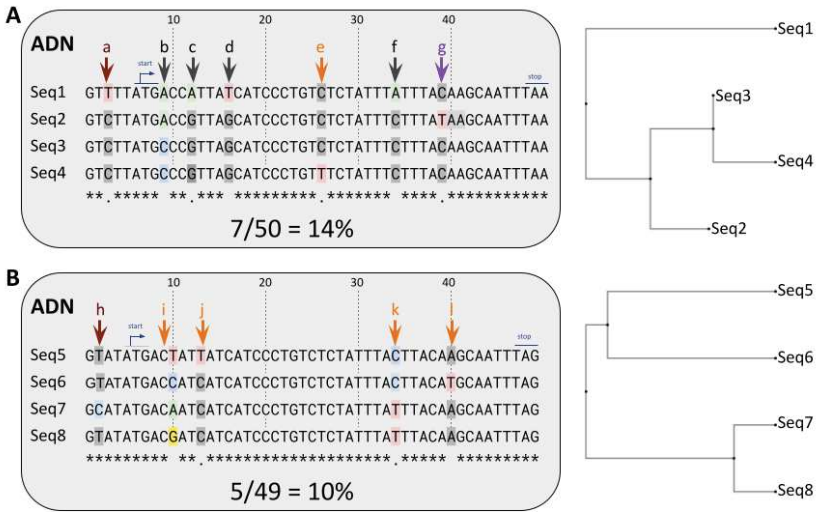


Figure 5. **Illustration de la diversité génétique.**
Les séquences d'ADN comportant le début d'un codon d'initiation de la transcription en 5eme position (codon start, en bleu) et un codon entraînant un arrêt de la transcription pour les 3 derniers nucléotides (codon stop, en bleu). L'alignement de 8 séquences d'ADN dont les mutations sont mises en évidence avec des lettres, de *a* à *l.* L'impact de ces mutations est catégorisé par des couleurs : en marron des mutations qui se trouvent en dehors de la séquence codante (*a*, *h*), en noir des mutations entrainant un changement d'acide aminée sur la séquence codante (*b*, *c*, *d*, *f*), en orange les mutations n'entraînant pas de changement d'amino acide, (*e*, *i*, *j*, k, *l*) en violet la mutation entraine un codon stop (*g*). L'arbre résultant de l'alignement a été construit à droite avec l'ensemble des nucléotides dans les quatres séquences. Les séquences en (**A**) montrent une plus grande diversité génétique comparées aux séquences en (**B**).

Résumé

- *Dans les interactions hôte-pathogène il y a un phénomène de coévolution si l'hôte et le pathogène ont des changements génétiques adaptatifs et réciproques entre eux.*
- *Il existe plusieurs niveaux de coévolution, le plus simple est la coévolution antagoniste spécifique.*
- *Les phénomènes de co-évolutions ont des dynamiques différentes, les deux principales dans l'interaction hôte-pathogène sont la dynamique de course à l'armement, et celle de guerre des tranchées.*
- *La dynamique de guerre des tranchées se fait majoritairement par la sélection d'équilibrage via des processus de sélection fréquence dépendante.*
- *La dynamique de course à l'armement se produit majoritairement par la sélection directionnelle grâce à des processus de balayages sélectifs.*
- *Ces phénomènes de coévolution peuvent être observés grâce aux signatures génomiques laissées dans le génome des protagonistes.*
- *Ces traces peuvent entraîner de petites modifications, comme de simples SNP , ou via des variations structurelles importantes pouvant avoir un fort impact fonctionnel sur des protéines.*
- *L'ensemble de ces modifications forment la diversité génétique et celle-ci peut être caractérisée par l'alignement de séquence nucléotidique.*

Les mécanismes de diversification génétique

Au cours de l'évolution, différents mécanismes ont été sélectionnés pour permettre d'augmenter, diminuer et ainsi contrôler la diversité génétique. Ces mécanismes peuvent être très différents selon le type d'organisme. Ainsi nous pouvons conceptualiser différentes stratégie de diversification génétique en fonction de l'espérance de vie, la reproduction et le support de l'information génétique des différents organismes. Mieux comprendre ces stratégies nous éclaire sur l'évolution de la résistance aux médicaments et de la virulence à la conception des vaccins et à l'émergence de nouvelles maladies. Ici, nous cherchons à présenter des mécanismes importants chez des organismes relativement proches de nos modèles. C'est pourquoi cette partie ne se veut pas exhaustive mais se concentrera sur les organismes eucaryotes diploïdes et les virus.

La diversité génétique chez les organismes eucaryotes pluricellulaire à reproduction sexuée

Chez les organismes eucaryotes sexués, la diversité génétique se fait principalement lors de la transmission du patrimoine génétique à la descendance au moment de la méiose des cellules germinales. Seules les mutations se produisant dans la lignée germinale : ovules et les spermatozoïdes peuvent être transmises à la génération suivante. En plus des modifications du génome décrit précédemment il y a également des phénomènes de **recombinaisons** homologues ou de ségrégation, qui vont permettre d'augmenter la diversité génétique. Ainsi, après la fin du développement embryonnaire d'un organisme, l'ensemble des cellules somatiques ont un patrimoine génétique presque identique. Des changements génomiques peuvent apparaître lorsqu'une cellule se divise, en mitose. Ce processus est très encadré et régulé au risque de favoriser l'apparition de cancer. De nombreux mécanismes sont là pour s'assurer de la bonne intégrité du génome afin de limiter au maximum la variabilité génétique. Il existe cependant d'autres mécanismes supplémentaires pour générer de la diversité génétique parmi certaines cellules somatiques d'un individu lorsqu'ils sont pluricellulaire. Ces mécanismes sont principalement activés dans les cellules impliquées dans la reconnaissance et le maintien de l'homéostasie, les cellules du système immunitaire.

Chez les vertébrés

Chez les vertébrés, la reconnaissance des agents pathogènes potentiels repose sur un ensemble de récepteurs immunitaire. Parmi ces récepteurs, certains sont capables de se lier spécifiquement à des motifs d'un pathogène particulier, il s'agit des récepteurs des cellules B (*BCR*) ou T (*TCR*). Grâce à ces récepteurs, plus grande est la diversité des molécules de reconnaissance dont un organisme est équipé contribue, plus grandes sont ses chances de survivre à une maladie infectieuse. Chez la souris et l'homme, on pense que le répertoire TCR/BCR, comprend environ 10e7 et 10e8 récepteurs distincts, respectivement (Haeryfar et al., 2008). Si on prend l'exemple du BCR et par prolongement des immunoglobulines, ces molécules font environ 2000 nucléotides, l'information génétique pour en générer 10e7 distinct nécessite une quantité d'ADN dédiée à tous les récepteurs lymphocytaires qui dépasse la taille du génome de l'organisme. Il existe donc des mécanismes permettant de dépasser cette contrainte en générant de la diversité génétique. Pour illustrer ces propos, nous allons nous intéresser aux recombinaisons V(D)J dans l'ontogénie des lymphocytes et à l'hypermutation somatique des cellules B lors d'une réponse immunitaire active.

L'une des stratégies de génération de diversité lors du développement des lymphocytes est la **recombinaison V(D)J** (*V(D)J recombination*) par des enzymes, dont RAG (*recombination activating genes*), qui permettent l'association combinatoire de multiples segments de gènes appelés V (variable) et J (jonction), qui codent ensemble la région variable de la chaîne légère de la molécule d'Ig, et V, D (diversité) et J qui codent la région variable de la chaîne lourde permettant la génération d'un répertoire important à partir d'un pull de gène limité (Hozumi and Tonegawa, 1976; Tonegawa, 1983). S'ajoute une étape pour augmenter la diversité génétique en incorporant aléatoirement des nucléotides grâce à l'enzyme désoxynucléotidyl-Transférase Terminale exprimée également lors du développement lymphocytaire (Desiderio et al., 1984; Tonegawa, 1983), les nucléotides ajoutés le sont dans la région CDR3 (une région

déterminante pour la complémentarité de la reconnaissance de l'antigène) pour accroître la diversité. Cette génération de diversité a un coût important pour l'organisme car pour ne pas endommager son propre organisme un processus de sélection positive, par exemple dans le thymus pour les lymphocytes T, doit avoir lieu et entraîne des pertes cellulaires importantes.

Une autre stratégie de génération de diversité intervient dans la maturation des lymphocytes lors de la réponse immunitaire adaptative dans les centres germinatifs pour diversifier les récepteurs des cellules B (Weigert et al., 1970). Cela permet une augmentation de l'affinité des anticorps sécrétés en réponse à un antigène (Griffiths et al., 1984). Il s'agit de **l'hypermutation somatique** (*Somatic hypermutation*) qui altère les séquences codant le site de liaison à l'antigène des gènes de régions variables d'Ig (Bernard et al., 1978), grâce à une enzyme spécifique appelée AID (*Activation-Induced cytidine Deaminase*). Cette enzyme désamine une cytidine dans l'ADN, ce qui crée un mésappariement uridine-guanosine (U-G). La réparation erronée de cette lésion introduit une mutation sur le brin complémentaire de celui portant la cytosine désaminée par AID. Ces mutations surviennent à un rythme 10e−3/paire de base /génération (Rajewsky et al., 1987). Là encore, le mécanisme a un coût pour l'organisme qui à travers des interactions cellulaires va enclencher une sélection positive dirigée par l'antigène induisant, ou non, l'apoptose de ces cellules en fonction de leur affinité.

Chez les invertébrés

Les invertébrés ne possèdent pas le système immunitaire adaptatif à base de lymphocytes, mais des mécanismes de diversification des molécules impliqués dans la défense immunitaire ont été observés. Des travaux ont révélé que l'escargot *Biomphalaria glabrata* possède une famille diversifiée de protéines appelées protéines apparentées au fibrinogène ou **FREP** (Fibrinogen-RElated Proteins) ayant un rôle dans les défenses immunitaires. Ces polypeptides sont constitués d'un ou deux domaines de la superfamille d'immunoglobulines (IgSF) amino-terminaux et d'un domaine de fibrinogène carboxyl-terminal (Adema et al., 1997; Zhang et al., 2004). Ces molécules ont montré une grande variabilité dans certaines de leurs régions (Léonard et al., 2001; Zhang et al., 2001) avec une diversité 3 à 4 fois supérieure à d'autres gènes (Zhang et al., 2004). L'analyse des séquences parentales et descendantes ayant montré une diversité différente, il a été suggéré pour expliquer ce modèle que l'ADN de la lignée germinale contient des séquences sources de FREP qui sont diversifiées dans toutes les cellules somatiques à une faible fréquence (Portet et al., 2017; Zhang et al., 2004). Cette spécificité de reconnaissance immunitaire a d'autant plus de sens qu'un processus co-évolutif impliquant les FREP a lieu entre *B. glabrata* et son parasite *Schistosoma mansoni* appelé **polymorphisme de compatibilité** (*compatibility polymorphism*) (Mitta et al., 2012). La compatibilité réside dans une concordance de phénotypes génétiquement déterminés dans l'escargot et le schistosome (Basch, 1976, 1975). Les mécanismes moléculaires qui sous-tendent le haut niveau de polymorphisme de compatibilité observé dans ces interactions sont basés sur la formation de complexes immuns entre les FREPs de *B. glabrata* et les mucines polymorphes de *S. mansoni* (SmPoMucs) (Dheilly et al., 2015; Mitta et al., 2012; Moné et al., 2010; Portet et al., 2017). Chez *B. glabrata* plusieurs sous-familles différentes de FREP ont été identifiées, et comprendre la fonction de telles molécules polymorphes ou diversifiées chez les invertébrés est toujours un axe de recherche en cours.

L'impact de la diversité génétique chez les organismes eucaryotes pluricellulaire à reproduction sexuée

Ces mécanismes de génération de diversité permettent non seulement de créer un répertoire immunitaire pour la reconnaissance de l'immense variété des antigènes étrangers mais également de réagir aux nouveaux variants capables d'échapper au système immunitaire au travers du phénomène d'**adaptation à la variabilité antigénique**. Cette adaptation va permettre de cibler et d'optimiser une réponse immunitaire. Cela peut entraîner soit l'élimination de l'agent pathogène ou soit la coexistence de nombreux variants au sein d'un même hôte avec une évolution intra-hôte. C'est le cas des infections causées par le virus de l'immunodéficience humaine (VIH), ou le virus de l'hépatite C (VHC), avec une «course aux armements» co évolutionnaire entre l'hôte et l'agent pathogène intense (Grenfell, 2004).

Les bases de la diversité génétique chez les virus

Les **virus** sont de petits fragments d'acides nucléiques protégés par une couche protéique ou lipoprotéique. Ils utilisent les ressources de leur hôte pour se répliquer (Owen et al., 2014). Bien que ce soient des parasites cellulaires absolus, ils ne sont pas toujours nocifs pour leurs hôtes (Gauthier et al., 2018; Roossinck, 2011). Ces entités sont connus comme présentant une grande diversité génétique et sont pertinents à étudier du fait de leur temps de génération court.

Les virus à ARN

Les mécanismes de génération de diversité génétique chez les virus à ARN

Les **virus à ARN** sont, comme leurs noms l'indiquent, des virus dont le support du matériel génétique est constitué de molécules d'ARN. Ces virus ont, en général, une réplication sujette aux erreurs due à l'absence de d'activité de correction de leurs ARN polymérases ARN-dépendantes (Coffey et al., 2011; Jin et al., 2011; Levi et al., 2010). Ces erreurs peuvent être mesurées à travers le **taux de mutation** (*mutation rate*), le taux de mauvaise incorporation de nucléotides par nucléotide copié. L'étude des bactériophages à ARN a permis d'estimer le taux de mutation avec une 10^{-4} mutation par nucléotide copié (Batschelet et al., 1976). Ce taux de mutation était 10^4 fois plus élevé que les valeurs qui, à l'époque, avaient été estimées pour certains bactériophages à ADN et 10^6 à 10^8 fois plus élevé que les valeurs pour les bactéries et les champignons (Drake, 1969). Les taux de mutation des virus à ARN mesurés par des procédures génétiques et biochimiques (Batschelet et al., 1976; Drake and Holland, 1999; Sanjuán et al., 2010) impliquent que la plupart des ARN matrices copiés dans un brin complémentaire doivent contenir entre 0,1 et 2 mutations. Ainsi chez de nombreux virus à ARN, les taux de mutation sont tels qu'il est peu probable qu'il produise à l'intérieur d'une cellule infectée une molécule d'ARN viral de descendance identique à sa matrice parentale immédiate (Domingo et al., 2012). Ce taux de mutation élevé conduient à la génération rapide de variants viraux et à partir d'un faible nombre de virus, un ensemble hétérogène et dynamique de mutants apparaît très rapidement (Arias et al., 2001; Carrasco et al., 2007). Pour qualifier cet ensemble de mutant, les virologues utilisent le terme de **quasi-espèces virales** (*viral quasispecies*) pour désigner les distributions de génomes non identiques mais apparentés

soumis à un processus continu de variation génétique, de compétition et de sélection et qui agissent comme une unité de sélection, on parle également de «nuages mutants» ou «d'essaim» (Domingo et al., 2012; Lauring and Andino, 2010; Perales et al., 2010). L'utilisation de ce terme est controversée (Fitzsimmons et al., 2018; Holmes, 2010) mais la théorie des quasi-espèces est appuyée par des preuves théoriques et expérimentales (Eigen, 1971). L'organisation génétique des populations est souvent représentée à l'aide du concept d'**espace de séquence** (*sequence space*), une représentation géométrique de toutes les séquences possibles où la distance physique reflète la similitude génétique. Un génome viral donné se répliquera et produira des centaines de descendants, lesquels généreront une distribution mutante plus complexe avec des variantes plus éloignées les unes des autres dans l'espace de séquence (Domingo et al., 2012; Lauring and Andino, 2010). Les travaux sur la modification de la fidélité de la polymérase virale agissant sur le taux d'erreur ont montré que le taux de mutation joue un rôle clé dans l'hétérogénéité observée du spectre mutant correspondant. Une hétérogénéité limitée peut entraîner un désavantage sélectif pour le virus, l'amplitude du spectre mutant peut être un caractère stable maintenu après plusieurs passages en culture cellulaire (Domingo et al., 2012; Herr et al., 2011; Levi et al., 2010; Wainberg et al., 1996; Whitney et al., 2002). Toutefois, l'hétérogénéité des populations est une arme à double tranchant, les virus à ARN sont soumis à la **mutabilité catastrophique** (*error catastrophe*). Ce phénomène correspond à la fois à la perte de la supériorité de la séquence majoritaire, et à une dérive dans l'espace des séquences conduisant à des génomes viraux non fonctionnels, et donc à l'extinction à la suite d'un taux de mutations excessif (Domingo et al., 2012).

L'impact de la diversité génétique chez les virus à ARN

L'une des conséquences des quasi-espèces virales est que la compétition et la complémentation entre les génomes mutants offre un gain de valeur sélective lorsque de grandes populations de virus à ARN sont autorisées à se répliquer dans un environnement défini (Arias et al., 2001). Cela leur confère plusieurs propriétés tel qu'une mémoire moléculaire des génomes viraux (Arias et al., 2001; Briones et al., 2003) c'est-à-dire un enregistrement du passage de génomes dominants dans les phases précédentes d'une même lignée (Ruiz-Jarabo et al., 2000); le dépassement de barrière génétique et phénotypique (Domingo et al., 2012) ; des phénomènes de complémentation sous-tenant le maintien de variants de faible valeur sélective (Aaskov et al., 2006; Gelderblom et al., 2008) ou à l'inverse d'interférence entre les variants (Roux et al., 1991; Torre and Holland, 1990). Toutes ces propriétés ont des implications importantes tant dans la biologie des virus que dans les interactions avec leurs hôtes (Domingo et al., 2012). Les mutations qui conduisent à des substitutions d'acides aminés au niveau des résidus externes des protéines de surface peuvent entraîner une modification du tropisme de la cellule hôte (Baranowski et al., 2001; Boyd et al., 1993; Matloubian et al., 1993; Qi et al., 2009; Vahlenkamp et al., 1997; Wu et al., 2006). D'autres variations génétiques peuvent permettre aux virus d'échapper au système immunitaire de l'hôte (Emu et al., 2018; Farci, 2011; Fausther-Bovendo et al., 2009; Mas et al., 2018; Phillips et al., 1991; Richman et al., 2003; Simões et al., 2020) ou d'échapper aux mécanismes d'action de molécules antivirales (Feigelstock et al., 2011; Scheidel et al., 1987; Young et al., 2003). Ces échappements au contrôle du système immunitaire sont donc un facteur important pour la virulence des virus. L'exemple classique illustrant l'importance de la diversité de la population virale coinfectante comme déterminant de la

virulence concerne un modèle de souris transgénique pour l'infection par le poliovirus (**Fig.6. p17**). En inoculant une population génétiquement homogène de poliovirus mutés avec une polymérase de plus haute fidélité chez la souris, ils ne sont pas capables d'envahir le système nerveux central contrairement au virus non muté avec une population très hétérogène (Vignuzzi et al., 2006). La diversité génétique d'une population fait donc partie des **déterminants de la virulence** de certains virus à ARN, facteur qui augmente soit la gravité de la maladie, soit la propension à provoquer la maladie.

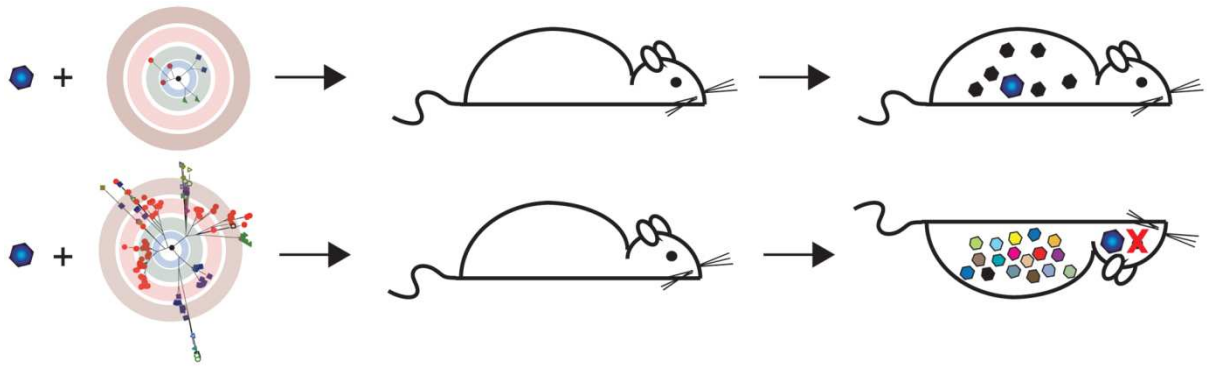


Figure 6. **Diversité des populations virales peut être un déterminant de la virulence**.
Les résultats des expériences décrites dans Vignuzzi et al. 2006. Un clone neuro-virulent du poliovirus a été isolé du cerveau de souris qui avaient été infectées par une souche de type sauvage. Des souris naïves ont ensuite été ré-infectées avec ce clone dans le cadre d'une population soit génétiquement contrainte (en haut), soit diversifiée (en bas). Bien que toutes les souris aient reçu le clone neuro-virulent, seules celles infectées par une quasi-espèce diverse ont développé la maladie. Des sous-populations au sein des diverses quasi-espèces ont coopéré avec le clone neuro-virulent pour faciliter son entrée dans le SNC. Figure extraite de (Lauring and Andino, 2010).

Les virus à ADN

Les mécanismes de génération de diversité génétique chez les virus à ADN et en particulier les Herpèsvirus

Avec la démocratisation de l'utilisation du séquençage à haut débit, l'évaluation de la variation des génomes à ADN a été effectuée. Contrairement aux virus à ARN, la plupart des **virus à ADN** (*DNA virus*), pour lesquelle le support du matériel génétique est constitué de molécules d'ADN, ont des polymérases faisant des erreurs d'incorporation de nucléotides avec une fréquence plus faible que pour les virus à ARN, ainsi qu'une activité de correction des erreurs d'incorporation dans le brin néoformé. Au cours de la réplication de l'ADN, les bases incorporées de manière erronée sont éliminées par une exonucléase 3'-5 ', une fonction enzymatique hautement conservée de l'ADN réplicatif mais pas des ARN polymérases (Trimpert et al., 2019). Dans le cas particulier des Herpèsvirus, des observations suggèrent que ces virus abritent une vaste diversité génétique à la fois au sein d'un seul individu (variation intrahôte) mais aussi entre plusieurs individus (variation interhôte) (Renner and Szpara, 2017; Renzette et al., 2017, 2015, 2014, 2013, 2011). Le taux de mutation chez l'herpès simplex de type 1 (HSV-1) est estimé à 5x10e-8 substitutions par nucléotide par infection cellulaire (Lu et al., 2002;

Sanjuán et al., 2010). Un taux faible comparé par exemple au poliovirus, un virus à ARN, qui est de 9x10e-5 substitutions par nucléotide copié (Drake and Holland, 1999; Sanjuán et al., 2010). Toutefois des taux de mutations plus élevés ont été observés pour l'HSV-1 avec 10e-5 taux de mutations en substitutions par nucléotide par an (Firth et al., 2010), les auteurs du papier soulignant que "Le mécanisme le plus probable par lequel un tel taux de substitution gonflé pourrait se produire est la sélection positive forte et continue". Une autre étude basée sur les transferts de plaques a également montré une fréquence moyenne de mutation de 2,1 × 10e−4 substitutions par nucléotide par transfert de plaque pour la région génomique de 2 gènes, codant pour la glycoprotéine gD et thymidine kinase, de l'HSV-1 alors qu'aucune mutation n'est observée dans le gène de l'ADN polymérase du virus (Jaramillo et al., 2013). Des recherches sont toujours en cours pour expliquer la diversité observée chez les virus à ADN, bien que la littérature offre plusieurs mécanismes non mutuellement exclusifs qui pourraient jouer un rôle. Des modélisations ont suggéré que la diversité virale chez les hôtes infectés résultait d'une combinaison de sources telles que : la sélection positive ; la réinfection ; la recombinaison entre souches et dans une moindre mesure, entre espèces ; les goulots d'étranglement (Pokalyuk et al., 2017; Renner and Szpara, 2017; Renzette et al., 2016, 2015).

L'impact de la diversité génétique chez les virus à ADN

À l'heure actuelle, nous avons une compréhension limitée du rôle que peut jouer la diversité génétique des virus à ADN dans l'évolution et la sévérité d'une maladie. Toutefois, plusieurs études ont associé la diversité génétique avec une augmentation de la virulence. Des résultats ont par exemple récemment montré que des facteurs intrinsèques de l'hôte ne suffisaient pas pleinement à expliquer par l'évolution de maladie invasive à l'herpès simplex virus alors qu'en parallèle plusieurs variantes de protéines virales pourraient avoir un impact sur la virulence (Akhtar et al., 2019). Dans une autre étude, des goulots d'étranglements génétiques ont été artificiellement produits par des transferts de plaque à plaque de HSV-1, ces transferts ont conduit à une accumulation de mutations dans les génomes viraux. Après une inoculation intracérébrale de souris avec ces génomes viraux, la dose létale était plus élevée pour les clones soumis à des transferts de plaque à plaque que pour leurs clones parentaux correspondants (Jaramillo et al., 2013). Il existe deux domaines spécifiques où la diversité génétique peut modifier la pathogenèse chez les virus à ADN : la dissémination et l'évasion immunitaire (Renzette et al., 2014). Par exemple, une meilleure dissémination d'un virus influencée par la diversité génétique a pu être observé lors de l'étude sur la distribution et le tropisme cellulaire de différents types de glycoprotéines virales de l'HCMV (Meyer-König et al., 1998). Pour l'évasion immunitaire, des recherches ont montré que l'évolution de populations de lymphocyte T réactives de variants du HCMV étaient modelées par les variants majoritaires chez les individus exposés (Corey et al., 2014). Cependant d'autres résultats sont en contradictions et montrent que le répertoire des cellules T dirigées contre HCMV ou EBV est stable dans le temps (Klarenbeek et al., 2012). L'ampleur et les conséquences biologiques de la variation génétique des virus à ADN complexes sont en grande partie un problème non résolu (Jaramillo et al., 2013). Ainsi la manière dont la variabilité génétique affecte les virus à ADN par rapport aux virus à ARN est une question non résolue en virologie évolutive (Firth et al., 2010; Jaramillo et al., 2013).

Résumé

- *Chez les organismes eucaryotes à il existe des moyens d'augmenter la diversité génétique au travers de mécanismes spécifiques.*
- *Ces mécanismes sont induits dans les cellules immunitaires principalement pour dépasser la contrainte du nombre de genes.*
- *Chez les virus les mécanismes d'augmentation de la diversité génétique sont nombreux et peuvent être des déterminants importants de la virulence.*
- *Le cas extrême de diversité génétique se retrouve chez les virus à ARN avec toutes les propriétés des quasi-espèces.*
- *Les recherches sur les mécanismes et les conséquences d'une forte diversité génétique chez les virus à ADN sont encore mal compris.*

Le virus Ostreid Herpèsvirus 1

Le virus **Ostreid Herpèsvirus 1** (**OsHV-1**) est un virus à ADN de l'ordre des *Herpesvirales* qui fait partie de la famille des Malacoherpesviridae et de la seule espèce du genre Ostreavirus (Davison et al., 2009). Ce virus est reconnu comme étant impliqué dans des épisodes de mortalité sporadiques chez plusieurs mollusques bivalves (Moss et al., 2007; Renault et al., 2001).

La découverte de l'OsHV-1

Des virus herpetiques semblent avoir été observé pour la première fois lors d'expérience de stress thermique avec l'espèce d'huître *C. virginica* (Farley et al., 1972). Des cellules contenaient des corps d'inclusion intranucléaire comparables à ceux associés aux infections aux Herpèsvirus chez d'autres animaux. Ce type de virus a ensuite été suspecté d'être la cause d'épisodes de mortalité importantes observées dans des écloseries de larves de *Crassostrea gigas* lors de l'année 1991 en Nouvelle-Zélande et en France (Hine et al., 1992; Nicolas et al., 1992). En 2005, la première version du génome de l'OsHV-1 chez l'huître a été publiée, le seul herpesvirus connu pour infecter des mollusques à cette époque (Davison, 2005). On le retrouve maintenant à l'échelle mondiale bien que toutes les zones de productions ne sont pas concernées par des épidémies massives de mortalité (**Fig.7. p20**). Il a été détecté en Europe : en France (Renault et al., 1994a, 1994b), en Espagne (Barbosa Solomieu et al., 2015; da Silva et

al., 2008; Elandaloussi et al., 2009), en Grèce et Irlande du nord (da Silva et al., 2008), en Italie (Dundon et al., 2011), en grande Bretagne ("Scientific Opinion on the increased mortality events in Pacific oysters, Crassostrea gigas," 2010), aux Pays-Bas (Gittenberger et al., 2014), en Suède et Norvège (Mortensen et al., 2016). Sur le continent Américains : aux USA (Burge et al., 2006; Friedman et al., 2005), au Mexique (Vásquez-Yeomans et al., 2010), au Brésil (Mello et al., 2017), et en Argentine (Barbieri et al., 2019). En Asie : au Japon (Shimahara et al., 2012), en Chine (Renault et al., 2012), en Corée du Sud (Hwang et al., 2013). Enfin il a été détecté en Océanie : en Nouvelle-Zélande (Hine et al., 1992), en Australie (Paul-Pont et al., 2013a).

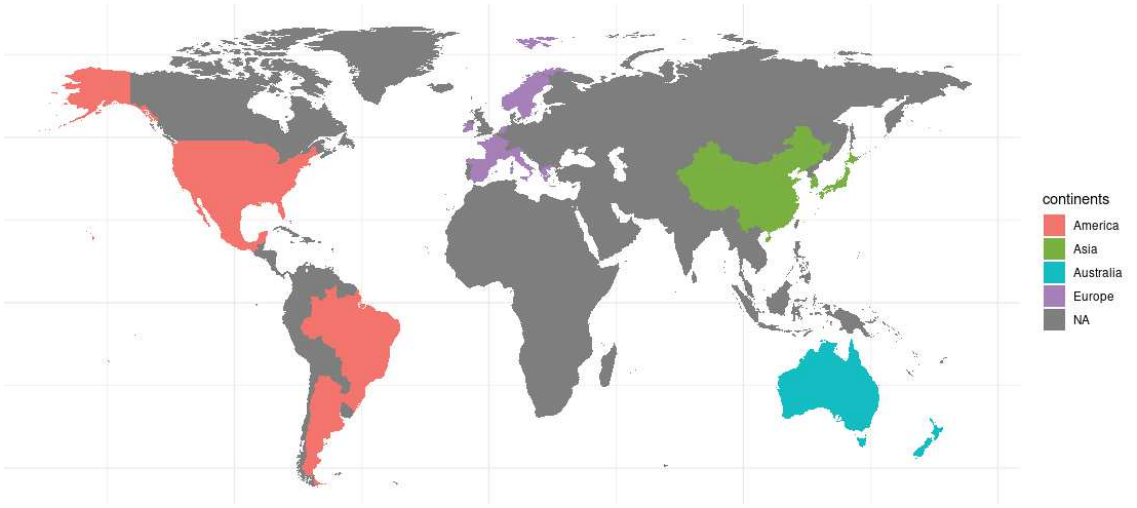


Figure 7. **Répartition mondiale des lieux où l'OsHV-1 a été détecté**.
Les 17 pays associés à la détection de l'ADN d'herpesvirus chez des mollusques ont été colorier en function du continents où ils sont localisés avec le continent Américain (en rouge), l'Europe (en violet), l'Asie (en vert) et l'Océanie (en bleu).

La structure de l'OsHV-1

Comme les autres virus de l'ordre des *Herpesvirales*, le virus OsHV-1 est composé d'un double brin d'ADN, une capside et d'une enveloppe (**Fig.8. p21**). La capside de l'OsHV-1 est similaire en apparence générale à celles des autres Herpèsvirus avec un diamètre estimé entre 70 nm à 140 nm selon le traitement utilisé (Burioli et al., 2018; Chongming et al., 2002; Davison, 2005; Renault 2007). La reconstruction de la capside de l'OsHV-1 a révélé une structure icosaédrique constituée de 162 capsomères de 150 hexons et 12 pentons (Davison, 2005). Dans cette étude il est fait mention de plusieurs caractéristiques supplémentaires qui sont unique à l'OSHV-1. Premièrement une double couche de densité en périphéries de la capside qui diffère des capsides d'autre Herpèsvirus, qui ont tendance à avoir des périphéries dentelées générées par des saillies externes vues en vue latérale. Deuxièmement, des grands trous aux sommets (mis en évidence par les flèches sur la Figure 8, marquant l'absence de pentons de ces sites.

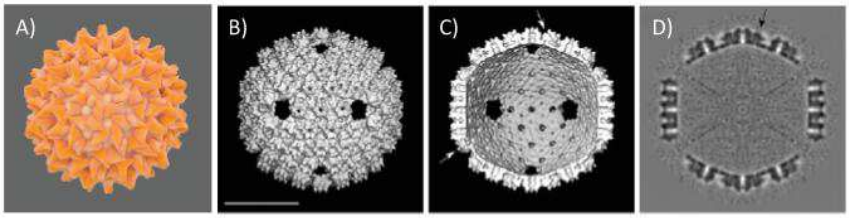


Figure 8. **Illustration de la structure de l'OsHV-1**.
(**A**) Dessin illustrant le virus OsHV-1 dans la thèse. Reconstruction au microscope cryoélectronique de la capside de l'OsHV-1 vue le long d'un double axe de symétrie: surface extérieure (**B**), surface intérieure (**C**) et section centrale (**D**). Les flèches indiquent la densité correspondant aux triplex. Barre, 50 nm. Les images de microscopie électronique sont issues (Davison, 2005)

Le génome de l'OsHV-1

Le séquençage du génome de référence OsHV-1 a été réalisé *via* une méthode de clonage. Le génome du virus OsHV-1 (GenBank AY509253) est formé d'une molécule d'ADN double brins linéaire d'une taille avoisinant 200 kbp avec un 38,7 % de GC. Comme pour d'autres herpèsvirus, le génome de l'OsHV-1 n'est pas un simple ADN double brins mais il est structuré de régions répétées inversées et de parties uniques non dupliquées. Le génome se compose de deux régions uniques, désignées par unique long (UL) en rouge, et unique short (US) en bleu, et trois régions de répétition, appelées répétition longue (Rl) en jaune, répéter court (RS) en violet et région X (X) en orange. Les deux premières régions de répétition ont chacune des copies présentes à un locus terminal (TRL, TRS) et un locus interne (IRL, IRS) dans le génome (**Fig.9. p21**). Au moment de sa découverte la structure du génome a été décrite comme : TRL – UL – IRL – X – IRS – US – TRS correspondant à la forme prédominante. Dans son article de 2005, Davison et coll. indiquent que des formes minoritaires avec 2 régions X ou sans X étaient retrouvées. Par la suite, les technologies de séquençage à haut débit ont permit de mieux caractériser la structure du génome en montrant un profondeur de séquençage deux fois supérieur au région unique indiquant une structure TRL – UL – IRL – X – IRS – US – TRS – X (Burioli et al., 2017). Même si le placement des deux copies des régions X n'est pas entièrement compris, la structure ci dessous est celle actuellement admise comme correspondant à l'organisation génomique du génome de l'OsHV-1.

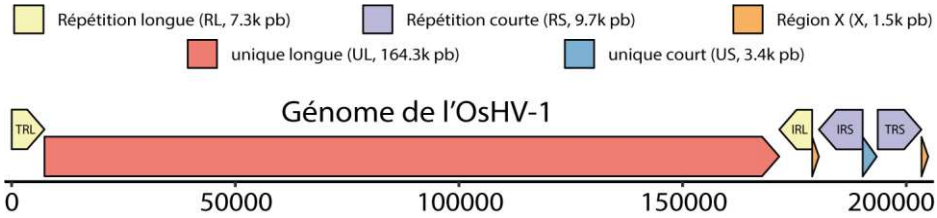


Figure 9. **Schéma illustrant la structure du genome actuelle de l'OsHV-1**.
Chacune des regions sont représentés avec des couleurs. Entre parenthèse la denomination anglaise ainsi que l'ordre de grandeur de en paire de base (pb) est décrit.

L'Hôte et le tropisme de l'OsHV-1

A ce jour le virus OsHV-1 semble avoir été détecté dans plus de 20 espèces de mollusques (**Annexe .1 p129**). À ce jour, on ne sait pas si les autres espèces ne sont que des porteurs passifs d'OsHV-1 et / ou sont susceptibles d'être affectées par des infections virales entraînant des événements de mortalité élevés (Alfaro et al., 2018). Il est associé à des épisodes de mortalité chez au moins 7 de ces espèces (Alfaro et al., 2018) dont l'huître plate européenne (*Ostrea edulis*) (Renault et al., 2000), la palourde de Manille (Ruditapes philippinarum) (Renault et al., 2001), le grand pétoncle (*Pecten maximus*)(Arzul et al., 2001b), l'huître portugaise (*Crassostrea angulata*) (Batista et al., 2014), la palourde (*Scapharca broughtonii*) le pétoncle chinois (*Chlamys farrieri*) (Bai et al., 2014; Xia et al., 2015) et l'huître du pacifique (*Crassostrea gigas*). Dans le cas du virus OsHV-1 le tropisme cellulaire n'est pas totalement compris mais plusieurs résultats ont montré que les hémocytes, des cellules circulantes au sein de l'hémolymphe siège de la réponse immunitaire, pourraient être le lieux de réplication pour le virus (Hine and Thorne, 1997; Martenot et al., 2017; Morga et al., 2017; Renault et al., 2011).

La diversité génétique de l'OsHV-1

Alors que l'analyse de la diversité génétique a pendant longtemps été réalisée à l'aide de courts fragments (**Encart 1 p24**) il existe maintenant 7 génomes apparentés à l'OsHV-1 qui ont été caractérisés grâce aux progrès des méthodes de séquençage (**Fig.10. p23**). Quatre d'entre eux sont des génomes d'OsHV-1 issus d'extractions de l'huître *C. gigas* : le génome de OsHV-1 de référence (AY509253) (Davison, 2005) , et des génomes d'un autre génotype : les **micro-variants** (µVar). Trois génomes de µVar ont été assemblés : les OsHV-1 µVar A/B (µVar A KY242785.1 et µVar B KY271630.1) (Burioli et al., 2017) et dernièrement l'OsHV-1-PT (MG561751.2) (Abbadi et al., 2018). Par rapport au génome de référence la présence de 5 grandes délétions ont été caractérisées dans des échantillons apparentés aux µVar. Une délétion de 605 bp dans les ORFs 36-37 (Renault et al., 2012) et une partie de l'ORF 38 ; une délétion de 1385 pb dans l'ORF 11 ; une délétion de 599 bp impactant l'ORF 48 ; une délétion de 3549 bp qui se trouve dans une région codant pour les ORFs 62-63 (Martenot et al., 2013) et enfin une délétion de 408 pb (175020-175427) qui impactent l'ORF 114 (Martenot et al., 2013). Trois insertion on également observe chez le µVar : une insertion de 86 pb dans la répétition inversée TRL / IRL, qui était présente deux fois (1655-181357), et une troisième insertion de *2671 pb située dans l'UL (60740)* (Burioli et al., 2017). Trois autres virus apparentés à OsHV-1 ont également étaient séquencés à partir d'autres espèces bivalves : l'OsHV-1-ZK118 (MF509813) associé à *Chlamys farreri* (Bai et al., 2018) , l'OsHV-1-SB associé à *Scapharca broughtonii* (KY271630) (Xia et al., 2015) et enfin HaHV-1 (KU096999) associé aux mollusques de type Haliotis (Savin et al., 2010).

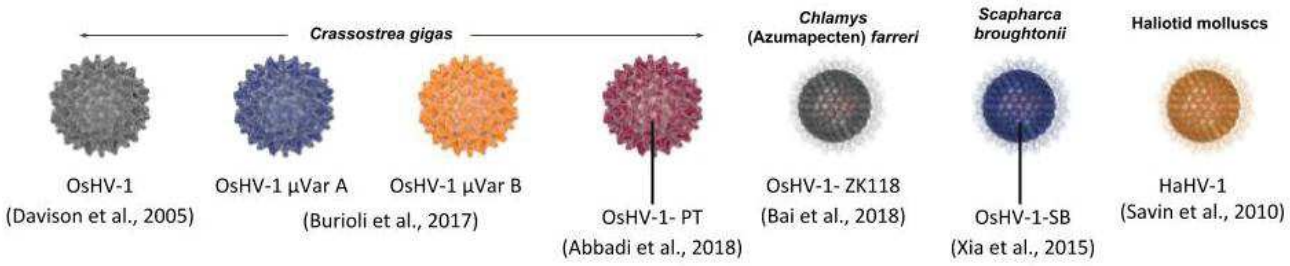


Figure 10. **Génomes complets de l'OsHV disponibles dans les bases de données**.
L'hôte pour lequel le virus a été assemblé est indiqué en haut. Les quatre virus à gauche sont issus de *C .gigas*.

Parallèlement à ces analyses génomiques, des analyses de transcriptomique ont permis d'observer des variants du génotype OsHV-1 dans des données de transcriptomique issues de de *Crassostrea gigas* (Rosani and Venier, 2017). Les premières analyses de SNPs dans des transcrits d'OsHV-1 avaient identifié 664 positions variables en utilisant le génome de référence, 86 % d'entre elles étaient retrouvées dans des gènes et 75% entraînant des substitutions d'acides aminés (Rosani and Venier, 2017). Les changements d'acides aminés dans les protéines virales sont connus pour faire évoluer la nature pathogène de certains virus. Si la plupart des SNP concernaient des différences entre les virus échantillonnés et l'OsHV-1 de référence, 20 SNPs présentaient une fréquence inférieure à 95% suggérant la présence de plus d'un variant viral dans et entre les échantillons de séquençage à ARN et donc que l'OsHV-1 pourrait former des populations hétérogènes.

Résumé

- *Le virus Ostreid Herpèsvirus 1 (OsHV-1) est associé à des épisodes de mortalité à l'échelle mondiale.*
- *L'OsHV-1 est un herpèsvirus composé d'un double brin d'ADN, d'une capside et d'une enveloppe.*
- *Ce virus présente une structure génomique particulière des herpèsvirus avec des régions répétées inverses*
- *D'autres virus apparentés à l'OsHV-1 ont été séquencés chez d'autres mollusques.*
- *La diversité génétique des populations virales d'OsHV-1 est encore très mal connue.*
- *Les génotypes µVar ont des modifications aux niveaux des ORFs avec de larges délétions et insertions.*

Encadré 1: Focus sur l'analyse de la diversité génétique du virus avant l'utilisation de génome complet

Les premières observations de diversité génétique du virus datent de la découverte du variant OsHV-1 Var, identifié chez Pecten maximus et Ruditapes philippnarum dans des écloseries françaises lors d'épidémies de mortalité en 1997 (Arzul et al., 2001). A cette époque et jusque dans les années 2010, l'analyse et les relations phylogénétique de la diversité génétique du virus OsHV-1 étaient caractérisées en séquençant des amplicons de PCR de la région C (Fig encadré 1). Ce fragment a été caractérisé sur une zone relativement restreinte du génome (>1000bp) . D'autres petites variations de ce fragment ont été observées (Friedman et al., 2005; Arzul et al., 2001c; Moss et al., 2007; Batista et al., 2005; Batista et al., 2007). En 2008, le génotype appelé OsHV-1 µVar a été observé à l'aide de cette technique lors d'épisodes de mortalité plus importants. Il est caractérisé par une délétion de 12 nucléotides consécutifs suivis d'une délétion d'une adénine dans la région C (Commission Regulation (EU) No 175/2010, 2010). Par la suite, d'autres variations associées à ce génotype ont pu être observées (Martenot et al., 2011; Martenot et al., 2013). Dix-huit ans après la mise au point de la PCR (Renault & lipart 1998) c'est près de 91 séquences qui étaient utilisées pour comprendre les relations entre les isolats séquencés (Burioli 2016). L'étude des relations phylogénétiques semblait montrer une diversité génétique importante et la coexistence de plusieurs génotypes au sein d'un même hôte a même était suggéré (Burioli 2016).

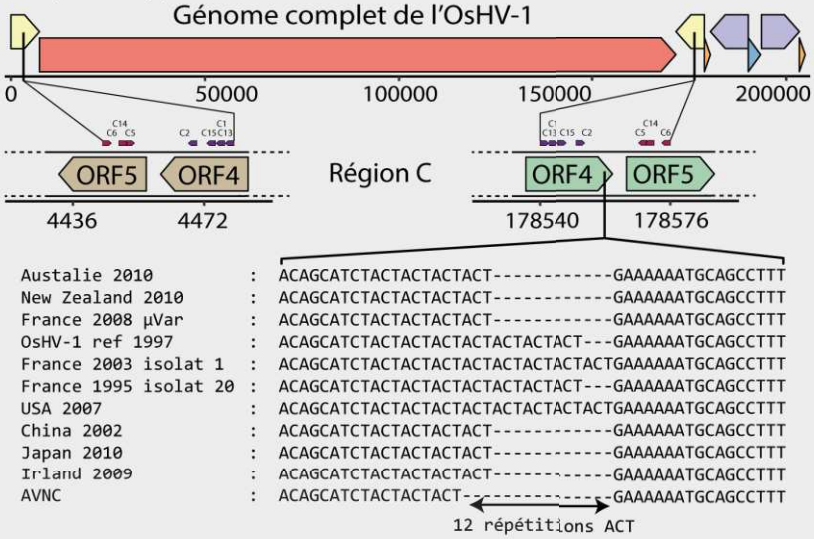


Figure encart 1. **Représentation de la Région C sur le génome de OsHV**. Représentation schématique du génome pour localiser la position génomique des amplicons séquencés et la région microsatellite constituée de la répétition ACT qui définit les micro variants. Illustration de la structure du génome de l'OsHV-1 avec la région C et ses amorces de PCR (C2 à C15). L'amplicon de PCR fait maximum 892 bp (C13-C6) et il se retrouve dans les régions répétées longues (en jaune). Enfin en bas, la figure est inspirée de la Figure 4 de l'étude de Jenkins et al. 2013 montrant l'alignement de la région microsatellite de 12 bp (dans la séquence C2 / C6) des séquences de différents variants : OsHV-1 de référence, OsHV-1 µVar, AVNV provenant de différentes zones.

La biologie et écologie de l'huître

L'**huître** *Crassostrea gigas* (Thunberg, 1793) (NCBI:txid29159) est considérée actuellement comme l'hôte principal du virus OsHV-1. C'est également l'espèce la plus sensible aux épidémies déclenchées par ce virus avec le déclenchement d'événements de mortalité récurrent. Cette huître doit initialement son nom taxonomique à sa morphologie avec "crass" qui signifie "épais" et "ostrea" signifiant "huître" en latin et du grec "gígās" pour "géant", on retrouve également *C. gigas* sous plusieurs noms communs comme l'huître creuse du Pacifique ou encore l'huître japonaise. Toutefois, on retrouve également cet organisme sous la dénomination controversée de *Magallana gigas* (Bayne et al., 2017; Salvi and Mariottini, 2016). C'est un organisme original pour l'étude des interactions hôtes et pathogènes du fait de son placement phylogénétique et de son mode de vie. En effet, l'huître *C. gigas* est un organisme invertébré marin qui appartient au mollusque de la classe Bivalvia, à la sous classe Pteriomorphia, et à la famille Ostreidae ; ce qui le place très loin des modèles vertébrés classiques. Ainsi, il n'y pas de cellules lymphocytaires chez *C. gigas* ce qui permet de s'affranchir de la réponse immunitaire médiée par ces cellules adaptatives et de mieux s'intéresser à des mécanismes ancestraux de l'immunité innée. L'huître japonaise est également intéressante dans d'autres domaines d'étude comme la génétique et l'écologie marine. Enfin, cet animal present un intérêt économique puisque c'est de loin l'espèce d'huître la plus vendue/produite au monde (Botta et al., 2020). En effet, son élevage, sa production et la vente d'huîtres est des activités dans plus de 40 pays, dont la France. Nous allons maintenant décrire (i) le mode de vie de l'huître de creuse *C. gigas* ; (ii) son cycle de vie et de reproduction; (iii) son anatomie et (iv) l'activité de culture des huîtres.

Le mode de vie de l'huître

C. gigas est un invertébré qui vit en majorité dans les eaux peu profondes des zones intertidales des baies, des lagunes et des estuaires. Cette espèce vit à l'état naturel fixée sur un substrat rocheux mais peut également être trouvée sur les coquilles d'autres animaux. Les larves se déposent souvent sur la coquille des adultes et de grandes masses d'huîtres peuvent se développer ensemble pour former des récifs d'huîtres. L'huître se nourrit en filtrant l'eau et absorbant des microalgues et du phytoplancton. L'huître a plusieurs prédateurs naturels comme par exemple l'huîtrier pie (ou Pie de mer), différentes espèces de crabes et de poissons (raies, brèmes, dorades) et des bigorneaux.

La reproduction et le cycle de vie de l'huître

Comme la plupart des bivalves, *C. gigas* présente un cycle de vie biphasique qui comprend des stades larvaires pélagiques qui nagent activement et des animaux adultes benthiques sédentaires. Lorsque les huîtres sont fixées, la maturité sexuelle est atteinte entre 3 et 12 mois, selon les conditions de croissance et l'individu. On parle alors d'huître au **stade juvénile** (animaux de plus d'un an et de moins de 18 mois). Les gonades vont alors grossir du fait de la prolifération des spermatozoïdes chez le mâle et des ovocytes matures pédonculés chez la femelle. En effet, l'huître est un animal hermaphrodite successif à tendance protandre (Buroker., 1982). Lorsque la température de l'eau est proche de 18 °C, les gamètes sont finalement libérés dans l'eau pour une fécondation externe (Mann, 1979). Chaque individu est

capable de rejeter des dizaines de millions de gamètes et l'huître se reproduit une ou plusieurs fois durant l'été, c'est à cette période que l'huître est laiteuse.

Une fois dans l'eau ces gamètes vont se disperser et lors de la fécondation externe l'union des gamètes mâles et femelles va entraîner une succession de mitose qui aboutiront à la formation d'une larve trochophore microscopique entre seize et vingt-quatre heures plus tard après fécondation. Après plusieurs étapes de développement, la larve trochophore va se transformer en larve véligère capable de nager. Celle-ci va passer par différents stades (larve D, umbo, oeillée) entre 2 et 14 jours après fécondation, jusqu'au stade pédivéligère entre 14 et 20 jours après fécondation. Le développement d'un pied va alors apparaître, lorsque la larve atteint une taille comprise entre 300 et 380 µm, et permettre à l'huître de se fixer sur un substrat. Après quelques semaines de vie planctonique, la larve va se fixer sur un substrat et se métamorphoser progressivement. La larve va alors devenir une « huître vraie » d'un point de vue morphologique et entame sa phase benthique où elle restera fixée à son support le reste de sa vie. Après la transition de la phase planctonique à benthique où s'effectue une métamorphose on parle alors de **naissain** (animaux de plus de 30 jpf et de moins d'un an) (**Fig.11. p26**). Une fois leur première maturité sexuelle (entre 3 et 12 mois), les huîtres seront considérées comme étant juvéniles et par la suite on parle d'animaux adultes. Toutefois, le passage du stade naissain au stade juvénile et adulte est très influencé par les conditions environnementales et les points de temps peuvent varier.

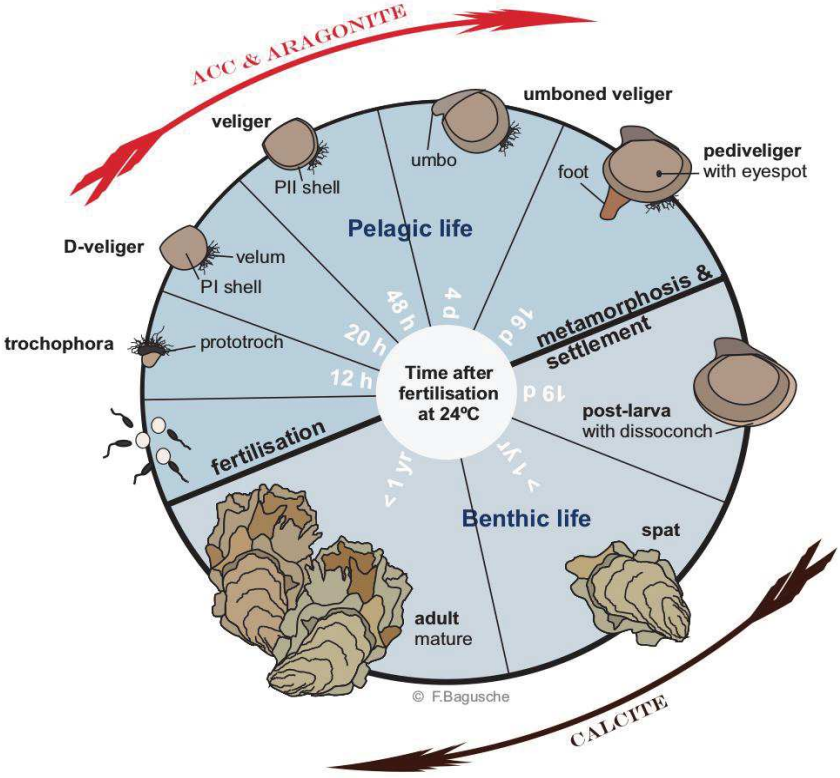


Figure 11. **Schéma représentant le cycle de vie de l'huître creuse**.
Les différents stades du cycle de vie sont représentés avec une dichotomie entre les phases pélagique (haut) et benthique (bas). Figure extraite tirée de la thèse de (Bagusche, 2013).

L'anatomie de l'huître

D'un point de vue morphologique, l'huître *C.gigas* est protégée par une coquille de carbonate de calcium formée de 2 valves asymétriques, l'une plate (supérieure) et l'autre creuse (inférieure) qui peuvent atteindre une taille entre 20 et 30 cm (Miossec et al., 2009). À l'intérieur de cette coquille se trouve un corps mou, composé par plusieurs organes (**Fig.12. p27**). Le manteau (*Mantle*) qui correspond à un tissu conjonctif qui contient des muscles, des nerfs et des vaisseaux hémolymphatiques. Un muscle adducteur (*adductor muscle*) reliant les deux valves qui permet à l'huître de s'ouvrir et de se fermer hermétiquement pour se protéger des prédateurs ou préserver l'eau à l'intérieur lorsqu'elle n'est pas immergée. Comme d'autres animaux marins, l'huître respire et se nourrit à l'aide de branchies (*gills*). Les particules continuent jusqu'à l'estomac (*stomach*) qui forme un long cæcum avec une évagination tubulaire et l'intestin. La présence de gonade qui recouvre la glande digestive assure les fonctions de reproduction. C'est un organe non permanent dont la structure varie au cours du cycle de reproduction (Naimi, 2009). L'huître a un système nerveux réduit avec deux ganglions et ces animaux ne possèdent pas de système nerveux central. L'huître possède également deux reins, des artères, des veines et un cœur constitué par deux oreillettes et un ventricule, qui se situe juste au-dessus du muscle adducteur, permettant faire circuler l'hémolymphe. L'hémolymphe comprend des cellules, les haemocytes et des micro-organismes (elle n'est pas stérile), de l'eau et des sels inorganiques dissous, des protéines, et de l'oxygène dissous dans l'eau. Ce système circulatoire intervient dans les fonctions de transport des nutriments, de l'oxygène, des déchets du métabolisme et dans la défense immunitaire, il est dit semi-ouvert du fait qu'il peut s'infiltrer dans l'ensemble des tissus de l'huître. Enfin les huîtres sont des animaux poïkilothermes, c'est-à-dire ayant une température corporelle qui varie avec celle de leur milieu, elles sont donc obligées de faire face à d'énormes changements dans les conditions environnementales biotiques (Lokmer and Mathias Wegner, 2015) et abiotiques.

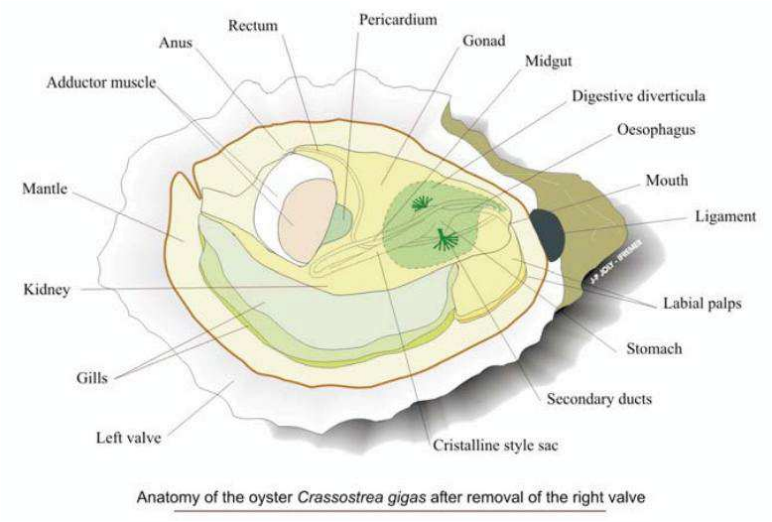


Figure 12. **Représentation schématique de l'anatomie de l'huître creuse** (valve inférieure).
Figure tirée de (Miossec et al., 2009).

L'ostréiculture

L'élevage des huîtres

L'huître joue un rôle culturel important dans lequel elle est considérée comme un produit de luxe (Buestel et al., 2009). Jusqu'au XIXe siècle, la vente des huîtres consistait principalement dans l'extraction des bancs naturels. De nos jours toute l'industrie conchylicole s'est développée et l'**ostréiculture** qui est l'activité d'élevage des huîtres s'est accrue. En ce qui concerne *C. gigas* ses origines sont en Asie de l'est, où elle a été cultivée pendant des siècles, elle a fait l'objet de plusieurs introductions notamment aux Etats-Unis dans les années 20 et en France en 1966. L'ostréiculture est basée sur le captage et l'élevage avant la vente des huîtres (**Fig.13. p29**).

- L'approvisionnement en naissains d'huîtres du Pacifique provient soit de captage naturel, soit d'écloserie. Dans le cas de **captage** naturel, les ostréiculteurs utilisent des collecteurs pour capter les larves après la ponte. Ce type de captage naturel est important dans le bassin de Marennes-Oléron grâce aux eaux estuariennes issues de l'embouchure de la Seudre, ou dans le bassin d'Arcachon. Autrement les naissains sont produits dans des écloseries en conditions contrôlées ou l'eau de mer utilisée y est filtrée et purifiée.

- Plusieurs techniques d'**élevage** existent selon la nature du sol, et le coefficient des marées. L'élevage en suspension sous tables d'élevage dû à l'absence de marée (en Méditerranée); L'élevage « à plat » : au sol émergent, en eau profonde ou surélevée. Les huîtres seront alors gardées dans des poches et seront retournées régulièrement par l'ostréiculteur afin que toutes les huîtres puissent croître dans de bonnes conditions et avec une forme régulière. Dans certain cas, les huîtres adultes sont affinées, par exemple, elles seront placées dans des bassins d'affinage (par exemple, elles peuvent être places dans des claires, où l'iode a tendance à se réduire) pour modifier la taille ou la couleur, ou les qualités gustatives de l'huître.

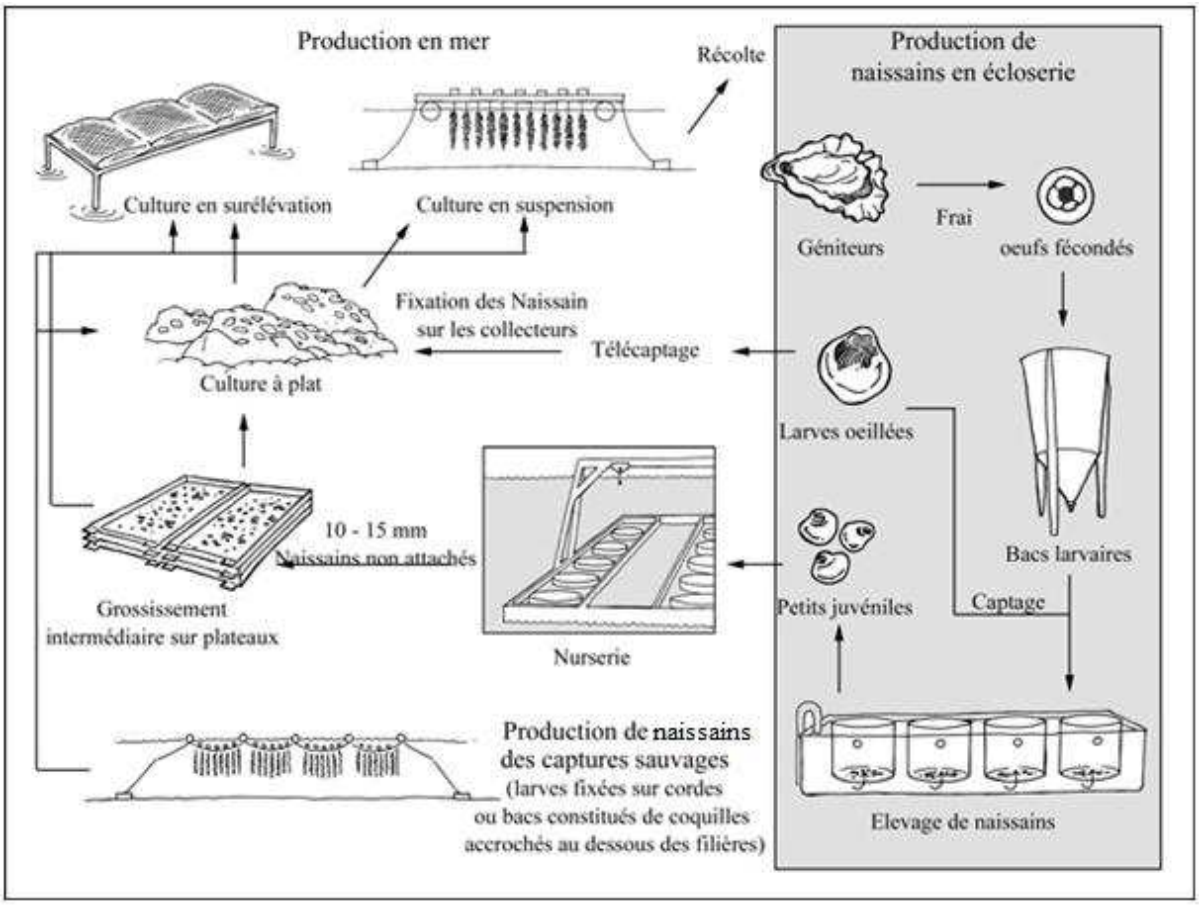


Figure 13. **Cycle de production de *Crassostrea gigas***.
Figure extraite de "Technical Guidance for the Development of the Growing Area Aspects of Bivalve Mollusc Sanitation Programmes" (FAO and WHO, 2018). Le terme "graine" de la figure original a été remplacé par le terme 'naissains' plus approprié.

L'impact économique

Comme pour l'aquaculture en général, la production ostréicole a augmenté rapidement au cours des dernières décennies (**Fig.14. p30**). Actuellement, les mollusques sont la deuxième plus grande catégorie de produits de la mer d'élevage en quantité et en valeur, représentant 21% de toute la production aquacole mondiale en poids en 2016 ("FAO, FAO Online Query Panels," 2018). Alors que l'ostréiculture est fortement dominée par l'Asie, qui représentait 95% de la production mondiale de l'ostréiculture en poids et 88% en valeur en 2016, la France fait partie des 6 plus grands producteurs d'huîtres dans le monde. Elle est le plus grand producteur ostréicole du monde jusque dans les années 1950 et reste le premier d'Europe (Botta et al., 2020; "FAO, FAO Online Query Panels," 2018).

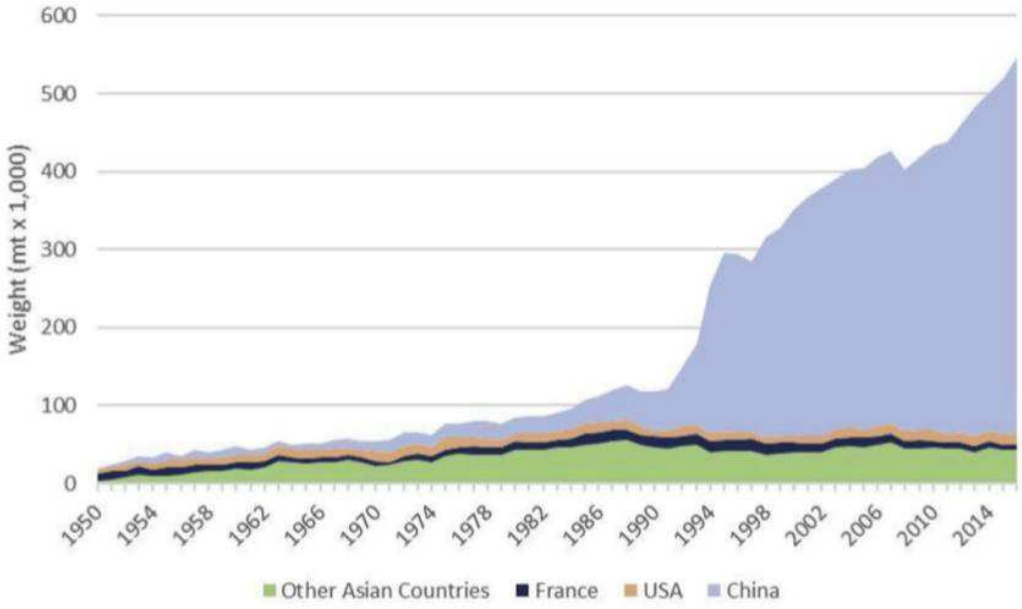


Figure 14. **Production aquacole ostréicole en termes de poids de production**.
Chaque couleur représente une région géographique. Figure tirée de (Botta et al., 2020) qui a été réalisée à partir des données de la FAO.

L'ostréiculture a toujours connu des épisodes de mortalité récurrente (Mori, 1979; Takeuchi et al., 1960). Plusieurs pathogènes sont connus pour être responsable de maladie chez *C. gigas* : des parasites protozoaires unicellulaires, des bactéries pathogènes et enfin des virus (Bower et al., 1994). L'une de ces pathologies est responsable de dommages économiques importants (estimés à 8 millions d'euros par an) à travers le monde : les épisodes de mortalité dus au **syndrôme de mortalité des huîtres du Pacifique** (*Pacific Oyster Mortality Syndrome*, POMS) (Paul-Pont et al., 2013b). Cette maladie a conduit, au fil du temps, à des problèmes d'approvisionnement en naissains dans la plupart des bassins conchylicoles. En effet, les deux sources de naissains, d'écloserie et de captage naturel, sont touchées par ces événements de mortalité (Bernard and Lapègue, 2012). Du fait des conséquences économiques dramatiques dans le secteur de l'ostréiculture, de nombreux projets de recherche ont émergé pour mieux caractériser la maladie des POMS.

Résumé

- *Crassostrea gigas est considérée actuellement comme l'hôte principal du virus OsHV-1.*
- *Cet animal est un mollusque invertébré qui vit fixé dans les eaux peu profondes et se nourrit en filtrant du phytoplancton.*
- *Les huîtres ont un cycle de vie biphasique, elles vont se métamorphoser lors de leur stade larvaire jusqu'à se fixer pour être des naissains et devenir adultes.*
- *L'ostréiculture est l'activité d'élevage des huîtres, pour s'approvisionner les ostréiculteurs font du captage ou achètent des naissains directement dans des écloseries.*
- *L'industrie ostréicole produit plusieurs millions d'euros de bénéfice, mais elle est menacée par des épisodes de mortalité récurrents.*

Le syndrôme de mortalité des huîtres du Pacifique (Pacific Oyster Mortality Syndrome-POMS)

Les mortalités de masse touchant *C. gigas* due à une quantité massive de l'OsHV-1 ont été défini en 2013 par Paul Pont et coll. comme le **syndrôme de mortalité des huîtres du Pacifique**. De nombreux travaux tendrai a montré que cette pathologie est d'**origine multifactorielle**. Ces épisodes de mortalité semblent survenir dans toutes les classes d'âge et de taille (Alfaro et al., 2018; Azéma et al., 2017; Hick et al., 2018; Petton et al., 2015). Cette maladie n'est pas zoonotique, c'est à dire qu'elle n'est pas transmissible à l'homme.

La cause

Ce syndrôme est causé par un ensemble de facteurs de l'hôte, de l'environnement et des agents pathogènes. Nous pouvons la conceptualiser à l'aide du concept de «**triade épidémiologique**» (Snieszko, 1974) qui fait intervenir des facteurs propres à l'hôte, à l'environnement, et aux agents pathogènes (**Fig.15. p32**). Sans l'un des trois composants principaux, l'hôte peut rester en bonne santé et asymptomatique. En l'absence d'agents pathogènes, des conditions environnementales non idéales peuvent provoquer un stress chez l'hôte. Dans des conditions environnementales favorables, les infections par des agents pathogènes peuvent entraîner des infections subcliniques ou asymptomatiques, où l'hôte est porteur du pathogène mais ne présente pas de symptômes. Enfin, les environnements sans hôte approprié ne peuvent pas entraîner d'hôte malade, mais ces conditions posent un environnement potentiellement infectieux si des hôtes appropriés deviennent disponibles.

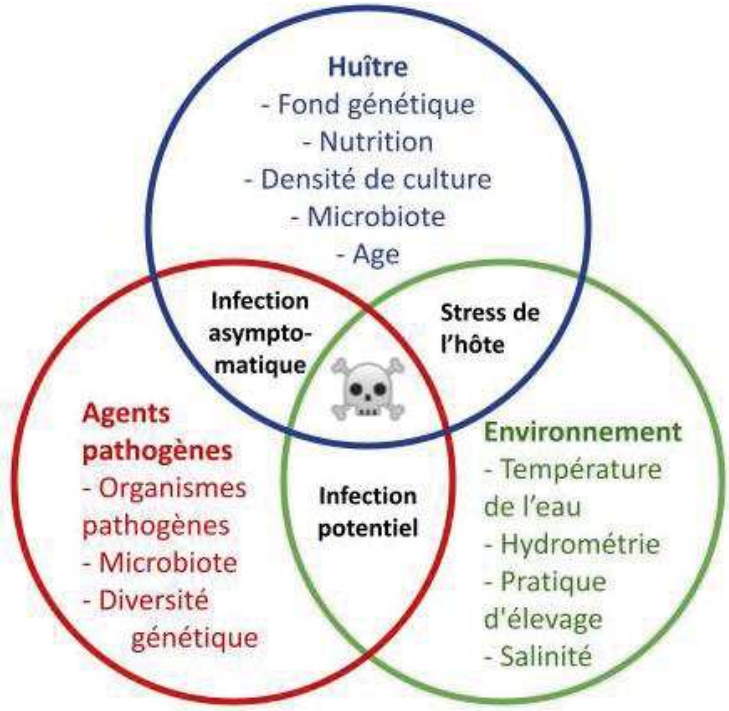


Figure 15. **Triade épidémiologique adaptée pour l'huître**.
Chaque ensemble représente l'une des composantes d'une maladie multifactorielle. Figure inspirée de (Gallagher, 2020) qui a illustré le concept de «la triade épidémiologique» de Sniezko en 1974.

Nous savons maintenant que le virus OsHV-1 est l'agent qui déclenche ces événements de mortalité (**Encart 2 p35**). En effet, une étude holistique récente a permis de révéler la succession des événements initiant la pathogenèse (de Lorgeril et al., 2018). Lorsque la température de l'eau dépasse les 16°C mais est inférieure à 24 °C (Pernet et al., 2012) et que les facteurs de l'hôte le permettent, pour une raison encore mal connue, le virus est capable d'infecter les huîtres. Lorsque l'huître possède un phénotype sensible, elle n'est pas en mesure de contrôler la prolifération du virus OsHV-1 (de lorgeril 2018). Cette prolifération va altérer l'efficacité du système immunitaire de son hôte. En l'absence de régulations sur les populations bactériennes il va y avoir une dysbiose qui permettra à des bactéries opportunistes de proliférer. Cette prolifération va entraîner la mort de l'animal (**Fig.16. p33**). Malgré cette avancée significative dans le domaine, de nombreuses questions sur le rôle de chacun des facteurs illustrés dans la triade épidémiologique (**Fig.15. p32**) sont toujours en cours d'investigation.

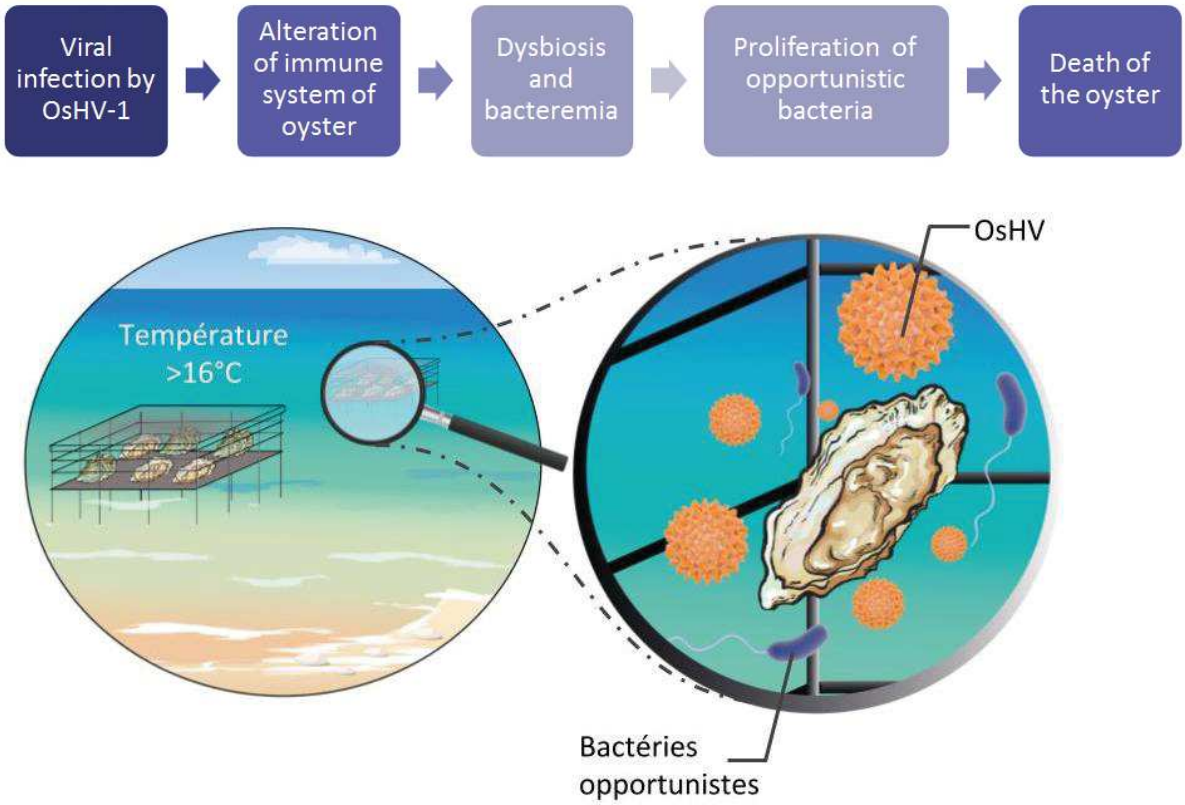


Figure 16. **Succession des événements entraînant la mortalité des huîtres lors des POMS**.
En haut, description des événements entraînant la mort des huîtres, en bas une illustration de ces épisodes de mortalité sur le terrain une fois la température de l'eau de mer supérieure à 16°C cela entraîne la prolifération du virus OsHV-1 et des bactéries opportunistes.

L'histoire des episodes de mortalité des huîtres

La première description scientifique d'un épisode de mortalité élevé d'huîtres associé à l'observation du virus OsHV-1 remonte aux années 1990 (Hine et al., 1992; Nicolas et al., 1992). Depuis de tels événements ont affecté les stocks d'élevage d'huîtres de manière récurrente à l'échelle mondiale. Dans un premier temps cette pathologie a été dénommée par le terme de "mortalité estivale" car elle a eu lieu principalement dans les périodes de printemps et d'été. Depuis les années 2008, un nouveau schéma de mortalité massive des huîtres du Pacifique a eu lieu en France et dans d'autres pays (**Fig.17. p34**). Ce sont ces épisodes qui sont nommés POMS (Paul-Pont et al., 2013b), lorsqu'ils surviennent ces événements de surmortalité se distinguent à la fois par leur généralisation, leur intensité et la détection récurrente d'un nouveau variant du virus OsHV-1, appelé OsHV-1 µVar (Segarra et al., 2010).

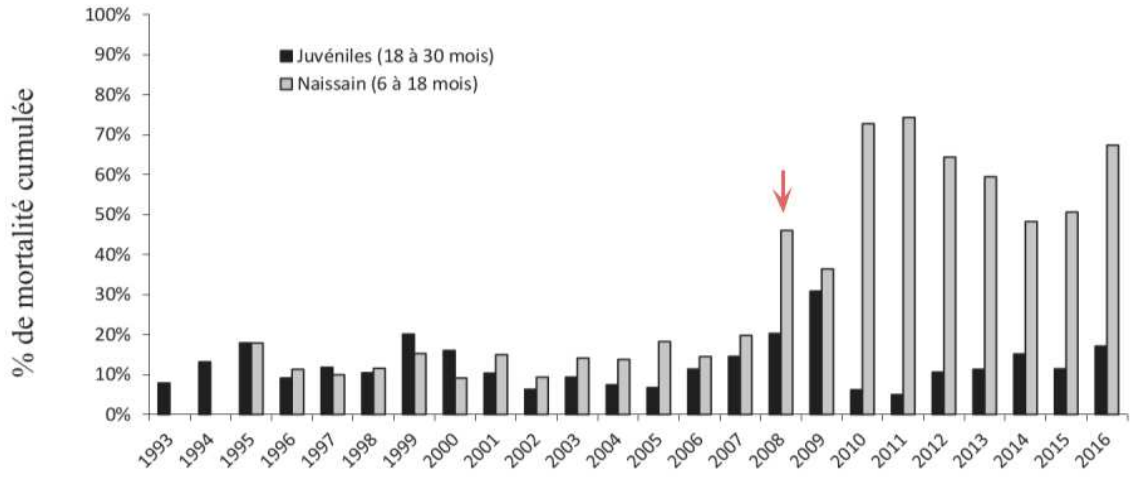


Figure 17. **Impact de mortalité cumulé sur l'élevage des huîtres**.
Pourcentage de mortalité cumulée chez les juvéniles et les naissains d'huîtres de C. gigas entre 1993 et 2016. La flèche en rouge met en évidence l'année 2008 avec une hausse des épisodes de mortalité. Source Ifremer RESCO, 2016

La manifestation de la maladie

Les signes cliniques de la maladie sont difficiles à observer sur le terrain. Chez les larves infectées les signes cliniques constatés sont une réduction des activités de la nage et de l'alimentation (Burge and Friedman, 2011; Hine et al., 1992; Nicolas et al., 1992). D'un point de vue histologique, des auteurs ont signalé des lésions tissulaires avec différents degrés dans les épithéliums. L'infection virale entraîne des lésions infectieuses du naissain ou des larves observées dans les tissus conjonctifs dans lesquels les cellules de type fibroblastique présentent des noyaux élargis avec de la chromatine périnucléaire (Hine et al., 1992; Hine and Thorne, 1997; Martenot et al., 2017; Nicolas et al., 1992; Renault et al., 1994a, 1994b). Une architecture tissulaire anormale et des changements cellulaires y compris la métaplasie cuboïdale, la dilatation des tubules et une diapédèse a également été notée au bord du manteau de certaines huîtres (Burge et al., 2006). Une hémocytose sévère a été observée dans le tissu conjonctif du manteau et à la base des filaments branchiaux, avec un degré d'invasion marqué (Burioli et al., 2018). Celle-ci a été notée à plusieurs reprises comme étant associée au tube digestif, bien qu'aucune pathologie n'ait été notée dans le tube digestif sus-jacent (Keeling et al., 2014). Enfin une érosion branchiale a été observée sur et dans des individus lors d'événements de mortalité associés à l'OsHV-1 au Mexique (Vásquez-Yeomans et al., 2010).

Encadré 2: Mise en évidence du virus OsHV-1 comme l'agent prépondérant des POMS

Lorsque les premières mortalités estivales ce sont déclarées, plusieurs micro-organismes ont été suspectés d'être responsable du phénomène. En plus des facteurs abiotiques comme la température, le virus OsHV-1 (Hine et al., 1992; Nicolas et al .,1992) et des bactéries du genre Vibrio (Paillard et al., 2004) ont rapidement été identifiés comme suspects. Par la suite une accumulation d'études ont a confirmé que le virus était un suspect crédible dans la pathologie:

a. La première infection expérimentale sur des larves d'huître "axénique", avec des le développement du premier filtrats viral (Le Deuff et al., 1994)
b. Les premières analyses PCR conventionnelles ont renforcé l'hypothèse (Burge et al., 2006; Friedman et al., 2005; Renault et al., 2000)
c. L'analyse PCR en temps réel (qPCR) en dégageant une corrélation suite à l'étude de la variance de la charge virale avec des animaux sensibles et résistants (Sauvage et al., 2009).
d. Suivi épidémiologique sur 9 ans qui permet de mettre en évidence que OsHV est l'agent causal (Garcia et al,. 2011)

L'apparition des POMS en 2008 qui ont affecté les naissains et juvéniles en parallèle de la détection du nouveau variant du virus OsHV-1, appelé OsHV-1 µVar (Segarra et al., 2010) a encore renforcé cette hypothèse. Par la suite entre 2010-2015 des études ont montré des résultats contradictoire avec l'observation d'une souche de Vibrio pathogénique pour les huîtres (Roux et al., 2009), son invalidation (Petton et al., 2015), la dynamique du microbiote de l'hémolymphe des huîtres du Pacifique (Crassostrea gigas) (Lokmer and Wegner, 2014), et enfin une réversion du phénotype de sensibilité par du polyI:C par T.Green et C.Montagnani (2013). C'est finalement la démonstration de la succession des événements initiant la pathogenèse qui confirmera le rôle prépondérant du virus OsHV-1 dans les POMS (de Lorgeril et al., 2018).

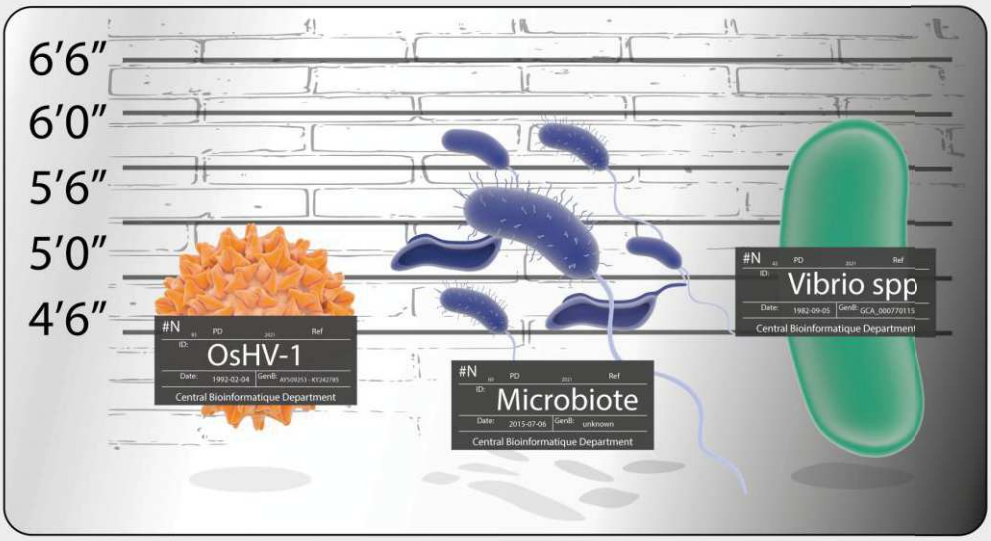


Figure encart 2. **Représentation schématique des microorganismes suspectés d'être les agents responsables de la maladie**. En abscisse la taille en pouce des suspects est représentée.

Les traitements

A ce jour, il n'existe pas de traitement applicable dans le milieu naturel lorsque survient le syndrôme de mortalité des huîtres du Pacifique. Toutefois du fait de la forte composante du fond génétique de l'hôte, plusieurs programmes de sélection pour la résistance des huîtres aux événements de mortalité massive ont été menés en France. Ces programmes ont permis une

amélioration importante du phénotype de résistance des huîtres à l'infection par OsHV-1. La survie des huîtres est passée de 34,5% après une génération de sélection à 69% sur quatre générations comparativement à 7,3% du groupe témoin (Degremont et al., 2015). Des études de transcriptomique ont montré que ces familles d'huîtres résistantes présentent des variations dans l'expression des gènes immunitaires (de Lorgeril et al., 2020, 2018; Renault et al., 2011; Segarra et al., 2014). La question de savoir si le fonds génétique de l'hôte est le reflet d'une évolution par rapport au virus n'est pas encore connue.

Résumé

- *Le syndrôme de mortalité des huîtres du Pacifique est une maladie d'origine multifactorielle touchant l'huître Crassostrea gigas.*
- *C'est une maladie multifactorielle causée par des facteurs de l'hôte, des agents pathogènes, et des facteurs environnementaux.*
- *Le schéma des épisodes de mortalité vient d'être déterminé et lorsque la température de l'eau augmente, le virus OsHV-1 µVar initie une réplication intense qui permet à des bactéries opportunistes d'entraîner une septicémie fatale.*
- *La pathologie entraîne des signes cliniques chez l'huître et des dégâts tissulaires.*
- *Il n'existe pas de traitement à ce jour, mais il existe des huîtres sélectionnées pour être résistantes à la maladie.*

Conclusion sur l'introduction générale

L'introduction générale de ce manuscrit a permis de montrer plusieurs points. (i) La compréhension de la diversité génétique peut se faire à l'aide de cadre conceptuel de la biologie évolutive comme la théorie synthétique de l'évolution. (ii) Dans le cadre des maladies infectieuses, l'interaction entre un hôte et un agent pathogène sont capables d'engager un processus de coévolution. Ces processus possèdent des propriétés spécifiques pouvant être observées à travers les signatures génomiques dans le génome des protagonistes. (iii) Les stratégies évolutives sont différentes selon le type d'organisme et ses contraintes et leurs impacts peuvent être très importants. L'état des connaissances sur l'importance de la diversité génétique comme déterminant de la virulence est encore mal compris chez les virus à ADN, et plus particulièrement les Herpèsvirus, contrairement aux virus à ARN. (iv) L'Herpèsvirus OsHV-1 est un virus à ADN associé à plusieurs événements de mortalité chez différentes espèces de bivalves à travers le monde et sa diversité génétique n'est pas suffisamment caractérisée. (v) Les huîtres sont les hôtes les mieux étudiés de ce virus et présentent un intérêt économique et scientifique important. (vi) L'Herpèsvirus OsHV-1 est le protagoniste initial du processus de pathogenèse du syndrôme de mortalité des huîtres juvéniles responsable de mortalité importante à travers le monde.

L'objectifs de la thèse

Jusqu'à présent, trop peu d'attention a été accordée au rôle que pourrait jouer la diversité génétique du virus OsHV-1 dans la mortalité massive qui affecte les huîtres juvéniles. Dans ce contexte, l'objectif de ce projet de thèse est de mieux caractériser la diversité génétique du virus OsHV-1 et commencer à évaluer son rôle comme déterminant de la virulence dans les événements de mortalité qui affectent l'huître *C.gigas*.

Pour atteindre ce but, nous allons dans un premier temps nous intéresser à la diversité virale suite à des expériences d'infections expérimentales en mésocosme (IEM) effectuées sur des familles d'huîtres ayant des phénotypes contrastés de sensibilités aux POMS. Au cours de ces expériences nous avons constaté que la sensibilité de certaines familles variait en fonction de l'origine géographique dont est issue la maladie. Nous avons utilisé les données de dual transcriptomique (c'est à dire la quantification des transcriptions de l'ARN cellulaire de l'hôte et son pathogène) à notre disposition pour étudier la diversité génétique du virus en fonction des environnements et dans les différentes familles. Ces résultats nous ont permis de montrer que les populations virales issues des deux environnements étaient différentes. Dans un second temps, nous nous sommes attachés à décrire la diversité virale de façon plus approfondie (au niveau du génome entier) dans les trois principales zones ostréicoles françaises. Ces résultats nous ont permis de mieux comprendre la source et la dispersion des populations du virus OsHV-1 en France.

Chapitre II Contribution de la diversité génomique virale à la sensibilité des huîtres dans le syndrôme de mortalité des huîtres du Pacifique

Le contexte des recherches

Comme nous l'avons vu, le rôle des variations génétiques des Herpèsvirus et leurs conséquences biologiques restent une question ouverte aujourd'hui. Dans le cas du virus OsHV-1, très peu d'études se sont intéressées à ce sujet. Lorsqu'en 2008 un nouveau génotype, l'OsHV-1 µVar, a été caractérisé simultanément aux événements de surmortalité massive apparus en France, cela a mis en évidence un manque de connaissance sur la caractérisation de la diversité génétique du virus (Segarra et al., 2010). Ainsi, des modifications génétiques du virus OsHV-1 µVar ont été caractérisés avec 5 délétions et 3 insertions d'ORFs par rapport au génome de référence (Abbadi et al., 2018; Burioli et al., 2017; Martenot et al., 2013; Renault et al., 2012; Rosani et al., 2014). Ces variations génétiques pouvaient potentiellement expliquer l'augmentation de la virulence du virus observée en 2008. L'ensemble de ces résultats a révélé que le virus OsHV-1 présente une diversité génétique non négligeable, mal caractérisée, et dont la signification biologique est encore mal comprise. Il était donc nécessaire de développer les outils pour décrire cette diversité à l'échelle du génome et en particulier sur l'impact que pouvait avoir celle-ci sur la structure des protéines du virus. De cette façon, avec des huîtres ayant des phénotypes au POMS variable il serait possible de mieux appréhender l'impact de la diversité génétique du virus.

Au début de ce travail, le laboratoire IHPE était en train de finaliser l'étude sur le projet DECIPHER ayant démarré en janvier 2015 ("Déchiffrage des maladies multifactorielles," 2015). Son objectif visait à contribuer à la compréhension de la dimension multifactorielle (agents infectieux, microbiote, génétique) des POMS affectant les huîtres et de proposer des stratégies de lutte adéquates et efficaces. Différentes familles d'huîtres ayant des sensibilités contrastées à la maladie ont donc été produites, des IEM ont été mis en place et de nombreuses données de métatranscriptomique étaient disponibles (**Fig.18. p39**). Ce projet a permis de mettre en évidence la succession des événements de la pathogénèse et le rôle prépondérant du virus OsHV-1 au début de la maladie (de Lorgeril et al., 2018). L'analyse des données de métatranscriptomique pouvait donc faire progresser nos connaissances sur la diversité génétique de l'OsHV-1. D'un point de vue plus fondamental, ces recherches permettent d'améliorer nos connaissances sur les conséquences biologiques de la diversité génétique des Herpèsvirus sur le développement d'une maladie infectieuse.

L'objectif de l'étude

L'objectif de ce travail était donc d'utiliser les données de dual métatranscriptomique obtenues sur plusieurs familles d'huîtres présentant des phénotypes de sensibilité au POMS variable en fonction de l'origine de la maladie pour caractériser la diversité génétique des populations virales.

L'illustration du matériel et méthode

Le plan expérimental de l'étude est illustré ci-dessous (**Fig.18. p39**). Ce matériel biologique initialement conçu pour répondre aux problématiques du projet DECIPHER nous a permis d'explorer l'impact de la diversité génétique d'OsHV-1 µVar sur les résultats de la maladie pour différents phénotypes contrastés de famille d'huîtres. Le pipeline d'analyse bioinformatique peut être retrouvé dans l'annexe 2 (p128) et le code à cette adresse https://github.com/IHPE/DivOsHV .

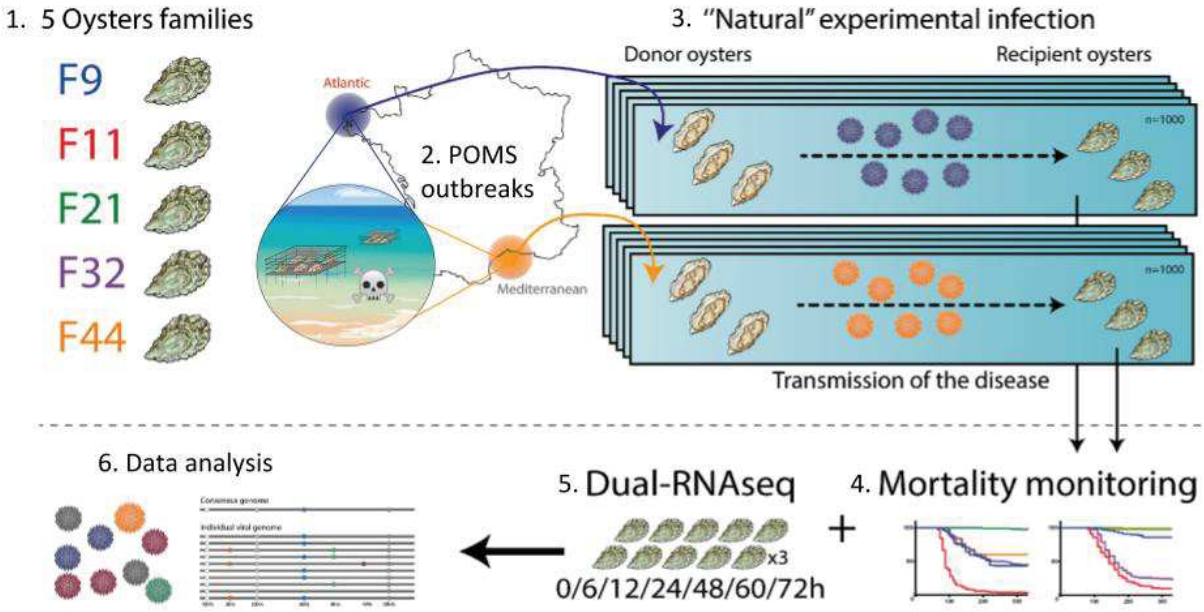


Figure 18. **Illustration du plan expérimental utilisé dans l'article I**.
Dans le cadre du projet DECIPHER cinq familles biparentales d'huîtres juvéniles présentant des phénotypes contrastés vis à vis du POMS ont été produites à partir de géniteurs collectés dans différents environnements (zones conchylicole et non conchylicoles en Atlantique et en Méditerranée). (**1**) Un mélange de 5 familles d'huîtres ayant des phénotypes contrastés vis-à-vis des POMS ont été produites et sont considérées comme les huîtres donneuses. (**2**) Ces huîtres donneuses ont été placées dans l'environnement des fermes ostréicoles de l'Atlantique et de la Méditerranée lors d'événements de mortalité pour capter la maladie. (**3**) Ce mélange d'huître donneuse a été mis en contact de chacune des familles individuellement, lesquelles sont considérées comme des huîtres receveuses, au cours d'une infection expérimentale. (**4**) Lors de cette infection expérimentale en mésocosme, le suivi de la mortalité des huîtres a été réalisé. (**5**) Pour cinq de ces familles, des échantillons de 10 animaux ont été recueillis à 7 temps pour extraire les ARN et faire du séquençage à haut débit. (**6**) Enfin, ces données ont permis la caractérisation de la diversité des populations virales à l'aide d'analyses bioinformatiques.

Article 1 : Contribution of Viral Genomic Diversity to Oyster Susceptibility in the Pacific Oyster Mortality Syndrome

Article publié dans *Frontiers in Microbiology* en 2020, doi : 10.3389/fmicb.2020.01579

Delmotte Jean., Chaparro Cristian, Galinier Richard., De Lorgeril Julien., Petton Bruno., Stenger Pierre-Louis., Vidal-Dupiol Jérémy., Destoumieux-Garzon Delphine., Gueguen Yannick., Montagnani Caroline., Escoubas Jean-Michel., Mitta Guillaume.

Cette publication a également fait l'objet de deux présentations :

- "**Rôle de la diversité génétique de l'Herpèsvirus OsHV-1 dans les mortalités de l'huître du Pacifique**" - présentation poster - Journées Francophones de Virologie en 2019

- "**Characterisation of oshv-1 µvar genotypic diversity during poms outbreaks in two french aquaculture area**" - présentation orale - 19th International Conference on Diseases of Fish and Shellfish en 2019

frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 July 2020
doi: 10.3389/fmicb.2020.01579

Contribution of Viral Genomic Diversity to Oyster Susceptibility in the Pacific Oyster Mortality Syndrome

Jean Delmotte[1], Cristian Chaparro[1], Richard Galinier[1], Julien de Lorgeril[1], Bruno Petton[2], Pierre-Louis Stenger[1], Jeremie Vidal-Dupiol[1], Delphine Destoumieux-Garzon[1], Yannick Gueguen[1], Caroline Montagnani[1], Jean-Michel Escoubas[1]* and Guillaume Mitta[1]*

[1] IHPE, Université de Montpellier, CNRS, Ifremer, Université de Perpignan Via Domitia, Montpellier, France, [2] LEMAR UMR 6539, Université de Bretagne Occidentale, CNRS, IRD, Ifremer, Argenton-en-Landunvez, France

OPEN ACCESS

Edited by:
Manuel Martinez Garcia,
University of Alicante, Spain

Reviewed by:
Colleen A. Burge,
University of Maryland, Baltimore County, United States
Jose A. Fernandez Robledo,
Bigelow Laboratory for Ocean Sciences, United States

***Correspondence:**
Cristian Chaparro
cristian.chaparro@univ-perp.fr
Jean-Michel Escoubas
jean.michel.escoubas@ifremer.fr

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

Received: 05 March 2020
Accepted: 17 June 2020
Published: 10 July 2020

Citation:
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 and Mitta G (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

Juvenile Pacific oysters (*Crassostrea gigas*) are subjected to recurrent episodes of mass mortalities that constitute a threat for the oyster industry. This mortality syndrome named "Pacific Oyster Mortality Syndrome" (POMS) is a polymicrobial disease whose pathogenesis is initiated by a primary infection by a variant of an Ostreid herpes virus named OsHV-1 μVar. The characterization of the OsHV-1 genome during different disease outbreaks occurring in different geographic areas has revealed the existence of a genomic diversity for OsHV-1 μVar. However, the biological significance of this diversity is still poorly understood. To go further in understanding the consequences of OsHV-1 diversity on POMS, we challenged five biparental families of oysters to two different infectious environments on the French coasts (Atlantic and Mediterranean). We observed that the susceptibility to POMS can be different among families within the same environment but also for the same family between the two environments. Viral diversity analysis revealed that Atlantic and Mediterranean POMS are caused by two distinct viral populations. Moreover, we observed that different oyster families are infected by distinct viral populations within a same infectious environment. Altogether these results suggest that the co-evolutionary processes at play between OsHV-1 μVar and oyster populations have selected a viral diversity that could facilitate the infection process and the transmission in oyster populations. These new data must be taken into account in the development of novel selective breeding programs better adapted to the oyster culture environment.

Keywords: *Crassostrea gigas*, herpesvirus diversity, genotype-genotype interactions, oyster genetic background, viral populations

INTRODUCTION

Bivalve mollusk culture is an important sector of world aquaculture, representing approximately a fifth of the global production with 15.7 million tons harvested in 2017 (FAO and WHO, 2018). Given the economic scale and growing importance as a food source for the human population of bivalve farming, much interest has been devoted to investigate infectious diseases, representing a main limitation for aquaculture expansion (Stentiford et al., 2012;

Pernet et al., 2016; Burge et al., 2018). Among pathogens, viruses represent major risks for the sustainable management of this sector, as some viruses are highly infectious, easily transmissible, and bear a high extinction potential (Guo and Ford, 2016). In particular, in the 1970s, an infection with an iridovirus completely decimated the population of *Crassostrea angulata* oysters, then cultivated on all European coasts (Comps et al., 1976; Renault and Novoa, 2004; Arzul et al., 2017). Currently, the Ostreid-1 herpesvirus (OsHV-1) is recognized as being responsible for sporadic mortalities in several bivalve mollusks but it is one of the main mortality threats for the most cultivated bivalve species, the Pacific oyster *Crassostrea gigas* (Arzul et al., 2001a; Renault et al., 2001; Davison et al., 2005; da Silva et al., 2008; Solomieu et al., 2015; Xia et al., 2015; Burge et al., 2018). In this bivalve species, OsHV-1 triggers the Pacific oyster mortality syndrome (POMS) that has plagued the oyster production worldwide, from Europe to north and south America and Asia, for more than a decade (Pernet et al., 2012; Hwang et al., 2013; Jenkins et al., 2013; Bai et al., 2015; Mortensen et al., 2016; Prado-Alvarez et al., 2016; Caceres-Martinez et al., 2018). This syndrome, affecting juveniles, leads to mass mortalities that can reach 100% within days (Segarra et al., 2010). Research efforts have revealed a series of factors contributing to the disease, including infectious agents interacting with seawater temperature and oyster genetics (Segarra et al., 2010; Pernet et al., 2012; Petton et al., 2013; EFSA, 2015; Petton et al., 2015; Le Roux et al., 2016; Azema et al., 2017). Recently, holistic molecular approaches uncovered the pathogenis associated to POMS (de Lorgeril et al., 2018). These studies showed that an infection by a variant of OsHV-1 (Oyster herpesvirus type 1 variant µVar) is the critical step of the infectious process leading to an immune-compromised state followed by a microbiota destabilization that "opens the door" to bacterial pathogens (e.g., vibrios) that target haemocytes to induce their lysis (Rubio et al., 2019). The infectious process is completed with subsequent bacteraemia ultimately leading to oyster death (de Lorgeril et al., 2018).

Although the first description of herpes-like viral infection in oyster (*Crassostrea virginica*) was reported in 1972 by Farley and collaborators (Farley et al., 1972), it was not until the 2000s that partial genome characterization allowed the identification of Ostreid herpesvirus 1 (OsHV-1), a new viral species infecting Pacific oyster larvae and spats (Le Deuff et al., 1994). The first variants of the virus identified, called OsHV-1 var, were associated with mortalities in *C. gigas* and other mollusks; they notably carried a deletion of 2.8 kpb as compared to the previous, not fully assembled, OsHV-1 genome (Arzul et al., 2001c). The genome of OsHV-1 has been fully sequenced from oyster larvae in 2005 and contains a double strand DNA genome of ∼207 kbp encoding approximately 124 open reading frames (ORFs) (Davison et al., 2005). Then in 2008, an emergent genotype, called OsHV-1 µVar, was characterized. It was associated with higher mortalities (Segarra et al., 2010) and several variants of µVar genotype have been identified over the last decades (reviewed in Solomieu et al., 2015).

The OsHV-1 µVar genotype has a 12 bp deletion in a microsatellite locus upstream of ORF4, indel/substitution in ORFs 4/42/43, and a complete loss of ORFs 36/37 as well as partial loss of ORF 38 (Segarra et al., 2010; Renault et al., 2012; Martenot, 2013). Over the past 5 years, with the increasing use of sequencing technologies, more and more genotypes were identified such as OsHV-1-SB (Xia et al., 2015) OsHV-1 µVar A and B genotypes with insertion of 4 ORFs including one coding a membrane protein (Burioli et al., 2017) OsHV-1-PT from the North Adriatic Sea (Abbadi et al., 2018) and OsHV-1 isolate ZK0118 (Bai et al., 2019). In addition, PCR amplification of a subset of ORFs allowed the screening of new genotypes (Burioli et al., 2018). This expanding number of OsHV-1 genotypes and number of carrier species (Arzul et al., 2001a,b, 2002; Batista et al., 2007; Bai et al., 2015) raises the question of the impact of this genetic diversity on spreading and severity of POMS.

Accumulating research over the past decades have shown that a large number of viruses exhibit a high level of genetic diversity. This is especially true for viruses with RNA genomes, including medically important viruses such as HIV, hepatitis C virus, and influenza; this is largely due to the low fidelity of their viral RNA-dependent RNA polymerase (RdRp) (Drake and Holland, 1999; Parvez and Parveen, 2017). Extensive study of this diversity has led to the development of the quasispecies concept that questioned our current understandings of viral diseases and their evolution (Batschelet et al., 1976; Vignuzzi et al., 2006; Lauring and Andino, 2010). It seems that many viruses have the ability to produce diverse genetically linked mutants that can be defined as viral populations or quasispecies maintained by mutation-selection equilibrium (Perales et al., 2015; Poirier and Vignuzzi, 2017). This concept hypothesizes that dynamic viral populations generate beneficial interactions and cooperation that allows them to rapidly adapt to changing environments, including various hosts and their different tissues, relevant for viral fitness and virulence (Xiao et al., 2016; Brooke, 2017). Studies have shown that the level of genetic diversity within viral populations likely influences viral pathogenicity, dissemination and host immune evasion (Mao et al., 2007). Conversely, limited diversity could strongly attenuate virulence (Pfeiffer and Kirkegaard, 2005). DNA viruses also bear genomic variability that rivals that of many RNA viruses (Renzette et al., 2011; Renner and Szpara, 2018). Recent work has shown a significant inter-host and intra-host genetic divergence across tissue compartments and time of infection by herpesviruses (Renzette et al., 2013, 2014). In addition, the evolution of a disease is sometimes not fully explained by intrinsic factors of the host, but rather by the variability of herpesvirus (Akhtar et al., 2019).

Up to now, far too little attention has been paid to the role of OsHV-1 genetic diversity in POMS. Here, we analyzed a set of biparental oyster families displaying contrasting susceptibility phenotypes to POMS in two different infectious environments. This biological material gave us the unique opportunity to explore the impact of the genetic diversity of OsHV-1 µVars on the outcomes of the disease for different oyster genotypic backgrounds. We found that susceptibility phenotypes of oyster families varied from one infectious environment to another. Single Nucleotide Polymorphism (SNP) exploration from RNAseq data revealed different viral populations in the

different infectious environments; these viral populations were different from the reference OsHV-1 μVar genome. Moreover, in a same infectious environment, the different oyster families were shown to be infected by different OsHV-1 variants. The genetic diversity of viruses between oyster families and environments suggests a host-pathogen co-evolutionary process. Interestingly, viral populations contain protein-coding variations that might impact viral pathogenesis, as revealed by SNPs localization. This work advances our knowledge of OsHV-1 μVar diversity and constitutes a step toward future investigation on the role of viral genetic variation, its relationship with phenotype, host genetics, and impact on disease outcome.

MATERIALS AND METHODS

Production of Biparental Oyster Families

Biparental oyster families were produced by *in vitro* fertilization from wild genitors sampled in farming and non-farming areas as previously described (de Lorgeril et al., 2018, 2020). Briefly, wild genitors used to produce Atlantic families F9 and F11 were collected at two sites approximately 20 km apart, in Logonna Daoulas (lat.: 48.335263 long.: -4.317922, farming areas) and Dellec (lat.: 48.353970 long.: -4.566123, non-farming areas), respectively (**Figure 1A**). Wild genitors used to produce Mediteranean families F32 and F44 were collected at two sites approximately 40 km apart, in Vidourle (lat.: 43.553906 long.: 4.095175, non-farming areas) and Thau lagoon (lat.: 43.418736 long.: 3.622620, farming areas), respectively. Genitors coming from aquaculture areas were assumed to be exposed to stronger selection pressure due to mass mortality outbreaks occurring annually at these sites. The last family, F21, was generated from a pair of broodstocks derived from mass selection conducted in the field during four generations in an aquaculture area at the Atlantic coast (La Tremblade, lat 45.781741 long -1.12191) (Degremont et al., 2015).

Experimental Infections

Our experimental infection protocol consists of a cohabitation between *C. gigas* oysters ("donors") carrying the disease and "pathogen-free" *C. gigas* oysters ("recipients") (Petton et al., 2013; de Lorgeril et al., 2018). The "pathogen-free" status of the animals was confirmed by (i) the absence of OsHV-1 DNA detection by qPCR and (ii) a low Vibrio presence ($\sim$10 cfu^{-1} tissue) determined by isolation on selective culture medium (thiosulfate-citrate-bile salts-sucrose agar, TCBS). Oysters were observed to remain free of any abnormal mortality throughout the larval stage until the beginning of the experimental infections. A first experimental infection used donors previously exposed to the infectious environment of Atlantic origin. The donors were "pathogen-free" oysters (mixture of 116-day-old oysters from the 15 families, 17,700 g with a mean individual weight of 1.1 g) were first deployed in a farming area (Logonna Daoulas, lat 48.335263, long - 4.317922) during the infectious period until the first mortalities occurred. Then, donor oysters were transferred back to the laboratory and placed in contact with "pathogen-free" recipient oysters in a controlled environment (**Figure 1B**). The experiment was conducted using the same biomass (1120 g) of donors in cohabitation in independent tanks, each containing one of the recipient oyster families (1120g with a mean individual weight of 1.1 g) which were previously acclimatized in these structures for 2 weeks. The Atlantic experimental infection began on 17 July 2015 and ended on 31 July 2015. Similarly, a second experimental infection used donors previously exposed to the infectious environment of Mediterranean origin in a farming area (Thau lagoon, lat 43.418736, long 3.622620), except that donors deployed were a mixture of 176-day-old oysters from the 15 families (26,500 g with a mean individual weight of 1.7 g) and that the biomass of donors and the biomass of recipients in each tank was 1760 g each (recipient oysters with a mean individual weight of 1.73 g). The Mediterranean experimental infection began on 21 September 2015 and ended on 6 October 2015. In parallel and for each experimental infection, a control cohabitation experiment was performed under identical conditions but using donors that had not spent time in the farming areas. Mortality was monitored in laboratory tanks. During the experimental infection, 10 oysters in triplicate were randomly sampled without blinding protocols from each tank and at each time point (0, 6, 12, 24, 48, 60, and 72 h). The shell was removed, and pools of 10 oysters were flash frozen in liquid nitrogen.

DNA and RNA Extractions

Oyster pools (10 individuals per pool) were ground in liquid nitrogen in 50 mL stainless steel bowls with 20 mm diameter grinding balls (Retsch MM400 mill). The obtained powders (stored at −80°C) were then used for extraction of DNA and RNA as previously described in de Lorgeril et al. (2018). Briefly, DNA and RNA were extracted using the NucleoSpin® Tissue Genomic extraction kit (Macherey-Nagel) and the Direct-Zol RNA Miniprep kit (Proteigene), respectively, according to the manufacturer's protocol. Nucleic acids concentration and purity were checked using a NanoDrop ND-1000 spectrometer (Thermo Scientific), and their integrity was analyzed by capillary electrophoresis on a BioAnalyzer 2100 (Agilent).

Viral Load Quantification

Quantification of OsHV-1 has been carried out using quantitative PCR (qPCR) and used as a proxy for infection. All amplification reactions have been analyzed using the Roche LightCycler 480 Real-Time thermocycler in three technical replicates (qPHD-Montpellier GenomiX platform, Montpellier University, France). The total qPCR reaction volume was 1.5 μL, consisting of 0.5 μL of DNA (40 ng/μL) and 1 μL of LightCycler 480 SYBR Green I Master mix (Roche) containing 0.5 μM of PCR primer (Eurogenetec SA). Virus-specific primer pairs target the region of the OsHV-1 genome predicted to encode a catalytic subunit of DNA polymerase (ORF100, AY509253): Fw-5′-ATTGATGATGTGGATAATCTGTG-3′ and Rev-5′-GGTAAATACCATTGGTCTTGTTCC-3′ (Webb et al., 2007). A Labcyte Acoustic Automated Liquid Handling Platform (ECHO) has been used for pipetting into the 384-well plate (Roche). A LightCycler® 480 Instrument (Roche) was used for qPCR with the following program: enzyme activation at 95° C for 10 min followed by 40 denaturation cycles (95° C, 10

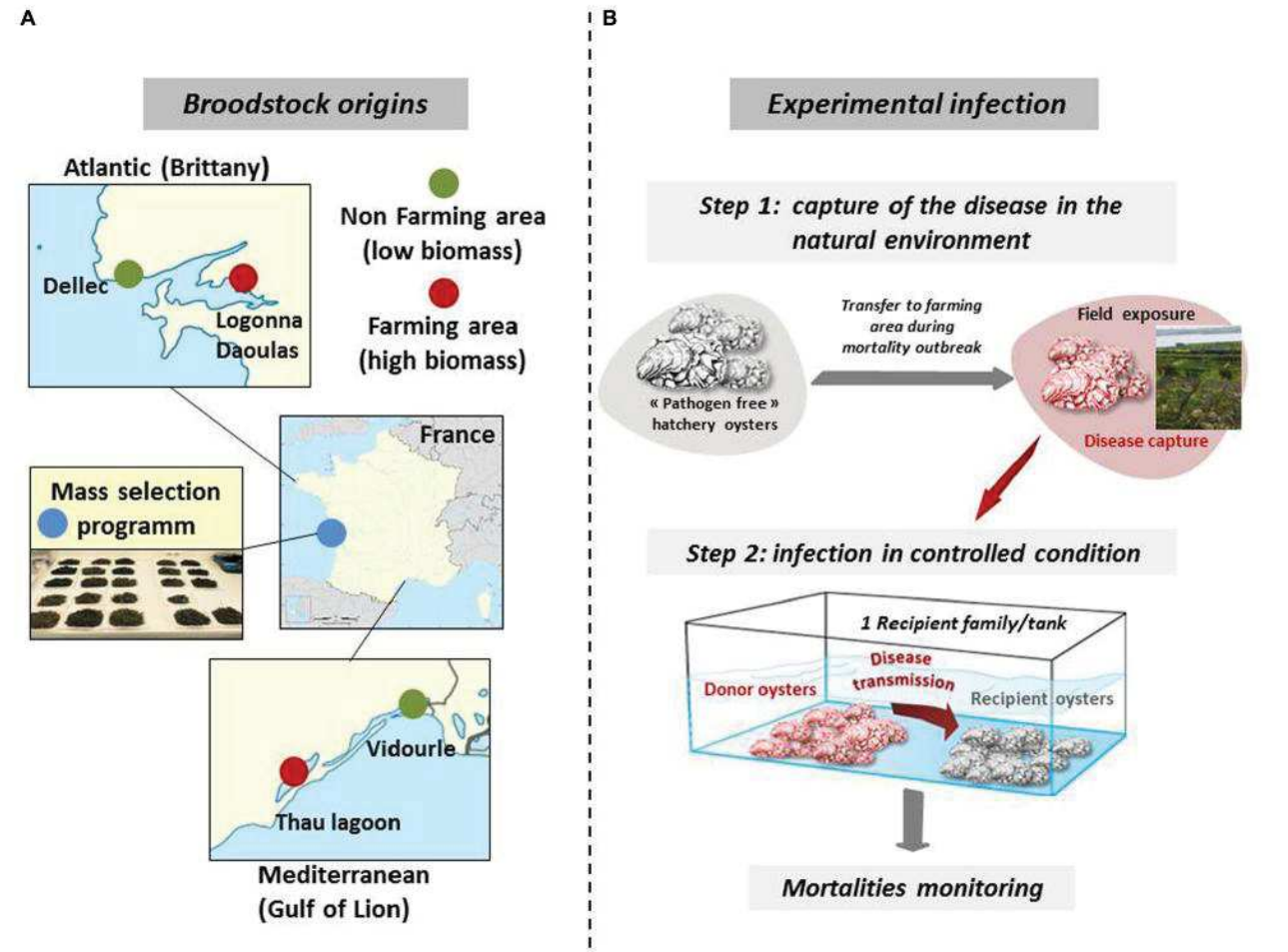


FIGURE 1 | Broodstock origins for the production of biparental oyster families and schematic of experimental infection. **(A)** Wild stocks were sampled in farming (red) and non-farming (green) sites in two geographic areas (Atlantic and Mediterranean coasts). Mass selected oysters (blue) originated from the Ifremer hatchery of La Tremblade. Image source: commons.wikimedia.org. **(B)** For experimental infection, pathogen free oysters were deployed in the natural environment in a farming area during disease outbreaks and brought back to a controlled environment to transfer the disease to each oyster families under controlled conditions. Experimental infections were performed with infectious environments from Atlantic and Mediterranean origin. This figure was adapted from de Lorgeril et al. (2020) (https://doi.org/10.1186/s12864-020-6471-).

s), hybridization (60° C, 20 s) and elongation (72° C, 25 s). To check the specificity of the amplification, a subsequent melting temperature curve of the amplicon was performed. For absolute quantification, DP amplification products were cloned into the pCR4-Topo vector and replicated in *Escherichia coli* DH5α (Invitrogen). Plasmids were extracted using the Wizard Plus SV miniprep DNA purification system (Promega) and standard curves of known concentration of plasmid generated according to the following formula [amount of plasmid DNA (ng) × $6.022.10^{23}$]/[length of DNA template (bp) × 10^9 × 660]. Absolute quantification of viral DNA copies was then calculated by comparing the reported Cq values to the standard curve constituted of 8 serial dilution points of the circular plasmid (10^9–10^2 copies per μL) in three technical replicates (R^2 = 0.993, efficiency = 1.91 and equation: y = -3.5416x + 40.02). Samples were considered positive when superior to 10^2 copies per μL (corresponding to 30 cycles) and but also when inferior to 10^2 copies per μL if the melting temperature of amplicon was correct (between 30 and 35 cycles).

RNA-Seq Analyses

RNA-Seq library construction and sequencing were performed by the Fasteris Company (Switzerland). Directional cDNA libraries were constructed using a TruSeq mRNA Stranded kit (Illumina) and sequenced on a Hiseq in paired-end reads of 2 × 75 bp. Adapters were cleaned using trimmomatic 0.39 (Bolger et al., 2014). The quality of the sequences was evaluated by FastQC v0.11.8 (Andrews, 2010) and reads with a mapping quality above 20 were retained. PCR duplicates from sequencing where remove using picard v2.18.14 ("Picard toolkit," 2019)[1] with the MarkDuplicates option. Then reads were aligned to the OsHV-1 μVar A genome (GenBank: KY242785.1) (Burioli et al., 2017)

[1]http://broadinstitute.github.io/picard/

using Bowtie 2 v2.3.4.3 (Langmead and Salzberg, 2012). The abundance of OsHV-1 μVar reads was counted using samtools v1.9 (Li et al., 2009) and normalized by multiplying raw mapped reads on the OsHV-1 μVar A genome by a library normalization factor (calculated as the average library size for all times and controls divided by the library size for the specific time point) (**Supplementary Table S1**). The abundance of OsHV-1 μVar reads was used as a proxy for viral replication. Data treatments were carried out under a High-performance computing (HPC) cluster using a self-scripted pipeline (code including parameters used are available on https://github.com/IHPE/DivOsHV).

SNP Calling and Quantification

SNP detection and quantification were performed on families F11 and F32 in both the Atlantic and Mediterranean experimental infections, while families F9 and F44 were only analyzed in the Atlantic experimental infections due to the lower read coverage in the Mediterranean infections. Sequencing reads were mapped to OsHV-1 μVar A genome as described previously. Because the sequencing depths differ between time-points and according to the sensitivities of infected oysters, we first concatenated the three replicates of the three time-points containing the highest viral loads together (48, 60, and 72 h) using samtools merge v1.9 (Li et al., 2009). Then pysamstats v1.1.2[2] was used to count the nucleotide composition on all genomic positions. SNPs were detected with an *ad hoc* script[3], using the R language R Core Team, applied to each pysamstats output. Nucleotide changes were called "SNP" if (i) the alternative base frequency was greater than 5%. (ii) the sequencing depth was greater than 180 reads, to insure an average coverage of 60 for the triplicates. (iii) the depth of the alternative base was greater than an intrinsic parameter to the library (mean + standard deviation of all alternative bases) which is used to eliminate all low abundant alternative bases. The outputs were stored in VCF format which allow us to determine the impact of each SNP with SnpEff and SnpSift (Cingolani et al., 2012). Finally, to compare the environments, we combined the data according to the experimental infection origin. SNP counts were summed by family and allelic frequency recalculated using R. All the graphics were made using the ggplot2 package (Wickham, 2016) and Venn diagrams were made with the Venn Diagram package (Chen and Boutros, 2011).

Statistical Analysis

OsHV-1 colonization was assessed before the first mortalities appear, at 12/24/48/60/72 h post-infection. For the statistical analyzes, the time points were subdivided into two groups, the first one corresponded to the early stage of the infection (12/24 h) that is characterized by an intense increase of the viral load, whereas the second one corresponded to the late stage of the infection (48/60/72 h) when viral load reached a plateau. Mann–Whitney *U*-tests were used (GraphPad Prism 8.4.2) to compare OsHV-1 load and replication rate between families at late phase of the infection (48/60/72 h) (**Supplementary Table S2**). Mann–Whitney *U*-tests were also used to compare OsHV-1 viral load within each family between early and late phase (**Supplementary Table S2**).

[2]http://github.com/alimanfoo/pysamstats

[3]https://github.com/IHPE/DivOsHV

RESULTS

Susceptibility of Oyster Families to POMS Is Associated With Viral Load and Varies According to Disease Origin

In this study, we used five biparental families of juvenile oysters with contrasting phenotypes toward POMS. These families were produced from genitors that had experienced different selective filters (see "Materials and Methods" section). The 5 families were subjected to an experimental infection mimicking disease transmission in nature. For that we performed cohabitation experiments with oysters previously immersed in Atlantic and Mediterranean oyster farms during POMS events (see **Figure 1B** and "Materials and Methods" section).

A high variability in survival rates was observed between families within each experimental infection, ranging from 1 to 98% and 10 to 99.7% in Atlantic and Mediterranean experimental infection, respectively (**Figures 2A,B**). Two families, F11 and F21, had the same phenotypes, susceptible (S, percentage of survival $< 25\%$) and resistant (R, percentage of survival $> 75\%$), respectively, whatever the experimental infection they faced (Atlantic or Mediterranean). Conversely, the phenotype of the three other families changed with experimental infections. Oyster families F9, F32, and F44, which had an intermediate phenotype (I, $25\% \leq$ percentage of survival $\geq 75\%$) in the Atlantic experimental infection (Atl), were found susceptible (F32) or resistant (F9 and F44) in the Mediterranean experimental infection (Med). These phenotypic changes cannot be attributed to a decrease in the severity of POMS in Med since (i) the F11 family had almost the same survival rate in both experimental infections and (ii) the F32 family has a lower survival rate in Med than in Atl.

The load and transcriptional activity of OsHV-1 μVar were monitored within the first 72 h post-infection in the five oyster families and two experimental infections (Atl and Med) using qPCR and RNA-Seq, respectively (**Figures 2C–F**). In all conditions we observed that the viral loads significantly increased between the early (12/ 24 h) and the late (48/60/72 h) phase of infection (**Figures 2C,D**; Mann–Whitney *U*-test, $P < 0.01$ in Atl and $P < 0.001$ in Med; details in **Supplementary Table S2**). We also detected viral transcripts in both experimental infections for all oyster families (**Figures 2E,F**). These results show that OsHV-1 was able to infect all oyster families, indicating all were tolerant to infection. However, the amount of viral nucleic acid varied with disease expression: susceptible families accumulated 10^2–10^3 times more viral transcripts than resistant ones (**Supplementary Table S3**). Taken together, these results show that all families were infected with the OsHV-1 μVar virus but the most resistant ones controlled viral replication.

Noteworthy, oyster families with an intermediate phenotype in Atl showed an R or S phenotype in Med. As the major difference between these two experiments is the geographic

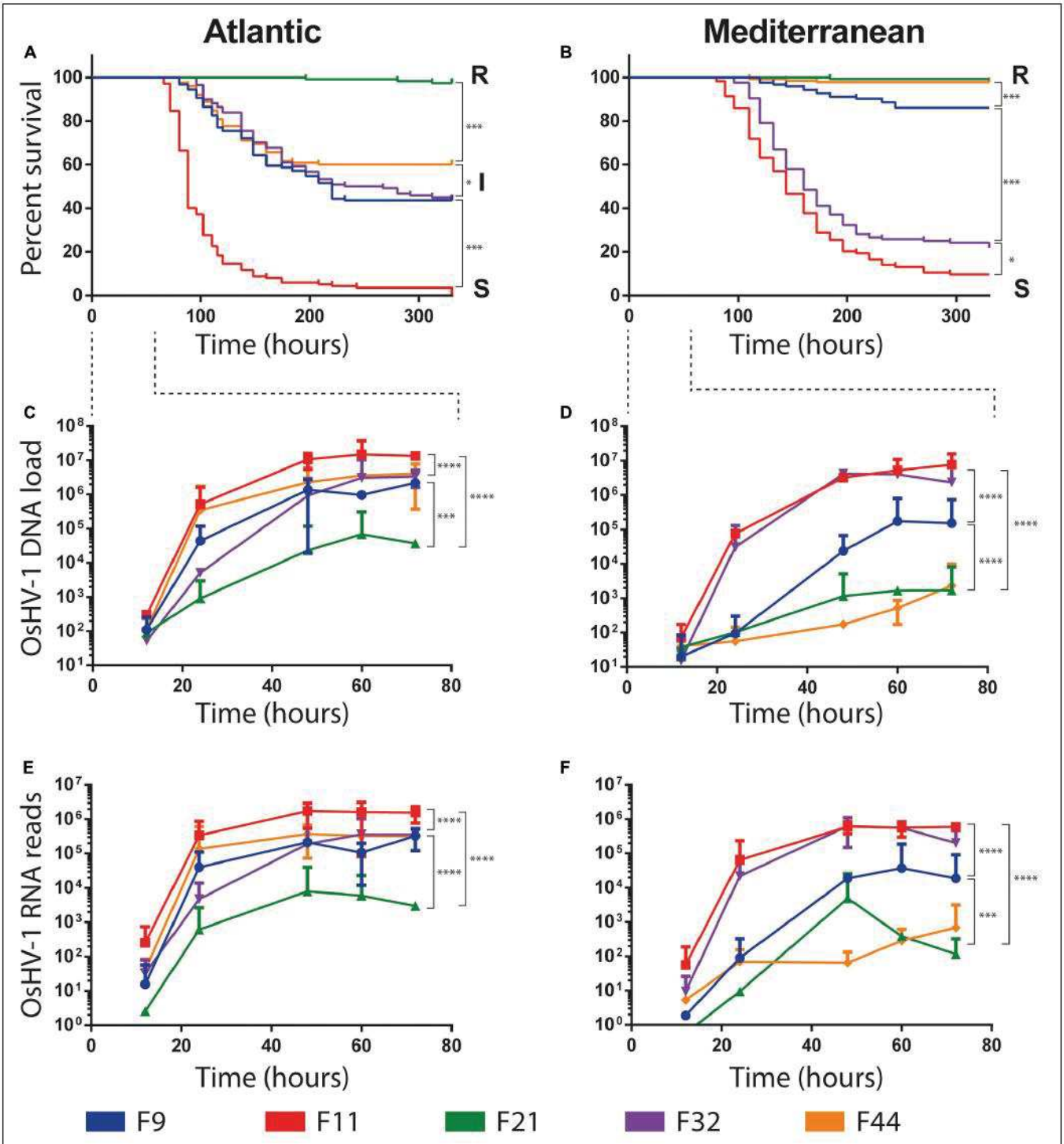


FIGURE 2 | Survival curves and OsHV-1 colonization of five biparental oyster families submitted to experimental infections using donors oysters deployed in Atlantic and Mediterranean farming area during infectious period. **(A,B)** Represent Kaplan-Meier survival curves of the 5 families of recipient oysters during the Atlantic and Mediterranean natural experimental infection, respectively. Resistant, intermediate and susceptible oyster families are indicated by R, I and S, respectively. The stars indicate *P*-values according to Mantel-Cox Log-rank test (*$P \leq 0.05$, ***$P \leq 0.001$). OsHV-1 colonization was assessed at the early time point of experimental infection at 12/24/48/60/72 h. **(C,D)** OsHV-1 load was quantified by qPCR and expressed as viral gene copy number per ng of oyster DNA. **(E,F)** Viral replication was estimated by the total number of RNA-Seq reads mapped on the OsHV-1 µVar A genome (KY242785). The stars indicate *P*-value according to Mann-Whitney multiple comparisons *U*-tests of viral load **(C,D)** and viral replication rate **(E,F)** between families at late phase of infection (48/60/72 h) (***$P \leq 0.001$ ****$P \leq 0.0001$).

origin of the infection, it suggests that the POMS occurring on the Atlantic coast is different from that of the Mediterranean coast. This difference in survival phenotypes is also visible at the late phase of infection (48/60/72 h) both in terms of viral load and viral replication rate. Indeed, in Atl, viral load and viral replication rate are not significantly different for the three intermediate phenotype families (F9, F32, and F44, Mann–Whitney U-test, $P > 0.05$) (**Figures 2C–E** and **Supplementary Table S2**). Conversely, in Med, the viral load and replication rate are significantly different between the three families (Mann–Whitney U-test, $P < 0.001$) (**Figures 2D–F** and **Supplementary Table S2**). These observations led us to hypothesize that oysters encountered different OsHV-1 μVar populations in the two environments.

Two Distinct Viral Populations in Atlantic and Mediterranean Infections Revealed by SNPs

Viral populations present in the Atlantic and Mediterranean environments, were analyzed by merging RNA-Seq data according to environments. This included the RNA-Seq reads mapping against OsHV-1 μVar A genome in the 3 replicates and the 3 time points containing the highest viral loads (48, 60, and 72 h), both in intermediate and susceptible families (resistant families were not included in the analyses due to the low number of viral reads). Then, single nucleotide polymorphisms (SNPs) were analyzed. We identified 436 variable positions along OsHV-1 μVar A genome: 268 and 59 were specific to Atl and Med, respectively, and 109 were common to both experimental infections as compared to OsHV-1 μVar A (**Figure 3A** and **Supplementary Tables S4**, **S5**).

Variants were binned (5% intervals) according to their frequency (**Figures 3B,C**). Variants specific to Atl may be subdivided into three groups, minor (47%; frequency $< 25\%$) intermediate (48.1%; $25\% \leq$ frequency $\leq 75\%$) and major variants (4.9%; frequency $> 75\%$) (**Figure 3B**, upper panel). Variants specific to Med segregated in 2 main clusters, minor (62.7%) and major variants (35.6%) (**Figure 3B**, lower panel). This difference in distribution of variant frequency from Atl and Med was also observed for variants common to both experimental infections (**Figure 3C**). Indeed, the percentages of minor and intermediate variants were higher in Atl than in Med (minor: 23.9% and 17.4%, respectively, intermediate: 49.5 and 21.1%, respectively). Conversely, the percentage of major variants was lower in Alt than in Med (26.6 and 60.5%, respectively).

Finally, we compared the variant frequency between Atl and Med for each of the common SNPs on the 109 positions spread all over the OsHV-1 μVar genome (**Figure 3D** and **Supplementary Table S6**). On these 109 common SNPs, 49.5% (54 SNPs) had the same variant frequency (difference in variant frequency $< 5\%$) and 50.5% (55 SNPs) had different variant frequency (ranging from 5.1 to 88.2%; median of 42.9%).

Altogether, these results show that the viral populations collected in the two environments (Atlantic and Mediterranean) are different from each other but also different from the OsHV-1 μVar A genome used as reference. Interestingly, the viral population from Med have a reduced genetic diversity compared to the one from Atl. Indeed, whereas viral population from Med had less SNPs, 52.3% of them correspond to major variants with 9.5% of fixation (variant frequency 100%) whereas in Atl, there were only 11.1% of major variants with 1.3% of fixation (**Supplementary Tables S4**, **S5**).

Atlantic and Mediterranean OsHV-1 μVar Viral Populations Bear Different Membranes Proteins

We further examined the distribution of all variants that occurred in ORFs and we found that 299 SNPs (out of 436) were located in ORFs. Among them, 99 correspond to synonymous SNPs and 200 to nonsynonymous SNPs (nsSNPs) that could be divided into 191 missense variants, 5 stop lost, 4 stop gained, and 2 start lost (**Figure 4**). The comparison of specific SNPs for each experimental infection revealed that for the Atlantic viral population, 76 ORFs (59.4%) carry at least one nsSNPs compared to only 53 ORFs (41.4%) for the viral populations from the Mediterranean.

We then focused on the nature of the proteins encoded by the ORFs carrying nsSNPs. The most represented ORF group, carrying 55.4% of the nsSNPs, encodes proteins of unknown functions (44 ORFs). The second most represented ORF group encodes membrane-related proteins (membrane and transmembrane proteins, glycosylated or not) and 15 out of 19 ORFs encoding this type of proteins carry 23.3% of the nsSNPs (**Table 1**). Some ORFs carry numerous nsSNPs (ORF32, ORF54 and ORF103) and were differentially mutated in the two experimental infections. For instance, ORF54 which is predicted to encode a transmembrane glycoprotein of 807 aa carries 8 nsSNPs, 7 of which are specific to Atl. Among these 7 snSNPs, 2 induce a stop_lost, which increases the protein size by 39 aa, and 5 correspond to missense-variants. These mutations in ORFs could impact protein functions.

More interestingly, two ORFs carry stop_gained mutations at high frequency. The first one, found specifically in Med at a frequency of 76.1%, truncates the ORF16, predicted to encode a transmembrane protein, by replacing the 31st amino acid by a stop codon removing 58.4% of the protein. The second one, found specifically in Atl at a frequency of 84.6%, truncates the ORF114 by replacing the 15th amino acid by a stop codon removing 95.3% of the protein. Amino acid sequence analysis predict that ORF114 encodes a transmembrane glycoprotein with a transmembrane helix (aa 13–35) and a glycosylated extracellular domain (aa 36–322), both of which are truncated by the mutation (**Supplementary Figure S1**). Interestingly, size variations of these ORFs have already been reported for the different OsHV-1 genomes available. The most striking size variations are observed for the protein encoded by the ORF114 with 494 aa in OsHV-1 (Davison et al., 2005) 322 in OsHV-1 μVar A (Burioli et al., 2017) and 282 in OsHV-1-PT (Abbadi et al., 2018). Surprisingly, whereas the three proteins are almost 100% identical on their C-terminal part (281 aa), solely the protein encoded by OsHV-1 μVar A

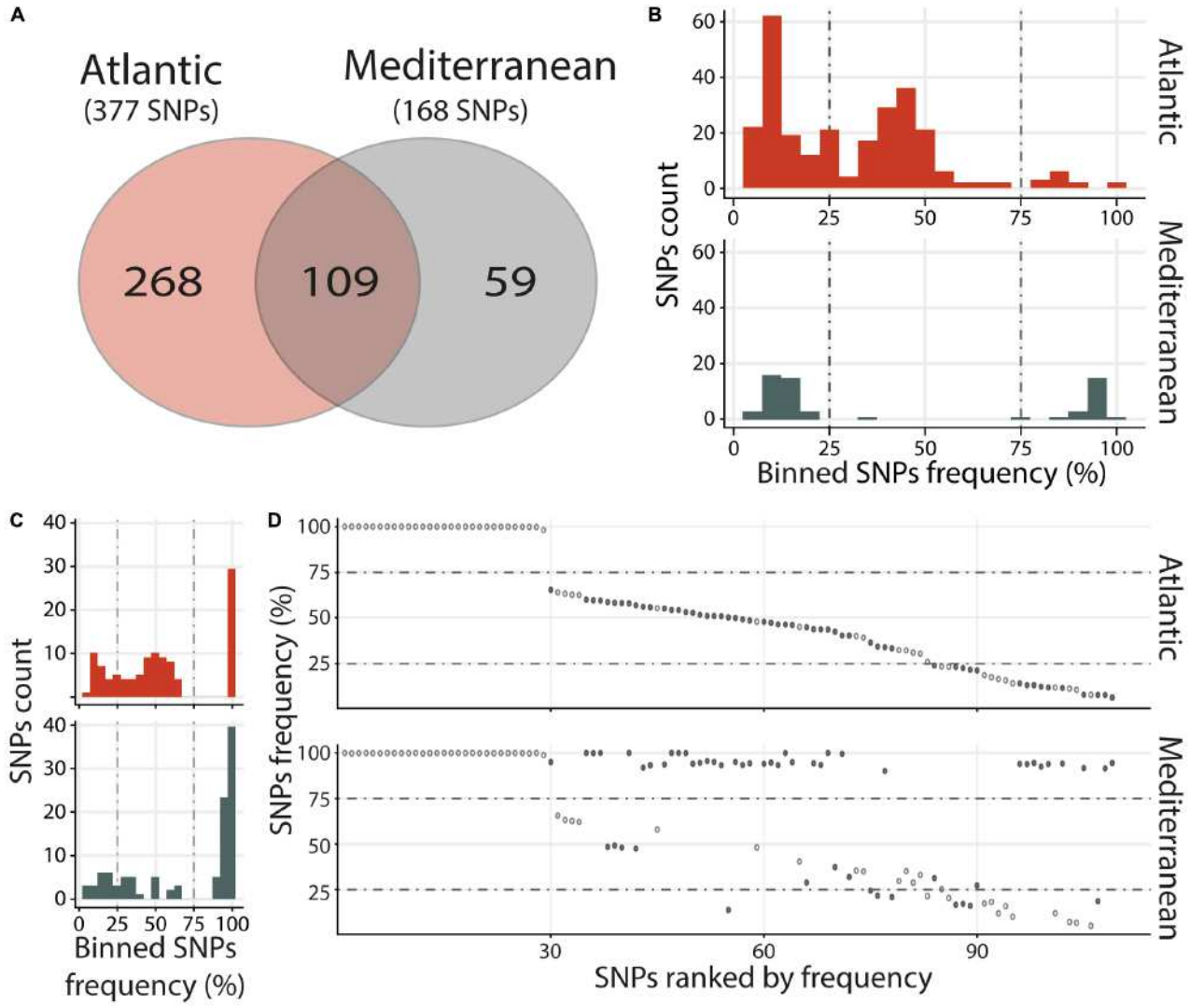


FIGURE 3 | Distinct heterogeneous viral populations within Atlantic and Mediterranean experimental infections. **(A)** Venn diagram representation of the number of specific and common SNPs within each experimental infection. The frequency distribution of SNPs for specific **(B)** and common **(C)** SNPs between Atlantic and Mediterranean experimental infections is illustrated in histograms. SNPs are binned at 5% intervals along the *x*-axis, and the number of SNPs in each bin is indicated on the *y*-axis. **(D)** Common SNPs are ordered according to decreasing variant frequency in Atlantic experimental infection and plotted for the same position in Mediterranean experimental infection. Black and white dots correspond to SNPs with different or similar variant frequencies (difference < 5%), respectively.

contains a transmembrane helix (**Supplementary Figure S2**; Burioli et al., 2017).

Oyster Families Are Infected by Distinct Viral Populations in a Same Environment

In order to determine if the different oyster families are infected or not by the same viral populations within each experimental infection, we compared, independently in experimental infection, the number and frequency of SNPs between oyster families with intermediates and/or susceptible phenotypes.

For the Atl, we compared the four oyster families with susceptible (F11) and intermediate (F9, F32, and F44) phenotypes. We observed that families with intermediate phenotypes contained a similar number of variants whereas the susceptible family F11 contained a larger number of variants (**Figure 5A** and **Supplementary Table S7**). This difference could be related to the fact that susceptible families have a higher viral load than the families with intermediate phenotype. The higher number of viral reads obtained in susceptible oysters could have facilitated the identification of viral variants. However, each family with intermediate phenotypes (F9, F32, and F44) contained specific SNPs which were not found in F11, which suggests viral populations differ between oyster families. We also identified 148 SNPs common to the 4 families. The analysis of F11/32/44 revealed that their variant frequency are equally scattered with 12–15% of minor variants, 68–71% of intermediate variants and 17% of major variants (**Figure 5B** and

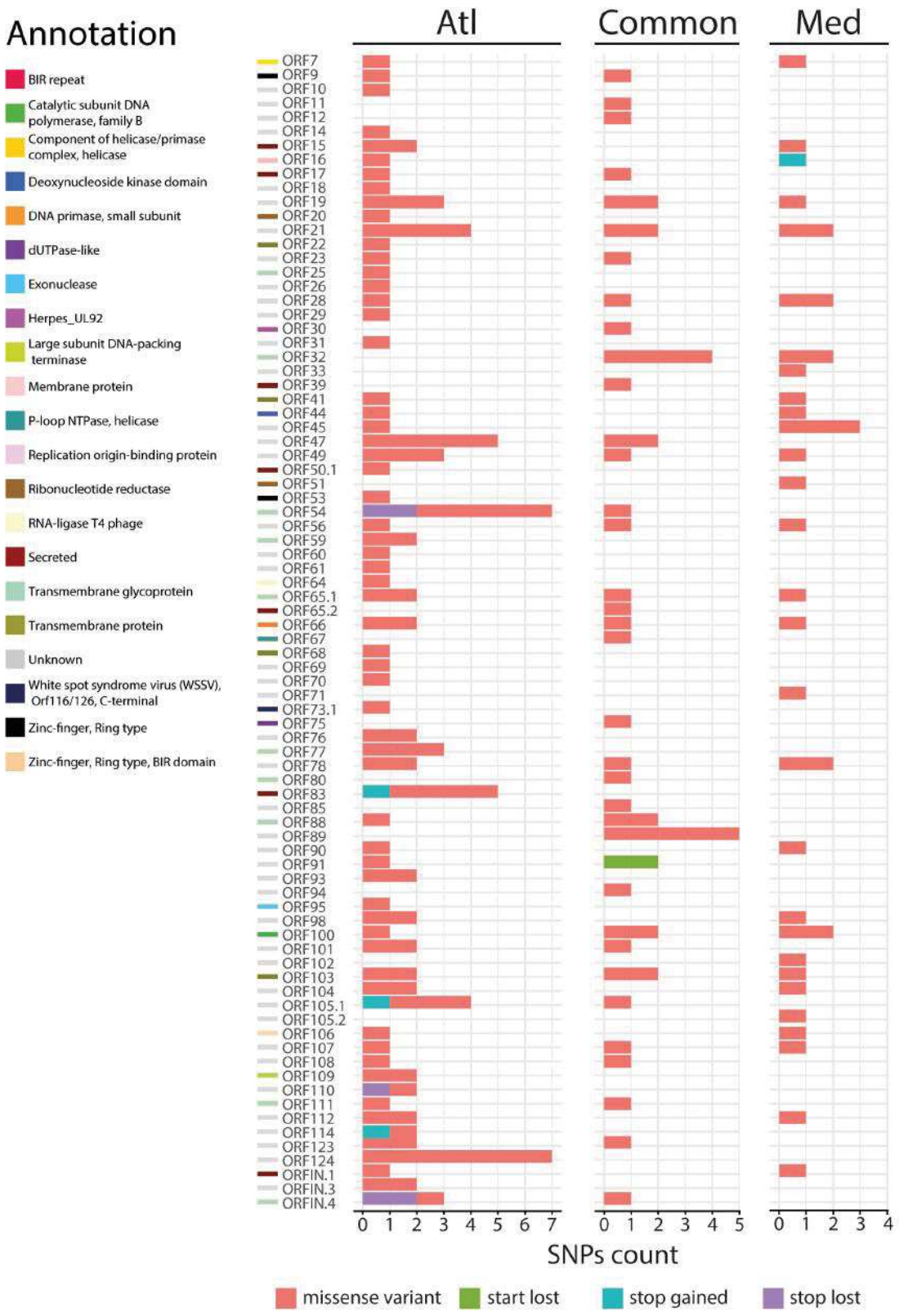


FIGURE 4 | Impact of the 200 nonsynonymous SNPs on the sequence of OsHV-1 µVar ORFs. Stacked barplot indicates the number of nonsynonymous SNPs (*x*-axis) located in each OsHV-1 ORF (*y*-axis). Each color indicates a specific mutation effect as indicated at the bottom of the figure. Each line is related to one of the expressed OsHV-1 ORFs. ORF colors (left side) correspond to a functional annotation of the proteins that they encode.

TABLE 1 | ORFs encoding membrane related proteins carrying nsSNPs.

ORF (GenBank accession #)	ORF size (in bp)	Protein size (in aa)	Protein putative function	SNPs position	Type of variation	Allel frequency (in %)		Experimental infection	Impact on protein
						Atl.	Med.		
ORF16 (ASK05544.1)	**231**	**76**	**Membrane protein**	21,892	missense_variant	43.00	0.00	Atlantic	His22Gln
				21,917	**stop_gained**	0.00	**76.09**	**Mediterranean**	**deletion 45 aa**
ORF22 (ASK05550.1)	4899	1632	Transmembrane protein, 1 helix	29,343	missense_variant	21.46	0.00	Atlantic	Val64Leu
ORF25 (ASK05553.1)	666	221	Transmembrane glycoprotein, 1 helix	39,333		41.52	0.00	Atlantic	Arg216Lys
ORF32	1659	552	Transmembrane glycoprotein, 1 helix	47,348	missense_variant	51.63	94.55	Common	Ser10Phe
				48,888		14.04	10.18	Common	Leu523Phe
				48,884		10.54	6.78	Common	Trp522Leu
				48,885		11.18	7.28	Common	Trp522Phe
				48,220		0.00	95.11	Mediterranean	Asp301Asn
				48,283		0.00	11.20	Mediterranean	Asp322Asn
ORF41 (ASK05566.1)	2922	973	Transmembrane protein, 1 helix	56,584	missense_variant	48.05	0.00	Atlantic	Asp866Asn
				54,968		0.00	19.09	Mediterranean	Val327Asp
ORF54 (ASK05581.1)	2424	807	Transmembrane glycoprotein, 1 helix	80,895	missense_variant	100.00	99.75	Common	Ser202Gly
				82,296		7.61	0.00	Atlantic	Asp669Asn
				80,297		45.50	0.00	Atlantic	Asn2Lys
				80,299		42.72	0.00	Atlantic	Thr3Lys
				80,301		15.96	0.00	Atlantic	Val4Ile
				80,302		11.41	0.00	Atlantic	Val4Asp
				82,714	stop_lost	25.36	0.00	Atlantic	addition 39 aa
				82,715		26.20	0.00	Atlantic	addition 39 aa
ORF59 (ASK05586.1)	3243	1080	Transmembrane glycoprotein, 1 helix	87,730	missense_variant	8.58	0.00	Atlantic	Asp332Asn
				87,319		5.97	0.00	Atlantic	Glu195Lys
ORF65.1	1263	420	Transmembrane glycoprotein, 1 helix	95,619	missense_variant	59.47	100.00	Common	Thr201Arg
				95,887		8.90	0.00	Atlantic	Glu112Lys
				96,057		9.74	0.00	Atlantic	Asp55Gly
				95,820		0.00	93.58	Mediterranean	Ser134Leu
ORF68 (ASK05592.1)	2082	693	Transmembrane protein, 1 helix	103,523	missense_variant	41.37	0.00	Atlantic	Asp394Tyr
ORF77 (ASK05600.1)	3795	1264	Transmembrane glycoprotein, 2 helix	115,467	missense_variant	40.22	0.00	Atlantic	Arg3Gln
				117,243		7.41	0.00	Atlantic	Gln595Arg
				115,875		7.46	0.00	Atlantic	Thr139Met
ORF80 (ASK05603.1)	351	116	Transmembrane glycoprotein, 1 helix	123,511	missense_variant	99.93	99.96	Common	Tyr77Phe
ORF88 (ASK05611.1)	2247	748	Transmembrane glycoprotein, 1 helix	130,979	missense_variant	99.95	99.90	Common	Val381Gly
				130,997		99.92	99.86	Common	Thr387Met
				131,902		7.63	0.00	Atlantic	Asp689Asn

(Continued)

TABLE 1 | Continued

ORF (GenBank accession #)	ORF size (in bp)	Protein size (in aa)	Protein putative function	SNPs position	Type of variation	Allel frequency (in %)	Experimental infection	Impact on protein
ORF103 (ASK05626.1)	1272	423	Transmembrane protein, 4 helix	153,070	missense_variant	40.57	Atlantic	Glu43Lys
				153,223		47.84	Common	Gly94Arg
				153,224		48.52	Common	Gly94Glu
				153,848		22.79	Atlantic	Ser302Leu
				153,151		0.00	Mediterranean	Glu70Lys
ORF111 (ASK05633.1)	870	289	Transmembrane glycoprotein, 5 helix	168,846	missense_variant	8.04	Atlantic	Ser51Leu
				168,853		43.81	Common	Glu49Lys
ORF114 (ASK05636.1)	**969**	**321**	**Transmembrane glycoprotein, 1 helix**	171,530	**stop_gained**	**84.63**	**Atlantic**	**Deletion 307 aa**
				171,532	p.Ile16Lys	89.40	Atlantic	Ile16Lys
ORFIN.4 (ASK05572.1)	981	326	Transmembrane glycoprotein, 1 helix	61,244	missense_variant	99.96	Common	Phe166Ile
				61,725	missense_variant	9.08	Atlantic	Ser326Leu
				61,728	stop_lost	11.40	Atlantic	Addition 3 aa
				61,729		12.40	Atlantic	Addition 3 aa

The ORFs in bold correspond to variants of the stop_gained type.

Supplementary Table S7). For the F9 we observed a majority of minor variant (45%), followed by 30 and 26% of intermediate and major variants, respectively (**Figure 5B** and **Supplementary Table S7**). We also compared the variant frequency between the four oyster families for each of the common SNPs on the 148 positions spread all over the OsHV-1 µVar genome (**Figure 5C**). On these 148 common SNPs, only 12.2% (32) have the same variant frequency in the four families (difference in variant frequency $< 5\%$). A two by two comparison of the four families revealed that the percentage of SNPs having the same variant frequency ranked from 29% (43 SNPs for F9/F11) to 46% (68 SNPs for F32/F44) (**Supplementary Table S8**). These results show that the four families of oyster do not contain the same viral populations.

For the Med, we compared the two oyster families with susceptible phenotypes (F11 and F32) and observed that they contain a similar number of variants with 111 common and 18 and 39 specific to F11 and F32, respectively (**Figure 6A**). Variant frequency analysis of the 111 common SNPs revealed that they are equally distributed in the three groups of variants (**Figure 6B** and **Supplementary Table S9**). Indeed, there was 6.3 and 8.1% of minor variants in F11 and F32, respectively, 16.2 and 15.3% of intermediate variants in F11 and F32, respectively, and 77.5 and 76.6% of major variants in F11 and F32, respectively. We also compared the variant frequency between the two oyster families for each of the common SNPs on the 111 positions spread all over the OsHV-1 µVar genome (**Figure 6C** and **Supplementary Table S10**). On these 111 common SNPs, 47 have the same variant frequency (difference in variant frequency $< 5\%$). For the other SNPs, it is interesting to note that the variant frequency of 91% of them (58 out of 64) is higher in F11 than in F32. In addition, the number of fixed SNPs (variant frequency 100%) is almost twice higher in F11 (41) than in F32 (24). These results also showed that F11 and F32 do not contain the same viral populations in this environment.

Altogether, these results show that in as same environment, each oyster family is infected by distinct viral populations.

DISCUSSION

In the present study, we analyzed the OsHV-1 diversity in oysters with different genetic backgrounds exposed to two different environments (Atlantic and Mediterranean oyster farming areas during an episode of POMS). The two infection experiments led to contrasting survival phenotypes between oyster families within and between environments. As OsHV-1 infection is the first and necessary step to induce POMS (de Lorgeril et al., 2018) we focused our study on the diversity of OsHV-1 populations infecting oysters.

Viral sequences found in oyster tissues differed between environments with an overall number of 436 SNP variants identified. Sequences differed from the OsHV-1 µVar A genome (originating from the English Channel, Normandy coast, France) used here as reference (Burioli et al., 2017). Moreover, viral populations from the Atlantic and the Mediterranean differed (i) by 268 SNPs specific to the Atlantic environment, 59

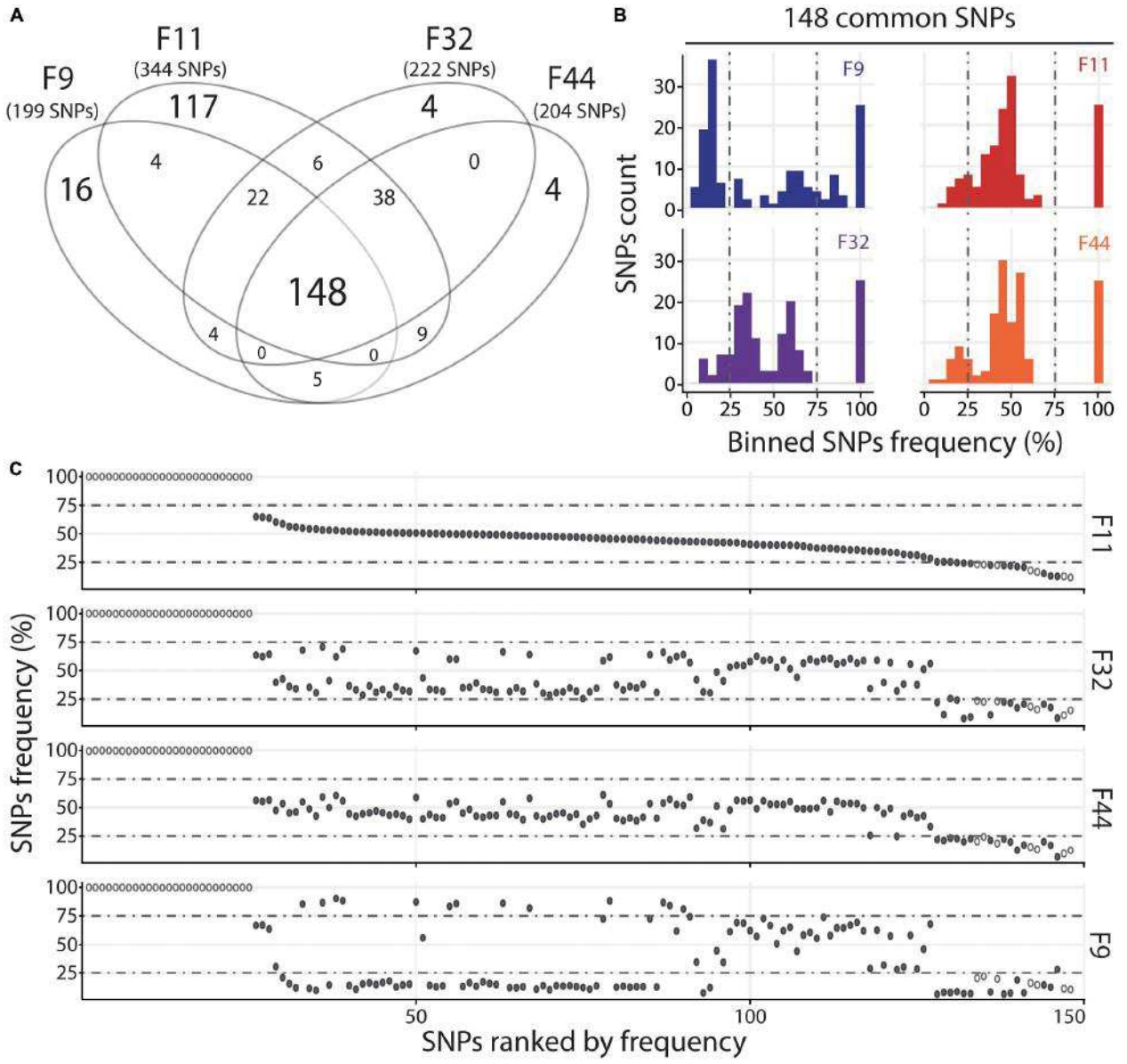


FIGURE 5 | Distinct heterogeneous viral populations infect oyster families during Atlantic experimental infection. **(A)** Venn diagram representation of the number of specific and common SNPs in viral populations infecting each susceptible or intermediate oyster families (F9/11/32/44). **(B)** The frequency distribution of common SNPs across the 4 infected families is illustrated by a histogram. SNPs are binned at 5% intervals along the *x*-axis, and the number of SNPs in each bin is indicated on the *y*-axis. **(C)** Common SNPs are ordered according to decreasing variant frequency in F11 family and plotted for the same position in F32, F44, and F9 families. Black and white dots correspond to SNPs with different or similar variant frequencies (difference < 5%), respectively.

specific to the Mediterranean environment, and also (ii) by variant frequencies within the 109 SNPs common to both environments. SNP abundance was twice higher in the Atlantic environment with a fixation rate seven times lower compared to the Mediterranean environment, which suggests that the viral populations infecting oysters in the Atlantic are more diversified than in the Mediterranean Sea. One possible environmental reason for this difference is that in the Atlantic, oyster farms are located in the Brest harbor, which is colonized by a huge population of wild oysters cohabitating with farmed oysters. In contrast in the Mediterranean Thau lagoon, farmed oysters represent the majority of the oyster population (Lejart, 2009; Pernet et al., 2012). The wild oyster populations of the Atlantic are likely to be reservoirs of viruses with a higher genetic diversity than the cultured populations of the Thau lagoon, as indicated by previous studies showing a lower viral diversity in hatchery−reared oysters relative to wild individuals (Vendrami et al., 2019). In full support of these observations, a higher host genetic diversity is expected to generate a higher viral diversity since viruses

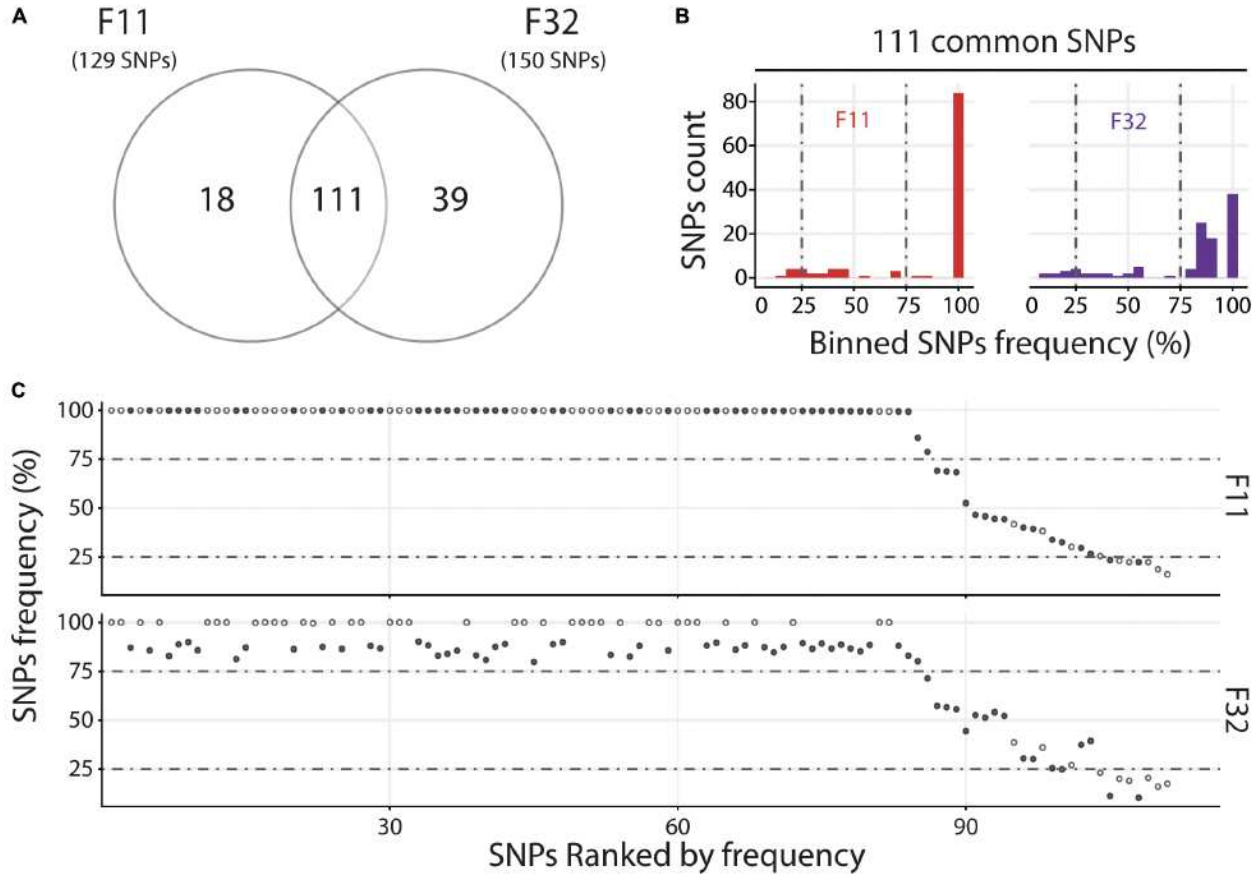


FIGURE 6 | Distinct heterogeneous viral populations infect oyster families during Mediterranean experimental infection. **(A)** Venn diagram representation of the number of specific and common SNPs in viral populations infecting each susceptible oyster families (F11/32). **(B)** The frequency distribution of common SNPs across the 2 infected families is illustrated by a histogram. SNPs are binned at 5% intervals along the *x*-axis, and the number of SNPs in each bin is indicated on the *y*-axis. **(C)** Common SNPs are ordered according to decreasing variant frequency in F11 family and plotted for the same position in F32 family. Black and white dots correspond to SNPs with different or similar variant frequencies (difference < 5%), respectively.

are obligatory cellular parasites, which implies that the virus and the host partners reciprocally affect each other's evolution (Kajan et al., 2019). Regardless of the origin of this diversity, our analysis revealed that the OsHV-1 μVar diversity along French coasts is higher than previously described (Burioli et al., 2018). This result is consistent with the high diversity also observed for this virus along the Italian Adriatic Sea coast (Burioli et al., 2016).

To further analyze the different viral populations infecting the different oyster families, we compared the OsHV-1 diversity between each oyster genetic background in a same environment (Atl or Med). We observed that (i) each oyster family had its own subset of specific variants and (ii) the majority of common variants displayed distinct frequencies between families. Therefore different families (with intermediates and/or susceptible phenotypes) are infected by distinct viral populations. This suggests that each biparental oyster family, with its own genetic background, is compatible with a subset of the viral variants. In other words, each oyster genotype could constitute a transmission bottleneck that may retain a part of the viral diversity present in the viral population. Although the role of genetic bottlenecks in virus transmission is well documented, their impact on viral genetic diversity and host-virus coevolution are difficult to predict (McCrone and Lauring, 2018). Indeed, it depends on numerous parameters like the virus type, the size of the initial viral population, the size of the bottleneck, as well as the relative impact of within and between host processes (Zwart and Elena, 2015). Genome-wide analyses recently highlighted the role played by inter- and intra-host herpesvirus genomic diversity in colonization of the host (new organ or host shift) and disease manifestations (Loncoman et al., 2017; Renner and Szpara, 2018; Akhtar et al., 2019). Defining the oyster and OsHV-1 factors that determine the transmission bottleneck would be an important step to understand the different levels of infectivity observed between the different oyster families. To reach this end, it would be necessary to study, within each individual oyster, the dynamics of OsHV-1 diversity.

In addition to the presence of distinct viral populations in the two environments that differentially infect each oyster

genetic backgrounds, we also observed that oyster families may display a different survival phenotype according to the infectious environment they experienced. Three oyster families F9, F32, and F44, which had an intermediate phenotype during the Atlantic experimental infection, were found susceptible or resistant during the Mediterranean experimental infection. This observation suggests that some oyster genotypes are more susceptible to some viral variants than to others. Several parameters such as the environment, viral genomic feature and host genetics might explain these differences in disease outcome.

By investigating SNPs found in viral populations specific to each environment, we uncovered some differences in viral proteins and more specifically in membrane-related proteins. In mammalian *herpesviridae* (i.e., cytomegalovirus) membrane-related proteins and particularly glycoproteins play a critical role in virus entry into host cell or escape of the host immunity and it has been proposed that the intense selection playing on these proteins could be driven by the interaction with the host immune system (Gardner and Tortorella, 2016; Lassalle et al., 2016). Here we found that 15 out of 19 ORFs encoding this type of proteins in the OsHV-1 genome, contain at least one nsSNP and most of these mutations are specific to Atlantic or Mediterranean viral populations. Two of these ORFs (ORF16 and ORF114) are particularly interesting since (i) they correspond to stop-gain mutations inducing large protein deletion and (ii) each of them is specific to a viral population. Previous studies already pointed out mutations in membrane proteins and suggested that they could be involved in the higher virulence of the μVar genotype compared to the reference OsHV-1 genotype (Martenot, 2013; Burioli et al., 2017; Burioli et al., 2018). Interestingly, two other ORFs encoding membrane-related proteins (ORF 25 and ORF 41), which are differentially mutated in Atlantic and Mediterranean viral populations, appeared to be involved in interaction between OsHV-1 and host cells (Martenot et al., 2019). We hypothesize that the mutations in membrane related proteins described in this study could have functional consequences on host-virus interactions and could be linked to the difference in susceptibility observed for the families F9, F32, and F44 when they were confronted to disease originated from Atlantic or Mediterranean environments.

The observed phenotypic heterogeneity could also be related to the genetic background of oysters. Numerous studies already showed the genetic basis for resistance to OsHV-1μVar in oyster (Azema et al., 2015, 2017; Camara et al., 2017; Gutierrez et al., 2018). Recent data further showed that oyster genetic selection for survival to one variant of OsHV-1 may not greatly increase survival to another variant (Divilov et al., 2019) suggesting that the genetic architecture of resistance to OsHV-1 infection in oyster is differentially structured according to viral genotypes involved.

CONCLUSION

In conclusion, the results obtained revealed the presence of OsHV-1 populations, distinct from the OsHV-1μVar A genome, are different in the Atlantic and Mediterranean environments but also differentially infect oyster families. These inter-environment and inter-family genetic variations might be indicative of a host-pathogen co-evolutionary process. Although OsHV-1 genotypic variations have been extensively described in the literature, it is the first time to our knowledge, that a potential genotype-genotype interaction is highlighted in oyster-OsHV-1 interactions. The next challenges will be to combine phenotypic measures of viral impact in experimental infections with whole viral genomes sequencing and comparative genomics techniques to make headway toward a better understanding of how specific genetic differences in OsHV-1 and oyster influence the outcome of infection and how this may shape the distribution of genetic variation within and between viral populations. This study highlights the need for future investigations on co-evolutionary dynamics of oyster-OsHV-1 interactions taking into account viral diversity, viral mutation rate through space and time, mechanisms of adaptation in changing environments that will have to be considered in future strategies to mitigate POMS and shape next oyster selective breeding programs.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the SRA database BioProject accession number PRJNA423079 with submission ID from SAMN08382938 to SAMN08382973, SAMN13818404 to SAMN13818448 and from SAMN13818479 to SAMN13818553.

AUTHOR CONTRIBUTIONS

JL, BP, CM, J-ME, CC, and GM were involved in the study concept and design. BP, YG, CM, JL, J-ME, and GM were involved in the collection of samples and in the experimental work. JD, RG, P-LS, and CC were involved in bioinformatic and statistical analyses. JD, RG, CM, DD-G, GM, and J-ME drafted the manuscript. All authors contributed to critical revisions and approved the final manuscript.

FUNDING

The present study was supported by the ANR projects DECIPHER (ANR-14-CE19-0023) by Ifremer; CNRS, Université de Montpellier; Université de Perpignan via Domitia. JD was supported by grant from the Université de Montpellier. This work was also supported by the labEx CeMEB, through the use of the "genomic environnementale" platform (http://www.labex-cemeb.org/fr/genomique-environnementale-1) and the exploratory research project HaploFit. We also thank the "Bio-Environnement" technological UPVD platform (Région Occitanie, CPER 2007-2013 Technoviv, CPER 2015-2020 Technoviv2). This study is set within the

framework of the "Laboratoires d'Excellences (LABEX)" TULIP (ANR-10-LABX-41).

ACKNOWLEDGMENTS

We thank the staff of the Ifremer station of Argenton (LPI, PFOM) and Sète (LER), and the Comité Régional Conchylicole de Méditerranée (CRCM) for technical support in the collection of oyster genitors, reproduction and transplantation experiments. We also thank Lionel Dégremont (Ifremer, SGMM) for providing us the genitors used to generate the F21 family. We also thank Marc Leroy for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01579/full#supplementary-material

REFERENCES

Abbadi, M., Zamperin, G., Gastaldelli, M., Pascoli, F., Rosani, U., Milani, A., et al. (2018). Identification of a newly described OsHV-1 mu var from the North Adriatic Sea (Italy). *J. Gen. Virol.* 99, 693–703. doi: 10.1099/jgv.0.001042

Akhtar, L. N., Bowen, C. D., Renner, D. W., Pandey, U., Della Fera, A. N., Kimberlin, D. W., et al. (2019). Genotypic and phenotypic diversity of herpes simplex virus 2 within the infected neonatal population. *mSphere* 4:e00590-18. doi: 10.1128/mSphere.00590-18

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arzul, I., Corbeil, S., Morga, B., and Renault, T. (2017). Viruses infecting marine molluscs. *J. Invertebr. Pathol.* 147, 118–135. doi: 10.1016/j.jip.2017.01.009

Arzul, I., Nicolas, J. L., Davison, A. J., and Renault, T. (2001a). French scallops: a new host for ostreid herpesvirus-1. *Virology* 290, 342–349. doi: 10.1006/viro.2001.1186

Arzul, I., Renault, T., and Lipart, C. (2001b). Experimental herpes-like viral infections in marine bivalves: demonstration of interspecies transmission. *Dis. Aquat. Organ.* 46, 1–6. doi: 10.3354/dao046001

Arzul, I., Renault, T., Lipart, C., and Davison, A. J. (2001c). Evidence for interspecies transmission of oyster herpesvirus in marine bivalves. *J. Gen. Virol.* 82, 865–870. doi: 10.1099/0022-1317-82-4-865

Arzul, I., Renault, T., Thebault, A., and Gerard, A. (2002). Detection of oyster herpesvirus DNA and proteins in asymptomatic *Crassostrea gigas* adults. *Virus Res.* 84, 151–160. doi: 10.1016/S0168-1702(02)00007-2

Azema, P., Lamy, J. B., Boudry, P., Renault, T., Travers, M. A., and Degremont, L. (2017). 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.* 49:23. doi: 10.1186/s12711-017-0297-2

Azema, P., Travers, M. A., De Lorgeril, J., Tourbiez, D., and Degremont, L. (2015). Can selection for resistance to OsHV-1 infection modify susceptibility to *Vibrio aestuarianus* infection in *Crassostrea gigas*? First insights from experimental challenges using primary and successive exposures. *Vet. Res.* 46:139. doi: 10.1186/s13567-015-0282-0

Bai, C. M., Morga, B., Rosani, U., Shi, J., Li, C., Xin, L. S., et al. (2019). 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., Wang, C. M., Xia, J. Y., Sun, H. L., 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

Batista, F. M., Arzul, I., Pepin, J. F., Ruano, F., Friedman, C. S., Boudry, P., et al. (2007). Detection of ostreid herpesvirus 1 DNA by PCR in bivalve molluscs: a critical review. *J. Virol. Methods* 139, 1–11. doi: 10.1016/j.jviromet.2006.09.005

Batschelet, E., Domingo, E., and Weissmann, C. (1976). Proportion of revertant and mutant phage in a growing population, as a function of mutation and growth-rate. *Gene* 1, 27–32. doi: 10.1016/0378-1119(76)90004-4

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Brooke, C. B. (2017). Population diversity and collective interactions during influenza virus infection. *J. Virol.* 91:e01164-17.

Burge, C. A., Shore-Maggio, A., and Rivlin, N. D. (2018). "Ecology of emerging infectious diseases of invertebrates," in *Ecology of Invertebrate Diseases*, ed. A. E. Hajek (Hoboken, NJ: John Wiley & Sons Ltd), 587–625. doi: 10.1002/9781119256106.ch16

Burioli, E. A. V., Prearo, M., and Houssin, M. (2017). Complete genome sequence of Ostreid herpesvirus type 1 mu Var 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

Burioli, E. A. V., Varello, K., Lavazza, A., Bozzetta, E., Prearo, M., and Houssin, M. (2018). A novel divergent group of *Ostreid herpesvirus* 1 mu Var variants associated with a mortality event in Pacific oyster spat in Normandy (France) in 2016. *J. Fish Dis.* 41, 1759–1769. doi: 10.1111/jfd.12883

Caceres-Martinez, J., Vasquez-Yeomans, R., Danigo, P., and Reyes-Roel, C. (2018). Histological alterations in Pacific oysters *Crassostrea gigas* that survived a summer mortality event in Baja California, Mexico. *J. Aquat. Anim. Health* 30, 31–38. doi: 10.1002/aah.10006

Camara, M. D., Yen, S., Kaspar, H. F., Kesarcodi-Watson, A., King, N., Jeffs, A. G., et al. (2017). Assessment of heat shock and laboratory virus challenges to selectively breed for ostreid herpesvirus 1 (OsHV-1) resistance in the Pacific oyster, *Crassostrea gigas*. *Aquaculture* 469, 50–58. doi: 10.1016/j.aquaculture.2016.11.031

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. doi: 10.1186/1471-2105-12-35

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., 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. doi: 10.4161/fly.19695

Comps, M., Bonami, J. R., Vago, C., and Campillo, A. (1976). Virus-disease of portuguese oyster (Crassostrea-Angulata Lmk). *Comptes Rendus Hebdomadaires Seances Acad. Sci. Ser. D* 282, 1991–1998.

da Silva, P. M., Renault, T., Fuentes, J., and 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. Organ.* 78, 181–188. doi: 10.3354/dao01874

Davison, A. J., Trus, B. L., Cheng, N. Q., 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

de Lorgeril, J., Lucasson, A., Petton, B., Toulza, E., Montagnani, C., Clerissi, C., et al. (2018). Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. *Nat. Commun.* 9:4215. doi: 10.1038/s41467-018-06659-3

de Lorgeril, J., Petton, B., Lucasson, A., Perez, V., Stenger, P. L., Degremont, L., et al. (2020). Differential basal expression of immune genes confers *Crassostrea gigas* resistance to Pacific oyster mortality syndrome. *BMC Genomics* 21:63. doi: 10.1186/s12864-020-6471-x

Degremont, L., Lamy, J. B., Pepin, J. F., Travers, M. A., and Renault, T. (2015). New insight for the genetic evaluation of resistance to ostreid herpesvirus infection, a worldwide disease, in *Crassostrea gigas*. *PLoS One* 10:e0127917. doi: 10.1371/journal.pone.0127917

Divilov, K., Schoolfield, B., Morga, B., Degremont, L., Burge, C. A., Cortez, D. M., et al. (2019). First evaluation of resistance to both a California OsHV-1 variant and a French OsHV-1 microvariant in Pacific oysters. *BMC Genet.* 20:96. doi: 10.1186/s12863-019-0791-3

Drake, J. W., and Holland, J. J. (1999). Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13910–13913. doi: 10.1073/pnas.96.24.13910

EFSA (2015). P. O. A. H. W. Oyster mortality. *EFSA J.* 13, 4122–4123.

FAO, and WHO (2018). *Technical Guidance for the Development of the Growing Area Aspects of Bivalve Mollusc Sanitation Programmes*. Rome: FAO.

Farley, C. A., Banfield, W. G., Kasnic, G. Jr., and Foster, W. S. (1972). Oyster herpes-type virus. *Science* 178, 759–760. doi: 10.1126/science.178.4062.759

Gardner, T. J., and Tortorella, D. (2016). Virion glycoprotein-mediated immune evasion by human cytomegalovirus: a sticky virus makes a slick getaway. *Microbiol. Mol. Biol. Rev.* 80, 663–677. doi: 10.1128/MMBR.00018-16

Guo, X., and Ford, S. E. (2016). Infectious diseases of marine molluscs and host responses as revealed by genomic tools. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150206. doi: 10.1098/rstb.2015.0206

Gutierrez, A. P., Matika, O., Bean, T. P., and Houston, R. D. (2018). Genomic selection for growth traits in Pacific oyster (*Crassostrea gigas*): potential of low-density marker panels for breeding value prediction. *Front. Genet.* 9:391. doi: 10.3389/fgene.2018.00391

Hwang, J. Y., Park, J. J., Yu, H. J., Hur, Y. B., Arzul, I., Couraleau, Y., et al. (2013). Ostreid herpesvirus 1 infection in farmed Pacific oyster larvae *Crassostrea gigas* (Thunberg) in Korea. *J. Fish Dis.* 36, 969–972. doi: 10.1111/jfd.12093

Jenkins, C., Hick, P., Gabor, M., Spiers, Z., Fell, S. A., Gu, X., et al. (2013). Identification and characterisation of an ostreid herpesvirus-1 microvariant (OsHV-1 micro-var) in *Crassostrea gigas* (Pacific oysters) in Australia. *Dis. Aquat. Organ* 105, 109–126. doi: 10.3354/dao02623

Kajan, G. L., Affranio, I., Bistyak, A. T., Kecskemeti, S., and Benko, M. (2019). An emerging new fowl adenovirus genotype. *Heliyon* 5:e01732. doi: 10.1016/j.heliyon.2019.e01732

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–U354.

Lassalle, F., Depledge, D. P., Reeves, M. B., Brown, A. C., Christiansen, M. T., Tutill, H. J., 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

Lauring, A. S., and Andino, R. (2010). Quasispecies Theory and the Behavior of RNA Viruses. *PLoS Pathog.* 6:e1001005. doi: 10.1371/journal.ppat.1001005

Le Deuff, R. M., Nicolas, J.-L., Renault, T., and Cochenec, N. (1994). Experimental transition of a Herpes-like virus to axenic larvae of the Pacific oyster *Crassosrtrea gigas*. *Bull. Eur. Ass. Fish Pathol.* 14:69.

Le Roux, F., Wegner, K. M., and Polz, M. F. (2016). Oysters and vibrios as a model for disease dynamics in wild animals. *Trends Microbiol.* 24, 568–580. doi: 10.1016/j.tim.2016.03.006

Lejart, M. (2009). *Etude Du Processus Invasif de Crassostrea gigas en Bretagne: Etat des Lieux, Dynamique Et Conséquences Écologiques*. Thesis, Universite de Bretagne Occidentale, Brest.

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

Loncoman, C. A., Vaz, P. K., Coppo, M. J., Hartley, C. A., Morera, F. J., Browning, G. F., et al. (2017). Natural recombination in alphaherpesviruses: insights into viral evolution through full genome sequencing and sequence analysis. *Infect. Genet. Evol.* 49, 174–185. doi: 10.1016/j.meegid.2016.12.022

Mao, Z. Q., He, R., Sun, M., Qi, Y., Huang, Y. J., and Ruan, Q. (2007). The relationship between polymorphisms of HCMV UL144 ORF and clinical manifestations in 73 strains with congenital and/or perinatal HCMV infection. *Arch. Virol.* 152, 115–124. doi: 10.1007/s00705-006-0826-8

Martenot, C. (2013). Variants of the ostreid herpesvirus-1 (OsHV-1) in the *Crassostrea gigas* oyster. *Virologie* 17, 81–87.

Martenot, C., Faury, N., Morga, B., Degremont, L., Lamy, J. B., Houssin, M., et al. (2019). Exploring first interactions between ostreid herpesvirus 1 (OsHV-1) and its host, *Crassostrea gigas*: effects of specific antiviral antibodies and dextran sulfate. *Front. Microbiol.* 10:1128. doi: 10.3389/fmicb.2019.01128

McCrone, J. T., and Lauring, A. S. (2018). Genetic bottlenecks in intraspecies virus transmission. *Curr. Opin. Virol.* 28, 20–25. doi: 10.1016/j.coviro.2017.10.008

Mortensen, S., Strand, A., Bodvin, T., Alfjorden, A., Skar, C. K., 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

Parvez, M. K., and Parveen, S. (2017). Evolution and emergence of pathogenic viruses: past, present, and future. *Intervirology* 60, 1–7. doi: 10.1159/000478729

Perales, C., Moreno, E., and Domingo, E. (2015). Clonality and intracellular polyploidy in virus evolution and pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8887–8892. doi: 10.1073/pnas.1501715112

Pernet, F., Barret, J., Le Gall, P., Corporeau, C., Degremont, L., Lagarde, 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. *Aquac. Environ. Interact.* 2, 215–237. doi: 10.3354/aei00041

Pernet, F., Lupo, C., Bacher, C., and Whittington, R. J. (2016). Infectious diseases in oyster aquaculture require a new integrated approach. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150213. doi: 10.1098/rstb.2015.0213

Petton, B., Bruto, M., James, A., Labreuche, Y., Alunno-Bruscia, M., and Le Roux, F. (2015). *Crassostrea gigas* mortality in France: the usual suspect, a herpes virus, may not be the killer in this polymicrobial opportunistic disease. *Front. Microbiol.* 6:686. doi: 10.3389/fmicb.2015.00686

Petton, B., Pernet, F., Robert, R., and Boudry, P. (2013). Temperature influence on pathogen transmission and subsequent mortalities in juvenile Pacific oysters *Crassostrea gigas*. *Aquac. Environ. Interact.* 3, 257–273. doi: 10.3354/aei00070

Pfeiffer, J. K., and Kirkegaard, K. (2005). Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog.* 1:e11. doi: 10.1371/journal.ppat.0010011

Poirier, E. Z., and Vignuzzi, M. (2017). Virus population dynamics during infection. *Curr. Opin. Virol.* 23, 82–87. doi: 10.1016/j.coviro.2017.03.013

Prado-Alvarez, M., Darmody, G., Hutton, S., O'Reilly, A., Lynch, S. A., and Culloty, S. C. (2016). Occurrence of OsHV-1 in *Crassostrea gigas* cultured in Ireland during an exceptionally warm summer. Selection of less susceptible oysters. *Front. Physiol.* 7:492. doi: 10.3389/fphys.2016.00492

Renault, T., Lipart, C., and Arzul, I. (2001). A herpes-like virus infects a non-ostreid bivalve species: virus replication in *Ruditapes philippinarum* larvae. *Dis. Aquat. Organ* 45, 1–7. doi: 10.3354/dao045001

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., and Novoa, B. (2004). Viruses infecting bivalve molluscs. *Aquat. Living Resour.* 17, 397–409. doi: 10.1051/alr:2004049

Renner, D. W., and Szpara, M. L. (2018). 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

Renzette, N., Gibson, L., Bhattacharjee, B., Fisher, D., Schleiss, M. R., Jensen, J. D., et al. (2013). Rapid intrahost evolution of human cytomegalovirus is shaped by demography and positive selection. *PLoS Genet.* 9:e1003735. doi: 10.1371/journal.pgen.1003735

Renzette, N., Gibson, L., Jensen, J. D., and Kowalik, T. F. (2014). Human cytomegalovirus intrahost evolution-a new avenue for understanding and controlling herpesvirus infections. *Curr. Opin. Virol.* 8, 109–115. doi: 10.1016/j.coviro.2014.08.001

Rubio, T., Oyanedel, D., Labreuche, Y., Toulza, E., Luo, X., Bruto, M., et al. (2019). Species-specific mechanisms of cytotoxicity toward immune cells determine the successful outcome of Vibrio infections. *Proc. Natl. Acad. Sci. U.S.A.* 116, 14238–14247. doi: 10.1073/pnas.1905747116

Segarra, A., Pepin, 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

Solomieu, V. B., Renault, T., and Travers, M. A. (2015). Mass mortality in bivalves and the intricate case of the Pacific oyster, *Crassostrea gigas*. *J. Invertebr. Pathol.* 131, 2–10. doi: 10.1016/j.jip.2015.07.011

Stentiford, G. D., Neil, D. M., Peeler, E. J., Shields, J. D., Small, H. J., Flegel, T. W., et al. (2012). Disease will limit future food supply from the global

crustacean fishery and aquaculture sectors. *J. Invertebr. Pathol.* 110, 141–157. doi: 10.1016/j.jip.2012.03.013

Vendrami, D. L. J., Houston, R. D., Gharbi, K., Telesca, L., Gutierrez, A. P., Gurney-Smith, H., 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. *Evol. Appl.* 12, 519–534. doi: 10.1111/eva.12736

Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E., and Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439, 344–348. doi: 10.1038/nature04388

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, 126–139. doi: 10.1016/j.aquaculture.2007.07-224

Wickham (2016). *Create Elegant Data Visualisations Using the Grammar of Graphics*. Available online at: https://www.rdocumentation.org/packages/ggplot2

Xia, J. Y., Bai, C. M., Wang, C. M., Song, X. L., 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

Xiao, Y. H., Rouzine, I. M., Bianco, S., Acevedo, A., Goldstein, E. F., Farkov, M., et al. (2016). RNA recombination enhances adaptability and is required for virus spread and virulence. *Cell Host Microbe* 19, 493–503. doi: 10.1016/j.chom.2016.03.009

Zwart, M. P., and Elena, S. F. (2015). Matters of size: genetic bottlenecks in virus infection and their potential impact on evolution. *Annu. Rev. Virol.* 2, 161–179. doi: 10.1146/annurev-virology-100114-055135

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 © 2020 Delmotte, Chaparro, Galinier, de Lorgeril, Petton, Stenger, Vidal-Dupiol, Destoumieux-Garzon, Gueguen, Montagnani, Escoubas and Mitta. 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.

Discussion

Dans cette étude, nous nous sommes concentrés sur l'analyse de la diversité virale lors d'infections expérimentales en mésocosme issus de deux environnements infectieux. Nous avons pu montrer que ces populations virales issues de milieux différents sont différentes entre elles. Ces populations possèdent des variations génétiques correspondant à la définition du génotype du µVar mais leurs génomes montrent des différences avec les génomes des OsHV-1 µVar A/B . Des polymorphismes nucléotidique unique (SNP) ont notamment été caractérisés dans des protéines virales annotées présentant des fonctions putatives de protéine membranaire ce qui pourrait avoir un impact sur la pathogenèse virale et pourrait être la signature génomique d'une coévolution avec l'huître.

Une stratégie d'analyse non optimale pour accéder à la diversité génétique des variants viraux

L'étude de la diversité génétique peut se faire avec plusieurs approches selon la question scientifique à poser. Dans notre cas, nous avons choisi l'approche d'**appel de variants** (*variant calling*) dans les données de HTS contre un génome de référence. Cette méthode est communément effectuée en bio-informatique et de nombreux outils sont disponibles depuis maintenant plus de 10 ans (Garrison and Marth, 2012; Koboldt et al., 2009; McKenna et al., 2010; Wang et al., 2010). Outre le fait qu'il s'agissait de données d'ARNm, l'un des problèmes était la difficulté à paramétrer la ploïdie en haploïde, c'est-à-dire que les fréquences alléliques ne sont pas des variables discrètes de 0, 0.5 et 1 mais dans notre cas, des variables continues de 0 à 1. Nous avons donc développé nous-même un outil d'appel de variants. Pour cela, nous avons donc utilisé l'outil pysamstats v1.1.2 (Miles, 2019) afin d'obtenir l'occurrence des nucléotides (A,C,T,G) pour chaque position du génome. En analysant ces données brutes, nous avons pu obtenir les fréquences alléliques des bases alternatives à travers la fonction `Variant_calling_pysamstats` du fichier *Figure 02ABCD-03 viral heterogeneous pop.Rmd*. Du fait, qu'il s'agissait d'échantillons de RNAseq et que les familles présentaient des phénotypes contrastés, le taux de réplication virale était différent entre les échantillons et l'ensemble des séquençages n'avait pas la même profondeur. Nous avons utilisé un facteur interne à la library utilisant une condition pour une profondeur supérieure à 10 reads et également supérieure à la moyenne additionnée à l'écart type de toutes les bases du génome. Ce type d'approche n'a malheureusement pas été correctement testé à l'aide de données simulées, cependant un contrôle semi-automatique des données brut d'alignement contre le génome a été réalisé aux positions ou il y avait des SNPs à l'aide du logiciel IGV et d'un script R pour limiter les cas de faux positif.

Une analyse impactée par la sensibilité de l'hôte

De nombreuses recherches ont mises en avant le fond génétique de l'hôte pour contrôler l'infection (de Lorgeril et al., 2018; Degremont et al., 2005; Samain et al., 2007; Sauvage et al., 2009). Lors des événements de mortalité chez les huîtres infectées, la quantité de données de séquençage est dépendante de la quantité d'ARN viraux dans les échantillons. Cette quantité de transcrits est elle-même dépendante du phénotype de l'hôte pour contrôler l'infection comme on peut le voir sur la **figure 1.E.F** de l'article. Nous nous sommes demandés si cette quantité de transcrits viraux impactait notre analyse de diversité génétique. En particulier lorsque l'on s'attarde sur la **figure 5.A** de l'étude présentant le nombre de mutations partagées ou non entre

les 4 familles F11/9/32/44. Nous observons que 117 sur 344 SNPs sont uniquement retrouvé dans la famille d'huître F11, la plus sensible, alors que pour les autres familles les SNPs uniques retrouvé est beaucoup moins important avec 16, 4 et 4 pour respectivement F9, F32 et F44. On peut donc se poser la question si dans le cas de cette famille avec un phénotype sensible le virus se réplique intensément et la diversité génétique mesuré et le reflet d'un processus biologique. Ou si nous n'avons pas séquencé assez profondément pour décrire la diversité virale de façon exhaustive. Pour répondre à cette question, nous avons par la suite adapté une analyse dite "de raréfaction génomique" à l'image des courbes de raréfaction en metabarcoding. Nous avons implémenté une méthode de tirage aléatoire incrémentative du nombre de reads alignés sur le génome de OsHV-1 µVar A suivi du variant calling. On s'attend alors à ce que le nombre de mutations converge vers un maximum. L'analyse des temps 24/48/60/72h de chaque famille dans les 2 IEM a permis de tracer ces courbes de raréfaction (**Fig.19. p58**).

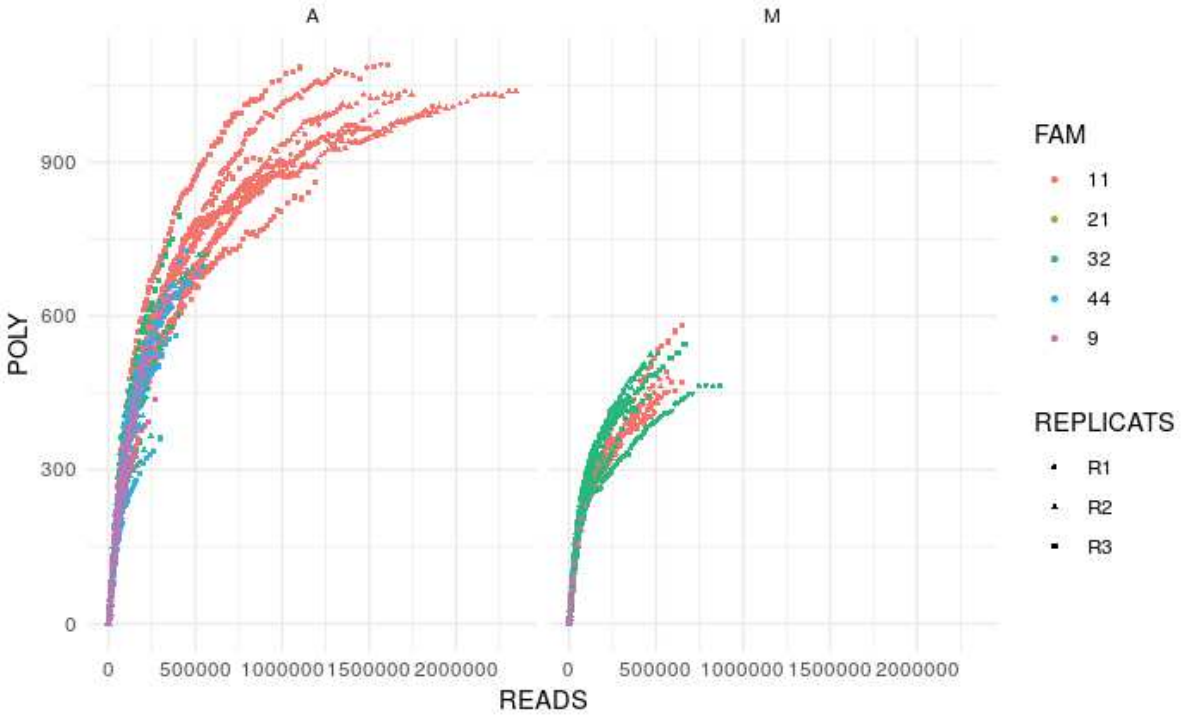


Figure 19. **Courbe de raréfaction virale**.
Les courbes de raréfaction ont été tracées pour les populations virales infectant les 5 familles. Le variant calling a été réalisé avec FreeBayes (Garrison and Marth, 2012) et l'axe des ordonnées représente l'ensemble des occurrences des SNPs. Chaque point est coloré en fonction des familles : F9 en violet, F11 en rouge, F21 en kaki, F32 en vert et F44 en bleu, et ceux dans les deux infections expérimentales en mésocosme, Atl à gauche et Med à droite. L'ensemble des temps a été traité simultanément. Les répliqua sont représentées par les différentes formes mais ne sont pas appariées.

Les résultats de cette figure montrent que nous n'atteignons le plateau pour aucun échantillon (**Fig.19. p58**). Par contre, nous pouvons observer que pour une même quantité de lectures, la diversité virale mesurée est plus faible dans les IEM issues de la zone Méditerranéenne que les IEM issues de la zone Atlantique. Ce résultat tend à montrer que la

diversité virale est plus restreinte dans cette zone. De plus, la quantité plus élevée de mutations observées dans la famille F11 en Atl par rapport aux autres échantillons pourrait s'expliquer par une profondeur de séquençage plus importante bien que les courbes n'atteignent pas le plateau. Lors de l'analyse de la diversité, nous avions fait le choix de mettre en commun les 3 derniers temps de la cinétique et les 3 répliquats pour augmenter le nombre de reads viraux. Cependant, il n'est pas certain que cette approche nous permette de capturer toute la diversité dans nos analyses comme le montrent les courbes précédentes. Ainsi, comparer les familles entre elles comme nous l'avons fait est biaisé par cette variation de profondeur de séquençage entre les échantillons. Ainsi la conclusion dans laquelle nous indiquons que l'abondance des SNPs était deux fois plus élevée dans le milieu atlantique avec un taux de fixation sept fois plus faible que dans le milieu méditerranéen, suggérant que les populations virales infectant les huîtres dans l'Atlantique sont plus diversifiées qu'en mer Méditerranée est donc à prendre avec parcimonie.

Les variations génétiques des transcrits viraux ne semblent pas façonner par le phénotype des huîtres

Avec les données issues du projet DECIPHER, nous avions l'opportunité de voir si les familles d'huîtres avec des phénotypes contrastés pouvaient influencer l'évolution des populations virales lors de l'infection. La meilleure approche pour mettre en évidence qu'elles peuvent être les mutations d'intérêt est l'étude statistique d'association (Bartha et al., 2013). Cette approche n'étant pas disponible, une analyse d'ensemble était envisageable. Pour se faire, en confrontant l'ensemble échantillons, l'intersection des mutations retrouvées dans 2 familles de 2 NEI différentes permettait de mettre en évidence les mutations façonnées par le fond génétique de l'hôte (**Fig.20**. **p60**). Seules les familles d'huîtres familles 11 (sensible) et 32 (sensible/Intermédiaire) pouvaient être analysées de la sorte. Nous constatons donc que la famille d'huître F11 induit l'apparition de 9 polymorphismes chez les populations virales issue des 2 environnements, tandis que la famille 32 n'induit pas l'apparition de polymorphisme spécifique. L'influence du fond génétique de l'hôte semble impacter faiblement les populations virales. Ce résultat ne remet pas en question la conclusion principale de l'étude qui est que les populations virales d'OsHV-1 issues d'environnements différents sont distinctes mais diminue la portée des hypothèses sur l'influence de la famille d'huîtres sur la diversité virale.

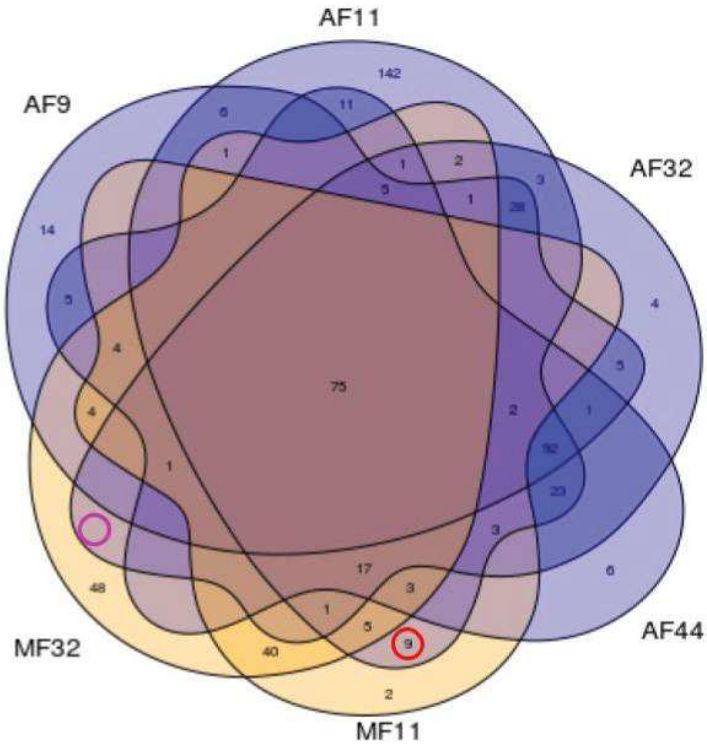


Figure 20. **Analyse d'ensemble complète pour mettre en évidence l'effet famille.**
Diagramme de Venn à 6 ensembles de variations génétiques observées par rapport au virus OsHV-1 µVar aux temps 48/60/72h avec la concaténation des répliqua. Les ensembles en beige représentent les infections expérimentales en mésocosme réalisées en Méditerranée (F11 et F32) et en bleu, ce sont les ensembles des infections expérimentales en mésocosme réalisées en Atlantique (F9/11/32/44). Les cercles représentent l'intersection des polymorphismes retrouvés dans une même famille d'huître pour les 2 environnements différents avec en rouge la famille F11 et en violet la F32.

L'origine des géniteurs des huîtres pourrait-il expliquer la différence de phénotype observé ?

Lors des infections expérimentales en mésocosme, nous avons observé que les phénotypes de résistance des huîtres de même fond génétique sont différents selon l'origine de la maladie. Trois types de phénotypes ont été observés : un phénotype sensible avec moins de 25% d'huîtres survivantes à la fin de l'IEM, d'autre un phénotype intermédiaire entre 25% et 75% de survivantes et enfin un phénotype résistant avec plus de 75% d'huîtres survivantes. Or le phénotype pouvait être différent pour les deux IEM. Nous avons donc essayé de comprendre quel facteur pourrait être à l'origine de ce changement de phénotype. Cette différence pourrait s'expliquer par l'âge des individus. Des études antérieures ont montré que la résistance à la maladie augmente avec l'âge et les deux IEM ont 2 mois de décalage (Azéma et al., 2017; de Lorgeril et al., 2020; Petton et al., 2015). La première IEM issue de la zone Atlantique a eu lieu en juillet 2015 alors que la seconde IEM en Méditerranée a eu lieu de fin septembre à octobre 2015. Toutefois, nos données indiquent que des familles deviennent plus sensibles à la maladie comme la F32 en Méditerranée suggérant un autre effet (**Figure 2B article**). Une autre possibilité pourrait être le fond génétique des géniteurs. Chaque famille d'huître utilisée dans cette étude est issue de fécondation *in vitro* à partir de géniteurs sauvages (un mâle et une femelle) collectés soit dans des zones conchylicoles soit dans des zones non conchylicoles (**Fig.21. p61**).

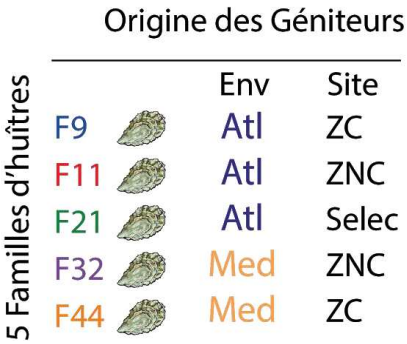


Figure 21. **Fonds génétiques des géniteurs utilisés pour la production des familles d'huîtres**.
Les cinq familles d'huîtres produites dans l'étude sont issues de géniteurs provenant de l'environnement Atlantique (Atl, bleu) ou de l'environnement Méditerrané (Med, orange). Ces géniteurs ont été recueillis dans des sites issus de zones conchylicoles (ZC), de zones non conchylicoles (ZNC) ou sont issus de programmes de sélections massales (Selec).

Nous pouvons faire l'hypothèse que les géniteurs issus de zones conchylicoles sont soumis à de fortes pressions de sélections dues au taux de mortalité récurrent et aux densités d'élevages élevées dans les fermes ostréicoles. À l'inverse, les géniteurs issus de zones non conchylicoles sont probablement moins en contact avec le virus OsHV-1 et donc moins susceptibles d'avoir passés des filtres de sélection imposés par le virus. Dans l'hypothèse où il y a une coévolution entre des variants du virus et le génotype de l'hôte, on peut faire deux hypothèses. Si (i) l'effet de la pression de sélection imposée par la culture intensive dans les zones conchylicoles permet d'augmenter la résistance de l'hôte (ii) et que des variants viraux existent et sont défavorisés dans la confrontation OsHV-1 / huître ; lorsque les huîtres sont issues de géniteurs du même environnement (sous-entendu que les huîtres sont mieux adaptées pour survivre à l'infection lorsque le virus vient du même endroit)[1] ; alors, on s'attend à ce que les huîtres issues des zones conchylicoles soient plus résistantes. Cette résistance serait encore améliorée lorsque les huîtres rencontrent des virus issus du même environnement que leurs géniteurs. Ce scénario correspond à ce qui peut être observé dans l'infection expérimentale de l'environnement méditerranéen.

Dans la IEM Med (**Figure 2B article**), on observe bien que les huîtres issues de géniteurs présents en zones conchylicoles (F44/F9) montrent un phénotype de résistance plus important et que ce phénotype est renforcé lorsqu'elles sont confrontées aux virus issus du même environnement que leurs géniteurs F44 (Med/ZC). Les deux familles ayant des géniteurs issus de zones non conchylicoles (F11/32) sont les plus sensibles et cette sensibilité est renforcée pour la famille F11 (Atl/ZNC).

Dans la IEM Atl (**Figure 2A article**), les deux hypothèses précédentes ne sont pas vérifiées on s'attendait à ce que la F9 (Atl/ZC) soit plus résistante que la F44 (Med/ZC), que la F11 (Atl/ZNC) soit moins sensible que la F32 (Med ZNC) or ce n'est pas le cas. Cela pourrait s'expliquer par un mécanisme plus complexe qu'une adéquation génotype/génotype comme l'implication d'autres facteurs impliqués dans la maladie.

La nature de nos données ne nous permet pas de statuer sur le rôle que peut avoir la diversité génétique du virus sur la pathogenèse. Les changements observés tels que la modification des séquences d'acides aminés des glycoprotéines de membranes peuvent être de bons candidats de signatures génomiques de coévolution entre le virus OsHV-1 et l'huître. Là encore, si ces variations sont le reflet d'une course évolutive, cela sous-entend un contact étroit entre le virus et son hôte. En effet, pour que cette course évolutive ait lieu il faut que l'hôte et le pathogène soient physiquement en contact, or le réservoir des populations virales n'est toujours pas connu à ce jour.

Conclusion

Ce premier travail de recherche a fourni une première image globale de l'étude de la diversité virale du virus OsHV-1 lors d'événements de mortalité. La production des données initiales sont celles qui ont permis de démontrer le rôle du virus dans la pathogenèse et n'étaient donc pas idéales pour caractériser la diversité génétique du virus. Ce travail a été l'occasion de développer de nombreux outils nécessaires à une meilleure évaluation de l'hétérogénéité des populations virales. Il a été un moteur important d'innovation dans l'analyse des données de séquençage et l'ensemble des remarques méthodologiques présentées ci-dessus sont retrouvées dans le second article (en préparation). Les résultats obtenus ont fait progresser nos connaissances sur la diversité de l'OsHV-1 µVar et constituent une étape vers de futures recherches sur le rôle de la variation génétique virale, sa relation avec le phénotype, la génétique de l'hôte et son potentiel impact sur l'issue de la maladie.

Chapitre III Aperçu de la diversité génomique de l'OsHV-1 en France lors d'événements de mortalité dans différentes zones de production

Le contexte des recherches

Avec l'essor croissant des technologies de séquençage, l'étude des génomes viraux constitue un apport déterminant pour comprendre les épidémies virales. Grâce à des approches d'épidémiologie moléculaire, l'origine, l'histoire évolutive et la dissémination géographique de plusieurs virus ont ainsi pu être reconstruites. Ces avancées ont considérablement augmentées les connaissances de virus émergent comme le virus de la grippe H1N1 (Rambaut et al., 2008), le virus de la dengue (Allicock et al., 2012), le virus ZIKA (Thézé et al., 2018), le virus du SIDA (Faria et al., 2014) ou dernièrement le Sars-CoV-2 (Popa et al., 2020). Ces approches sont pertinentes car les mutations s'accumulent dans le génome des populations virales, il est donc possible de faire des liens sur la propagation des événements qui ont façonné une épidémie et de dresser une carte de la propagation du virus. Jusqu'à présent aucune étude n'a regardé spécifiquement ces liens sur la répartition géographique d'OsHV-1 en France au niveau du génome entier. Les résultats précédents sur les relations phylogénétiques du virus ont principalement été conduits sur de courts fragments du génome du virus (**Encart 1 p22**) et s'intéressaient à la caractérisation des échantillons comme appartenants, ou non, au groupe des µVar (Burioli et al., 2016; Jenkins et al., 2013; Lynch et al., 2012; Martenot et al., 2015; Mineur et al., 2014; Peeler et al., 2012; Renault et al., 2012). Jusqu'à présent, quatre génomes d'OsHV-1 infectants les huîtres *C. gigas* ont été assemblés, à savoir le génome de référence (AY509253), les genomes de l'OsHV-1 µVar (µVar A KY242785.1 et µVar B KY271630.1) avec dernièrement l'OsHV-1-PT considéré comme un nouveau microvariant par ses auteurs (MG561751.2) (Abbadi et al., 2018; Burioli et al., 2017; Davison, 2005). Ces génomes montrent une diversité non négligeable, toutefois, ils ne nous renseignent pas sur la répartition géographique et sur la dissémination du virus. Dans la discussion de l'article précédent, nous avons fait l'hypothèse que les populations d'huîtres sont mieux adaptées aux virus issus du même environnement infectieux. Or, pour qu'il y ait une coévolution entre les agents pathogènes et leur hôte, il est nécessaire que les huîtres soient soumises à des filtres de sélection de virus issue du même

environnement régulièrement et vice versa (Kaján et al., 2019). Cela sous-entend que le virus OsHV-1 pourrait être endémique dans les zones conchylicoles, c'est-à-dire que le virus sévit de manière permanente à cet endroit. Certaines observations vont dans ce sens : la source du virus pourrait être les huîtres parentales et/ou les huîtres sauvages dans une localisation donnée et ainsi maintenir un réservoir (Barbosa-Solomieu et al., 2005; Corbeil et al., 2015; da Silva et al., 2008; López-Sanmartín et al., 2016; Petton et al., 2015; Renault et al., 2014). Toutefois d'autres formes de réservoirs ont été envisagées comme d'autres espèces hôtes (Arzul et al., 2001b; O' Reilly et al., 2017; Renault et al., 2001) et les sites conchylicoles peuvent avoir des spécificités dans cette maintenance de réservoirs. Une analyse plus approfondie de la diversité génétique de l'OsHV-1 avec des approches d'épidémiologie moléculaire pourrait donc nous fournir des informations sur l'origine et la dissémination géographique du virus.

Ce travail s'inscrit à nouveau dans l'étude de la diversité génétique du virus OsHV-1, le laboratoire IHPE a bénéficié du financement du projet HaploFit. Son objectif était de déchiffrer le rôle joué par les assemblages d'haplotypes d'OsHV-1 µVar sur l'aptitude globale du virus. Grâce aux données générées avec ce projet, et les innovations méthodologiques qui découlent du projet précédent, nous avons utilisé, pour la première fois, des données de séquençage d'ADN viral dans des huîtres infectées prises individuellement (**Fig.22. p64**). Ainsi, la stratégie et les méthodologies d'analyse sont beaucoup plus robustes, toutefois, l'analyse est beaucoup plus complexe. De cette manière, nous pouvons quantifier la variation entre les échantillons et entre les environnements. Cela nous a permis d'obtenir des informations sur le réservoir géographique des populations virales en France avec des approches inspirées de l'épidémiologie moléculaire. D'un point de vue plus fondamental ces résultats nous donnent l'occasion d'évaluer l'approche quasi exclusive de l'étude phylogénétique au travers des séquences moléculaires consensus en y intégrant l'analyse des variants minoritaires. C'est à dire la coexistence dans une meme population, à côté du virus dominant, de virus génétiquement distinct avec une plus faible fréquence allélique (inférieur à 50%).

L'objectif de l'étude

L'objectif de l'étude est de caractériser la diversité génomique des populations virales infectant les huîtres, au niveau individuel, et ainsi améliorer la compréhension de la dissémination du virus dans les zones conchylicoles françaises.

L'illustration du matériel et méthode

Le plan expérimental de l'étude est illustré ci-dessous (**Fig.22. p64**). Dans le cadre de la caractérisation de la diversité génétique du virus, nous avons utilisé des naissains standardisés IFREMER (NSI) qui n'ont pas un fond génétique homogène. Elles ne sont pas issues de fécondations biparentales mais de fécondations massales avec une 50aine de géniteurs afin de ne pas introduire un goulot d'étranglement sur la prolifération des populations virales. Les animaux ont été introduits dans trois zones conchylicoles en période d'événements de mortalité en 2018. Le séquençage de l'ADN des virus a été réalisé sur les échantillons individuellement pour pouvoir analyser les variations inter-individuelles et de mieux caractériser la diversité génétique du virus. Le pipeline d'analyse bioinformatique peut être retrouvé dans dans l'annexe 3 (p129) et une partie du code dans l'annexe 4 (p139).

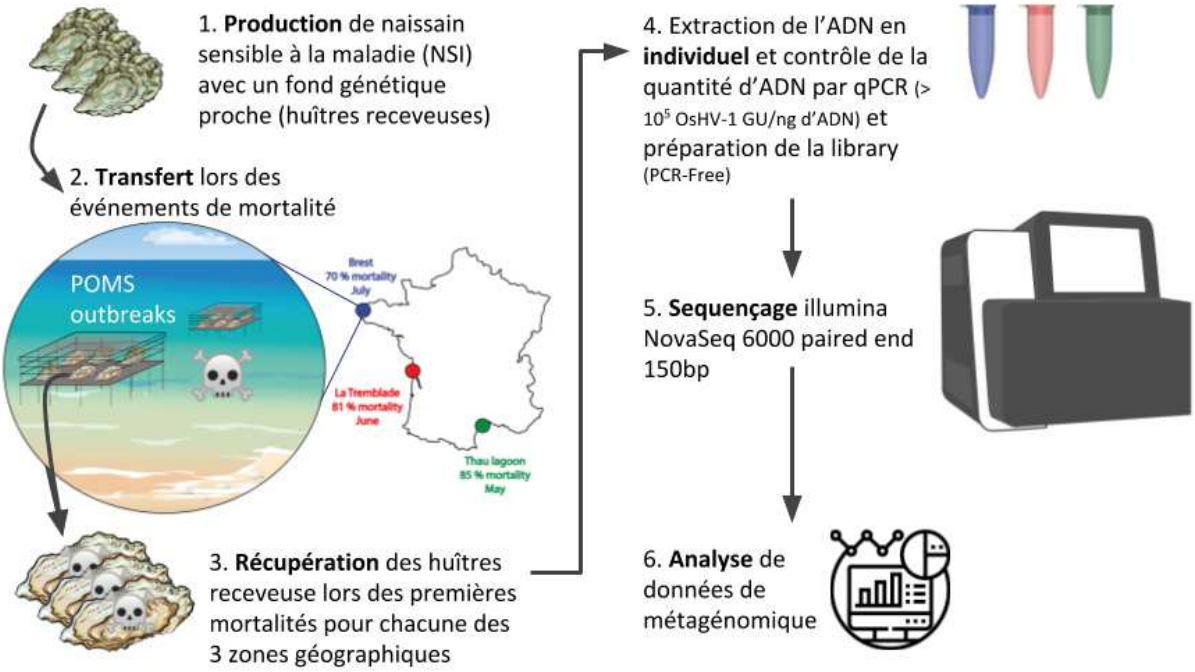


Figure 22. **Illustration du plan expérimental utilisé dans l'article II**.
(**1**) Des huîtres NSI ayant un fond génétique hétérogène ont été produites dans les écloseries ifremer. (**2**) Ces huîtres ont été placées en milieu naturel (dans l'environnement) dans des (fermes) parcs ostréicoles de Brest (bleu), de Marennes-Oléron (rouge) et de l'étang de Thau (vert) lors d'événements de mortalité. (**3**) Les huîtres moribondes ont ensuite été ramenées au laboratoire. (**4**) L'ADN a été extrait du manteau de plusieurs huîtres moribondes et la charge virale de chaque huître a été quantifiée par qPCR. (**5**) Lorsque la charge virale était suffisamment élevée (≥ 10^5 UG/ng d'ADN), les échantillons ont été envoyés à la plateforme de séquençage et séquencés en technologie Illumina NovaSeq 6000. (**6**) Enfin ces données ont permis l'analyse de la diversité des populations virales individuellement entre les 3 zones conchylicoles.

Article 2 : Aperçu de la diversité génomique de l'OsHV-1 en France lors d'événements de mortalité dans différentes zones de production

Article soumis à *Journal of Virology* et déposé dans bioRxiv, doi : 10.1101/2021.04.30.442107

Jean Delmotte, Camille Pelletier, Benjamin Morga, Richard Galinier, Bruno Petton, Jean-Baptiste Lamy, Oliver Kaltz, Jean-Christophe Avarre, Caroline Montagnani, Jean-Michel Escoubas

Contributions des auteurs : la collecte d'échantillons et les travaux expérimentaux ont été réalisés par J-ME et CM en collaboration avec l'équipe IFREMER SG2M de La Tremblade. L'ensemble de la stratégie d'analyse et la conception initiale de la trame du papier a été réalisé par JD. Le manuscrit a été rédigé principalement par JD et J-ME avec le soutien de CM et JCA. Tous les auteurs ont contribué aux révisions critiques et ont approuvé le manuscrit final.

Phylogeography of infectious disease: genetic diversity and connectivity of the Ostreid herpesvirus 1 population in France

Running title: Phylogeography of *Ostreid herpesvirus 1*

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], Caroline Montagnani[1]#, Jean-Michel Escoubas[1]#

[1]IHPE, Univ. Montpellier, CNRS, Ifremer, UPVD, F-34095 Montpellier, France

[2]Ifremer, RBE-SG2M-LGPMM, Station La Tremblade, F-17390 La Tremblade, France

[3]IHPE, Univ. Montpellier, CNRS, Ifremer, UPVD, F-66000 Perpignan, France

[4]Ifremer, LEMAR UMR 6539 (Université de Bretagne Occidentale, CNRS, IRD, Ifremer), F-29840 Argenton-en-Landunvez, France

[5]ISEM, IRD, CNRS, EPHE, Univ. Montpellier, F-34095 Montpellier, France

[#]Corresponding authors: Caroline.Montagnani@ifremer.fr; Jean.Michel.Escoubas@ifremer.fr

Abstract word count: 208 Importance word count : 150 Text word count: 4104

Abstract

The genetic diversity of viral populations is a key to understanding ther phylogeographic and dissemination history of viruses, but studying the diversity of whole genomes from natural populations remains a challenge. Molecular ecology approaches are commonly used for RNA viruses harboring small genomes, but have only rarely been applied to DNA viruses with large genomes. Here, we used the Pacific oyster mortality syndrome (POMS, 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 new bioinformatics methodology, we *de novo* assembled 21 OsHV-1 genomes. Combining whole-genome comparisons with phylogenetic analysis and quantification of major and minor variants, 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. Here, we demonstrate that molecular ecology approaches can be applied to large-genome viruses to determine the extent of their genetic diversity and better understand the spread of viral populations in natural environments.

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 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 (dsDNA) genomes generally have lower nucleotide diversity than RNA viruses. In this study, by combining the use of both major and minor variants with phylogeographic analyses of the oyster herpesvirus OsHV-1, we highlight genealogical relationships that have not been 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.

Keywords Herpesvirus; OsHV-1; cupped oyster; phylogeography; non-redundant genomes; genetic diversity; viral spread; minor variant

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 (1), 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 variants, known as viral quasispecies. This genetic organization currently encumbers our understanding of viral diseases and their evolution and impedes the straightforward characterization of virus population structure (2, 3). Studies have shown that the level of genetic diversity within these viral populations likely influences viral pathogenicity, dissemination and host immune evasion, and 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 certain large DNA viruses, such as herpesviruses, whose genomic variability rivals that of many RNA viruses (4-6). For example, human cytomegalovirus shows considerable inter-host and intra-host genetic divergence across tissue compartments and times of infection (7, 8). In addition, the evolution of a disease is sometimes not fully explained by intrinsic host factors, but by the genotypic diversity of herpesviruses (9).

Aquaculture is one of the fastest growing food-producing sectors, representing an important animal protein supply for human 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 (10-12). One illustration is the Pacific oyster mortality syndrome (POMS), which threatens global *Crassostrea gigas* oyster production, a main sector in aquaculture worldwide (reviewed in (13)). Since 2008, this syndrome has caused mass mortality in cultivated oysters around the world, from Europe to America and Asia (14-22). 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) (23-27). Two major genetic factors seem to affect the severity of POMS: the genetic background of the oysters (28-30) and OsHV-1 genetic diversity (31, 32). Since the characterization of the emergent genotype OsHV-1 µVar, associated with the 2008 high mortality events (27), several variants of the µVar genotype have been described (reviewed in (16)). 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 (33). 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 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 (34) 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 (35). To date, no study has specifically looked at the links between the geographic distribution of OsHV-1 and its genetic diversity at the level of the whole genome. To date, the entire 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) (14, 34, 35). These four assemblages show significant genomic diversity; however, they tell us very little about the geographic distribution and dissemination of the virus.

As shown with the H1N1 influenza virus (36), the dengue virus (37), the Zika virus (38), the AIDS virus (39) and more recently SARS-CoV-2 (40), 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 of OsHV-1 µVar in natural viral populations both within and between host individuals, to better understand the dissemination and evolution of OsHV-1 µVar populations. Using a deep-sequencing approach conducted at the individual level and novel bioinformatics analyses, we identified 21 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. This phylogeographic approach led us to propose a scenario explaining the phylogenetic relationships between viral populations encountered in the three farming areas. This analysis

strategy, combined with epidemiological information, holds promise for understanding OsHV-1 evolution, for revealing how the genetic relationships shape epidemic dynamics and for identifying transmission patterns between oyster-producing regions. These data are ultimately useful for developing novel disease management strategies.

Results

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. The cumulative moralities for the three batches of oysters were 85%, 81% and 70% for Thau (Th), Marennes-Oléron (Mo) and Brest (Br), respectively (Fig. 1A). 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. and Table S1). To evaluate OsHV-1 sequencing depth in each sample, sequencing reads were mapped on the OsHV-1 µVar A genome (KY242785.1) (35). 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%) for Mo ind3 to 11.94 M reads (13.21%) for Th ind10 (Fig.1B, Table S1). Therefore, OsHV-1 genome coverage varied from 480X to 8745X for Mo ind3 and Th ind10, respectively.

A rarefaction analysis was run 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 X (Fig. 1C). Since the number of mapped reads of each sample was clearly above this threshold,

144 sequencing depth was considered sufficient to capture OsHV-1 genomic diversity.
145 Rarefaction curves indicate that viral diversity was higher in infected Mo oysters than in the
146 other two areas. Comparison of the average number of individual polymorphisms per
147 environment showed that i) the number of SNPs in viruses from Mo (170 ± 27.7) was
148 significantly higher than that of viruses from Br (136.4 ± 5.9; $P \leq 0.001$) and Th (143.8 ± 5.8;
149 $P \leq 0.05$) and ii) the number of InDels in viruses from Mo (35.9 ±3.1 SD) was also
150 significantly higher than that of viruses from Th (25.7 ± 0.9; $P \leq 0.0001$) (Fig. 1D, Table S2).
151 In addition, in Mo, viral diversity greatly varied from one individual to the other; for example
152 Mo ind3 contained 222 SNPs and 42 InDels, whereas Mo ind4 contained only 144 SNPs and
153 33 InDels.

154 **2) Phylogenetic analysis of non-redundant OsHV-1 genomes reveals that all three**
155 **OsHV-1 populations belong to the µVar genotype.**

156 We assembled non-redundant genomes so as to reduce the complexity of the OsHV-1
157 genome and keep only one copy of each repeated region in the virus sequence. The OsHV-1
158 µVar A genome contains two unique regions and three repeats: the unique long region (UL,
159 164,268 bp), flanked by two inverted repeats (repeat long [RL], 7338 bp), and the unique
160 short region (US, 3370 bp), flanked by two inverted repeats (repeat short [RS], 9777 bp), and
161 two X regions (1510 bp) (Fig 2A, upper panel). Considering the high sequence identity
162 within each set of repeats (between 99.9 and 100%), removing one of the repeats makes it
163 possible to reliably map the reads from these repeated regions without affecting the overall
164 genome composition. Removing copy of each repeated region led to a non-redundant OsHV-
165 1 µVar A genome (NR-genome) of 186,262 bp, which represents 91% of the full-length
166 genome and contains all the coding sequences (Fig 2A, lower panel). Using this approach, 21
167 NR-genomes were assembled from the sequencing data generated during this study, as well

as 7 other published genomes belonging to the *Malacoherpesviridae* family: six *Ostreid herpesvirus 1* (genus *Ostreavirus*) and one *Haliotid herpesvirus 1* (abalone herpesvirus (AbHV-1), genus *Aurivirus*).

A phylogenetic analysis of these 28 NR-genomes revealed three main polyphyletic clusters (Fig. 2B). AbHV-1, the sole herpesvirus isolated from a gastropod, was identified as an outgroup, whereas the two herpesviruses isolated from bivalves other than oysters (OsHV-1-SB isolated from the clam *Scapharca broughtonii* and OsHV-1 ZK0118 isolated from the scallop *Azumapecten farreri*) formed two separate branches. Among the viruses isolated from *C. gigas*, the 21 new OsHV-1 isolates were phylogenetically closer to OsHV-1 µVar A and B isolates (Fig. 2B cluster 3) than to the OsHV-1 reference genome (AY509253.2), suggesting that they belong to the µVar genotype. Th and Br isolates grouped into two distinct clusters reflecting their geographic origin (Fig. 2B, cluster 1 and 2). Conversely, Mo OsHV-1 isolates were distributed in these two clusters: five in cluster 1 and two in cluster 2. The phylogenetic distance within the Th or Br isolates (average $5.33\ 10^{-6} \pm 1.99\ 10^{-6}$ SD and $2.16\ 10^{-5} \pm 1.43\ 10^{-5}$ SD, respectively) was lower than that observed within the Mo isolates (average $6.16\ 10^{-4} \pm 5.29\ 10^{-4}$), indicating greater viral diversity in Mo than in the other two farming areas.

3) Comparative genomics of the 21 non-redundant genomes confirms that OsHV-1 diversity is higher in Marennes-Oléron than in the other two farming areas.

To compare the genetic diversity at the whole genome scale between the 21 samples, we used the 21 NR-genomes to construct a consensus non-redundant genome (C-NR-genome). Comparison of the 21 NR-genomes to the C-NR-genome identified a total of 399 variable positions (SNPs and InDels), hereafter called variants. The average number of variants per genome was significantly higher in samples from Br (159.8 ± 4.9) than from Th (125 ± 5.6; *P*

191 < 0.01), whereas the average number of variants in Mo (139.4 ± 20.8) was not significantly
192 different from the other two farming areas (Fig. 3A, Table S3).
193 However, viruses from Mo showed the highest diversity, with 321 variants accounting for
194 80.5% of the total diversity. Among the variants, 151 (47%) were specific to Mo, and 87
195 were observed in one sample only (singletons) (Fig. 3B, Table S3). None of these 151
196 variants were shared by all the samples from this area. The majority of Mo singleton variants
197 were observed in two samples: ind10 (49) and ind4 (13) (Fig. 3C). For the Th and Br farming
198 areas, 134 and 168 variants were identified, among which 51 and 24 were area-specific,
199 respectively (Fig. 3B, Table S3). In contrast to Mo, 70.8% (17/24) and 94.1% (48/51) of the
200 variants were common to all the samples of Br and Th, respectively (Fig. 3B, Table S3).
201 The three variants common to Br and Th were present in almost all individuals of the two
202 farming areas (Fig. 3C, green arrows). Namely, among the 90 variants shared by Br and Mo,
203 78 (86.7%) were shared by two Mo individuals (ind10 and ind4) and all Br individuals (Fig.
204 3C, purple arrow). These two Mo individuals also shared six variants with all Th individuals
205 (Fig. 3C, red arrows).

206 **4) Diversity of minor OsHV-1 variants is higher in Marennes-Oléron than in the**
207 **other two farming areas.**

208 To fully characterize OsHV-1 genetic diversity, we performed variant calling analyses,
209 using the C-NR-genome as the reference, to quantify both major (frequency >50%) and
210 minor (frequency <50%) variants in the 21 individual samples. Among the 399 variants
211 identified as major variants, 117 had variable frequencies (above or below 50%) across
212 samples; moreover, 129 additional minor variants were identified (Fig. 4A).
213 Frequencies of the 129 strictly minor variants ranged from 5.1% to 45.9% (mean =
214 8.75% ± 5.75% SD). The number of minor variants was lower in oysters from Br (39) than in

oysters from Th (71) or Mo (75) (Fig. 4B, Table S4). Twenty-five of these 129 minority variants (19.3%) were found in all three farming areas. Moreover, the majority of site-specific variations were found in only one individual oyster (singletons).

A hierarchical clustering analysis applied to the minor variants showed that individuals did not cluster according to their geographic origin, except for individuals 4, 6 and 10 from Th (Fig. 4C). In some cases, we even observed two or three individuals from different farming areas clustering together (cluster 1, 2 and 3 in Fig. 4C), suggesting that some OsHV-1 found in these individuals had a common origin. Figure 4C also revealed that most of the 25 variants shared by three farming areas were found in all sampled individuals, suggesting that they had not arisen within individual oysters, but had been transmitted horizontally to other oysters.

5) Some minor OsHV-1 variants from the Marennes-Oléron farming area are major variants in the other two farming areas.

The 116 variants that had variable frequencies across samples (Fig. 4A) were present as minor variants 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 114 of these minor variants. Sample Mo ind10 contained 29 minor variants (with a frequency ranging from 5 to 10%; Fig. 5B) that were major variants in samples Mo ind1/3/7/8/9. For instance, the variant at position 65,016 (T>A) of the C-NR-genome had a frequency of 7% in Mo ind10, whereas its frequency varied from 76% to 99% in Mo ind1/3/7/8/9 (Fig. 5C). Similar fluctuations in variant frequency were observed for the other 28 variants (Table S5). These 29 variants highlighted a clear genetic link 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 85 minor variants that corresponded to major variants in Mo ind10, 38 that corresponded to major variants in Mo ind4 and all Br samples, and 4 that corresponded to major variants in all Th samples (Fig. 5 A and E). For instance, the variant at position 57,091 (T>C) in the C-NR-genome had a frequency of 24% in Mo ind3, whereas its frequency was 93% in Mo ind10 and 99% in all Th samples (Fig. 5F, Table S5). Another example is the variant at position 65,624 (C>T) in the C-NR-genome: its frequency grew from 25% in Mo ind3 to 93% in Mo ind10, and 99% in Mo ind4 and all Br samples (Fig. 5F, Table S5). Again, minor variants highlighted genetic links between samples from different farming areas and from phylogenetic clusters 1 and 2 (Fig. 5G, Table S5).

Altogether, the analysis of minor variants 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 strongly suggest that all the OsHV-1 variants characterized in this study originated from the Mo farming area.

Discussion

The accurate description of viral genetic diversity (major and minor variants) 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 (41). 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 moribund oysters. To optimize the characterization of OsHV-1 variants, 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 ≈3000X. However, because OsHV-1 coverage varied among samples (ranging from 480X to 8745X), 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 the species richness in a community (42), 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 (43, 44).

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% of the genome of OsHV-1 µVar A. To solve this problem, we used the approach proposed by Morse *et al.* (45): contigs generated by *de novo* assembly were ordered using a reference-based approach to ultimately assemble non-redundant genomes in which only one copy of each repeated element was kept. These non-redundant 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 variants 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 herpesviruses, e.g. the Cyprinid herpesvirus 3 (4). They also suggest that all OsHV-1 characterized during this study share a common ancestor with OsHV-1 µVar characterized by

287 Burioli *et al.* in Normandy (France), reinforcing the idea that genotypes closely related to
288 µVar have replaced the reference genotype and are predominant in oyster-farming areas
289 along the French coasts (35, 46). Our findings also showed that OsHV-1 µVar populations
290 are much more diverse than previously thought. Phylogenetic analyses corroborated results
291 from Delmotte *et al.* showing that OsHV-1 µVar from the Bay of Brest and Thau Lagoon
292 constitute two distinct viral populations (31). Clearly, it is now necessary to change the
293 common practice of relying on the generation of consensus genomes to study the
294 epidemiology of DNA viruses, because they do not accurately depict the actual genetic
295 diversity (5). For this reason, we developed this novel approach, applicable to any other DNA
296 virus model.

297 Our approach demonstrated that OsHV-1 diversity in Marennes-Oléron Bay was
298 higher than in the Bay of Brest and Thau Lagoon. This variability may be related to the fact
299 that the Marennes-Oléron Bay has become one of the largest French oyster-farming areas,
300 with intensive shellfish cultivation and chronic overstocking (47). Moreover, a network
301 analysis of oyster movements revealed that oysters are moved up to 9 times during their
302 production cycle, with peaks of transfers in spring and autumn. These aquacultural practices
303 likely contribute to the spread of OsHV-1 infection, and in turn increase its genetic diversity
304 (48).

305 Our approach also made it possible to identify major and minor variants, and variants
306 with frequencies varying dramatically among sampled oyster individuals. Thus, ultra-deep
307 sequencing can support the reconstruction of the true structure of a viral population, and can
308 even highlight potential genetic links between individuals from the same farming area or
309 between individuals from the three farming areas (Fig. 5 D and G, respectively). These
310 genetic links, combined with whole-genome phylogeny, can be used to infer the dynamics of
311 dispersal of OsHV-1 populations across the three farming areas (Fig. 6). Samples from

312 Marennes-Oléron have the most heterogeneous viral populations with the fastest evolution
313 rate. Moreover, major variants characterized in the other two farming areas were also present
314 in Marennes-Oléron, as minor variants. This suggests that Marennes-Oléron farming area is
315 the main source of OsHV-1 viral diversity, from where variants are dispersed to the other
316 farming areas. This hypothesis is in agreement with what is known about oyster transfers.
317 Indeed, Marennes-Oléron is the main area of *C. gigas* spat collection, and juvenile oysters are
318 then transferred to growing sites along the coasts of Brittany, Normandy and also along the
319 Mediterranean (47). Importantly, the three studied sites are separated by long distances (>500
320 km) and therefore are not directly connected through ocean currents. OsHV-1 thus likely
321 disseminates via "travel during transportation of infected oysters, similar to the spread of
322 other diseases during livestock transfer" (*e.g.*, (49)). Viral populations originating from the
323 Marennes-Oléron Bay are then subsequently shaped by local evolutionary and ecological
324 processes in the different farming areas. The next challenge is to characterize these
325 evolutionary and ecological processes that affect OsHV-1 genetic diversity and, in turn, their
326 impact on OsHV-1 virulence in the different farming areas. Given that the fate of new genetic
327 variants is largely determined by host selection, one hypothesis is that oyster populations
328 located in the Bay of Brest and Thau Lagoon exert a bottleneck on viral diversity imported
329 from Marennes-Oléron. However, several studies conducted in Europe to document oyster
330 genetic variability and population structure have revealed a high level of genetic diversity,
331 but no genetic differentiation between French populations (50-52). This lack of
332 differentiation suggests that selection pressure exerted by oysters is not sufficient to explain
333 differences in viral genetic diversity between farming areas. Genetic drift also influences the
334 probability and rate by which alleles increase or decrease in frequency in a viral population;
335 drift results in the loss of genetic diversity, because only a subset of the population
336 contributes to the next generation (41). Furthermore, the relationship between environmental

337 factors and mortality events in Pacific oysters has been well documented (reviewed in (53)).
338 Environmental factors such as temperature, food availability, water quality and salinity or UV
339 radiation have been identified as risk factors that directly or indirectly affect POMS
340 dynamics. Given the variation in these factors among the studied oyster farming areas,
341 identifying the effects of these factors on the structure of OsHV-1 µVar populations may
342 shed new light on the processes of variant selection.

343 Combining data on OsHV-1 viral population structure using our new genomics
344 pipeline with robust epidemiological information from the different farming areas can offer
345 valuable insight into the dynamics of POMS infection (routes of transmission, potential
346 reservoirs) and thus ultimately help implement effective integrated viral disease management
347 strategies (54).

348

349 **Figure legends**

350 **Fig. 1: Variability in sequencing depth does not impair the analysis of OsHV-1 diversity**.
351 (**A**) Sampling locations of the Ostreid herpesvirus-1 (OsHV-1) infected oysters during the
352 Pacific oyster mortality syndrome (POMS) event in 2018. Each color represents a farming
353 area, with Brest (Br) in blue (5 individual samples), Marennes-Oléron (Mo) in red (7
354 individual samples) and Thau Lagoon (Th) in green (9 individual samples). Below each
355 farming area, the percentage of mortality within the site is given. (**B**) Quantification of the
356 number of reads in the library and the number of reads mapping to the OsHV-1 µVar A
357 genome. Numbers on the right side of the figure indicate the percentages of reads mapping to
358 the OsHV-1 µVar A genome. (**C**) Rarefaction curves using the OsHV-1 µVar A genome as
359 the reference sequence. The top panel shows the number of single nucleotide polymorphisms
360 (SNPs) and the bottom panel, the number of insertion-deletions (InDels). The dashed vertical
361 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, $*P \leq 0.05$, $**P \leq 0.01$ **** $P \leq 0.0001$).

Fig. 2: Phylogenetic relationship between the 21 *de novo* assembled non-redundant genomes and other *Malacoherpesviridae* viruses. (**A**) Full-length (upper panel) and non-redundant (NR) (lower panel) genomes of OsHV-1 µVar A (KY242785). The OsHV-1 µVar A genome consists of two unique regions, referred to as unique long (UL; in red) and unique short (US, in blue), and three repeat regions, referred to as repeat long (RL, in yellow), repeat short (RS, in purple) and X region (X, in orange). The first two repeat regions each have copies at a terminal locus (TRL, TRS) and an internal locus (IRL, 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 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**) Maximum-likelihood phylogenetic tree of the 28 OsHV-1 NR-genomes. Bootstrap values were obtained from 1000 replicates. Numbers at branch nodes indicate branch confidence (percentage of bootstrap replicates for that branch). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree, except for the outgroup, AbHV-1. The 21 isolates from this study are colored according to their geographic origin, Brest (Br) in blue, Marennes-Oléron (Mo) in red, and Thau Lagoon (Th) in green. Three clusters were defined based on the clustering of genetic distances (numbered on the right from 1 to 3).

Fig. 3: Comparative genomics of the 21 non-redundant genomes of OsHV-1 isolated from three oyster farming areas. (**A**) Distribution of the 399 variable positions (hereafter named variants) between the 21 individuals and the three farming areas. Significance was calculated according to Kruskal-Wallis and Dunn's multiple comparison tests (ns: non-significant, ** $P < 0.01$). **B)** Venn diagram summarizing the distribution of variants between the three farming areas. Color depends on the farming area with Brest (Br) in blue, Marennes-Oléron (Mo) in red and Thau Lagoon (Th) in green. The numbers of variants per farming area are indicated in parentheses. Singletons are variants that occur in only one sample, and the term "common" indicates variants present in all the sampled individuals at a given location. (**C**) UpSet plot of variant distribution among the 21 non-redundant (NR-) genomes. The bottom panel shows the specific combinations (or intersections), and the vertical bars indicate the number of variants within these combinations. Colored arrows on top indicate variants shared by almost all individuals from Br and Th (green arrows) and variants 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 variants per sample.

Fig. 4: Variant calling analysis and minor variant distribution in the three farming areas. (**A**) Quantification of both major and minor variants using a variant calling analysis (**B**) Venn diagram summarizing the distribution of the 129 minor variants (frequency <50%) between the three farming areas (Brest (Br) in blue, Marennes-Oléron (Mo) in red, and Thau Lagoon (Th) in green). The number of minor variants within each farming area is indicated in parentheses. Singletons correspond to variants that occur in only one oyster sample, whereas the term "common" corresponds to variants present in all the oyster samples within the dataset. (**C**) Heatmap of the allelic frequency of minor variants across the 21 samples with a

411 gradient from 0 (blue) to 0.5 (red). Hierarchical clustering using the Euclidean distance of the
412 allele frequency is displayed on top. Each line of the heatmap corresponds to a minority
413 variant. Variant sets on the right are derived from the Venn diagram. Samples are indicated
414 on the bottom, and colored according to their geographic origin.
415
416 **Fig. 5: Some minor OsHV-1 variants found in the Marennes-Oléron farming area are**
417 **major variants in the other two farming areas.** (**A**) Matrix of the 116 variants which were
418 major or minor, according to the sample. Rows display the four samples in which the 116
419 variants were minor, and columns show all 21 samples. Intersections give the number of
420 shared variants with minor (row) and major (column) frequency. For self-comparisons, no
421 number is given. (**B**) Dot plot of the 29 minor variants of Mo ind10 which were major
422 variants in Mo ind1/3/7/8/9. Positions on the C-NR-genome and the type of SNPs are
423 indicated on the x-axis. Variant frequency (in %) is indicated in ordinate. Black arrow
424 indicates the variant that is detailed in (C). (**C**) Pie charts of the frequency of the minority
425 variant found at position 65,016 of the NR-genome in Mo ind1/3/7/8/9/10. (**D**) Phylogenetic
426 tree of the 21 *de novo* assembled NR-genomes. The black line shows the link established
427 between Mo ind10 and Mo ind3/4/7/8/9 from variant frequency at position 65,016. (**E**) Dot
428 plot of the 85 minor variants of Mo ind3 that were major variants in Mo ind4/10 and in all
429 samples from Br and Th. Green and turquoise arrows indicate the two variants (57,091_T>C
430 and 65,264_C>T respectively) that are detailed in (F). (**F**) The minor variant found in Mo
431 ind3 at position 57,091 of the C-NR-genome corresponds to major variants in Mo ind10 and
432 all Th samples (upper panel). The minor variant found in Mo ind3 at position 65,624 of the
433 NR-genome corresponds to major variants in Mo ind4/10 and all Br samples (lower panel).
434 (**G**) Phylogenetic tree of the 21 *de novo* assembled NR-genomes. The turquoise lines show
435 the link established between Mo ind3 and Mo ind10 and all Th samples based on the study of

variant frequency at position 57,091. The purple lines show the link established between Mo ind3 and Mo ind4/10 and all Br samples based on the study of variant frequency at position 65,624.

Fig. 6. Graphical representation of OsHV-1 population connectivity in France.

Materials and Methods

OsHV-1 field infection and oyster sampling

All *Crassostrea gigas* oysters used in the present study were produced in August 2017 in the Ifremer experimental facilities located in Argenton (Brittany,France), as specific-pathogen-free (SPF) juveniles (55). To sample OsHV-1 diversity, SPF oysters originating from the same batch were transplanted into three oyster farming areas (~1000 individuals per farm; average weight 1.5 g/oyster) during a disease outbreak when seawater temperature was above 16°C: Thau Lagoon (Th) (Mèze, Lat. 43.378888 long. 3.571111) in May 2018, Marennes-Oléron Bay (Mo) (La Floride lat.: 45.802968 and long.: -1.153421) in June 2018 and the Bay of Brest (Br) (Logonna-Daoulas, lat.: 48.335263 long.: -4.317922) 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 (55, 56). For this reason, oysters were transferred back to laboratory facilities 5 days after transplantation. They were then maintained in seawater tanks at 20°C, and moribund oysters were collected daily and stored at -80°C until analysis. A control group consisted of oysters that had not been transplanted in natural environment. Monitoring was stopped after three consecutive days without oyster mortality. During transport, oysters were packed in polystyrene containers and kept moist by covering them with a damp cloth.

460 **DNA extraction, viral load quantification and sequencing**

461 DNA was extracted from moribund oysters using the MagAttract® HMW DNA kit
462 (Qiagen) according to the manufacturer's protocol. DNA purity and concentration were
463 checked using a Nano-Drop ND-1000 spectrometer (Thermo Scientific) and Qubit® dsDNA
464 HS assay kits (Molecular Probes Life Technologies), respectively. Quantification of OsHV-1
465 was carried out using quantitative PCR (qPCR). Amplification reactions were performed
466 using the Roche LightCycler 480 Real-Time thermocycler on three technical replicates
467 (qPHD-Montpellier GenomiX platform, Montpellier University, France). The total qPCR
468 reaction volume was 1.5 µL, consisting of 0.5 µL of DNA (40 ng/µL) and 1 µL of
469 LightCycler 480 SYBR Green I Master mix (Roche) containing 0.5 µM PCR primers
470 (Eurogenetec SA). The primers used were virus-specific and targeted the region of the
471 OsHV-1 genome predicted to encode a catalytic subunit of DNA polymerase (ORF100,
472 AY509253): Fw-5'-ATTGATGATGTGGATAATCTGTG-3' and Rev-5'-
473 GGTAAATACCATTGGTCTTGTTCC-3' (57). The following program was applied: enzyme
474 activation at 95° C for 10 min followed by 40 cycles of denaturation (95°C, 10 s), annealing
475 (60°C, 20 s) and elongation (72°C, 25 s). To check the specificity of the amplification, a
476 subsequent melting step was applied. Twenty-one samples with a viral load greater than or
477 equal to 10^5 genomic units (GU)/ng of DNA were selected for sequencing (5 from Br, 7 from
478 Mo and 9 from Th). DNA-Seq library preparation and sequencing were performed by the
479 Genome Quebec Company (Genome Quebec Innovation Center, McGill University,
480 Montreal, Canada) using the Shotgun PCR-free library preparation kit (Lucigen) and the
481 NovaSeq™ 6000 Sequencing system (Illumina®) (paired ends, 150 bp).

Raw read QC and selection of viral reads

Read quality was evaluated using FastQC v0.11.8 (58), and sequence adapters were removed using trimmomatic 0.39 (59). 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 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/). Then, viral reads were extracted using KrakenUniq 0.5.8 (60) and Seqkit v0.12.0 (61) 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 (62)) using Bowtie2 4.8.2 (63), and non-aligned viral reads were kept for subsequent analyses.

De novo assembly of non-redundant OsHV-1 genomes

Viral reads were used for *de novo* assembly using SPAdes with the --meta and --only-assembler options (64). Assembled scaffolds were aligned to the OsHV-1 µVar A genome (GenBank no. KY242785.1) using BLAST 2.9.0 (65) with an e-value of 0.00001. Scaffolds were then extended using SSPACE v3.0 (66). 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, (34)) and that the latter have a size varying from 160 to 9777 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 (61). The first group of scaffolds had a size between 2 kbp and 4 kbp and made it possible to select the US region. The second group was composed of sequences with a size between 4 kbp 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, to select the UL region. We used these three groups

of scaffolds to construct non-redundant genomes (NR-genomes) that contain only one copy of each repeated region (UL-IRL-X-IRS-US) (45). Because the repeat regions are inverted complements of each other, several adjustments were made to avoid assembling sequences in the wrong orientation. Using this strategy, 21 OsHV-1 NR-genomes were assembled from the 21 individual oyster samples.

Whole-genome comparisons and phylogenetic analyses

To assess the relationships between OsHV-1 genomes, the 21 NR-genome assemblies were aligned using MAFFT v7.455 with default parameters (67). A consensus non-redundant genome (C-NR-genome) was subsequently built using the "cons" argument from the EMBOSS suite v6.6.0.0 with the default settings (68). Then the 21 NR-genomes were compared to the C-NR-genome using MUMer4 4.0.0beta2 (69) with the option "nucmer -c 100 -l 15 –f" to identify all the variable positions between the 21 NR-genomes. A maximum-likelihood phylogenetic inference was conducted on the 21 NR-genomes and 7 NR-genomes derived from previously published OsHV-1 genomes (AY509253, KU096999, KP412538, KY242785, KY271630, MG561751, MF509813) using IqTree 1.6.7 (70), with 1000 bootstrap replications. The NR genome of the abalone herpesvirus (AbHV-1, KU096999) was used as an outgroup. The ModelFinder implementation (71) resulted in the following best-fit model according to Bayesian information criterion: Transversion model + empirical base frequencies + gamma model with 4 rate categories. Results were locally visualized using the R package ggTree v2.0.4 (72).

Variant calling analysis

To assess the genomic variability between the farming areas, both minor and major variants of each sequencing library (21) were called on the C-NR-genome previously generated using Freebayes v1.3.2-dirty (73), 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- (74), decomposed with vt v1.0.0 (75) and split with vcflib v1.0.0.. Considering the very low occurrence of multi-nucleotide polymorphisms, these latter were counted as InDels. Variable positions with a frequency of >50% 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 minority variants (with a frequency of <50%). A unique identifier composed of the position and sequence information (e.g. 128737_A>G) of each minority variant was created. Set analyses of nucleotide variations were performed with a multiplication matrix using R (76) operation and visualized with the Venn Diagram package (77), UpSetR v1.4.0 (78) or pheatmap v1.0.12.

Downstream analyses

All downstream analyses (tables, graphs, plot creation and edition) were performed locally in R 3.6 (76) on R studio IDE (79) with an extensive use of Dplyr v1.0.0 (80), and ggplot2 (81) from the tidyverse v1.3.0 package (82). Genome visualization was carried out using gggenes v0.4.0.

Statistical Analyses

For statistical analyses, the 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.

Data availability

The datasets generated from this study can be found in the SRA database BioProject accession number PRJNA681599 with submission ID SUB8642385. All the scripts generated

in this study are freely available at https://github.com/propan2one/OshV-1-molepidemio. Complementary information is available from the corresponding authors upon reasonable request.

Acknowledgments

The present study was supported by the EU project VIVALDI (H2020 program, no. 678589) led by Ifremer; the CNRS and the University of Montpellier and the University of Perpignan Via Domitia. JD was supported by grant from the University of Montpellier. This work also benefitted 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 project. 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.

JCA, OK, BM, JD and JME were involved in the study conception and design. BM, BP, CP, JD, and JME were involved in the collection of samples and in the experimental work. JD and JME were involved in bioinformatics and statistical analyses. JD, RG, CM and JME drafted the manuscript, and all authors revised and approved the final manuscript.

We thank the staff of the Ifremer stations at Argenton (LPI, PFOM) and La Tremblade (SGMM) and the Comité Régional Conchylicole de Méditerranée (CRCM) for technical support in the production of standard (NSI) oysters and transplantation experiments. We also thank Nicole Faury (Ifremer, SGMM) and Marc Leroy for technical assistance. We are also grateful to Eric Rivals (LIRMM) for fruitful discussions.

We declare that we have no competing interests.

578 **Reference**

579 1. Nkili-Meyong AA, Bigarre L, Labouba I, Vallaeys T, Avarre JC, Berthet N. 2016.
580 Contribution of Next-Generation Sequencing to Aquatic and Fish Virology.
581 Intervirology 59:285-300.

582 2. Lauring AS, Andino R. 2010. Quasispecies Theory and the Behavior of RNA Viruses.
583 Plos Pathogens 6.

584 3. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. 2006. Quasispecies
585 diversity determines pathogenesis through cooperative interactions in a viral
586 population. Nature 439:344-348.

587 4. Hammoumi S, Vallaeys T, Santika A, Leleux P, Borzym E, Klopp C, Avarre JC.
588 2016. Targeted genomic enrichment and sequencing of CyHV-3 from carp tissues
589 confirms low nucleotide diversity and mixed genotype infections. PeerJ 4:e2516.

590 5. Renner DW, Szpara ML. 2018. Impacts of Genome-Wide Analyses on Our
591 Understanding of Human Herpesvirus Diversity and Evolution. J Virol 92.

592 6. Renzette N, Gibson L, Bhattacharjee B, Fisher D, Schleiss MR, Jensen JD, Kowalik
593 TF. 2013. Rapid intrahost evolution of human cytomegalovirus is shaped by
594 demography and positive selection. PLoS Genet 9:e1003735.

595 7. Renzette N, Gibson L, Jensen JD, Kowalik TF. 2014. Human cytomegalovirus
596 intrahost evolution-a new avenue for understanding and controlling herpesvirus
597 infections. Curr Opin Virol 8:109-15.

598 8. Renzette N, Pokalyuk C, Gibson L, Bhattacharjee B, Schleiss MR, Hamprecht K,
599 Yamamoto AY, Mussi-Pinhata MM, Britt WJ, Jensen JD, Kowalik TF. 2015. Limits
600 and patterns of cytomegalovirus genomic diversity in humans. Proc Natl Acad Sci U
601 S A 112:E4120-8.

9. Akhtar LN, Bowen CD, Renner DW, Pandey U, Della Fera AN, Kimberlin DW, Prichard MN, Whitley RJ, Weitzman MD, Szpara ML. 2019. Genotypic and Phenotypic Diversity of Herpes Simplex Virus 2 within the Infected Neonatal Population. mSphere 4.

10. Pernet F, Lupo C, Bacher C, Whittington RJ. 2016. Infectious diseases in oyster aquaculture require a new integrated approach. Philos Trans R Soc Lond B Biol Sci 371.

11. Stentiford GD, Neil DM, Peeler EJ, Shields JD, Small HJ, Flegel TW, Vlak JM, Jones B, Morado F, Moss S, Lotz J, Bartholomay L, Behringer DC, Hauton C, Lightner DV. 2012. Disease will limit future food supply from the global crustacean fishery and aquaculture sectors. J Invertebr Pathol 110:141-57.

12. Burge CA, Shore-Maggio A, Rivlin ND. 2017. Ecology of Emerging Infectious Diseases of Invertebrates. *In* Hajek AE (ed), Ecology of Invertebrate Diseases doi:doi.org/10.1002/9781119256106.ch16. John Wiley & Sons Ltd.

13. Petton B, Destoumieux-Garzon 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. Front Immunol 12:630343.

14. 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 microvar from the North Adriatic Sea (Italy). J Gen Virol 99:693-703.

15. Barbieri ES, Medina CD, Vazquez N, Fiorito C, Martelli A, Wigdorovitz A, Schwindt E, Morga B, Renault T, Parreno V, Baron PJ. 2019. First detection of Ostreid herpesvirus 1 in wild Crassostrea gigas in Argentina. J Invertebr Pathol 166:107222.

627 16. Barbosa Solomieu V, Renault T, Travers MA. 2015. Mass mortality in bivalves and
628 the intricate case of the Pacific oyster, Crassostrea gigas. J Invertebr Pathol 131:2-10.
629 17. Friedman CS, Estes RM, Stokes NA, Burge CA, Hargove JS, Barber BJ, Elston RA,
630 Burreson EM, Reece KS. 2005. Herpes virus in juvenile Pacific oysters Crassostrea
631 gigas from Tomales Bay, California, coincides with summer mortality episodes. Dis
632 Aquat Organ 63:33-41.
633 18. Gittenberger A, Voorbergen-Laarman MA, Engelsma MY. 2016. Ostreid herpesvirus
634 OsHV-1 mu Var in Pacific oysters Crassostrea gigas (Thunberg 1793) of the Wadden
635 Sea, a UNESCO world heritage site. Journal of Fish Diseases 39:105-109.
636 19. Jenkins C, Hick P, Gabor M, Spiers Z, Fell SA, Gu X, Read A, Go J, Dove M,
637 O'Connor W, Kirkland PD, Frances J. 2013. Identification and characterisation of an
638 ostreid herpesvirus-1 microvariant (OsHV-1 micro-var) in Crassostrea gigas (Pacific
639 oysters) in Australia. Dis Aquat Organ 105:109-26.
640 20. Moss JA, Burreson EM, Cordes JF, Dungan CF, Brown GD, Wang A, Wu X, Reece
641 KS. 2007. Pathogens in Crassostrea ariakensis and other Asian oyster species:
642 implications for non-native oyster introduction to Chesapeake Bay. Dis Aquat Organ
643 77:207-23.
644 21. Peeler EJ, Reese RA, Cheslett DL, Geoghegan F, Power A, Thrush MA. 2012.
645 Investigation of mortality in Pacific oysters associated with Ostreid herpesvirus-1 mu
646 Var in the Republic of Ireland in 2009. Preventive Veterinary Medicine 105:136-143.
647 22. Vasquez-Yeomans R, Garcia-Ortega M, Caceres-Martinez J. 2010. Gill erosion and
648 herpesvirus in Crassostrea gigas cultured in Baja California, Mexico. Dis Aquat
649 Organ 89:137-44.

650 23. Paul-Pont I, Dhand NK, Whittington RJ. 2013. Spatial distribution of mortality in
651 Pacific oysters Crassostrea gigas: reflection on mechanisms of OsHV-1 transmission.
652 Diseases of Aquatic Organisms 105:127-138.
653 24. Pernet F, Barret J, Le Gall P, Corporeau C, Dégremont L, Lagarde F, Pépin J-F, Keck
654 N. 2012. Mass mortalities of Pacific oysters Crassostrea gigas reflect infectious
655 diseases and vary with farming practices in the Mediterranean Thau lagoon, France.
656 Acquaculture environmental interactions 2:215-237.
657 25. Petton B, Pernet F, Robert R, Boudry P. 2013. Temperature influence on pathogen
658 transmission and subsequent mortalities in juvenile Pacific oysters Crassostrea gigas.
659 Aquaculture Environment Interactions 3:257-273.
660 26. Renault T, Moreau P, Faury N, Pepin JF, Segarra A, Webb S. 2012. Analysis of
661 Clinical Ostreid Herpesvirus 1 (Malacoherpesviridae) Specimens by Sequencing
662 Amplified Fragments from Three Virus Genome Areas. Journal of Virology 86:5942-
663 5947.
664 27. Segarra A, Pepin JF, Arzul I, Morga B, Faury N, Renault T. 2010. Detection and
665 description of a particular Ostreid herpesvirus 1 genotype associated with massive
666 mortality outbreaks of Pacific oysters, Crassostrea gigas, in France in 2008. Virus Res
667 153:92-9.
668 28. de Lorgeril J, Lucasson A, Petton B, Toulza E, Montagnani C, Clerissi C, Vidal-
669 Dupiol J, Chaparro C, Galinier R, Escoubas J-M, Haffner P, Dégremont L, Charrière
670 GM, Lafont M, Delort A, Vergnes A, Chiarello M, Faury N, Rubio TP, Leroy MA,
671 Pérignon A, Régler D, Morga B, Alunno-Bruscia M, Boudry P, Le Roux F,
672 Destoumieux-Garzón D, Gueguen Y, Mitta G. 2018. Immune-suppression by OsHV-1
673 viral infection causes fatal bacteraemia in Pacific oysters. Nature Communications
674 9:4215.

675 29. de Lorgeril J, Petton B, Lucasson A, Perez V, Stenger PL, Degremont L, Montagnani
676 C, Escoubas JM, Haffner P, Allienne JF, Leroy M, Lagarde F, Vidal-Dupiol J,
677 Gueguen Y, Mitta G. 2020. Differential basal expression of immune genes confers
678 Crassostrea gigas resistance to Pacific oyster mortality syndrome. BMC Genomics
679 21:63.
680 30. Dégremont L, Nourry M, Maurouard E. 2015. Mass selection for survival and
681 resistance to OsHV-1 infection in Crassostrea gigas spat in field conditions: response
682 to selection after four generations. Aquaculture 446:111-121.
683 31. Delmotte J, Chaparro C, Galinier R, de Lorgeril J, Petton B, Stenger PL, Vidal-Dupiol
684 J, Destoumieux-Garzon D, Gueguen Y, Montagnani C, Escoubas JM, Mitta G. 2020.
685 Contribution of Viral Genomic Diversity to Oyster Susceptibility in the Pacific Oyster
686 Mortality Syndrome. Frontiers in Microbiology 11:1579.
687 32. Martenot C, Oden E, Travaille E, Malas JP, Houssin M. 2011. Detection of different
688 variants of Ostreid Herpesvirus 1 in the Pacific oyster, Crassostrea gigas between
689 2008 and 2010. Virus Research 160:25-31.
690 33. Burioli EA, Prearo M, Riina MV, Bona MC, Fioravanti ML, Arcangeli G, Houssin M.
691 2016. Ostreid herpesvirus type 1 genomic diversity in wild populations of Pacific
692 oyster Crassostrea gigas from Italian coasts. J Invertebr Pathol 137:71-83.
693 34. Davison AJ, Trus BL, Cheng N, Steven AC, Watson MS, Cunningham C, Le Deuff
694 RM, Renault T. 2005. A novel class of herpesvirus with bivalve hosts. J Gen Virol
695 86:41-53.
696 35. Burioli EAV, Prearo M, Houssin M. 2017. Complete genome sequence of Ostreid
697 herpesvirus type 1 microVar isolated during mortality events in the Pacific oyster
698 Crassostrea gigas in France and Ireland. Virology 509:239-251.

699 36. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. 2008.
700 The genomic and epidemiological dynamics of human influenza A virus. Nature
701 453:615-9.
702 37. Allicock OM, Lemey P, Tatem AJ, Pybus OG, Bennett SN, Mueller BA, Suchard
703 MA, Foster JE, Rambaut A, Carrington CV. 2012. Phylogeography and population
704 dynamics of dengue viruses in the Americas. Mol Biol Evol 29:1533-43.
705 38. Theze J, Li T, du Plessis L, Bouquet J, Kraemer MUG, Somasekar S, Yu G, de Cesare
706 M, Balmaseda A, Kuan G, Harris E, Wu CH, Ansari MA, Bowden R, Faria NR, Yagi
707 S, Messenger S, Brooks T, Stone M, Bloch EM, Busch M, Munoz-Medina JE,
708 Gonzalez-Bonilla CR, Wolinsky S, Lopez S, Arias CF, Bonsall D, Chiu CY, Pybus
709 OG. 2018. Genomic Epidemiology Reconstructs the Introduction and Spread of Zika
710 Virus in Central America and Mexico. Cell Host Microbe 23:855-864 e7.
711 39. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa
712 JD, Arinaminpathy N, Pepin J, Posada D, Peeters M, Pybus OG, Lemey P. 2014. HIV
713 epidemiology. The early spread and epidemic ignition of HIV-1 in human
714 populations. Science 346:56-61.
715 40. Martin MA, VanInsberghe D, Koelle K. 2021. Insights from SARS-CoV-2 sequences.
716 Science 371:466-467.
717 41. Sanjuán R, Domingo-Calap P. 2019. Genetic Diversity and Evolution of Viral
718 Populations. doi:10.1016/b978-0-12-809633-8.20958-8.
719 42. Chase JM, McGill BJ, McGlinn DJ, May F, Blowes SA, Xiao X, Knight TM,
720 Purschke O, Gotelli NJ. 2018. Embracing scale-dependence to achieve a deeper
721 understanding of biodiversity and its change across communities. Ecol Lett 21:1737-
722 1751.

723 43. Al Khatib HA, Benslimane FM, Elbashir IE, Coyle PV, Al Maslamani MA, Al-Khal
724 A, Al Thani AA, Yassine HM. 2020. Within-Host Diversity of SARS-CoV-2 in
725 COVID-19 Patients With Variable Disease Severities. Front Cell Infect Microbiol
726 10:575613.
727 44. Domingo E, Sheldon J, Perales C. 2012. Viral Quasispecies Evolution. Microbiology
728 and Molecular Biology Reviews 76:159-216.
729 45. Morse AM, Calabro KR, Fear JM, Bloom DC, McIntyre LM. 2017. Reliable
730 Detection of Herpes Simplex Virus Sequence Variation by High-Throughput
731 Resequencing. Viruses 9.
732 46. Burioli EAV, Varello K, Lavazza A, Bozzetta E, Prearo M, Houssin M. 2018. A
733 novel divergent group of Ostreid herpesvirus 1 muVar variants associated with a
734 mortality event in Pacific oyster spat in Normandy (France) in 2016. J Fish Dis
735 doi:10.1111/jfd.12883.
736 47. Buestel D, Ropert M, Prou J, Goulletquer P. 2009. History, status, and future of oyster
737 culture in France. Journal of Shellfish Research 28:813-820.
738 48. Lupo C, Ezanno P, Arzul I, Garcia C, Jadot C, Joly J-P, Renault T, Bareille N. How
739 network analysis of oyster movements can improve surveillance and control programs
740 of infectious diseases?, p. *In* (ed), Frontiers,
741 49. Keeling MJ, Woolhouse ME, Shaw DJ, Matthews L, Chase-Topping M, Haydon DT,
742 Cornell SJ, Kappey J, Wilesmith J, Grenfell BT. 2001. Dynamics of the 2001 UK foot
743 and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science
744 294:813-7.
745 50. Lapegue S, Heurtebise S, Cornette F, Guichoux E, Gagnaire PA. 2020. Genetic
746 Characterization of Cupped Oyster Resources in Europe Using Informative Single
747 Nucleotide Polymorphism (SNP) Panels. Genes (Basel) 11.

748 51. Rohfritsch A, Bierne N, Boudry P, Heurtebise S, Cornette F, Lapegue S. 2013.
749 Population genomics shed light on the demographic and adaptive histories of
750 European invasion in the Pacific oyster, Crassostrea gigas. Evol Appl 6:1064-78.
751 52. Vendrami DLJ, Houston RD, Gharbi K, Telesca L, Gutierrez AP, Gurney-Smith H,
752 Hasegawa N, Boudry P, Hoffman JI. 2019. Detailed insights into pan-European
753 population structure and inbreeding in wild and hatchery Pacific oysters (Crassostrea
754 gigas) revealed by genome-wide SNP data. Evolutionary Applications 12:519-534.
755 53. Alfaro AC, Nguyen TV, Merien F. 2018. The complex interactions of Ostreid
756 herpesvirus 1, Vibrio bacteria, environment and host factors in mass mortality
757 outbreaks of Crassostrea gigas. Reviews in Aquaculture 11:1148-1168.
758 54. Avarre JC. 2017. Editorial: Molecular Tracing of Aquatic Viruses: Where
759 Epidemiology Needs to Meet Genomics. Front Microbiol 8:1498.
760 55. Petton B, Bruto M, James A, Labreuche Y, Alunno-Bruscia M, Le Roux F. 2015.
761 Crassostrea gigas mortality in France: the usual suspect, a herpes virus, may not be
762 the killer in this polymicrobial opportunistic disease. Front Microbiol 6:686.
763 56. Dupont S, Lokmer A, Corre E, Auguet JC, Petton B, Toulza E, Montagnani C,
764 Tanguy G, Pecqueur D, Salmeron C, Guillou L, Desnues C, La Scola B, Bou Khalil J,
765 de Lorgeril J, Mitta G, Gueguen Y, Escoubas JM. 2020. Oyster hemolymph is a
766 complex and dynamic ecosystem hosting bacteria, protists and viruses. Animal
767 Microbiome 2.
768 57. Webb SC, Fidler A, Renault T. 2007. Primers for PCR-based detection of ostreid
769 herpes virus-1 (OsHV-1): Application in a survey of New Zealand molluscs.
770 Aquaculture 272:126-139.
771 58. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data.
772 *Available online at:* http://wwwbioinformaticsbabrahamacuk/projects/fastqc.

773 59. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina
774 sequence data. Bioinformatics 30:2114-20.
775 60. Breitwieser FP, Baker DN, Salzberg SL. 2018. KrakenUniq: confident and fast
776 metagenomics classification using unique k-mer counts. Genome Biology 19.
777 61. Shen W, Le S, Li Y, Hu F. 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for
778 FASTA/Q File Manipulation. PLoS One 11:e0163962.
779 62. Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H,
780 Xiong Z, Que H, Xie Y, Holland PW, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C,
781 Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A,
782 Hedgecock D, Xu Z, Liu Y, Domazet-Loso T, Du Y, Sun X, Zhang S, Liu B, Cheng
783 P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li
784 N, Wang J, et al. 2012. The oyster genome reveals stress adaptation and complexity of
785 shell formation. Nature 490:49-54.
786 63. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat
787 Methods 9:357-9.
788 64. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new
789 versatile metagenomic assembler. Genome Res 5:824-834.
790 65. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment
791 search tool. J Mol Biol 215:403-10.
792 66. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-
793 assembled contigs using SSPACE. Bioinformatics 27:578-9.
794 67. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid
795 multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res
796 30:3059-66.

797 68. Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology
798 Open Software Suite. Trends Genet 16:276-7.
799 69. Marcais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018.
800 MUMmer4: A fast and versatile genome alignment system. PLoS Comput Biol
801 14:e1005944.
802 70. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and
803 effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol
804 Biol Evol 32:268-74.
805 71. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017.
806 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods
807 14:587-589.
808 72. Yu G, Smith DK, Zhu H, Guan YY, Lam TTY, McInerny G. 2016. GGTREE : an R
809 package for visualization and annotation of phylogenetic trees with their covariates
810 and other associated data. Methods in Ecology and Evolution 8:28-36.
811 73. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read
812 sequencing. arXivorg.
813 74. Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. 2016.
814 BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from
815 next-generation sequencing data. Bioinformatics 32:1749-51.
816 75. Tan A, Abecasis GR, Kang HM. 2015. Unified representation of genetic variants.
817 Bioinformatics 31:2202-4.
818 76. R Core Team. 2018. R: A language and environment for statistical computing, R
819 Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
820 77. Chen H, Boutros PC. 2011. VennDiagram: a package for the generation of highly-
821 customizable Venn and Euler diagrams in R. BMC Bioinformatics 12:35.

78. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. 2014. UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput Graph 20:1983-92.

79. R-Studio-Team. 2015. RStudio: Integrated Development Environment for R., RStudio Inc., Boston, MA.

80. Wickham H, François R, Henry L, Müller K. 2018. dplyr: A Grammar of Data Manipulation, RStudio, https://dplyr.tidyverse.org/reference/dplyr-package.html.

81. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York,

82. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the Tidyverse. Journal of Open Source Software 4:1686.

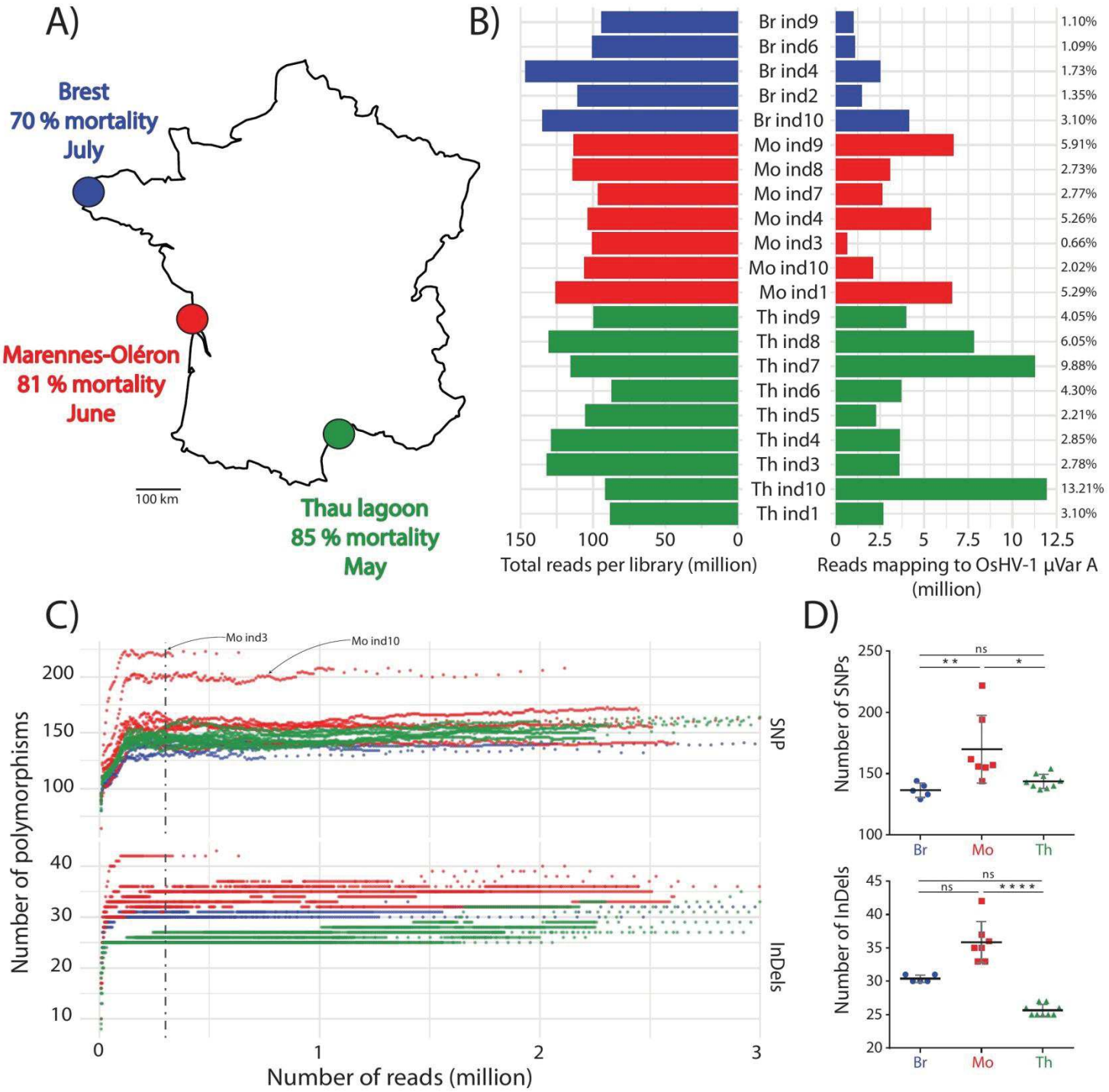


Fig. 1: Variability in sequencing depth does not impair the analysis of OsHV-1 diversity. (**A**) Sampling locations of the Ostreid herpesvirus-1 (OsHV-1) infected oysters during the Pacific oyster mortality syndrome (POMS) event in 2018. Each color represents a farming area, with Brest (Br) in blue (5 individual samples), Marennes-Oléron (Mo) in red (7 individual samples) and Thau Lagoon (Th) in green (9 individual samples). Below each farming area, the percentage of mortality within the site is given. (**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 using the OsHV-1 µVar A genome as the reference sequence. The top panel shows the number of single nucleotide polymorphisms (SNPs) and the bottom panel, the number of insertion-deletions (InDels). 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, *$P \leq 0.05$, **$P \leq 0.01$ **** $P \leq 0.0001$).

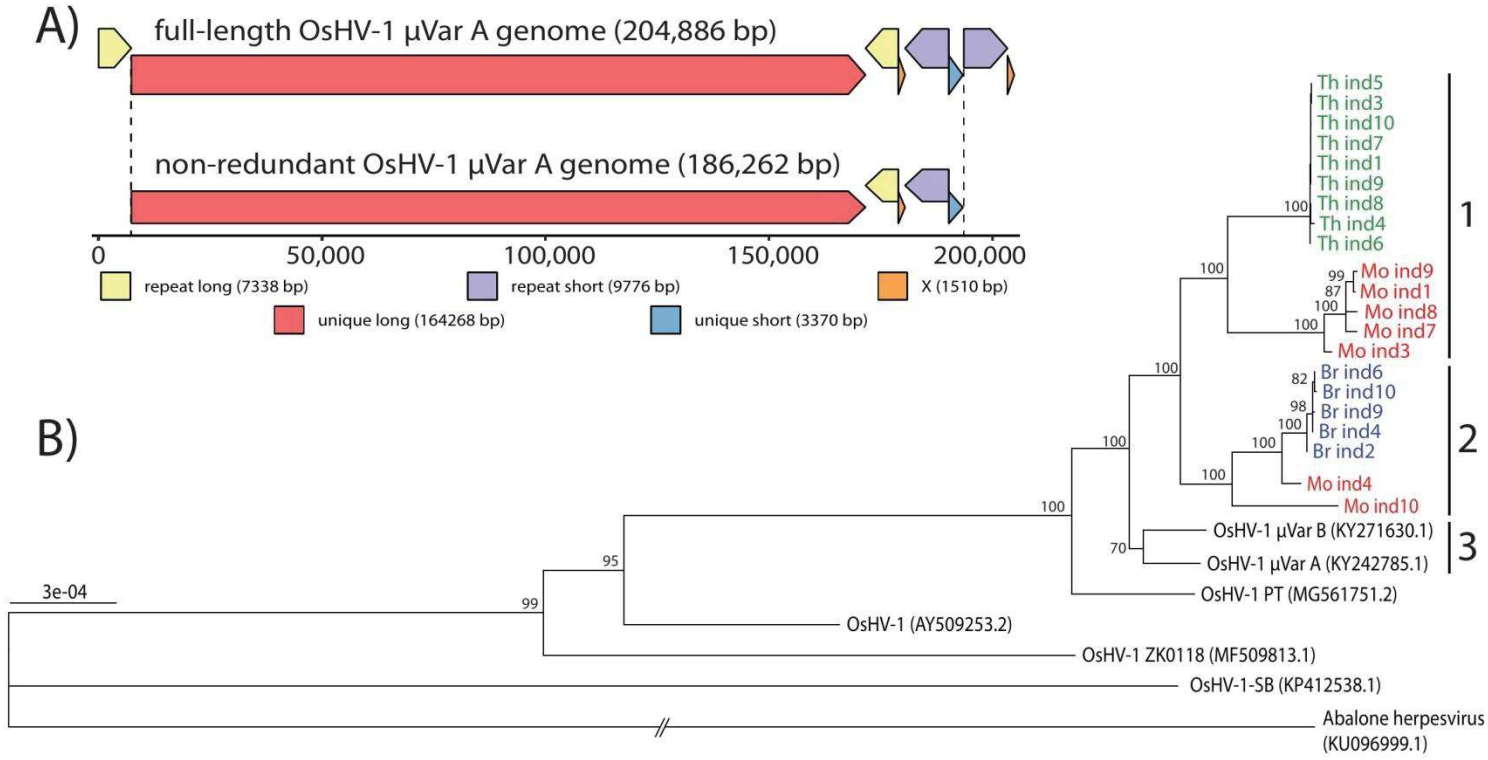


Fig. 2: Phylogenetic relationship between the 21 de novo assembled non-redundant genomes and other Malacoherpesviridae viruses. (**A**) Full-length (upper panel) and non-redundant (NR) (lower panel) genomes of OsHV-1 µVar A (KY242785). The OsHV-1 µVar A genome consists of two unique regions, referred to as unique long (UL; in red) and unique short (US, in blue), and three repeat regions, referred to as repeat long (RL, in yellow), repeat short (RS, in purple) and X region (X, in orange). The first two repeat regions each have copies at a terminal locus (TRL, TRS) and an internal locus (IRL, 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 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**) Maximum-likelihood phylogenetic tree of the 28 OsHV-1 NR-genomes. Bootstrap values were obtained from 1000 replicates. Numbers at branch nodes indicate branch confidence (percentage of bootstrap replicates for that branch). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree, except for the outgroup, AbHV-1. The 21 isolates from this study are colored according to their geographic origin, Brest (Br) in blue, Marennes-Oléron (Mo) in red, and Thau Lagoon (Th) in green. Three clusters were defined based on the clustering of genetic distances (numbered on the right from 1 to 3).

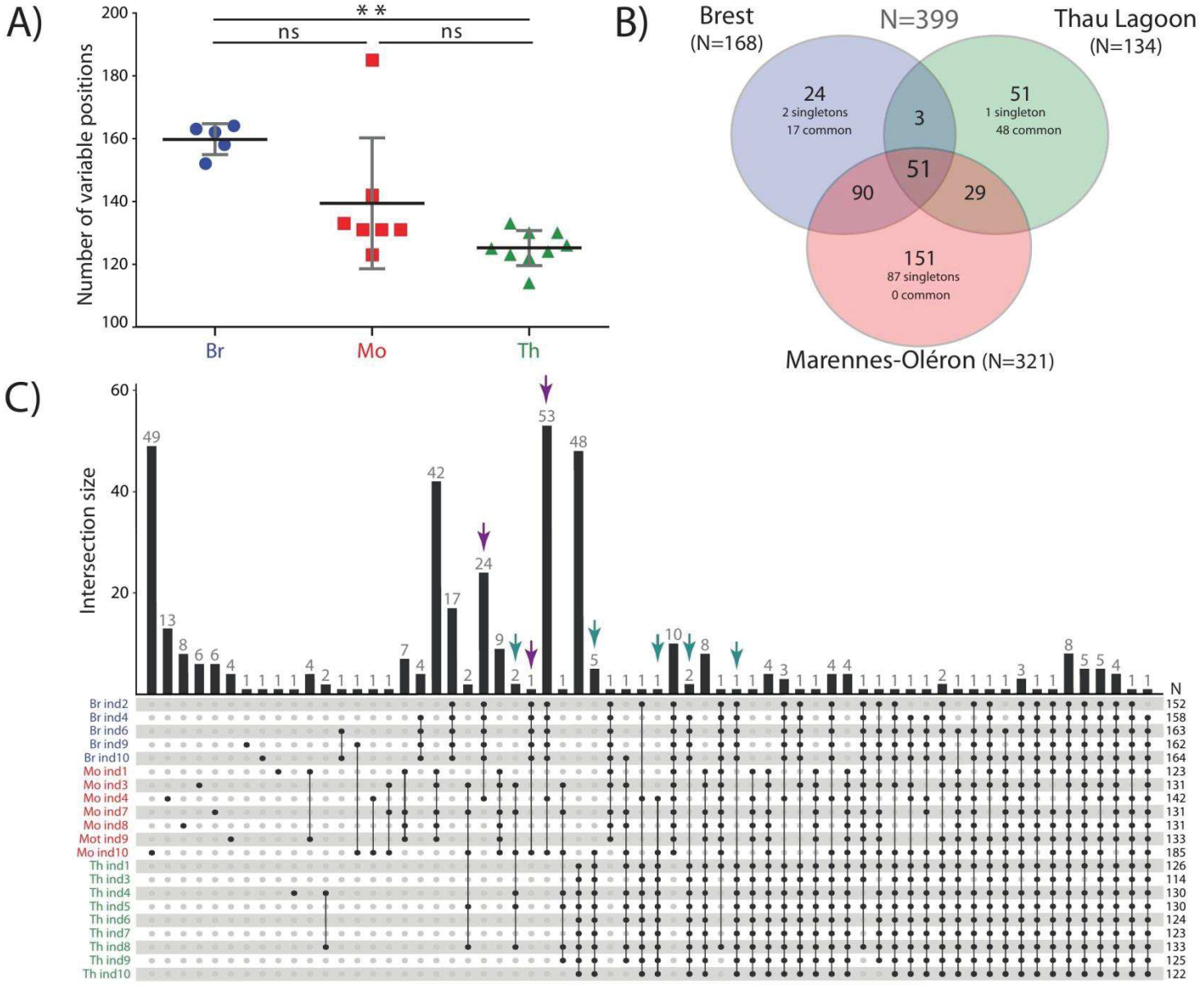


Fig. 3: Comparative genomics of the 21 non-redundant genomes of OsHV-1 isolated from three oyster farming areas. (**A**) Distribution of the 399 variable positions (hereafter named variants) between the 21 individuals and the three farming areas. Significance was calculated according to Kruskal-Wallis and Dunn's multiple comparison tests (ns: non-significant, ** $P < 0.01$). (**B**) Venn diagram summarizing the distribution of variants between the three farming areas. Color depends on the farming area with Brest (Br) in blue, Marennes-Oléron (Mo) in red and Thau Lagoon (Th) in green. The numbers of variants per farming area are indicated in parentheses. Singletons are variants that occur in only one sample, and the term "common" indicates variants present in all the sampled individuals at a given location. (**C**) UpSet plot of variant distribution among the 21 non-redundant (NR-) genomes. The bottom panel shows the specific combinations (or intersections), and the vertical bars indicate the number of variants within these combinations. Colored arrows on top indicate variants shared by almost all individuals from Br and Th (green arrows) and variants 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 variants per sample.

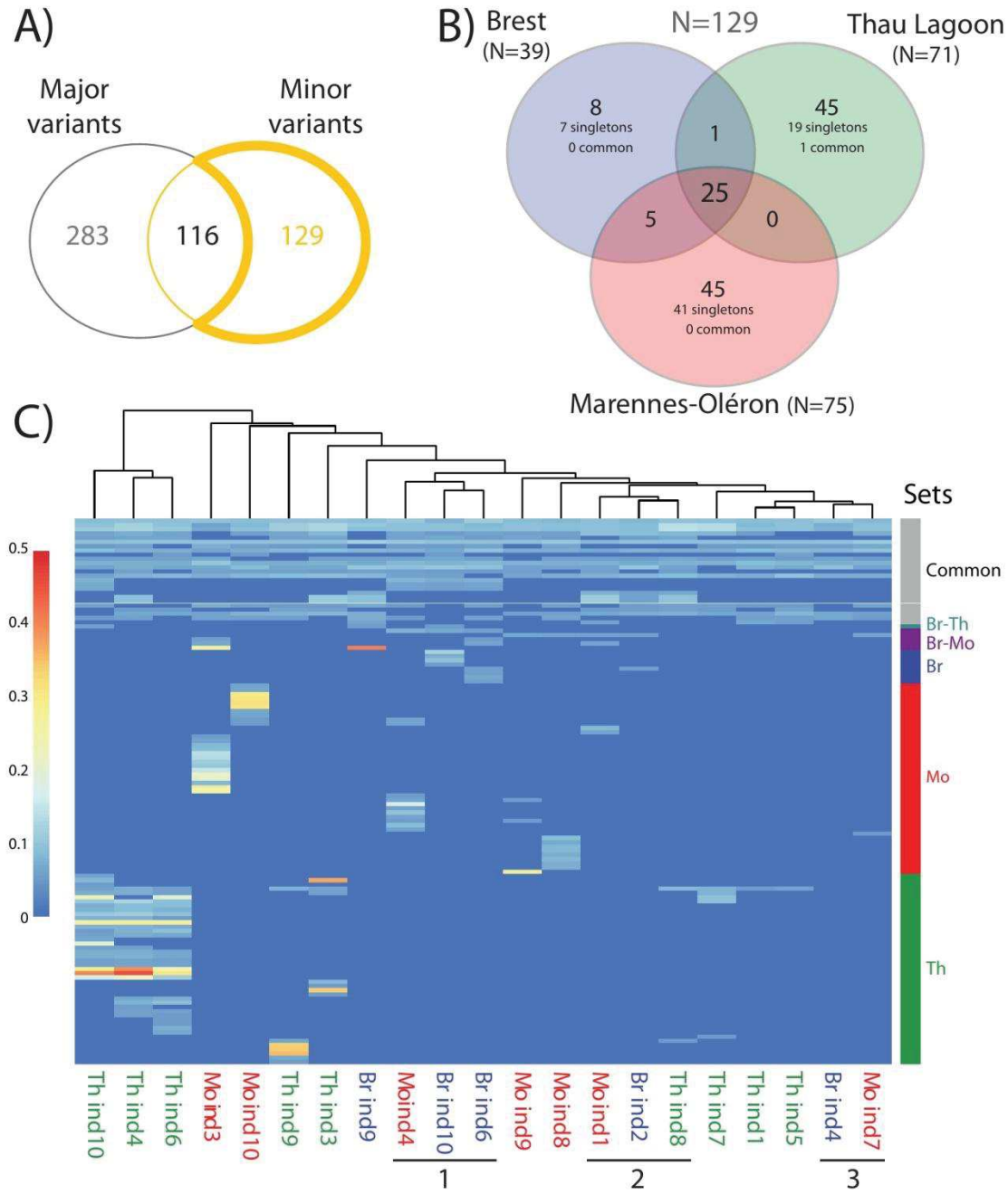


Fig. 4: Variant calling analysis and minor variant distribution in the three farming areas. (**A**) Quantification of both major and minor variants using a variant calling analysis (**B**) Venn diagram summarizing the distribution of the 129 minor variants (frequency <50%) between the three farming areas (Brest (Br) in blue, Marennes-Oléron (Mo) in red, and Thau Lagoon (Th) in green). The number of minor variants within each farming area is indicated in parentheses. Singletons correspond to variants that occur in only one oyster sample, whereas the term "common" corresponds to variants present in all the oyster samples within the dataset. (**C**) Heatmap of the allelic frequency of minor variants across the 21 samples with a gradient from 0 (blue) to 0.5 (red). Hierarchical clustering using the Euclidean distance of the allele frequency is displayed on top. Each line of the heatmap corresponds to a minority variant. Variant sets on the right are derived from the Venn diagram. Samples are indicated on the bottom, and colored according to their geographic origin.

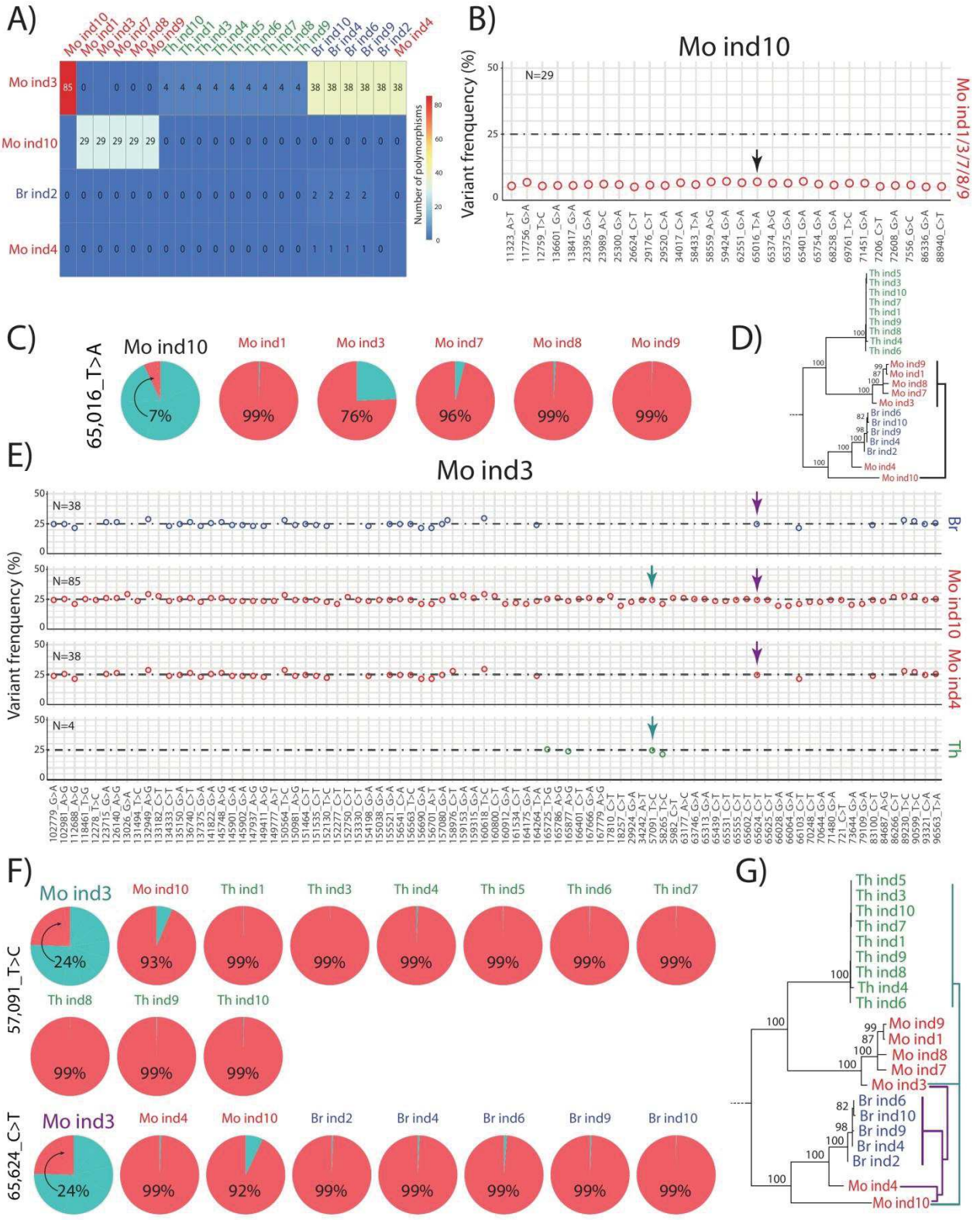
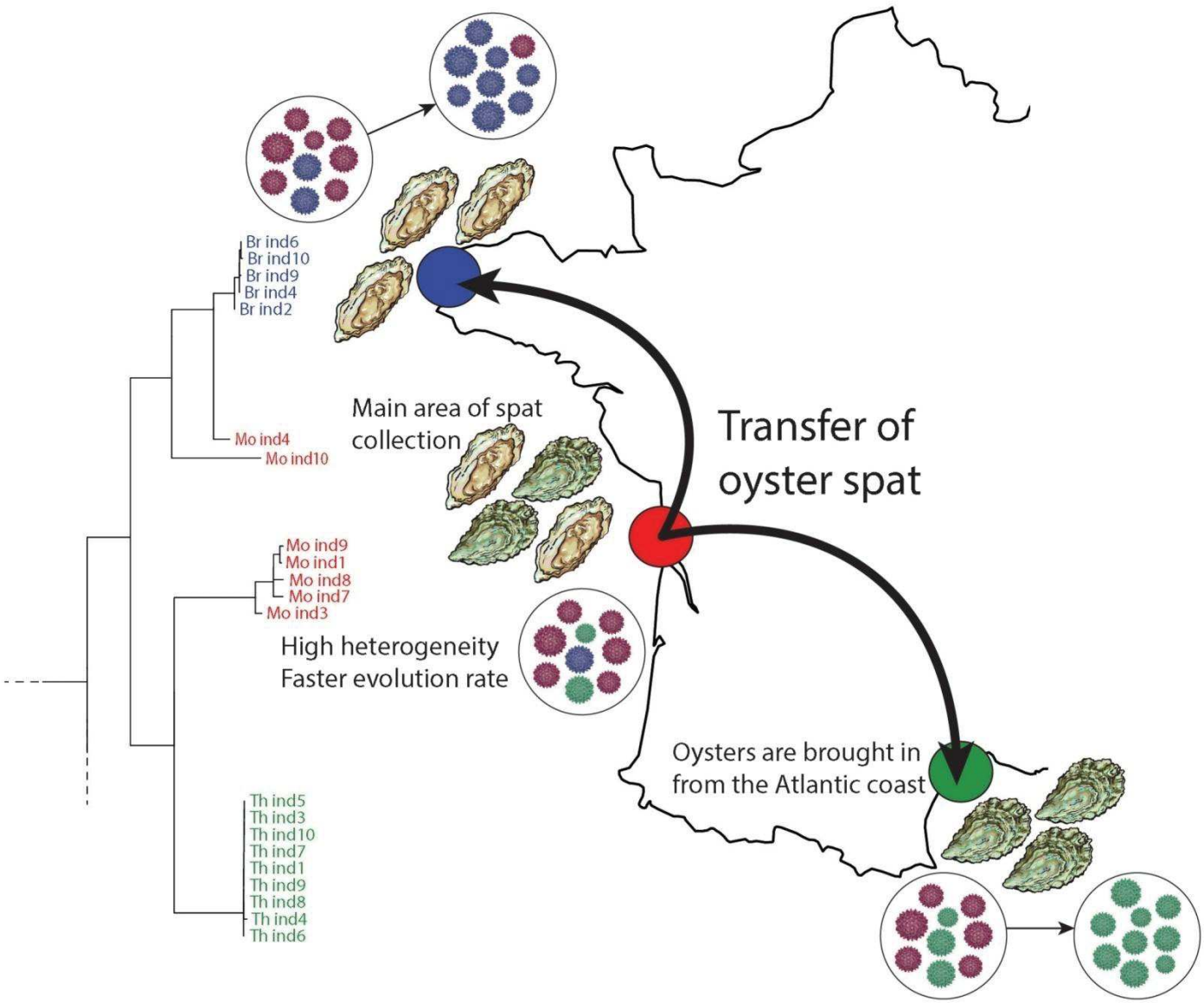


Fig. 5: Some minor OsHV-1 variants found in the Marennes-Oléron farming area are major variants in the other two farming areas. (**A**) Matrix of the 116 variants which were major or minor, according to the sample. Rows display the four samples in which the 116 variants were minor, and columns show all 21 samples. Intersections give the number of shared variants with minor (row) and major (column) frequency. For self-comparisons, no number is given. (**B**) Dot plot of the 29 minor variants of Mo ind10 which were major variants in Mo ind1/3/7/8/9. Positions on the C-NR-genome and the type of SNPs are indicated on the x-axis. Variant frequency (in %) is indicated in ordinate. Black arrow indicates the variant that is detailed in (C). (**C**) Pie charts of the frequency of the minority variant found at position 65,016 of the NR-genome in Mo ind1/3/7/8/9/10. (**D**) Phylogenetic tree of the 21 de novo assembled NR-genomes. The black line shows the link established between Mo ind10 and Mo ind3/4/7/8/9 from variant frequency at position 65,016. (**E**) Dot plot of the 85 minor variants of Mo ind3 that were major variants in Mo ind4/10 and in all samples from Br and Th. Green and turquoise arrows indicate the two variants (57,091_T>C and 65,264_C>T respectively) that are detailed in (F). (**F**) The minor variant found in Mo ind3 at position 57,091 of the C-NR-genome corresponds to major variants in Mo ind10 and all Th samples (upper panel). The minor variant found in Mo ind3 at position 65,624 of the NR-genome corresponds to major variants in Mo ind4/10 and all Br samples (lower panel). (**G**) Phylogenetic tree of the 21 de novo assembled NR-genomes. The turquoise lines show the link established between Mo ind3 and Mo ind10 and all Th samples based on the study of variant frequency at position 57,091. The purple lines show the link established between Mo ind3 and Mo ind4/10 and all Br samples based on the study of variant frequency at position 65,624.


Br ind6
Br ind10
Br ind9
Br ind4
Br ind2
Mo ind4
Mo ind10
Mo ind9
Mo ind1
Mo ind8
Mo ind7
Mo ind3
Th ind5
Th ind3
Th ind10
Th ind7
Th ind1
Th ind9
Th ind8
Th ind4
Th ind6
Main area of spat collection
Transfer of oyster spat
High heterogeneity
Faster evolution rate
Oysters are brought in from the Atlantic coast

Discussion

Dans la présente étude, nous avons assemblé 21 séquences génomiques complètes d'OsHV-1 à partir d'huîtres infectées au sein de l'environnement infectieux de 3 zones conchylicoles françaises lors d'événements de mortalité en 2018. En utilisant une combinaison d'analyses originales de phylogénétiques et d'appel des variants, nous avons pu observer pour la première fois la diversité génétique virale entre des huîtres au niveau individuel. Nos résultats suggèrent que les populations virales étaient différentes entre les huîtres infectées au sein des zones conchylicoles avec la plus forte hétérogénéité observée pour la zone d'élevage de Marennes-Oléron. En utilisant l'information génomique de chacun de ces échantillons nous avons pu reconstruire les relations phylogénétiques entre les souches et nos données montrent que les populations virales ont pour origine commune la zone d'élevage de Marennes-Oléron. Cette zone d'élevage est une importante zone de collecte de naissains d'huîtres en Europe et la dissémination du virus pourrait être une conséquence des pratiques de transferts d'huîtres. En conclusion, cette étude illustre le pouvoir de combiner l'analyse épidémiologique avec le séquençage profond du génome viral pour démêler la propagation de l'OsHV-1 et pour obtenir des informations fondamentales sur l'ampleur de la diversité génétique du virus.

Une approche non biaisée dans l'analyse de la diversité virale

Grâce à l'assemblage de génomes consensus de manière indépendante pour chacune des 3 zones conchylicoles nous avons assemblé 21 génomes d'OsHV-1. Ce genre d'approche est en adéquation avec un récent article analysant les variations génétiques de l'HSV-1 qui suggère que l'approche *de novo* est plus pertinente pour des isolats apparentés plus éloignés ou des isolats qui peuvent contenir une recombinaison entre génotypes (Morse et al., 2017). Jusqu'à présent l'assemblage *de novo* du virus OsHV-1 n'avait pas réussi à un génome entier avec des lectures courtes d'HTS. En effet, dans le cas de l'assemblage de l'OsHV-1 PT plusieurs positions du génome contiennent des ambiguïtés (N dans la séquence) probablement dues à une profondeur de séquençage trop faible. Une approche intéressante a été utilisée pour les génomes des µVar A/B à l'aide du génome de référence et d'un réassemblage local avec MindTheGap (Guyomar et al., 2019).

Les virus appartiennent au même génotype

Contrairement à une approche en alignant les lectures sur un génome de référence, l'assemblage de génome *de novo* offre la possibilité non seulement de comprendre les relations phylogénétiques, mais également, permet d'identifier de larges variations structurales. De telles variations ont déjà été observées chez d'autres virus à ADN comme le Cyprinid Herpèsvirus 3 et peuvent avoir un rôle important dans la virulence (Klafack et al., 2019). Dans le cas du virus OsHV-1, de telles variations structurales distinguent le génotype de référence par rapport au génotype µVar. En effet, en se référant à la séquence génomique, cinq délétions distinguent ces deux variants (Burioli et al., 2017; Martenot et al., 2013; Renault et al., 2012; Rosani et al., 2014). Ces cinq délétions typiques du génotype µVar sont également retrouvées pour les 21 génomes assemblés dans l'étude. Ce résultat est en adéquation avec le fait que les variants µVar sont devenus prédominants en Europe et constituent désormais 100% des isolats d'Herpèsvirus en France (Alfaro et al., 2018; Dégremont et al., 2015). Cependant nous n'avons pas observé de

larges variations structurales de l'ampleur de la différence entre les génotypes µVar et la référence entre nos 21 isolats.

Une propagation du virus OsHV via des transferts d'huîtres

Comprendre la source, les mécanismes de transmission et de dispersion de l'OsHV-1 est essentiel pour prendre des mesures efficaces de contrôle de l'infection (Paul-Pont et al., 2013a). Toutefois peu d'informations sont disponibles sur ce sujet. Nos résultats semblent indiquer que les populations virales d'OsHV-1 sont originaires de la zone conchylicole de Marennes-Oléron. Si tel est le cas, nous pouvons nous demander comment les populations virales se sont retrouvées dans les autres zones d'élevage en sachant que C. gigas est un animal fixé. Il semble peu vraisemblable qu'une transmission directe entre les 3 zones ait eu lieu. D'une part parce que les distances entre les 3 zones conchylicoles sont conséquentes et que la propagation du virus dans l'eau de mer est limitée par le taux d'excrétion et de survie du virus dans l'eau de mer (Alfaro et al., 2018; Mojica and Brussaard, 2014). Et d'autres part que la stabilité de l'OsHV-1 µVar ne serait probablement pas suffisante, des expériences dans des conditions de laboratoire démontrent que le virus ne survivra pas plus de deux jours à 20 ° C dans l'eau de mer (Alfaro et al., 2018; Hick et al., 2016). L'autre possibilité, plus probable, est une transmission indirecte. Cela a été suggéré récemment à l'aide de particule (e.g. du plancton (Paul-Pont et al., 2013a)), celle-ci serait préférable à la transmission directe (Whittington et al., 2018) dans laquelle les auteurs évoquent une capacité de transmission sur des kilomètres, sans toutefois préciser combien. Dans notre cas, la distance entre Marennes d'Oléron et l'étang de Thau est physiquement impossible pour que des particules, même des larves d'huître infectées, soient le véhicule du virus comme précédemment suggéré (Barbosa-Solomieu et al., 2005). Ainsi pour expliquer la présence de virus originaires de Marennes-Oléron dans les autres zones conchylicoles, nous faisons l'hypothèse d'une transmission indirecte liée aux transferts d'animaux. Les transferts de naissains infectés entre les zones de production pourraient être responsables de la dissémination du virus comme le suggèrent nos résultats. Ces naissains pourraient être des porteurs sains du virus, en effet des huîtres asymptomatiques pour la maladie ont déjà été décrites (Arzul et al., 2002; Dundon et al., 2011). Ces déplacements sont une source majeure de transmission de maladies en France et dans toute l'Europe (Degremont and Benabdelmouna, 2014). Il est possible que le virus puisse être présent à très faible charge virale ou sous forme latente, comme chez de nombreux Herpèsvirus (e.g. koi Herpèsvirus), mais dont l'hypothèse n'est pas claire chez les huîtres (Arzul et al., 2002; Barbosa-Solomieu et al., 2005; Dundon et al., 2011; Pepin et al., 2008; Segarra et al., 2014). L'augmentation du nombre de naissains cultivés est une des deux pratiques courantes développées par les ostréiculteurs pour faire face à la mortalité massive liée à l'OsHV-1 (Degremont and Benabdelmouna, 2014). Cet accroissement du nombre de naissains se fait notamment à travers le prélèvement d'huîtres sauvages. Or il existe 2 sites de captation majoritaire en France, l'un d'entre eux est le site de Marennes-Oléron (Degremont and Benabdelmouna, 2014; Goulletquer and Le Moine, 2002). L'autre pratique pour faire face aux pertes dues aux épisodes de mortalité massives lors des POMS est l'utilisation d'huîtres résistantes issues d'écloseries commerciales et semble une meilleure solution pour limiter la dissémination des populations virales (Degremont et al., 2015; Degremont and Benabdelmouna, 2014).

Le maintien des populations virales dans le bassin de Marennes-Oléron

La source endémique du virus est à ce jour inconnue, mais certaines pistes ont été explorées. Ainsi, plusieurs auteurs suggèrent une source environnementale du virus comme source des populations virales sans préciser laquelle (Whittington et al., 2018). Le rôle des huîtres sauvages comme réservoirs d'OsHV-1 a été suggérée, en particulier dans le bassin de Marennes-Oléron (Alfaro et al., 2018; Degremont and Benabdelmouna, 2014). La détection d'OsHV-1 dans des huîtres sauvages a été signalée dans le monde entier (Burioli et al., 2016; Degremont and Benabdelmouna, 2014; López-Sanmartín et al., 2016). L'hypothèse du maintien du virus OsHV-1 par l'intermédiaire des populations d'huîtres sauvages semble en partie cohérente avec nos observations. Dans la zone de Marennes-Oléron, les populations d'huîtres sauvages qui ont survécu à une infection par OsHV-1 pourraient transmettre verticalement le virus à différentes générations. Même si la transmission transgénérationnelle du virus n'est pas clairement démontrée à ce jour, beaucoup d'observations vont dans ce sens. Des traces d'ADN, d'ARNm et de protéines ont été observées dans les tissus gonadiques d'huîtres adultes (Alfaro et al., 2018; Arzul et al., 2002; Barbosa-Solomieu et al., 2005; Corbeil et al., 2015; Morga et al., 2017). L'hypothèse du maintien des populations virales via les huîtres sauvages est pertinente, elle permet également d'expliquer la diversité génétique plus importante observée dans les échantillons de Marennes-Oléron par rapport aux deux autres zones. Soit parce que la taille effective des populations virales serait plus importante et donc la dérive génétique plus forte (Retel et al., 2019). Soit parce qu'une plus grande hétérogénéité des populations d'huîtres sauvages serait susceptible d'engendrer une coévolution hôte-pathogène plus hétérogène du virus. Cette dernière remarque est possible, une étude récente montre que les populations d'huîtres présentent une grande diversité génétique au niveau individuel, mais pas de différenciation entre les différences populationnelles sur la côte française (Lapègue et al., 2020). Enfin, il ne faut pas non plus exclure que les populations virales d'OsHV-1 pourraient se maintenir en infectant d'autres espèces de mollusques.

Des populations virales locales supplantaient par l'apport des variants viraux issus de Marennes-Oléron

Nous pouvons nous poser la question si un phénomène de maintenance des populations virales par des huîtres sauvages pourrait également avoir lieu dans les autres zones conchylicoles, à savoir l'étang de Thau et ou la rade de Brest. Nos données ne permettent pas de répondre clairement à cette question parce que s'il existe bien des populations d'OsHV-1 endémique à cet endroit, elles entreront en compétitions avec des variants viraux issus du bassin de Marennes-Oléron. Avec l'analyse des échantillons d'huîtres individuellement, nous avons pu observer une perte importante d'hétérogénéité dans les échantillons de Brest et de l'étang de Thau. Cette perte de la variation génomique peut être interprétée comme illustrant un effet fondateur des populations virales provenant de Marennes-Oléron. Il est donc possible que des variants issus de Marennes-Oléron supplante régulièrement des populations virales locales. Soit parce que ces variants possèdent un avantage de valeur sélective par rapport à ses prédécesseurs du fait qu'ils sont déjà confrontés aux huîtres qui viennent d'être importées. Soit au travers d'événements stochastiques, comme une taille effective plus importante lors de l'importation sur des fermes ostréicoles, leurs permettant de se répliquer avec relativement peu de concurrence. Toutefois, il ne faut pas exclure que des populations virales locales si elles

existent amènent une certaine adaptation locale. En effet, la recombinaison dans les Herpèsvirus peut fournir une force motrice pour des changements évolutifs (Renner and Szpara, 2017). Déterminer quel scénario s'applique aux populations d'OsHV-1 nécessite que nous considérions l'adéquation comparative des différents génotypes lors de l'infection des huîtres ainsi que l'évolution des populations aux cours du temps. Les prochains défis seront donc de renouveler une campagne d'échantillonnage pour affiner les hypothèses grâce aux données génomiques temporelles en incluant également la zone du bassin d'Arcachon qui présente des caractéristiques physico-chimique similaire à la zone conchylicole de Marennes d'Oléron.

Conclusion

En résumé, l'utilisation de séquençage à haut débit et des analyses bioinformatiques complexes nous ont permis de mieux appréhender la diversité génétique d'OsHV-1 et de proposer un scénario de la dissémination du virus sur le littoral Français. L'interprétation de nos données fournit de nouvelles informations sur la propagation du virus et suggère que les virus sont principalement endémiques dans le bassin de Marennes-Oléron du fait de la présence importante de population d'huîtres sauvages. Des précautions doivent être prises lorsque les ostréiculteurs utilisent de tels animaux, car ils peuvent avoir été élevés dans une zone où OsHV-1 est présent et ainsi contaminé de nouvelles zones d'élevages conchylicoles. Ces pratiques contribuent ainsi à disséminer les populations virales.

Discussion générale

Du fait des conséquences économiques importantes pour la conchyliculture, de nombreuses études ont cherché à identifier les facteurs clés impliqués dans le déclenchement du syndrôme de mortalité massive « POMS ». Quelque temps avant l'initiation du projet de thèse, les recherches menées par le laboratoire IHPE ont montré le rôle prépondérant du virus OsHV-1 dans l'initiation de la pathogenèse. Plusieurs hypothèses avaient suggérées que la diversité génétique du virus OsHV-1 pouvait avoir un rôle important dans la maladie. Toutefois, la majorité des connaissances disponibles sur la diversité génétique du virus se limitaient aux travaux sur de courts fragments génomiques. Ainsi, au début de ma thèse, très peu de travaux avaient réussis à caractériser la diversité génétique du virus à l'échelle du génome entier. L'objectif premier de la thèse était donc de mieux caractériser la diversité génétique du virus OsHV-1. Les recherches conduites durant ce travail de thèse ont permis de développer l'ensemble des outils bioinformatiques pour caractériser presque exhaustivement la diversité génétique virale de l'OsHV-1. Grâce à ces recherches, nous savons maintenant que la sensibilité de certaines familles d'huîtres varient en fonction de l'origine géographique de la maladie et que les populations virales sont différentes dans les zones d'élevage en France. Parmi ces différences, nous avons noté des modifications d'ORFs codant des protéines membranaires qui pourraient impacter l'infectiosité du virus. L'analyse des variants majeurs et mineurs d'OsHV-1 dans les trois zones ostréicoles étudiées nous a conduits à proposer un scénario dans lequel le bassin de Marennes-Oléron constituerait la principale source de variants viraux qui seraient ensuite disséminés *via* les transferts d'animaux entre zones d'élevage. Nous allons maintenant nous risquer à pousser l'interprétation en confrontant les résultats obtenus dans les deux

études pour essayer de dépasser la simple caractérisation de la diversité virale et émettre des hypothèses sur son rôle adaptatif pour le virus OsHV-1 dans les POMS.

La comparaison des approches entre les deux études

Dans le cadre des deux études précédentes, il est intéressant de noter que les zones géographique d'élevage où la maladie a été captée sont très proches voir identiques. En effet, L'environnement Méditerranéen décrit dans la première étude correspond à la zone conchylicole de l'étang de Thau de l'étude II (lat 43.418736, long 3.622620 pour l'étude I et lat 43.378889, long 3.571111 pour l'étude II ; soit 3 km de différence). L'environnement Atlantique et la zone conchylicole de Brest correspondent exactement à la même localisation géographique (lat 48.335263, long -4.317922). Dans la suite de la discussion nous appellerons ces zones sous la dénomination du second article en ajoutant un chiffre romain pour signifier le numéro de l'étude (*e.g.* étang de Thau I pour la zone Méditerranée). Toutefois, avant de comparer les études et d'en tirer des hypothèses il faut souligner les importantes variations expérimentales de ces deux études. Pour commencer, il existe une différence temporelle de la date d'échantillonnage. Dans l'étude I la mise en place des huîtres lors d'événements de mortalité s'est effectué en 2015 et en 2018 dans l'étude II. Pour l'étude I il y a également une étape de transfert de la maladie supplémentaire dans des infections expérimentales en mésocosme. Ensuite il y a une différence sur le design expérimentale : dans l'étude I le matériel biologique est issu de 10 huîtres entières en triplicats mélangées à trois points de temps (soit 3*10*3 = 90 huîtres). Ces données risquent d'artificiellement amplifier l'observation de variations génétiques minoritaires communes aux individus et à l'inverse diluer l'observation de variations génétiques minoritaires singletons spécifiques à un individu. On peut considérer que ce design expérimental minimise la variance biologique. À l'inverse, le matériel biologique de l'étude II est un morceau du manteau d'un individu (>1 huître). Ces données vont peut-être favoriser l'observation individuelle de variations génétiques minoritaires et majoritaires spécifiques au manteau si jamais un phénomène de compartimentation à lieu chez OsHV-1. On peut considérer que, ici, le design expérimental maximise la variance biologique. Cette approche est plus compliquée à traiter expérimentalement et d'un point de vue analytique. Elle présente cependant l'avantage de quantifier la variation intra-individuelle et inter-individuelle en plus de la variation inter-site. Enfin il existe une différence dans la stratégie d'analyse bioinformatique : avec un approche de comparaison contre une référence dans l'étude I due au fait qu'il s'agit de donnée de transcriptomique contrairement à une approche d'assemblage *de novo* de données de métagénomique dans l'étude II. On peut noter cependant que le contrôle qualité initiale dans l'étude II (*i.e.* courbe de raréfaction) s'est fait contre le même génome de l'étude I pour permettre une comparaison succincte du nombre de polymorphismes. Toutes ces différences entre les deux études entraînent énormément de biais et empêchent de tirer des conclusions générales robustes. C'est pourquoi la majorité des assertions énoncées dans les prochains paragraphes sont davantage des pistes de réflexions. Pour finir, il semble important de mentionner la nécessité de choisir une méthodologie standardisée pour améliorer la compréhension de la diversité génétique du virus OsHV-1. Les futures études dans ce domaine gagneront à adopter une stratégie d'échantillonnage et d'analyse similaire à celle utilisée dans l'étude II. La recherche va dans ce sens, un pipeline de type nf-core : maeva, très proche de ce

qui a été utilisé dans l'étude II, est en cours de développement par l'équipe du département de bioinformatique SeBiMer de l'IFREMER pour faciliter l'analyse de la diversité génétique.

La caractérisation de la diversité génétique du virus OsHV-1

La caractérisation de la diversité génétique se fait systématiquement de manière relative, il est donc délicat de comparer directement le nombre de variations entre les études ou encore entre différents virus. En effet, il est nécessaire de prendre en compte plusieurs paramètres qui influencent directement la mesure de la diversité génétique tel que le nombre de génomes utilisé, la variabilité spatiale entre ces génomes, la variabilité temporelle d'échantillonnage et enfin la variabilité analytique (**Annexe.5 p144**). Bien souvent ces paramètres ne sont clairement pas mis en évidence dans la littérature et de nombreuses études se contentent de mettre en avant le % de variabilité du génome. Toutefois, à défaut de données plus appropriées actuellement, de telles comparaisons peuvent nous aider à mieux appréhender le rôle de la diversité génétique du virus dans les POMS. Dans les deux études précédentes le nombre de variations génétiques observées par rapport au génome de l'OsHV-1 µVar A était du même ordre de grandeur malgré les variations expérimentales décrites ci-dessus. Pour la zone de Brest 377 SNPs pour Brest I et 136 SNP / 30 INDEL en moyenne pour Brest II (**Figure 1D article 2**). Pour Thau 168 SNPs pour Thau I et en moyenne 144 SNPs / 26 INDEL pour Thau II. La 3e zone conchylicole de Marennes-Oléron présente encore davantage de polymorphisme avec en moyenne 170 SNPs / 36 INDEL par rapport au génome du µVar A. Contrairement à ce qui est observé dans les deux autres zones géographiques, pour la zone de Marennes-Oléron la variation interindividuelle est très importante donc ce nombre est probablement sous-estimé. Finalement, en se plaçant dans des conditions pour maximiser la variabilité génétique, il est peu probable que le nombre de positions génomiques soit très important. Avec les 0.27% de variations observés entre les 21 génomes assemblés on peut donc estimer que pour une même année dans l'espèce C. gigas, moins de 0.5% du génome du virus pourrait être variable (en cumulant majoritaire et minoritaire < 1000). Cette assertion semble être en désaccord avec les études antérieures menées sur OsHV-1 (Bai et al., 2018; Burioli et al., 2017; Xia et al., 2015) car contrairement à ce qui vient d'être fait, elle définit la variabilité au sein d'un type d'hôte particulier et à une date spécifique. Ces études sur la diversité génétique surévaluent la diversité en prenant toutes les séquences peu importe l'hôte, l'échantillonnage géographique et sa date (**Tableau.1. p112**). Il convient par exemple d'observer que les échantillons ayant aboutis à la purification de l'ADN et au séquençage du premier génome (AY509253) sont antérieurs aux années 2000 et la distance entre certains sites d'échantillonnages peut varier jusqu'à 8000 KM. On s'attend donc à ce que la dérive génétique avec ou sans une quelconque pression de sélection modifie la séquence génomique du virus. C'est pourquoi à l'avenir il est recommandé de prendre d'avantage en compte cet aspect spatio-temporel dans l'interprétation des résultats des prochaines recherches sur la diversité génétique du virus. En particulier lors de l'étude de la diversité génétique au sein d'une même espèce hôte pour ne pas surestimer l'interprétation des comparaisons génomiques.

N	Accession number	size (pb)	autre noms	assembler	gene caller	Année	lieux	ref
1	-	-	LT/Br/Th	metaSPAdes	Prokka? / ORFfinder?	2018	France (LT/Br/Th)	Etude II
2	MF509813	204 652	ZK0118	SOAPdenovo/Celera	GeneMarkS et GATU	2001	Chine (Shandong)	Bai et al. 2018
3	MG561751	203 983	OsHV-1-PT	IDBA-UD*	ORFfinder	2016	Italie (Porto Tolle)	Abbadi et al. 2018
4	KY242785	204 886	OsHV-1 µVar A	Bowtie2/MindTheGap	ORFfinder	2010	France (Meuvaines, Blainville, Géfosse)	Burioli et al., 2017
5	KY271630	204 897	OsHV-1 µVar B	Bowtie2/MindTheGap	ORFfinder	2010	France (Blainville, Géfosse, Trelande)	Burioli et al., 2017
6	KP412538	199 354	OsHV-1-SB / CDSB2012	DNASTAR**	ORFfinder	2012	Chine (Qingdao)	Xia et al., 2015
7	GQ153938	210 993	AVNV	DNASTAR**	ORFfinder	2007	Chine (Qingdao)	Ren et al., 2013
8	KU096999	199 102	HaHV-1/ AbHV-1	Genomics Workbench	Artemis	2005	Taiwan / Australie?	Savin et al., 2010
9	AY509253	207 439	OsHV-1 ref	Pregap4 Gap4	GCG Programs	1997	France (?)	Davison et al., 2005

Tableau 1. **Récapitulatif des génomes assemblés associé à OsHV-1**.
L'ensemble des virus sont classés dans les Ostreavirus excepté pour KU096999 qui correspond au genre des Aurivirus. Lorsqu'il y a des contradictions ou que les données sont manquantes un point d'interrogation est ajouté. * assemblage guidé par la référence ou en sélectionnant les reads contre la référence AY509253. ** basée sur la PCR, la technique de marche des amorces, en utilisant la séquence de référence OsHV-1 comme matrice. OsHV-1-SB est issu de *Scapharca broughtonii* (palourde).

Lors de nos recherches nous avons pu observer de potentielles signatures de coévolution dans l'article I. Pour rappel, des modifications des glycoprotéines membranaires putatives suggèrent ainsi que les zones de l'étang de Thau I et Brest I possèdent des variants différents qui pourraient être l'image d'une coévolution locale. Or comme nous l'avons vu dans l'article II les variants sont majoritairement originaires de la zone conchylicole de la baie de Marennes-Oléron. Cette assertion limite les possibilités de contact des populations virales et de leurs hôtes pour l'établissement d'une adaptation locale. Toutefois, dans l'analyse d'ensemble des zones de Brest II et Thau II, il existe un faible nombre de variations homogènes entre les

échantillons qui semble indiquer une sélection directionnelle. Nous ne pouvons pas déterminer si ces variations sont dues à une adaptation locale des populations virales introduite ou tout simplement de la sélection d'un variant suite au goulot d'étranglement lors du transfert des populations virales sur les deux autres sites. Cette dernière hypothèse semble moins probable car la force de la dérive génétique dépend principalement de la taille effective de la population, les petites populations subissant une dérive plus forte (Retel et al., 2019). Malheureusement, comme nous n'avons pas déterminé l'annotation des gènes dans nos échantillons, nous ne pouvons donc pas explicité si ces variations sont retrouvées ou non dans des séquences codantes. Nous ne pouvons pas exclure par exemple que ces variations soient les mêmes que celles modifiant la structure des protéines membranaires putative observées dans l'article I. Même si la quantité de mutation impliquée dans cette possible adaptation locale est faible, des travaux futurs devront réévaluer ces mutations avec l'annotation des gènes pour mieux comprendre s' il existe ou non une force de sélection locale exercée sur ces populations virales.

Enfin, dans les deux études le profil des fréquences allélique n'était pas tout à fait les mêmes avec des fréquences distribuées entre les deux extrêmes 0 et 1 pour les échantillons issus des 3 zones conchylicoles de l'étude II (**Figure 4C** article II) et pour la zone de l'étang de Thau I (**Figure 3 B & C** et **Figure 6B & C** article I). A l'inverse, les fréquences alléliques de Brest I sont plus distribuées autour de 0, 0.5 et 1 (**Figure 3**. **B & C** article I). Il est assez compliqué d'expliquer cette différence de résultat. Il est possible qu'à cause du phénotype contrasté des huîtres les fréquences alléliques des données de séquençage soient diluées artificiellement lors de la fusion des alignements. Les profondeurs étant globalement similaires pour Thau I ce phénomène est moins visible (**Fig.19. p58**). En se basant sur les fréquences alléliques des quelques loci qui pourraient être impliqué dans la coévolution avec l'huître nous pouvons faire l'hypothèse d'une dynamique co évolutionnaire de course à l'armement. Cette hypothèse semble en accord avec les travaux ayant abouti au programme de sélection massale réalisé sur les huîtres. Elle pourrait être testé à l'aide d'échantillons de populations virales issus de plusieurs années passées sur des huîtres actuelles ayant survécu comme par exemple les huîtres issues de programmes de sélections génétiques (Degremont et al., 2015). Une expérience récente adopte une approche similaire avec l'exposition d'huîtres résistantes à deux variants, l'un issu de Tomales Bay aux USA ou le µVar n'a pas encore été observé (et donc probablement proche de la référence), et l'autre issu d'épisodes de mortalité en france où sévit l'OsHV-1 µVar(Agnew et al., 2020). Les huîtres françaises issues de programmes de sélections semblent résistantes aux virus issus des deux origines confirmant d'une certaine façon cette dynamique de course à l'armement.

Si on compare la diversité génétique du virus OsHV-1 (ou des malaco herpesviridae) par rapport à d'autres espèces virales. Le virus OsHV-1 était présumé avoir une moindre diversité génétique par rapport aux virus à ARN, cela semble être confirmé avec nos recherches. Il est hautement improbable que la variabilité génomique soit de l'ordre de celle observée pour les virus à ARN tel que le VIH (Li et al., 2015). Au vu de la faible quantité de positions variables sur le génome des virus OsHV-1 et les données antérieures il est très probable qu'une enzyme polymérase similaire à celle des virus à ADN classique le prive de mécanisme de génération de diversité identique à celui des génomes à ARN. Par rapport aux virus à ADN, si l'on compare à des membres connus HSV-1 (*Alphaherpesvirinae*) ou HCMV (*Betaherpesvirinae*) qui infectent

l'homme, là encore la diversité génétique semble plus faible. En effet, de récente comparaisons à l'échelle du génome entier d'herpès simplex ou du Cytomegalovirus Humain ont révélés des variations nucléotidiques affectant 3% à 4% du génome (Akhtar et al., 2019; Bradley et al., 2009; Parsons et al., 2015; Renner and Szpara, 2017). Nous avons observé seulement 0.27% de variations chez OsHV indiquant une diversité génétique intra-individuelle et inter-individuelle de l'OsHV-1 inférieure à celle des herpèsvirus de vertébrés au sein d'un même hôte. Même si cette comparaison souffre des mêmes biais identifiés précédemment (**Annexe.5 p144**), si nos résultats sont confirmés par des recherches futures, il faudra comprendre pourquoi l'OsHV-1 présente un niveau d'hétérogénéité différent des herpèsvirus de vertébrés. Une des pistes envisagées ici est que la diversité des populations virales de l'hôte pourrait être influencée par le système immunitaire de l'hôte

L'influence du système immunitaire de l'huître sur la diversité virale

L'étude des mécanismes du système immunitaire de l'huître est une thématique de recherche active, nous savons que c'est une machinerie complexe qui présente tout le répertoire nécessaire à la mise en place d'une réponse immunitaire efficace avec des capacités d'adaptation importantes (Klein, 1997). Nous savons que le système immunitaire des vertébrés joue un rôle important dans l'évolution et la diversité génétique de certains virus. En effet, des mécanisme d'adaptation à la variabilité antigénique comme la recombinaison V(D)J ou l'hypermutation somatique des lymphocytes entraînent une pression de sélection immunitaire continue et forte, provenant soit d'anticorps neutralisants, soit de lymphocytes T cytotoxiques. A ce jour, il n'existe pas de mécanisme similaire mis à jour chez *C. gigas*. Nous pouvons donc faire deux hypothèses sur l'influence du système immunitaire dans l'évolution de la diversité génétique du virus OsHV-1.

Si le système immunitaire de l'huître ne possède pas de récepteur hypervariable et que le système immunitaire ne possède pas de réponse immunitaire spécifique à l'OsHV-1, il y a donc peu de potentiel pour que les mutants viraux soient avantagés et la diversité virales serait principalement le reflet de la dérive génétique. Cette assertion s'appuie sur l'observation d'un impact négatif sur la valeur sélective des populations d'Herpèsvirus lors d'un gain important de diversité. Un récent travail de Trimpert et coll. (2019) a ainsi montré la nécessité évolutive de l'activité exonucléase pendant la réplication du virus à ADN comme un paramètre sélectionné chez les populations virales d'Herpèsvirus . Dans ces recherches, le virus de la maladie de Marek (MDV), un alphaherpèsvirus du poulet, a permis la construction de virus possédant des mutations dans les domaines exonucléase de l'ADN polymérase virale. Les mutations de l'ADN polymérase viral étaient viables et ont compromis son activité de relecture. Lors de la confrontation du virus modifié avec son hôte la sévérité de la maladie a considérablement augmenté avec la présence de populations virales de MDV très diverses proches de ce qui peut être observé pour les quasi-espèces avec une augmentation importante de la virulence *in vivo* (Trimpert et al., 2019).

Si le système immunitaire de l'huître possède des récepteurs capables de se diversifier pour s'adapter à la variabilité antigénique, alors la dynamique de la dérive antigénique entraînera une «course aux armements» co évolutionnaire plus forte et donc une diversité génétique plus importante (Grenfell, 2004). A ce jour 190 séquences homologues à des FREPs ont été retrouvées dans le génome de *C. gigas* (Huang et al., 2015; Lafont, 2017; Zhang et al., 2012, 2015). Même si la région amino-terminal du domaine de la superfamille d'immunoglobulines (IgSF) n'a pas encore été recherchée ou identifiée nous ne pouvons pas exclure qu'un récepteur hypervariable ou d'autres mécanismes existent.

Les recherches futures qui pourront combiner l'analyse de l'hôte et du virus en parallèle à travers des études d'association à l'échelle du génome (GWAS) pour les deux protagonistes. Ces recherches pourraient aboutir à la caractérisation de loci très variables et potentiellement à l'identification de nouveau récepteur immunitaire chez l'huître. Dépendamment de la pression exercée par le système immunitaire de l'huître, la diversité génétique du virus progresse lentement et l'évolution pourrait être constante. Ces mécanismes de coévolution n'expliquent pas entièrement l'apparition du génotype µVar avec ses nombreuses variations.

L'espèce hôte, un facteur primordial pour expliquer la diversité virale de l'OsHV-1

Pour la plupart des Herpèsvirus de vertébrés le spectre d'hôte est restreint (McGeoch et al., 2006). Or dans le cas d'OsHV-1, le virus a pu être détecté dans une multitude d'hôtes, d'autres espèces d'huîtres, de palourdes, de coques, de pétoncles et de moules au cours des dernières décennies (Alfaro et al., 2018). Bien que nous ne savons pas si ces espèces où le virus ont été observées ne sont que des porteurs passifs du virus, cela sous-entend une gamme d'hôtes importante. Ainsi, en prenant en compte les séquences disponible entre les différents hôtes du virus Bai et coll (2018) ont récemment observés 1583 polymorphismes majoritaires (décrit comme étant des SNPs) dans la comparaison de OsHV-1 ref, OsHV-1 µVar A/B, OsHV-1-SB et les 6 scaffolds issues de leurs études sur le variant infectant *Scapharca broughtonii* et *Chlamys farreri*. Dans les conclusions de leurs recherches, ces auteurs suggèrent que l'origine géographique et les espèces hôtes jouent un rôle dans le processus de différenciation génétique des variants d'OsHV-1(Bai et al., 2018). Cette assertion semble particulièrement intéressante au regard de la propriété des mollusques bivalves à s'hybrider entre eux. En effet, la spéciation de deux espèces est un événement qui ne se déroule probablement jamais de manière allopatrique ou instantané (Abbott et al., 2013). L'hybridation interspécifique concerne souvent deux espèces du même genre et est considérée comme un moteur important de l'évolution du génome (Depotter et al., 2016). Une hybridation interspécifique peut se produire lorsque des gamètes d'espèces différentes fusionnent en un zygote viable pour des espèces à fécondation externe. De tels phénomènes d'hybridations peuvent conduire à l'expansion rapide des pathogènes émergents et entraîner des épidémies. Des travaux chez les plantes ont ainsi montré que l'hybridation a joué un rôle central dans la deuxième épidémie de maladie de graphiose de l'orme en Europe (Brasier, 2001; Depotter et al., 2016). Des organismes hôtes hybrides auraient facilité le flux génétique d'une espèce de champignons et supplanté l'espèce pathogène locale à terme (Bovers et al., 2006; Brasier, 2001; Depotter et al., 2016). Des processus d'hybridation interspécifique ont déjà été signalés parmi les espèces d'huîtres *Crassostrea ariakensis*, *Crassostrea angulata* et *Crassostrea gigas* (Leitão et al., 2007; Yao et al., 2015). Il est donc

envisageable que des hybrides interspécifiques de mollusque bivalve aient joué un rôle clé dans la diversité génétique du virus et plus particulièrement dans l'émergence du variant OsHV-1 µVar. L'infection d'une huître présentant un patrimoine génétique hybride avec un genre voisin pourrait être à la fois le lieu de rencontre privilégié pour l'échange de matérielle génomique grâce aux recombinaisons entre des variants infectants différents hôtes, mais également le lieu de sélection et d'amplification idéale de ces variants. Ce phénomène aurait pu favoriser l'émergence du µVar il y a longtemps, plusieurs travaux suggèrent qu'il a eu lieu en Asie de l'est avant d'être introduit dans d'autres régions à travers le monde (Mineur et al., 2014; Shimahara et al., 2012). Par la suite le virus OsHV-1 µVar s'est rapidement répandu dans toute l'Europe, l'Australie et la Nouvelle-Zélande, causant la mortalité des huîtres du Pacifique de tous les stades de leur cycle biologique, y compris les huîtres adultes (Jenkins et al., 2013; Keeling et al., 2014; Schikorski et al., 2010a, 2010b; Segarra et al., 2010; Whittington et al., 2018). L'émergence du variant comme l'OsHV-1 µVar est donc préoccupante dans des conditions d'élevage intensif, où différentes espèces de bivalves sont maintenues à proximité. Cela semble avoir été observé pour l'OsHV-1 Var, qui est différent de l'OsHV-1 'réf' et de l'OsHV-1 µVar (Arzul et al., 2001a; Renault et al., 2012). Une prospection plus globale sur des zones où sont localisés les différents hôtes du virus à l'aide de technologie de séquençage pourrait aider à identifier l'apparition de ces variants. De plus, ils pourraient ainsi être intéressants d'évaluer le degré d'hybridation des hôtes sur les trois sites conchylicoles que nous avons échantillonnés. Potentiellement à l'image des travaux de Burioli et coll (2016) qui ont essayé de discriminer des hybrides *C.gigas* x *C.angulata*. Pour finir, il faut souligner que le nombre d'événements d'hybridations entraînant l'apparition de tels variants devrait augmenter dans les prochaines années. Principalement en raison de la mondialisation, des transferts d'animaux, ainsi que des conditions d'élevage intensif permettant le contact physique entre des organismes géographiquement séparés.

Les facteurs de la triade épidémiologique expliquent davantage la variabilité phénotypique des huîtres lors des POMS

Les taux de mortalité affectant les huîtres sont connus pour avoir une origine multifactorielle. Lors de l'identification des événements de mortalité, les familles d'huîtres sensibles à l'OsHV-1 ont montré une perturbation de la structure de la communauté bactérienne lorsqu'elles cohabitent avec des huîtres injectées d'OsHV-1 (de Lorgeril et al., 2018). D'autres travaux plus récents ont également montré des différences sur la structure et la composition des populations bactériennes entre les familles d'huîtres résistantes et sensibles (Clerissi et al., 2020). De plus, cette composition de communauté bactérienne présente des variations adaptatives lorsqu'elles sont placées sur différentes zones géographiques (Pathirana et al., 2019) ce qui pourrait mieux expliquer le changement de phénotypes observés dans l'étude I. En plus de ces modifications du microbiote bactérien, la composante génétique de la résistance des huîtres est admise dans le domaine (Azéma et al., 2015; Degremont, 2011; Degremont et al., 2015; Sauvage et al., 2009). Les programmes de sélection massale permettent ainsi d'améliorer considérablement la survie des huîtres. Une étude d'association à l'échelle du génome de l'huître a montré que la résistance à OsHV-1 est polygénique, avec un groupe de liaison contenant 6 locus de caractères quantitatifs (*quantitative trait loci*) important pouvant jouer un rôle dans la résistance de l'hôte (Gutierrez et al., 2018). Des recherches ont comparé le niveau

transcriptomique basale (avant infection) entre des huîtres résistantes et sensibles et a mis en évidence une différence de statut immunitaire selon le phénotype de résistance avec une régulation à la hausse des des gènes immunitaires (de Lorgeril et al., 2020). Cet état immunitaire permet notamment une réponse antivirale plus rapide chez les huîtres résistantes par rapport aux d'huîtres sensibles (de Lorgeril et al., 2018). Cet état immunitaire peut également être amorcé artificiellement à l'aide de molécules immuno stimulante comme le polyI:C (de Lorgeril et al., 2018; Green et al., 2015) pour permettre un transfert immunitaire qui prépare l'hôte à faire face à l'infection, et lui confère une protection (Lafont et al., 2020). Même si tous les acteurs ont un rôle à jouer dans la manifestation de la maladie, les récentes avancés sur chacun de ces éléments semble indiquer que le fond génétique des huîtres et le microbiote sont plus à même d'expliquer la variabilité phénotypique des huîtres vis à vis des POMS que la diversité génétique du virus. En l'absence de changement important comme l'apparition ponctuelle de génotypiques comme l'émergence de l'OsHV-1 µVar, le scénario le plus probable est que l'infection virale amoindrie la réponse immunitaire de l'hôte quelque que soit son hétérogénéité. Lorsque celle-ci est mal préparée, la septicémie des communautés bactériennes influencée par plusieurs facteurs va aboutir à la mort de l'animal.

Perspectives

L'avènement des technologies de HTS rapides et bon marché depuis ces dernières années a fait du re-séquençage l'un des outils d'analyse de la diversité génétique les plus populaires, et il est raisonnable de s'attendre à ce que ce ne soit que le début. Dans cette section, nous abordons trois pistes de recherches différentes: l'harmonisation des annotations du génome, les analyses pan-génomiques et les approches en séquençage de longs fragments. Bien que cette liste ne soit pas exhaustive, nous avons cherché à sélectionner ceux qui, selon nous, sont les plus pertinents pour l'étude afin de mieux comprendre le rôle de la diversité génétique du virus OsHV-1 lors des POMS.

L'urgence d'harmoniser l'annotation des gènes d'OsHV-1

Pour comprendre quel peut être l'impact de la diversité génétique du virus il est nécessaire d'annoter les mutations observées dans la séquence codant pour des protéines et d'estimer des paramètres évolutifs. Initialement lors de la publication du 1er génome séquencé (Davison, 2005) 124 gènes uniques du virus codants pour les protéines ont été définis à l'aide des règles suivantes :

- Les ORF sont composés d'un codon d'initiation, suivi de minimum 100 acides aminés et d'un codon stop, sauf pour les ORF16, ORF36 et ORF37.

- Les ORF plus petits chevauchant les plus grands de plus de 60% de leur longueur ont ensuite été réduits, sauf pour ORF17 et ORF18 qui codent pour une protéine membranaire putative.

- Un ORF de 117 codons a été écarté car il enjambait la jonction X – RS, avait une teneur en G + C inhabituellement élevée et manquait d'un signal de polyadénylation potentiel.

Cependant, selon les outils d'annotation utilisés, la définition d'un ORF peut changer et l'annotation d'une même séquence d'ADN peut être différente selon l'outil. De plus, même si la majorité des outils bioinformatiques d'annotation des gènes sont automatiques, certains auteurs créent ou modifient eux même les fichiers d'annotation. Avec la multiplication des génomes et l'apport de nouveaux outils des discordances peuvent s'introduire dans l'annotation des génomes. Initialement les outils bioinformatiques conçus dans l'étude II, où 21 génomes ont été assemblés, avaient été conçus en prenant en compte l'annotation des ORFs du virus et l'analyse de l'impact des mutations sur les protéines virales. Nous voulions utiliser l'outil prokka (prodigal) pour annoter nos génomes (Hyatt et al., 2010; Seemann, 2014). Or en validant nos résultats nous avons pu voir que l'annotation des gènes du virus présente de nombreuses discordances. Les résultats préliminaires suivants ont pour objectif de mettre en évidence les incohérences d'annotations présentent dans les génomes d'OsHV-1 déjà publiés.

Dans un premier temps nous avons voulu vérifier les différences entre les fichiers d'annotations des génomes connus. Pour cela l'ensemble des fichiers d'annotation (GFF) ont été comparés en se basant uniquement sur le nom des ORFs viraux et non pas la séquence (**Fig.23. p119**). Nous pouvons observer qu'il existe des erreurs d'annotations probablement dûes à une annotation manuelle des ORFs 40 et 24 pour MF509813 (ZK0118 3eme colonne) (Bai et al., 2018). Pour éviter ce genre d'incohérences il serait souhaitable d'éviter au maximum d'entrer des annotations manuellement à l'avenir. Si jamais des changements sont nécessaires, nous encourageons l'utilisation d'un script et d'en faire mention dans le matériel et méthode afin d'avoir une analyse reproductible.

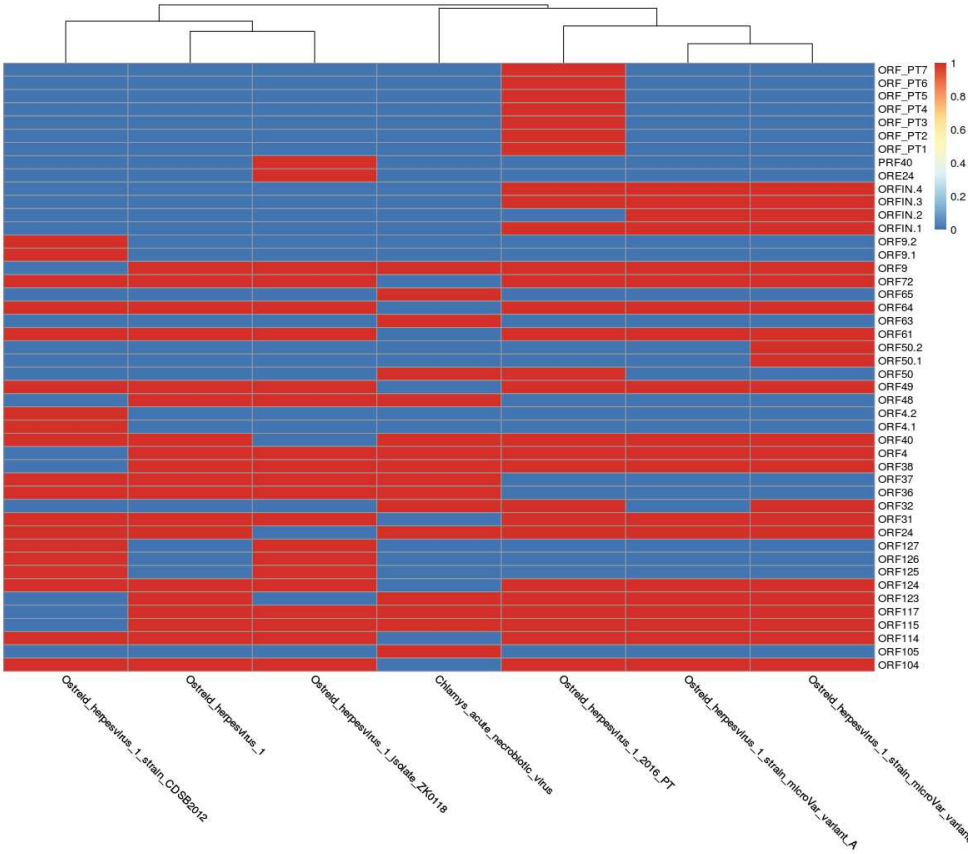


Figure 23. **Erreur d'annotation pour les génomes déjà disponible dans les bases de données**.
Les annotations des 7 génomes viraux associés à l'OsHV-1 ont été récupérées à partir de NCBI, dans l'ordre KP412538, AY509253, MF509813, GQ153938, MG561751, KY242785 et KY271630. Les données GFF3 ont simplement été transformées en tableau à l'aide du logiciel AGAT (Dainat, 2020) et visualisées sur une analyse sur R au travers d'une matrice binaire de présence (rouge) / absence (bleu). Les protéines ne présentant aucune anomalie dans le nom pour tous les virus ne sont pas affichées.

Dans un second temps, nous avons testé l'outil en ré-annotant la séquence d'ADN de l'OsHV-1 µVar A (KY242785) à l'aide de prokka (Hyatt et al., 2010; Seemann, 2014) et d'une base de données protéiques de toutes les séquences connues des OsHV-1 déjà publiés. Nous avons ensuite comparé cette annotation avec celle déjà existante issue de la publication. Nous avons pu observer 27 ORFs avec une variation par rapport à l'annotation existante (**Fig.24. p120**). Au total, sur ces 27 ORFs dont 17 présentaient de petites variations de taille, et 10 d'entre eux étaient absents de l'une ou de l'autre des annotations. Sur ces 10 ORFs, 2 ORFs ne sont pas correctement annotés par prokka : les ORFs 16 et 72, et 8 sont annotés dans prokka, mais absents de l'annotation de NCBI. Parmi ces 8 ORFs nous avons retrouvé les ORFs PT1-6 dans le génome de l'OsHV-1 µVar A. Pour aller plus loin, nous avons regardé rapidement ou ce situe ces ORFs issues de l'annotation de l'OsHV-1-PT. Dans certain cas comme l'ORF PT4 et l'ORF PT5, l'annotation est retrouvée à la même position que les ORF 73.1 et 73.2 qui sont annotés comme non fonctionnels dans l'OsHV-1 AY509253 et dans les OsHV-1 µVar A/B KY242785/KY271630 (**Fig.25. p121**). Ces résultats suggèrent que les outils d'annotation peuvent avoir un fort impact sur l'annotation des gènes du virus OsHV-1. L'observation des

gènes est interprétée comme spécifique dans l'OsHV-1-PT, ces derniers peuvent être retrouvés dans l'OsHV-1 µVar A. Or, ces ORFs sont suggérés comme de nouveaux ORFs dans l'article de recherche sur le génome de l'OsHV-1-PT, c'est sur cette base que les auteurs suggèrent que ce virus pourrait être un nouveau microvariant (Abbadi et al., 2018).

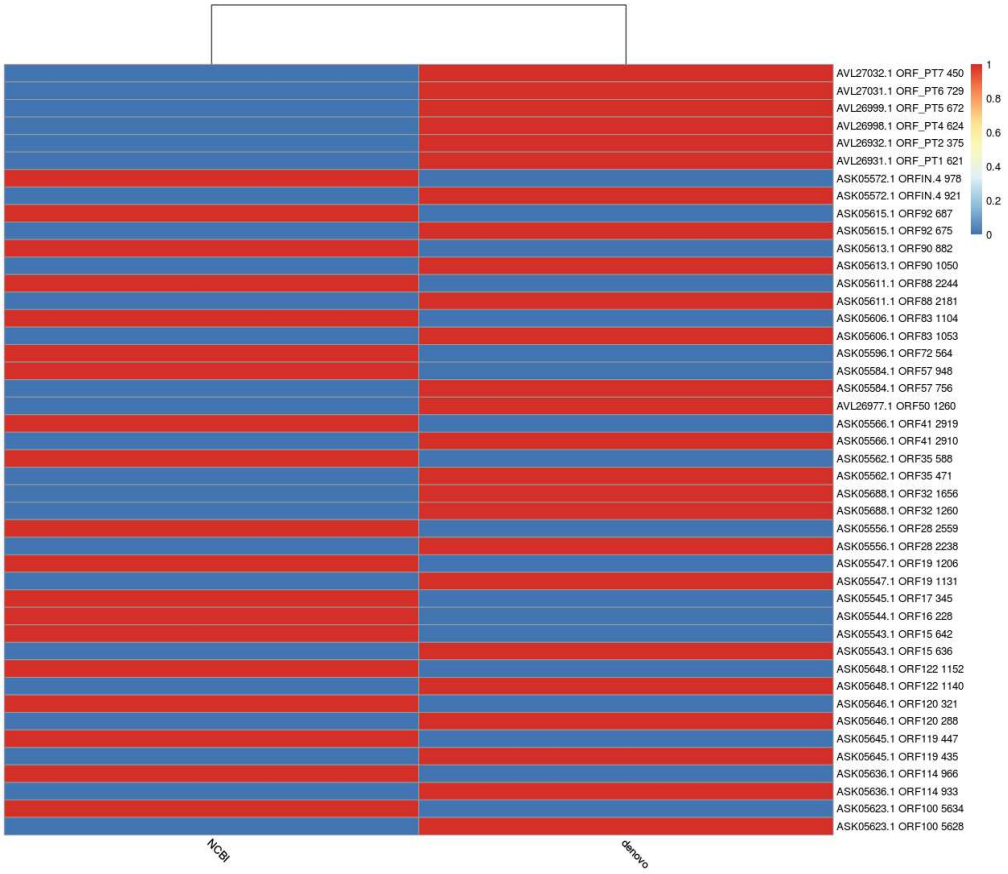


Figure 24. **Comparaison d'une même séquence avec l'annotation prokka avec celle existante**.
Les annotations de gènes de l'OsHV-1 µVar A (gauche, NCBI) ont été comparées à l'annotation de séquence de gènes issus de Prokka (droite, de novo). À droite de la figure, les annotations contiennent les noms des ORFs avec leurs tailles. Seuls les ORFs ayant des variations en tailles, ou différents entre les deux annotations (16/72 et ORF PT1 à 6) sont affichés. Les ORFs prédits comme hypotheticals protéines ne sont pas montrés.

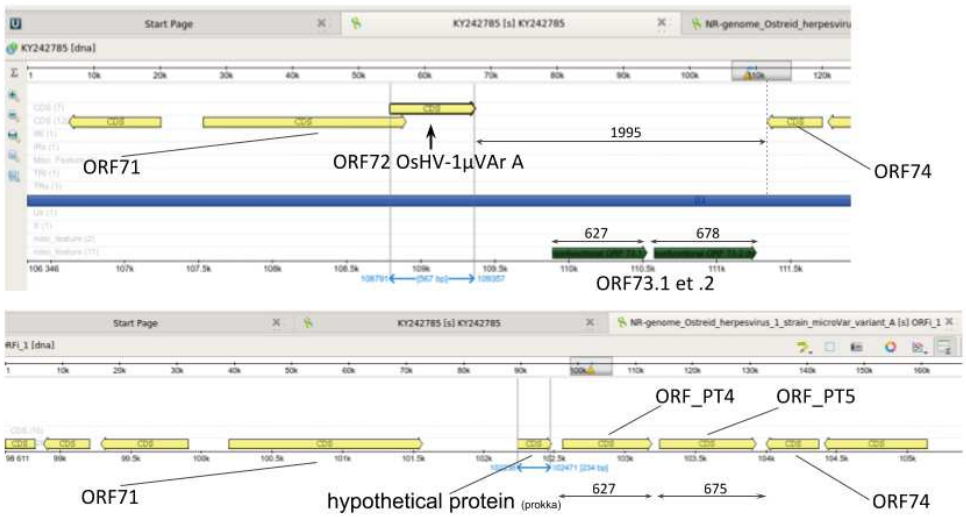


Figure 25. **Visualisation de l'erreur d'annotation de l'ORF PT4 et PT5**.
Les séquences de l'OsHV-1 µVar sont visualisées sur Ugene avec l'annotation disponible sur Ugene (haut) et l'annotation issue de Prokka (bas). Cette dernière a été construite en utilisant l'ensemble de séquences codant des ORFs parmi lesquelles les ORFs PT4 et PT5 sont retrouvés (issue de MG561751).

Ensemble ces résultats suggèrent qu'il est urgent de redéfinir les protéines du virus OsHV-1 avec des outils plus appropriés. Le choix de l'outil devra limiter au maximum les biais, c'est pourquoi il serait préférable d'utiliser un outil d'annotation *ab initio* qui utilise un ensemble d'apprentissage avec une annotation connue pour l'entraînement des paramètres en parallèle de données de transcriptomique pour valider ces prédictions. Cette harmonisation de l'annotation des ORFs du virus est obligatoire pour évaluer la présence de codons stop et de mutations non-synonymes dans des ORFs et ainsi dégager les paramètres évolutifs des protéines (e.g. le dn/ds, la diversité nucléotidique etc..).

L'approche pangénomique

Outre les paramètres évolutifs du virus, l'annotation correcte des gènes permettra de connaître s'il existe des variations du nombre de gènes entre différents génomes. Or pour appréhender ce type de question il est nécessaire de faire une analyse **pangénomique**, c'est-à-dire des analyses décrivant la gamme complète de gènes dans une espèce. Ce type d'analyse est essentiel pour mieux comprendre comment le virus OsHV-1 se comporte avec ses différents hôtes. Ce type d'approche devrait se développer à l'avenir et permettra de limiter les erreurs d'annotation comme pour le génome de l'OsHV-1-PT (Abbadi et al., 2018).

Pour aller dans ce sens, nous avons fait des recherches préliminaires sur l'analyse pangénomique des 27 génomes utilisés dans l'étude II (**Fig.26. p122**). Ces génomes ont été utilisés sous forme non redondante pour être annotés avec porkka comme décrit précédemment. Puis nous avons comparé les ORFs en fonction de l'homologie de séquence à l'aide de la suite Anvi'o (Eren et al., 2015). Nos résultats semblent indiquer qu'il y aurait la présence de 148 groupes de gènes chez OsHV-1 dont 101 sont uniques et présents dans tous les

génomes et 8 gènes sont retrouvés comme des singletons. Toutefois l'un des problèmes à surmonter dans le futur sera d'utiliser des génomes complets dans l'analyse. En effet, l'ORF 6 et l'ORF 114 sont respectivement dans l'intersection TRL-UL et UL-IRL biaisant notre pré-analyse.

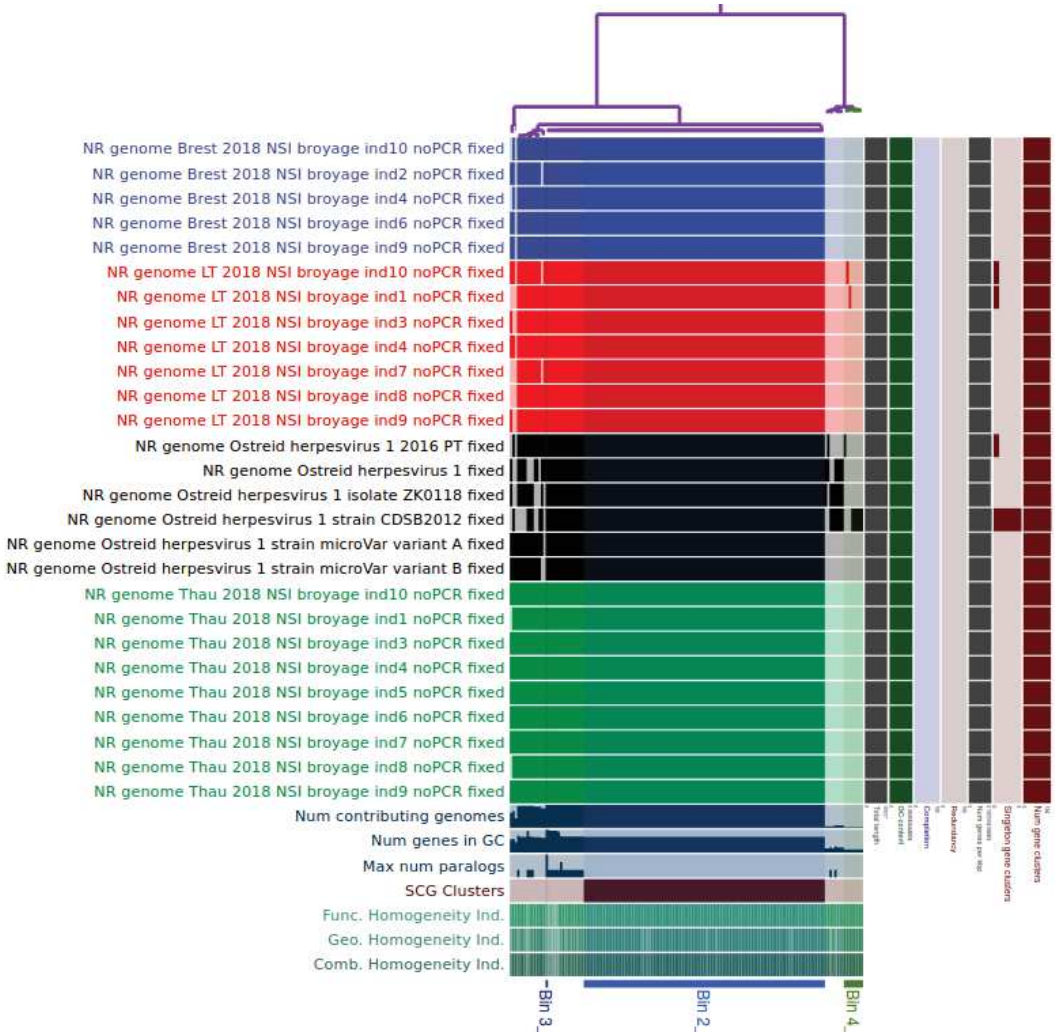


Figure 26. **Pangénome des virus connus d'OsHV-1**.
Les séquences protéiques des 27 génomes, sous formes non redondantes, du virus OsHV-1 ont été utilisées. Les isolats de l'étude II sont représentés avec des couleurs : Brest (bleu), de Marennes-Oléron (rouge) et étang de Thau (vert) ainsi que les 6 génomes déjà publiés (noir). Chacun des 148 groupes de gènes contient un ou plusieurs gènes apportés par un ou plusieurs génomes isolés. Les gènes présents en 1 exemplaire dans tous les génomes sont retrouvés dans le Bin 2 (en bleu) avec 101 groupes de gènes. 8 gènes sont retrouvés comme des singletons dans le Bin 4 (en vert).

Ces résultats préliminaires semblent indiquer que l'approche est pertinente pour de futures recherches comme il existe des variations géniques dans le temps, l'espace et les hôtes du virus OsHV-1.

L'assemblage de longs fragments du génome d'OsHV-1

Bien que le coût du séquençage des technologies de HTS ait chuté ces dernières années, ces approches sont chères et nécessitent souvent l'envoi d'échantillons à un tiers du fait de l'utilisation de grandes machines dans des plateformes dédiées. Les dernières améliorations techniques ont conduit à une nouvelle génération de dispositifs de séquençage, notamment le séquenceur portable *Oxford Nanopore MinION* permettant un séquençage en dehors des laboratoires moléculaires traditionnels (Quick et al., 2016). En plus, de permettre une stratégie portable de surveillance des POMS, ce type de séquençage détermine la séquence nucléotidique de longs fragments. Ces longs fragments permettent de faciliter l'assemblage de génomes complexes comme celui de l'OsHV-1. Cependant il demande une certaine quantité d'ADN de bonne qualité du fait de l'absence de lignée cellulaire viable pour cultiver l'OsHV-1. Une approche a été utilisée pour l'assemblage de l'OsHV-1 ZK0118 (MF509813) cependant une étape de PCR était nécessaire pour amplifier l'ADN du virus, mais introduit un biais conséquent dans l'analyse de la diversité virale (Bai et al., 2018). Dernièrement nous avons mis au point un protocole utilisant la filtration tangentielle pour extraire et purifier suffisamment d'ADN du virus sans amplification d'ADN. L'utilisation de cette technologie couplée à des campagnes d'échantillons dans le temps sera utile pour une meilleure compréhension de l'évolution des populations virales d'OsHV-1.

Conclusion

En conclusion, ce travail de thèse enrichit nos connaissances sur l'interaction virale hôte-pathogène chez les mollusques au travers de la création d'outils bioinformatiques pour mieux caractériser la diversité génétique du virus OsHV-1. Nous avons multiplié par quatre le nombre de génomes de malacoherpèsvirus dans les bases de données. Les résultats de nos recherches suggèrent que la diversité génétique du virus chez l'huître est principalement façonnée par les déplacements de l'hôte, et que l'apport des populations d'huîtres sauvages pourrait constituer une source d'hétérogénéité des populations. Ces dernières pourraient être le réservoir des populations virales. Nos recherches ont permis d'observer et quantifier la variabilité inter-individuelle ce qui nous a permis d'avoir une vision nouvelle sur la diversité du virus au sein d'une espèce hôte *C.gigas*. Toutefois, cette diversité était relativement faible comparée aux connaissances acquises sur la diversité génétique d'autres virus. Notamment chez les virus à ARN avec la théorie des quasi-espèces qui représente une dynamique d'adaptation importante. Pris ensemble, nos résultats suggèrent que le moteur de l'évolution du virus est principalement façonné par la dérive génétique. Bien que beaucoup de travaux de recherches soient encore nécessaires pour confirmer nos hypothèses, il semble que l'évolution du virus est relativement lente au sein de *C.gigas*. Les différents facteurs impliqués dans le syndrôme de mortalité des huîtres du pacifique semble mieux expliquer les épisodes de mortalité observés. Ainsi, bien que des facteurs génétiques de l'huître soient impliqués dans des phénomènes de résistance, ils sont vraisemblablement le reflet d'une meilleure adaptation au stress induit par la température de l'eau de mer qu'au virus. Toutefois la diversité génétique du virus pourrait s'accroître du fait de la présence d'un grand nombre d'hôtes disponibles pour le virus. Nous avons ainsi fait l'hypothèse que de tels réarrangements de la séquence génomique du virus

pourraient être issus de l'infection d'espèces hôtes hybrides. C'est dans ces conditions qu'un génotype du virus peut présenter un risque important d'épizootie comme l'a montrée l'apparition en 2008 du génotype du µVar. Du fait de la surpopulation d'huître dans les zones conchylicoles, d'une homogénéité de celle-ci avec des programmes de sélection massales et l'existence de culture d'autres mollusques à proximité, l'émergence de nouveaux génotypes pathogènes est susceptible d'avoir un impact significatif sur l'industrie ostréicole dans les prochaine années. Si cela arrive, les méthodes développées dans cette thèse ouvrent la voie à une caractérisation rapide de nouveaux génotypes, ainsi que l'étude et la surveillance de ceux déjà existants. Par conséquent, ces travaux offrent de nouvelles méthodes de prévention et de contrôle des POMS affectant les huîtres et pourraient être adoptées pour soutenir l'industrie ostréicole.

Référence

Aaskov, J., Buzacott, K., Thu, H.M., Lowry, K., Holmes, E.C., 2006. Long-Term Transmission of Defective RNA Viruses in Humans and Aedes Mosquitoes. Science 311, 236–238. https://doi.org/10.1126/science.1115030

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). J. Gen. Virol. 99, 693–703. https://doi.org/10.1099/jgv.0.001042

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Adema, C.M., Hertel, L.A., Miller, R.D., Loker, E.S., 1997. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. Proc. Natl. Acad. Sci. 94, 8691–8696. https://doi.org/10.1073/pnas.94.16.8691

Agnew, M.V., Friedman, C.S., Langdon, C., Divilov, K., Schoolfield, B., Morga, B., Degremont, L., Dhar, A.K., Kirkland, P., Dumbauld, B., Burge, C.A., 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. https://doi.org/10.3390/pathogens9121057

Akhtar, L.N., Bowen, C.D., Renner, D.W., Pandey, U., Della Fera, A.N., Kimberlin, D.W., Prichard, M.N., Whitley, R.J., Weitzman, M.D., Szpara, M.L., 2019. Genotypic and Phenotypic Diversity of Herpes Simplex Virus 2 within the Infected Neonatal Population. mSphere 4, e00590-18, /msphere/4/1/mSphere590-18.atom. https://doi.org/10.1128/mSphere.00590-18

Alfaro, A.C., Nguyen, T.V., Merien, F., 2018. The complex interactions of Ostreid herpesvirus 1 , Vibrio bacteria, environment and host factors in mass mortality outbreaks of Crassostrea gigas. Rev. Aquac. 11, 1148–1168. https://doi.org/10.1111/raq.12284

Alizon, S., Méthot, P.-O., 2018. Reconciling Pasteur and Darwin to control infectious diseases. PLOS Biol. 16, e2003815. https://doi.org/10.1371/journal.pbio.2003815

Allicock, O.M., Lemey, P., Tatem, A.J., Pybus, O.G., Bennett, S.N., Mueller, B.A., Suchard, M.A., Foster, J.E., Rambaut, A., Carrington, C.V.F., 2012. Phylogeography and Population Dynamics of Dengue Viruses in the Americas. Mol. Biol. Evol. 29, 1533–1543. https://doi.org/10.1093/molbev/msr320

Andras, J.P., Fields, P.D., Du Pasquier, L., Fredericksen, M., Ebert, D., 2020. Genome-wide association analysis identifies a genetic basis of infectivity in a model bacterial pathogen. Mol. Biol. Evol. https://doi.org/10.1093/molbev/msaa173

Arias, A., Lázaro, E., Escarmís, C., Domingo, E., 2001. Molecular intermediates of fitness gain of an RNA virus: characterization of a mutant spectrum by biological and molecular cloning. J. Gen. Virol. 82, 1049–1060. https://doi.org/10.1099/0022-1317-82-5-1049

Arzul, I., Lipart, C., Davison, A.J., Renault, T., 2001a. Evidence for interspecies transmission of oyster herpesvirus in marine bivalves. J. Gen. Virol. 82, 865–870. https://doi.org/10.1099/0022-1317-82-4-865

Arzul, I., Nicolas, J.-L., Davison, A.J., Renault, T., 2001b. French Scallops: A New Host for Ostreid Herpesvirus-1. Virology 290, 342–349. https://doi.org/10.1006/viro.2001.1186

Arzul, I., Renault, T., Thébault, A., Gérard, A., 2002. Detection of oyster herpesvirus DNA and proteins in asymptomatic Crassostrea gigas adults. Virus Res. 84, 151–160. https://doi.org/10.1016/S0168-1702(02)00007-2

Avolio, M.L., Beaulieu, J.M., Lo, E.Y.Y., Smith, M.D., 2012. Measuring genetic diversity in ecological studies. Plant Ecol. 213, 1105–1115. https://doi.org/10.1007/s11258-012-0069-6

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. Genet. Sel. Evol. 49, 23. https://doi.org/10.1186/s12711-017-0297-2

Azéma, P., Travers, M.-A., De Lorgeril, J., Tourbiez, D., Dégremont, L., 2015. Can selection for resistance to OsHV-1 infection modify susceptibility to Vibrio aestuarianus infection in Crassostrea gigas? First insights from experimental challenges using primary and successive exposures. Vet. Res. 46, 139. https://doi.org/10.1186/s13567-015-0282-0

Bagusche, F., 2013. Environmental effects on the physiology of calcification in the Pacific oyster Crassostrea gigas Thunberg, 1793. University of Southampton.

Bai, C., Wang, C., Xia, J., Sun, H., Zhang, S., Huang, J., 2014. Emerging and endemic types of Ostreid herpesvirus 1 were detected in bivalves in China. J. Invertebr. Pathol. 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., 2018. 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

Baranowski, E., Ruiz-Jarabo, C.M., Domingo, E., 2001. Evolution of Cell Recognition by Viruses. Science 292, 1102–1105. https://doi.org/10.1126/science.1058613

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. J. Invertebr. Pathol. 166, 107222. https://doi.org/10.1016/j.jip.2019.107222

Barbosa Solomieu, V., Renault, T., Travers, M.-A., 2015. Mass mortality in bivalves and the intricate case of the Pacific oyster, Crassostrea gigas. J. Invertebr. Pathol. 131, 2–10. https://doi.org/10.1016/j.jip.2015.07.011

Barbosa-Solomieu, V., Dégremont, L., Vázquez-Juárez, R., Ascencio-Valle, F., Boudry, P., Renault, T., 2005. Ostreid Herpesvirus 1 (OsHV-1) detection among three successive generations of Pacific oysters (Crassostrea gigas). Virus Res. 107, 47–56. https://doi.org/10.1016/j.virusres.2004.06.012

Bartha, I., Carlson, J.M., Brumme, C.J., McLaren, P.J., Brumme, Z.L., John, M., Haas, D.W., Martinez-Picado, J., Dalmau, J., López-Galíndez, C., Casado, C., Rauch, A., Günthard, H.F., Bernasconi, E., Vernazza, P., Klimkait, T., Yerly, S., O'Brien, S.J., Listgarten, J., Pfeifer, N., Lippert, C., Fusi, N., Kutalik, Z., Allen, T.M., Müller, V., Harrigan, P.R., Heckerman, D., Telenti, A., Fellay, J., for the HIV Genome-to-Genome Study and the Swiss HIV Cohort Study, 2013. A genome-to-genome analysis of associations between human genetic variation, HIV-1 sequence diversity, and viral control. eLife 2, e01123. https://doi.org/10.7554/eLife.01123

Basch, P.F., 1976. Intermediate host specificity in Schistosoma mansoni. Exp. Parasitol. 39, 150–169. https://doi.org/10.1016/0014-4894(76)90022-9

Basch, P.F., 1975. An interpretation of snail-trematode infection rates: Specificity based on concordance of compatible phenotypes. Int. J. Parasitol. 5, 449–452. https://doi.org/10.1016/0020-7519(75)90012-0

Batista, F.M., López-Sanmartín, M., Grade, A., Morgado, I., Valente, M., Navas, J.I., Power, D.M., Ruano, F., 2014. 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

Batschelet, E., Domingo, E., Weissmann, C., 1976. The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. Gene 1, 27–32. https://doi.org/10.1016/0378-1119(76)90004-4

Bayne, B.L., Ahrens, M., Allen, S.K., D'auriac, M.A., Backeljau, T., Beninger, P., Bohn, R., Boudry, P., Davis, J., Green, T., Guo, X., Hedgecock, D., Ibarra, A., Kingsley-Smith, P., Krause, M., Langdon, C., Lapègue, S., Li, C., Manahan, D., Mann, R., Perez-Paralle, L., Powell, E.N., Rawson, P.D., Speiser, D., Sanchez, J.-L., Shumway, S., Wang, H., 2017. The Proposed Dropping of the Genus Crassostrea for All Pacific Cupped Oysters and Its Replacement by a New Genus Magallana: A Dissenting View. J. Shellfish Res. 36, 545–547. https://doi.org/10.2983/035.036.0301

Bergelson, J., Kreitman, M., Stahl, E.A., Tian, D., 2001. Evolutionary Dynamics of Plant R-Genes. Science 292, 2281–2285. https://doi.org/10.1126/science.1061337

Bernard, I., Lapègue, S., 2012. Etude de faisabilité d'opérations de 'repeuplement orienté' dans deux sites français de captage de l'huître creuse (Crassostrea gigas).

Bernard, O., Hozumi, N., Tonegawa, S., 1978. Sequences of mouse immunoglobulin light chain genes before and after somatic changes. Cell 15, 1133–1144. https://doi.org/10.1016/0092-8674(78)90041-7

Bordier, A., 1888. Sixième conférence transformiste: Les microbes et le transformisme. Bull. Société Anthropol. Paris 11, 743–779. https://doi.org/10.3406/bmsap.1888.5406

Botta, R., Asche, F., Borsum, J.S., Camp, E.V., 2020. A review of global oyster aquaculture production and consumption. Mar. Policy 117, 103952. https://doi.org/10.1016/j.marpol.2020.103952

Bovers, M., Hagen, F., Kuramae, E.E., Diaz, M.R., Spanjaard, L., Dromer, F., Hoogveld, H.L., Boekhout, T., 2006. Unique hybrids between the fungal pathogens Cryptococcus neoformans and Cryptococcus gattii. FEMS Yeast Res. 6, 599–607. https://doi.org/10.1111/j.1567-1364.2006.00082.x

Bower, S.M., McGladdery, S.E., Price, I.M., 1994. Synopsis of infectious diseases and parasites of commercially exploited shellfish. Annu. Rev. Fish Dis. 4, 1–199. https://doi.org/10.1016/0959-8030(94)90028-0

Boyd, M.T., Simpson, G.R., Cann, A.J., Johnson, M.A., Weiss, R.A., 1993. A single amino acid substitution in the V1 loop of human immunodeficiency virus type 1 gp120 alters cellular tropism. J. Virol. 67, 3649–3652.

Brackney, D.E., Beane, J.E., Ebel, G.D., 2009. RNAi Targeting of West Nile Virus in Mosquito Midguts Promotes Virus Diversification. PLOS Pathog. 5, e1000502. https://doi.org/10.1371/journal.ppat.1000502

Bradley, A.J., Lurain, N.S., Ghazal, P., Trivedi, U., Cunningham, C., Baluchova, K., Gatherer, D., Wilkinson, G.W.G., Dargan, D.J., Davison, A.J., 2009. High-throughput sequence analysis of variants of human cytomegalovirus strains Towne and AD169. J. Gen. Virol. 90, 2375–2380. https://doi.org/10.1099/vir.0.013250-0

Brasier, C.M., 2001. Rapid Evolution of Introduced Plant Pathogens via Interspecific Hybridization: Hybridization is leading to rapid evolution of Dutch elm disease and other fungal plant pathogens. BioScience 51, 123–133. https://doi.org/10.1641/0006-3568(2001)051[0123:REOIPP]2.0.CO;2

Briones, C., Domingo, E., Molina-París, C., 2003. Memory in Retroviral Quasispecies: Experimental Evidence and Theoretical Model for Human Immunodeficiency Virus. J. Mol. Biol. 331, 213–229. https://doi.org/10.1016/S0022-2836(03)00661-2

Brisson, D., 2018. Negative Frequency-Dependent Selection Is Frequently Confounding. Front. Ecol. Evol. 6. https://doi.org/10.3389/fevo.2018.00010

Bubb, K.L., Bovee, D., Buckley, D., Haugen, E., Kibukawa, M., Paddock, M., Palmieri, A., Subramanian, S., Zhou, Y., Kaul, R., Green, P., Olson, M.V., 2006. Scan of Human Genome Reveals No New Loci Under Ancient Balancing Selection. Genetics 173, 2165–2177. https://doi.org/10.1534/genetics.106.055715

Buestel, D., Ropert, M., Prou, J., Goulletquer, P., 2009. History, Status, and Future of Oyster Culture in France. J. Shellfish Res. 28, 813–820. https://doi.org/10.2983/035.028.0410

Burge, C., Griffin, F., Friedman, C., 2006. Mortality and herpesvirus infections of the Pacific oyster Crassostrea gigas in Tomales Bay, California, USA. Dis. Aquat. Organ. 72, 31–43. https://doi.org/10.3354/dao072031

Burge, C.A., Friedman, C.S., 2011. Quantifying Ostreid Herpesvirus (OsHV-1) Genome Copies and Expression during Transmission. Microb. Ecol. 63, 596–604. https://doi.org/10.1007/s00248-011-9937-1

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., Prearo, M., Riina, M.V., Bona, M.C., Fioravanti, M.L., Arcangeli, G., Houssin, M., 2016. Ostreid herpesvirus type 1 genomic diversity in wild populations of Pacific oyster Crassostrea gigas from Italian coasts. J. Invertebr. Pathol. 137, 71–83. https://doi.org/10.1016/j.jip.2016.05.004

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. J. Fish Dis. 41, 1759–1769. https://doi.org/10.1111/jfd.12883

Buroker., 1982. Sexuality with respect to shell length and group size in the Japanese oyster Crassostrea gigas, in: Malacologia. Ann Arbor,Institute of Malacology, pp. 271–279.

Carrasco, P., Iglesia, F. de la, Elena, S.F., 2007. Distribution of Fitness and Virulence Effects Caused by Single-Nucleotide Substitutions in Tobacco Etch Virus. J. Virol. 81, 12979–12984. https://doi.org/10.1128/JVI.00524-07

Chongming, W., Xiuhua, W., Xiaoling, S., Jie, H., Weibo, S., 2002. Purification and ultrastructure of a spherical virus in cultured scallop Chlamys farreri. Shuichan Xuebao 26, 180—183.

Clerissi, C., de Lorgeril, J., Petton, B., Lucasson, A., Escoubas, J.-M., Gueguen, Y., Dégremont, L., Mitta, G., Toulza, E., 2020. Microbiota Composition and Evenness Predict Survival Rate of Oysters Confronted to Pacific Oyster Mortality Syndrome. Front. Microbiol. 11, 311. https://doi.org/10.3389/fmicb.2020.00311

Coffey, L.L., Beeharry, Y., Bordería, A.V., Blanc, H., Vignuzzi, M., 2011. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. Proc. Natl. Acad. Sci. U. S. A. 108, 16038–16043. https://doi.org/10.1073/pnas.1111650108

Corbeil, S., Faury, N., Segarra, A., Renault, T., 2015. Development of an in situ hybridization assay for the detection of ostreid herpesvirus type 1 mRNAs in the Pacific oyster, Crassostrea gigas. J. Virol. Methods 211, 43–50. https://doi.org/10.1016/j.jviromet.2014.10.007

Corey, S., Stephanie, G., Rebekah M, B., Nicola L, B., Sophie A, V., Kelly-Anne, T., Jacqueline M, B., John J, M., Daniel, C., Scott, B., Scott, C., Katherine, K., Scott R, B., Jamie, R., Rajiv, K., 2014. Molecular imprint of exposure to naturally occurring genetic variants of human cytomegalovirus on the T cell repertoire. Sci. Rep. 4, 3993–3993. https://doi.org/10.1038/srep03993

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. Organ. 78, 181–188. https://doi.org/10.3354/dao01874

Dainat, J., 2020. NBISweden/AGAT: AGAT-v0.2.3. https://doi.org/10.5281/zenodo.3714855

Darwin, C., 1859. On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life. London :John Murray, Albemarle Street,.

Davison, A.J., 2005. A novel class of herpesvirus with bivalve hosts. J. Gen. Virol. 86, 41–53. https://doi.org/10.1099/vir.0.80382-0

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. Arch. Virol. 154, 171–177. https://doi.org/10.1007/s00705-008-0278-4

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., Leroy, M.A., Pérignon, A., Régler, D., Morga, B., Alunno-Bruscia, M., Boudry, P., Le Roux, F., Destoumieux-Garzón, D., Gueguen, Y., Mitta, G., 2018. Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. Nat. Commun. 9, 4215. https://doi.org/10.1038/s41467-018-06659-3

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, 63. https://doi.org/10.1186/s12864-020-6471-x

Decaestecker, E., Gaba, S., Raeymaekers, J.A.M., Stoks, R., Van Kerckhoven, L., Ebert, D., De Meester, L., 2007. Host-parasite "Red Queen" dynamics archived in pond sediment. Nature 450, 870–873. https://doi.org/10.1038/nature06291

Déchiffrage des maladies multifactorielles: cas des mortalités de l'huître [WWW Document], 2015. . Agence Natl. Rech. URL https://anr.fr/Projet-ANR-14-CE19-0023 (accessed 8.28.20).

Degremont, L., 2011. Evidence of herpesvirus (OsHV-1) resistance in juvenile Crassostrea gigas selected for high resistance to the summer mortality phenomenon. Aquaculture 317, 94–98. https://doi.org/10.1016/j.aquaculture.2011.04.029

Degremont, 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, 213–229. https://doi.org/10.1016/j.aquaculture.2005.03.046

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. Aquac. Int. 22, 1767–1781. https://doi.org/10.1007/s10499-014-9781-7

Dégremont, L., Garcia, C., Allen, S.K., 2015. Genetic improvement for disease resistance in oysters: A review. J. Invertebr. Pathol., Pathogens and Disease Processes in Marine Molluscs 131, 226–241. https://doi.org/10.1016/j.jip.2015.05.010

Degremont, L., Nourry, M., 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, 111–121. https://doi.org/10.1016/j.aquaculture.2015.04.029

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. Front. Microbiol. 11, 1579. https://doi.org/10.3389/fmicb.2020.01579

Depotter, J.R., Seidl, M.F., Wood, T.A., Thomma, B.P., 2016. Interspecific hybridization impacts host range and pathogenicity of filamentous microbesa. Curr. Opin. Microbiol., Host-microbe interactions: parasites/fungi/viruses 32, 7–13. https://doi.org/10.1016/j.mib.2016.04.005

Desiderio, S.V., Yancopoulos, G.D., Paskind, M., Thomas, E., Boss, M.A., Landau, N., Alt, F.W., Baltimore, D., 1984. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. Nature 311, 752–755. https://doi.org/10.1038/311752a0

Dheilly, N.M., Duval, D., Mouahid, G., Emans, R., Allienne, J.-F., Galinier, R., Genthon, C., Dubois, E., Du Pasquier, L., Adema, C.M., Grunau, C., Mitta, G., Gourbal, B., 2015. A family of variable immunoglobulin and lectin domain containing molecules in the snail Biomphalaria glabrata. Dev. Comp. Immunol. 48, 234–243. https://doi.org/10.1016/j.dci.2014.10.009

Domingo, E., Sheldon, J., Perales, C., 2012. Viral Quasispecies Evolution. Microbiol. Mol. Biol. Rev. 76, 159–216. https://doi.org/10.1128/MMBR.05023-11

Drake, J.W., 1969. Spontaneous Mutation: Comparative Rates of Spontaneous Mutation. Nature 221, 1132–1132. https://doi.org/10.1038/2211132a0

Drake, J.W., Holland, J.J., 1999. Mutation rates among RNA viruses. Proc. Natl. Acad. Sci. 96, 13910–13913. https://doi.org/10.1073/pnas.96.24.13910

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, 49–52. https://doi.org/10.1016/j.aquaculture.2011.02.005

Ebert, D., 2008. Host–parasite coevolution: Insights from the Daphnia–parasite model system. Curr. Opin. Microbiol., Ecology and Industrial Microbiology Techniques 11, 290–301. https://doi.org/10.1016/j.mib.2008.05.012

Ebert, D., Fields, P.D., 2020. Host–parasite co-evolution and its genomic signature. Nat. Rev. Genet. https://doi.org/10.1038/s41576-020-0269-1

Eigen, M., 1971. Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58, 465–523. https://doi.org/10.1007/BF00623322

Elandaloussi, L., Carrasco, N., Andree, K., Furones, D., Roque, A., 2009. Esdeveniments de mortalitat de lostró del Pacific (Crassostrea gigas) en el delta del Ebre. Estudi de cas. Proceedings of the II Simposi d'aqüicultura de Catalunya.

Ellegren, H., Galtier, N., 2016. Determinants of genetic diversity. Nat. Rev. Genet. 17, 422–433. https://doi.org/10.1038/nrg.2016.58

Emu, B., Fessel, J., Schrader, S., Kumar, P., Richmond, G., Win, S., Weinheimer, S., Marsolais, C., Lewis, S., 2018. Phase 3 Study of Ibalizumab for Multidrug-Resistant HIV-1. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa1711460

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T.O., 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3, e1319. https://doi.org/10.7717/peerj.1319

FAO, FAO Online Query Panels, 2018. URL http://www.fao.org/fishery/topic/16140/en (accessed 9.22.20).

FAO, WHO, 2018. Technical Guidance for the Development of the Growing Area Aspects of Bivalve Mollusc Sanitation Programmes.

Farci, P., 2011. New insights into the HCV quasispecies and compartmentalization. Semin. Liver Dis. 31, 356–374. https://doi.org/10.1055/s-0031-1297925

Faria, N.R., Rambaut, A., Suchard, M.A., Baele, G., Bedford, T., Ward, M.J., Tatem, A.J., Sousa, J.D., Arinaminpathy, N., Pépin, J., Posada, D., Peeters, M., Pybus, O.G., Lemey, P., 2014. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. Science 346, 56–61. https://doi.org/10.1126/science.1256739

Farley, C.A., Banfield, W.G., Kasnic, G., Foster, W.S., 1972. Oyster Herpes-Type Virus. Science 178, 759. https://doi.org/10.1126/science.178.4062.759

Fausther-Bovendo, H., Sol-Foulon, N., Candotti, D., Agut, H., Schwartz, O., Debré, P., Vieillard, V., 2009. HIV escape from natural killer cytotoxicity: nef inhibits NKp44L expression on CD4+ T cells. AIDS Lond. Engl. 23, 1077–1087. https://doi.org/10.1097/QAD.0b013e32832cb26b

Feigelstock, D.A., Mihalik, K.B., Feinstone, S.M., 2011. Selection of hepatitis C virus resistant to ribavirin. Virol. J. 8, 402. https://doi.org/10.1186/1743-422X-8-402

Firth, C., Kitchen, A., Shapiro, B., Suchard, M.A., Holmes, E.C., Rambaut, A., 2010. Using Time-Structured Data to Estimate Evolutionary Rates of Double-Stranded DNA Viruses. Mol. Biol. Evol. 27, 2038–2051. https://doi.org/10.1093/molbev/msq088

Fitzsimmons, W.J., Woods, R.J., McCrone, J.T., Woodman, A., Arnold, J.J., Yennawar, M., Evans, R., Cameron, C.E., Lauring, A.S., 2018. A speed–fidelity trade-off determines the mutation rate and virulence of an RNA virus. PLOS Biol. 16, e2006459. https://doi.org/10.1371/journal.pbio.2006459

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. Dis. Aquat. Organ. 63, 33–41. https://doi.org/10.3354/dao063033

Gallagher, M., 2020. New Approaches to Characterise Viral Pathogens in Aquaculture.

Gandon, S., Buckling, A., Decaestecker, E., Day, T., 2008. Host–parasite coevolution and patterns of adaptation across time and space. J. Evol. Biol. 21, 1861–1866. https://doi.org/10.1111/j.1420-9101.2008.01598.x

Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. ArXiv12073907 Q-Bio.

Gauthier, J., Drezen, J.-M., Herniou, E.A., 2018. The recurrent domestication of viruses: major evolutionary transitions in parasitic wasps. Parasitology 145, 713–723. https://doi.org/10.1017/S0031182017000725

Gelderblom, H.C., Vatakis, D.N., Burke, S.A., Lawrie, S.D., Bristol, G.C., Levy, D.N., 2008. Viral complementation allows HIV-1 replication without integration. Retrovirology 5, 60. https://doi.org/10.1186/1742-4690-5-60

Gittenberger, A., Voorbergen-Laarman, M.A., Engelsma, M.Y., 2014. Ostreid herpesvirus OsHV-1 µVar in Pacific oysters Crassostrea gigas (Thunberg 1793) of the Wadden Sea, a UNESCO world heritage site. J. Fish Dis. 39, 105–109. https://doi.org/10.1111/jfd.12332

Goulletquer, P., Le Moine, O., 2002. Shellfish farming and Coastal Zone Management (CZM) development in the Marennes-Oléron Bay and Charentais Sounds (Charente Maritime, France): A review of recent developments. Aquac. Int. 10, 507–525. https://doi.org/10.1023/A:1023975418669

Green, T.J., Rolland, J.-L., Vergnes, A., Raftos, D., Montagnani, C., 2015. OsHV-1 countermeasures to the Pacific oyster's anti-viral response. Fish Shellfish Immunol. 47, 435–443. https://doi.org/10.1016/j.fsi.2015.09.025

Grenfell, B.T., 2004. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. Science 303, 327–332. https://doi.org/10.1126/science.1090727

Griffiths, G.M., Berek, C., Kaartinen, M., Milstein, C., 1984. Somatic mutation and the maturation of immune response to 2-phenyl oxazolone. Nature 312, 271–275. https://doi.org/10.1038/312271a0

Gutierrez, A.P., Bean, T.P., Hooper, C., Stenton, C.A., Sanders, M.B., Paley, R.K., Rastas, P., Bryrom, M., Matika, O., Houston, R.D., 2018. A Genome-Wide Association Study for Host Resistance to Ostreid Herpesvirus in Pacific Oysters (Crassostrea gigas). G3 GenesGenomesGenetics 8, 1273. https://doi.org/10.1534/g3.118.200113

Guyomar, C., Delage, W., Legeai, F., Mougel, C., Simon, J.-C., Lemaitre, C., 2019. Reference-guided genome assembly in metagenomic samples 9.

Haeryfar, S.M.M., Hickman, H.D., Irvine, K.R., Tscharke, D.C., Bennink, J.R., Yewdell, J.W., 2008. Terminal Deoxynucleotidyl Transferase Establishes and Broadens Anti-Viral CD8+ T Cell Immunodominance Hierarchies. J. Immunol. Baltim. Md 1950 181, 649–659.

Herr, A.J., Williams, L.N., Preston, B.D., 2011. Antimutator variants of DNA polymerases. Crit. Rev. Biochem. Mol. Biol. 46, 548–570. https://doi.org/10.3109/10409238.2011.620941

Hick, P., English, C., Evans, O., Whittington, R.J., Looi, R., 2016. Stability of Ostreid herpesvirus-1 (OsHV-1) and assessment of disinfection of seawater and oyster tissues using a bioassay. Aquac. Amst. Neth. 450, 412–421. https://doi.org/10.1016/j.aquaculture.2015.08.025

Hick, P.M., Evans, O., Rubio, A., Dhand, N.K., Whittington, R.J., 2018. Both age and size influence susceptibility of Pacific oysters (Crassostrea gigas) to disease caused by Ostreid herpesvirus -1 (OsHV-1) in replicated field and laboratory experiments. Aquaculture 489, 110–120. https://doi.org/10.1016/j.aquaculture.2018.02.013

Hine, P., Thorne, T., 1997. Replication of herpes-like viruses in haemocytes of adult flat oysters Ostrea angasi:an ultrastructural study. Dis. Aquat. Organ. 29, 189–196. https://doi.org/10.3354/dao029189

Hine, P., Wesney, B., Hay, B., 1992. Herpesviruses associated with mortalities among hatchery-reared larval Pacific oysters Crassostrea gigas. Dis. Aquat. Organ. 12, 135–142. https://doi.org/10.3354/dao012135

Ho, S.S., Urban, A.E., Mills, R.E., 2019. Structural variation in the sequencing era. Nat. Rev. Genet. https://doi.org/10.1038/s41576-019-0180-9

Holmes, E.C., 2010. The RNA Virus Quasispecies: Fact or Fiction? J. Mol. Biol. 400, 271–273. https://doi.org/10.1016/j.jmb.2010.05.032

Hozumi, N., Tonegawa, S., 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. Proc. Natl. Acad. Sci. U. S. A. 73, 3628–3632.

Hsu, E., Du Pasquier, L. (Eds.), 2015. Pathogen-Host Interactions: Antigenic Variation v. Somatic Adaptations, Results and Problems in Cell Differentiation. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-20819-0

Huang, B., Zhang, L., Li, L., Tang, X., Zhang, G., 2015. Highly diverse fibrinogen-related proteins in the Pacific oyster Crassostrea gigas. Fish Shellfish Immunol. 43, 485–490. https://doi.org/10.1016/j.fsi.2015.01.021

Huxley, J., 1942. Evolution, the modern synthesis, Science editions. Harper & brothers.

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. J. Fish Dis. 36, 969–972. https://doi.org/10.1111/jfd.12093

Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11, 119. https://doi.org/10.1186/1471-2105-11-119

Janzen, D.H., 1980. When Is It Coevolution? Evolution 34, 611–612. https://doi.org/10.1111/j.1558-5646.1980.tb04849.x

Jaramillo, N., Domingo, E., Munoz-Egea, M.C., Tabares, E., Gadea, I., 2013. Evidence of Muller's ratchet in herpes simplex virus type 1. J. Gen. Virol. 94, 366–375. https://doi.org/10.1099/vir.0.044685-0

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. Dis. Aquat. Organ. 105, 109–126. https://doi.org/10.3354/dao02623

Jin, Z., Deval, J., Johnson, K.A., Swinney, D.C., 2011. Characterization of the Elongation Complex of Dengue Virus RNA Polymerase. J. Biol. Chem. 286, 2067–2077. https://doi.org/10.1074/jbc.M110.162685

Kaján, G.L., Doszpoly, A., Tarján, Z.L., Vidovszky, M.Z., Papp, T., 2019. Virus–Host Coevolution with a Focus on Animal and Human DNA Viruses. J. Mol. Evol. https://doi.org/10.1007/s00239-019-09913-4

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. Dis. Aquat. Organ. 109, 231–239. https://doi.org/10.3354/dao02735

Kimura, M. (Ed.), 1983. Introduction, in: The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, pp. xi–xvi. https://doi.org/10.1017/CBO9780511623486.002

King, J.L., Jukes, T.H., 1969. Non-Darwinian Evolution. Science 164, 788–798. https://doi.org/10.1126/science.164.3881.788

Klafack, S., Fiston-Lavier, A.-S., Bergmann, S., Hammoumi, S., Schröder, L., Fuchs, W., Lusiastuti, A., Lee, P.-Y., Heredia, S., Master student consortium, Gosselin-Grenet, A.-S., Avarre, J.-C., 2019. Cyprinid herpesvirus 3 Evolves In Vitro through an Assemblage of Haplotypes that Alternatively Become Dominant or Under-Represented. Viruses 11, 754. https://doi.org/10.3390/v11080754

Klarenbeek, P.L., Remmerswaal, E.B.M., Berge, I.J.M. ten, Doorenspleet, M.E., Schaik, B.D.C. van, Esveldt, R.E.E., Koch, S.D., Brinke, A. ten, Kampen, A.H.C. van, Bemelman, F.J., Tak, P.P., Baas, F., Vries, N. de, Lier, R.A.W. van, 2012. Deep Sequencing of Antiviral T-Cell Responses to HCMV and EBV in Humans Reveals a Stable Repertoire That Is Maintained for Many Years. PLOS Pathog. 8, e1002889. https://doi.org/10.1371/journal.ppat.1002889

Klein, J., 1997. Homology between immune responses in vertebrates and invertebrates: does it exist? Scand. J. Immunol. 46, 558–564. https://doi.org/10.1046/j.1365-3083.1997.d01-164.x

Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., Ding, L., 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25, 2283–2285. https://doi.org/10.1093/bioinformatics/btp373

Laanto, E., Hoikkala, V., Ravantti, J., Sundberg, L.-R., 2017. Long-term genomic coevolution of host-parasite interaction in the natural environment. Nat. Commun. 8, 111. https://doi.org/10.1038/s41467-017-00158-7

Lafont, M., 2017. Mechanisms and specificity of antiviral immune priming in a Lophtrochozoan, the Pacific oyster Crassostrea gigas (Theses No. 2017PERP0041). Université de Perpignan.

Lafont, M., Vergnes, A., Vidal-Dupiol, J., de Lorgeril, J., Gueguen, Y., Haffner, P., Petton, B., Chaparro, C., Barrachina, C., Destoumieux-Garzon, D., Mitta, G., Gourbal, B., Montagnani, C., 2020. A Sustained Immune Response Supports Long-Term Antiviral Immune Priming in the Pacific Oyster, Crassostrea gigas. mBio 11, e02777-19, /mbio/11/2/mBio.02777-19.atom. https://doi.org/10.1128/mBio.02777-19

Laland, K.N., Uller, T., Feldman, M.W., Sterelny, K., Müller, G.B., Moczek, A., Jablonka, E., Odling-Smee, J., 2015. The extended evolutionary synthesis: its structure, assumptions and predictions. Proc. R. Soc. B Biol. Sci. 282, 20151019. https://doi.org/10.1098/rspb.2015.1019

Lapègue, S., Heurtebise, S., Cornette, F., Guichoux, E., Gagnaire, P.-A., 2020. Genetic Characterization of Cupped Oyster Resources in Europe Using Informative Single Nucleotide Polymorphism (SNP) Panels. Genes 11, 451. https://doi.org/10.3390/genes11040451

Lauring, A.S., Andino, R., 2010. Quasispecies Theory and the Behavior of RNA Viruses. PLoS Pathog. 6, e1001005. https://doi.org/10.1371/journal.ppat.1001005

Lefevre, T., Raymond, M., Thomas, F., 2016. Biologie évolutive. De Boeck Superieur.

Leffler, E.M., Gao, Z., Pfeifer, S., Ségurel, L., Auton, A., Venn, O., Bowden, R., Bontrop, R., Wall, J.D., Sella, G., Donnelly, P., McVean, G., Przeworski, M., 2013. Multiple Instances of Ancient Balancing Selection Shared Between Humans and Chimpanzees. Science 339, 1578–1582. https://doi.org/10.1126/science.1234070

Leitão, A., Chaves, R., Santos, S., Guedes-Pinto, H., Boudry, P., 2007. Interspecific hybridization in oysters: Restriction Enzyme Digestion Chromosome Banding confirms Crassostrea angulata×Crassostrea gigas F1 hybrids. J. Exp. Mar. Biol. Ecol. 343, 253–260. https://doi.org/10.1016/j.jembe.2006.12.007

Léonard, P.M., Adema, C.M., Zhang, S.M., Loker, E.S., 2001. Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail Biomphalaria glabrata. Gene 269, 155–165. https://doi.org/10.1016/s0378-1119(01)00444-9

Levi, L.I., Gnädig, N.F., Beaucourt, S., McPherson, M.J., Baron, B., Arnold, J.J., Vignuzzi, M., 2010. Fidelity Variants of RNA Dependent RNA Polymerases Uncover an Indirect, Mutagenic Activity of Amiloride Compounds. PLOS Pathog. 6, e1001163. https://doi.org/10.1371/journal.ppat.1001163

Li, G., Piampongsant, S., Faria, N.R., Voet, A., Pineda-Peña, A.-C., Khouri, R., Lemey, P., Vandamme, A.-M., Theys, K., 2015. An integrated map of HIV genome-wide variation from a population perspective. Retrovirology 12, 18. https://doi.org/10.1186/s12977-015-0148-6

Lokmer, A., Mathias Wegner, K., 2015. Hemolymph microbiome of Pacific oysters in response to temperature, temperature stress and infection. ISME J. 9, 670–682. https://doi.org/10.1038/ismej.2014.160

López-Sanmartín, M., Power, D.M., de la Herrán, R., Navas, J.I., Batista, F.M., 2016. Evidence of vertical transmission of ostreid herpesvirus 1 in the Portuguese oyster Crassostrea angulata. J. Invertebr. Pathol. 140, 39–41. https://doi.org/10.1016/j.jip.2016.08.012

Lu, Q., Hwang, Y.T., Hwang, C.B.C., 2002. Mutation Spectra of Herpes Simplex Virus Type 1 Thymidine Kinase Mutants. J. Virol. 76, 5822–5828. https://doi.org/10.1128/JVI.76.11.5822-5828.2002

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, 1526–1532. https://doi.org/10.1017/S0031182012000881

Majerus, M., Amos, W., Hurst, G., 1996. Evolution: the four billion year war. Evol. Four Billion Year War.

Mann, R., 1979. Some biochemical and physiological aspects of growth and gametogenesis in Crassostrea gigas and Ostrea edulis grown at sustained elevated temperatures. J. Mar. Biol. Assoc. U. K. 59, 95–110. https://doi.org/10.1017/S0025315400046208*

Martenot, C., Gervais, O., Chollet, B., Houssin, M., Renault, T., 2017. Haemocytes collected from experimentally infected Pacific oysters, Crassostrea gigas: Detection of ostreid herpesvirus 1 DNA, RNA, and proteins in relation with inhibition of apoptosis. PLOS ONE 12, e0177448. https://doi.org/10.1371/journal.pone.0177448

Martenot, C., Lethuillier, O., Fourour, S., Oden, E., Trancart, S., Travaillé, E., Houssin, M., 2015. Detection of undescribed ostreid herpesvirus 1 (OsHV-1) specimens from Pacific oyster, Crassostrea gigas. J. Invertebr. Pathol. 132, 182–189. https://doi.org/10.1016/j.jip.2015.10.005

Martenot, C., Travaillé, E., Lethuillier, O., Lelong, C., Houssin, M., 2013. Genome exploration of six variants of the Ostreid Herpesvirus 1 and characterization of large deletion in OsHV-1µVar specimens. Virus Res. 178, 462–470. https://doi.org/10.1016/j.virusres.2013.08.006

Mas, V., Nair, H., Campbell, H., Melero, J.A., Williams, T.C., 2018. Antigenic and sequence variability of the human respiratory syncytial virus F glycoprotein compared to related viruses in a comprehensive dataset. Vaccine 36, 6660–6673. https://doi.org/10.1016/j.vaccine.2018.09.056

Matloubian, M., Kolhekar, S.R., Somasundaram, T., Ahmed, R., 1993. Molecular determinants of macrophage tropism and viral persistence: importance of single amino acid changes in the polymerase and glycoprotein of lymphocytic choriomeningitis virus. J. Virol. 67, 7340–7349.

McGeoch, D.J., Rixon, F.J., Davison, A.J., 2006. Topics in herpesvirus genomics and evolution. Virus Res. 117, 90–104. https://doi.org/10.1016/j.virusres.2006.01.002

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110

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., 2017. First evidence of viral and bacterial oyster pathogens in the Brazilian coast. J. Fish Dis. 41, 559–563. https://doi.org/10.1111/jfd.12755

Mendel, G., 1865. Versuche Über Pflanzen-Hybriden. Proc. Nat. Hist. Soc. Brünn 3–47, 41.

Meyer-König, U., Vogelberg, C., Bongarts, A., Kampa, D., Delbrück, R., Wolff-Vorbeck, G., Kirste, G., Haberland, M., Hufert, F.T., Laer, D. von, 1998. Glycoprotein B genotype correlates with cell tropism in vivo of human cytomegalovirus infection. J. Med. Virol. 55, 75–81. https://doi.org/10.1002/(SICI)1096-9071(199805)55:1<75::AID-JMV12>3.0.CO;2-Z

Miles, A., 2019. Pysamstats.

Mineur, F., Provan, J., Arnott, G., 2014. Phylogeographical analyses of shellfish viruses: inferring a geographical origin for ostreid herpesviruses OsHV-1 (Malacoherpesviridae). Mar. Biol. 162, 181–192. https://doi.org/10.1007/s00227-014-2566-8

Miossec, L., Deuff, R.-M.L., Goulletquer, P., 2009. Alien species alert: Crassostrea gigas (Pacific oyster). ICES Coop. Res. Rep. 46.

Mitta, G., Adema, C.M., Gourbal, B., Loker, E.S., Theron, A., 2012. Compatibility polymorphism in snail/schistosome interactions: From field to theory to molecular mechanisms. Dev. Comp. Immunol. 37, 1–8. https://doi.org/10.1016/j.dci.2011.09.002

Mojica, K.D.A., Brussaard, C.P.D., 2014. Factors affecting virus dynamics and microbial host–virus interactions in marine environments. FEMS Microbiol. Ecol. 89, 495–515. https://doi.org/10.1111/1574-6941.12343

Moné, Y., Gourbal, B., Duval, D., Pasquier, L.D., Kieffer-Jaquinod, S., Mitta, G., 2010. A Large Repertoire of Parasite Epitopes Matched by a Large Repertoire of Host Immune Receptors in an Invertebrate Host/Parasite Model. PLoS Negl. Trop. Dis. 4, e813. https://doi.org/10.1371/journal.pntd.0000813

Morga, B., Faury, N., Guesdon, S., Chollet, B., Renault, T., 2017. Haemocytes from Crassostrea gigas and OsHV-1: A promising in vitro system to study host/virus interactions. J. Invertebr. Pathol. 150, 45–53. https://doi.org/10.1016/j.jip.2017.09.007

Mori, K., 1979. Effects of artificial eutrophication on the metabolism of the Japanese oyster Crassostrea gigas. Mar. Biol. 53, 361–369. https://doi.org/10.1007/BF00391619

Morse, A., Calabro, K., Fear, J., Bloom, D., McIntyre, L., 2017. Reliable Detection of Herpes Simplex Virus Sequence Variation by High-Throughput Resequencing. Viruses 9, 226. https://doi.org/10.3390/v9080226

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. Dis. Aquat. Organ. 117, 171–176. 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. Dis. Aquat. Organ. 77, 207–223. https://doi.org/10.3354/dao01829

Naimi, A., 2009. Le déterminisme du sexe de l'huître creuse Crassostrea gigas au cours de son développement et du cycle gamétogénétique adulte : recherche de déterminants moléculaires, Sex determinism in the pacific oyster Crassostrea gigas during its development and the adult gametogenetic cycle: search of molecular markers (Thesis). Université de Caen Basse-Normandie.

Nicolas, J., Comps, M., Cochennec, N., 1992. Herpes-like virus infecting Pacific oyster larvae, Crassostrea gigas. Bull. Eur. Assoc. Fish Pathol. EAFP 12, 11–13.

O' Reilly, A.J., Laide, C., Maloy, A., Hutton, S., Bookelaar, B., O' Sullivan, K., Lynch, S.A., Culloty, S.C., 2017. The role of the mussel Mytilus spp. in the transmission of ostreid herpesvirus-1 microVar. Parasitology 145, 1095–1104. https://doi.org/10.1017/S0031182017002244

Owen, J., Punt, J., Stranford, S., 2014. Immunologie - 7e édition: Le cours de Janis Kuby avec questions de révision. Dunod.

Papkou, A., Guzella, T., Yang, W., Koepper, S., Pees, B., Schalkowski, R., Barg, M.-C., Rosenstiel, P.C., Teotónio, H., Schulenburg, H., 2019. The genomic basis of Red Queen dynamics during rapid reciprocal host–pathogen coevolution. Proc. Natl. Acad. Sci. 116, 923–928. https://doi.org/10.1073/pnas.1810402116

Parsons, L.R., Tafuri, Y.R., Shreve, J.T., Bowen, C.D., Shipley, M.M., Enquist, L.W., Szpara, M.L., 2015. Rapid Genome Assembly and Comparison Decode Intrastrain Variation in Human Alphaherpesviruses. mBio 6, e02213-14. https://doi.org/10.1128/mBio.02213-14

Pathirana, E., Fuhrmann, M., Whittington, R., Hick, P., 2019. Influence of environment on the pathogenesis of Ostreid herpesvirus-1 (OsHV-1) infections in Pacific oysters (Crassostrea gigas) through differential microbiome responses. Heliyon 5, e02101. https://doi.org/10.1016/j.heliyon.2019.e02101

Paul-Pont, I., Dhand, N., Whittington, R., 2013a. Spatial distribution of mortality in Pacific oysters Crassostrea gigas: reflection on mechanisms of OsHV-1 transmission. Dis. Aquat. Organ. 105, 127–138. https://doi.org/10.3354/dao02615

Paul-Pont, I., Dhand, N.K., Whittington, R.J., 2013b. Influence of husbandry practices on OsHV-1 associated mortality of Pacific oysters Crassostrea gigas. Aquaculture 412–413, 202–214. https://doi.org/10.1016/j.aquaculture.2013.07.038

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. Prev. Vet. Med. 105, 136–143. https://doi.org/10.1016/j.prevetmed.2012.02.001

Pepin, J.F., Riou, A., Renault, T., 2008. Rapid and sensitive detection of ostreid herpesvirus 1 in oyster samples by real-time PCR. J. Virol. Methods 149, 269–276. https://doi.org/10.1016/j.jviromet.2008.01.022

Perales, C., Lorenzo-Redondo, R., López-Galíndez, C., Martínez, M.A., Domingo, E., 2010. Mutant spectra in virus behavior. Future Virol. 5, 679–698. https://doi.org/10.2217/fvl.10.61

Pernet, F., Barret, J., Le Gall, P., Corporeau, C., Dégremont, L., Lagarde, F., Pépin, J., Keck, N., 2012. Mass mortalities of Pacific oysters Crassostrea gigas reflect infectious diseases and vary with farming practices in the Mediterranean Thau lagoon, France. Aquac. Environ. Interact. 2, 215–237. https://doi.org/10.3354/aei00041

Petton, B., Boudry, P., Alunno-Bruscia, M., Pernet, F., 2015. Factors influencing disease-induced mortality of Pacific oysters Crassostrea gigas. Aquac. Environ. Interact. 6, 205–222. https://doi.org/10.3354/aei00125

Phillips, R.E., Rowland-Jones, S., Nixon, D.F., Gotch, F.M., Edwards, J.P., Ogunlesi, A.O., Elvin, J.G., Rothbard, J.A., Bangham, C.R., Rizza, C.R., 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature 354, 453–459. https://doi.org/10.1038/354453a0

Pokalyuk, C., Renzette, N., Irwin, K.K., Pfeifer, S.P., Gibson, L., Britt, W.J., Yamamoto, A.Y., Mussi-Pinhata, M.M., Kowalik, T.F., Jensen, J.D., 2017. Characterizing human cytomegalovirus reinfection in congenitally infected infants: an evolutionary perspective. Mol. Ecol. 26, 1980–1990. https://doi.org/10.1111/mec.13953

Popa, A., Genger, J.-W., Nicholson, M.D., Penz, T., Schmid, D., Aberle, S.W., Agerer, B., Lercher, A., Endler, L., Colaço, H., Smyth, M., Schuster, M., Grau, M.L., Martínez-Jiménez, F., Pich, O., Borena, W., Pawelka, E., Keszei, Z., Senekowitsch, M., Laine, J., Aberle, J.H., Redlberger-Fritz, M., Karolyi, M., Zoufaly, A., Maritschnik, S., Borkovec, M., Hufnagl, P., Nairz, M., Weiss, G., Wolfinger, M.T., Laer, D. von, Superti-Furga, G., Lopez-Bigas, N., Puchhammer-Stöckl, E., Allerberger, F., Michor, F., Bock, C., Bergthaler, A., 2020. Genomic epidemiology of superspreading events in Austria reveals mutational dynamics and transmission properties of SARS-CoV-2. Sci. Transl. Med. 12. https://doi.org/10.1126/scitranslmed.abe2555

Portet, A., Pinaud, S., Tetreau, G., Galinier, R., Cosseau, C., Duval, D., Grunau, C., Mitta, G., Gourbal, B., 2017. Integrated multi-omic analyses in Biomphalaria-Schistosoma dialogue reveal the immunobiological significance of FREP-SmPoMuc interaction. Dev. Comp. Immunol., Impact of high throughput sequencing on comparative immunogenomics 75, 16–27. https://doi.org/10.1016/j.dci.2017.02.025

Qi, X., Gao, H., Gao, Y., Qin, L., Wang, Y., Gao, L., Wang, X., 2009. Naturally occurring mutations at residues 253 and 284 in VP2 contribute to the cell tropism and virulence of very virulent infectious bursal disease virus. Antiviral Res. 84, 225–233. https://doi.org/10.1016/j.antiviral.2009.09.006

Quick, J., Loman, N.J., Duraffour, S., Simpson, J.T., Severi, E., Cowley, L., Bore, J.A., Koundouno, R., Dudas, G., Mikhail, A., Ouédraogo, N., Afrough, B., Bah, A., Baum, J.H.J., Becker-Ziaja, B., Boettcher, J.P., Cabeza-Cabrerizo, M., Camino-Sánchez, Á., Carter, L.L., Doerrbecker, J., Enkirch, T., Dorival, I.G.-, Hetzelt, N., Hinzmann, J., Holm, T., Kafetzopoulou, L.E., Koropogui, M., Kosgey, A., Kuisma, E., Logue, C.H., Mazzarelli, A., Meisel, S., Mertens, M., Michel, J., Ngabo, D., Nitzsche, K., Pallasch, E., Patrono, L.V., Portmann, J., Repits, J.G., Rickett, N.Y., Sachse, A., Singethan, K., Vitoriano, I., Yemanaberhan, R.L., Zekeng, E.G., Racine, T., Bello, A., Sall, A.A., Faye, Ousmane, Faye, Oumar, Magassouba, N., Williams, C.V., Amburgey, V., Winona, L., Davis, E., Gerlach, J., Washington, F., Monteil, V., Jourdain, M., Bererd, M., Camara, Alimou, Somlare, H., Camara, Abdoulaye, Gerard, M., Bado, G., Baillet, B., Delaune, D., Nebie, K.Y., Diarra, A., Savane, Y., Pallawo, R.B., Gutierrez, G.J., Milhano, N., Roger, I., Williams, C.J., Yattara, F., Lewandowski, K., Taylor, J., Rachwal, P., J. Turner, D., Pollakis, G., Hiscox, J.A., Matthews, D.A., Shea, M.K.O., Johnston, A.M., Wilson, D., Hutley, E., Smit, E., Di Caro, A., Wölfel, R., Stoecker, K., Fleischmann, E., Gabriel, M., Weller, S.A., Koivogui, L., Diallo, B., Keïta, S., Rambaut, A., Formenty, P., Günther, S., Carroll, M.W., 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature 530, 228–232. https://doi.org/10.1038/nature16996

Rajewsky, K., Forster, I., Cumano, A., 1987. Evolutionary and somatic selection of the antibody repertoire in the mouse. Science 238, 1088–1094. https://doi.org/10.1126/science.3317826

Rambaut, A., Pybus, O.G., Nelson, M.I., Viboud, C., Taubenberger, J.K., Holmes, E.C., 2008. The genomic and epidemiological dynamics of human influenza A virus. Nature 453, 615–619. https://doi.org/10.1038/nature06945

Renault, T., Bouquet, A.L., Maurice, J.-T., Lupo, C., Blachier, P., 2014. Ostreid Herpesvirus 1 Infection among Pacific Oyster (Crassostrea gigas) Spat: Relevance of Water Temperature to Virus Replication and Circulation Prior to the Onset of Mortality. Appl. Environ. Microbiol. 80, 5419–5426. https://doi.org/10.1128/AEM.00484-14

Renault, T., Cochennec, N., Le Deuff, R.-M., Chollet, B., 1994a. Herpes-like virus infecting Japanese oyster (Crassostrea gigas) spat. Bull. Eur. Assoc. Fish Pathol. EAFP 14, 64–66.

Renault, T. « Tenter de maîtriser les maladies infectieuses chez les mollusques : une clé pour une aquaculture durable », Mémoire d'Habilitation à diriger des recherches, La Tremblade, Université de la Rochelle, 2007, 269p

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. Dev. Comp. Immunol. 35, 725–735. 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. Dis. Aquat. Organ. 42, 173–183. https://doi.org/10.3354/dao042173

Renault, T., Le Deuff, R.-M., Cochennec, N., Maffart, P., 1994b. Herpesviruses associated with mortalities among Pacific oyster, Crassostrea gigas, in France-Comparative study. Rev. Med. Veterinaire 145, 735–742.

Renault, T., Lipart, C., Arzul, I., 2001. A herpes-like virus infects a non-ostreid bivalve species: virus replication in Ruditapes philippinarum larvae. Dis. Aquat. Organ. 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. J. Virol. 86, 5942–5947. https://doi.org/10.1128/JVI.06534-11

Renner, D.W., Szpara, M.L., 2017. Impacts of Genome-Wide Analyses on Our Understanding of Human Herpesvirus Diversity and Evolution. J. Virol. 92, e00908-17. https://doi.org/10.1128/JVI.00908-17

Renzette, N., Bhattacharjee, B., Jensen, J.D., Gibson, L., Kowalik, T.F., 2011. Extensive Genome-Wide Variability of Human Cytomegalovirus in Congenitally Infected Infants. PLoS Pathog. 7, e1001344. https://doi.org/10.1371/journal.ppat.1001344

Renzette, N., Gibson, L., Bhattacharjee, B., Fisher, D., Schleiss, M.R., Jensen, J.D., Kowalik, T.F., 2013. Rapid Intrahost Evolution of Human Cytomegalovirus Is Shaped by Demography and Positive Selection. PLoS Genet. 9, e1003735. https://doi.org/10.1371/journal.pgen.1003735

Renzette, N., Gibson, L., Jensen, J.D., Kowalik, T.F., 2014. Human cytomegalovirus intrahost evolution-a new avenue for understanding and controlling herpesvirus infections. Curr. Opin. Virol. 8, 109–115. https://doi.org/10.1016/j.coviro.2014.08.001

Renzette, N., Kowalik, T.F., Jensen, J.D., 2016. On the relative roles of background selection and genetic hitchhiking in shaping human cytomegalovirus genetic diversity. Mol. Ecol. 25, 403–413. https://doi.org/10.1111/mec.13331

Renzette, N., Pfeifer, S.P., Matuszewski, S., Kowalik, T.F., Jensen, J.D., 2017. On the Analysis of Intrahost and Interhost Viral Populations: Human Cytomegalovirus as a Case Study of Pitfalls and Expectations. J. Virol. 91, e01976-16, e01976-16. https://doi.org/10.1128/JVI.01976-16

Renzette, N., Pokalyuk, C., Gibson, L., Bhattacharjee, B., Schleiss, M.R., Hamprecht, K., Yamamoto, A.Y., Mussi-Pinhata, M.M., Britt, W.J., Jensen, J.D., Kowalik, T.F., 2015. Limits and patterns of cytomegalovirus genomic diversity in humans. Proc. Natl. Acad. Sci. 112, E4120–E4128. https://doi.org/10.1073/pnas.1501880112

Retel, C., Märkle, H., Becks, L., Feulner, P.G.D., 2019. Ecological and Evolutionary Processes Shaping Viral Genetic Diversity. Viruses 11, 220. https://doi.org/10.3390/v11030220

Richman, D.D., Wrin, T., Little, S.J., Petropoulos, C.J., 2003. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc. Natl. Acad. Sci. 100, 4144–4149. https://doi.org/10.1073/pnas.0630530100

Roossinck, M.J., 2011. The good viruses: viral mutualistic symbioses. Nat. Rev. Microbiol. 9, 99–108. https://doi.org/10.1038/nrmicro2491

Rosani, U., Varotto, L., Domeneghetti, S., Arcangeli, G., Pallavicini, A., Venier, P., 2014. Dual analysis of host and pathogen transcriptomes in ostreid herpesvirus 1-positive Crassostrea gigas Host-pathogen RNA-seq in OsHV-1 positive C. gigas. Environ. Microbiol. 17, 4200–4212. https://doi.org/10.1111/1462-2920.12706

Rosani, U., Venier, P., 2017. Oyster RNA-seq Data Support the Development of Malacoherpesviridae Genomics. Front. Microbiol. 8, 15. https://doi.org/10.3389/fmicb.2017.01515

Roux, L., Simon, A.E., Holland, J.J., 1991. Effects of defective interfering viruses on virus replication and pathogenesis in vitro and in vivo. Adv. Virus Res. 40, 181–211. https://doi.org/10.1016/s0065-3527(08)60279-1

Ruiz-Jarabo, C.M., Arias, A., Baranowski, E., Escarmís, C., Domingo, E., 2000. Memory in Viral Quasispecies. J. Virol. 74, 3543–3547. https://doi.org/10.1128/JVI.74.8.3543-3547.2000

Ryabov, E.V., Childers, A.K., Lopez, D., Grubbs, K., Posada-Florez, F., Weaver, D., Girten, W., vanEngelsdorp, D., Chen, Y., Evans, J.D., 2019. Dynamic evolution in the key honey bee pathogen deformed wing virus: Novel insights into virulence and competition using reverse genetics. PLOS Biol. 17, e3000502. https://doi.org/10.1371/journal.pbio.3000502

Salvi, D., Mariottini, P., 2016. Molecular taxonomy in 2D: a novel ITS2 rRNA sequence-structure approach guides the description of the oysters' subfamily Saccostreinae and the genus Magallana (Bivalvia: Ostreidae). Zool. J. Linn. Soc. https://doi.org/10.1111/zoj.12455

Samain, J.F., Dégremont, L., Soletchnik, P., Haure, J., Bédier, E., Ropert, M., Moal, J., Huvet, A., Bacca, H., Van Wormhoudt, A., Delaporte, M., Costil, K., Pouvreau, S., Lambert, C., Boulo, V., Soudant, P., Nicolas, J.L., Le Roux, F., Renault, T., Gagnaire, B., Geret, F., Boutet, I., Burgeot, T., Boudry, P., 2007. Genetically based resistance to summer mortality in the Pacific oyster (Crassostrea gigas) and its relationship with physiological, immunological characteristics and infection processes. Aquaculture 268, 227–243. https://doi.org/10.1016/j.aquaculture.2007.04.044

Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., Belshaw, R., 2010. Viral Mutation Rates. J. Virol. 84, 9733–9748. https://doi.org/10.1128/JVI.00694-10

Sauvage, C., Boudry, P., de Koning, D.-J., Haley, C.S., Heurtebise, S., Lapègue, S., 2009. QTL for resistance to summer mortality and OsHV-1 load in the Pacific oyster (Crassostrea gigas). Anim. Genet. https://doi.org/10.1111/j.1365-2052.2009.02018.x

Savin, K.W., Cocks, B.G., Wong, F., Sawbridge, T., Cogan, N., Savage, D., Warner, S., 2010. A neurotropic herpesvirus infecting the gastropod, abalone, shares ancestry with oyster herpesvirus and a herpesvirus associated with the amphioxus genome. Virol. J. 7, 308. https://doi.org/10.1186/1743-422X-7-308

Scheidel, L.M., Durbin, R.K., Stollar, V., 1987. Sindbis virus mutants resistant to mycophenolic acid and ribavirin. Virology 158, 1–7. https://doi.org/10.1016/0042-6822(87)90230-3

Schikorski, Faury, N., Pepin, J.F., Saulnier, D., Tourbiez, D., Renault, T., 2010a. 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 Res. 155, 28–34. https://doi.org/10.1016/j.virusres.2010.07.031

Schikorski, Renault, T., Saulnier, D., Faury, N., Moreau, P., Pépin, J.-F., 2010b. Experimental infection of Pacific oyster Crassostrea gigas spat by ostreid herpesvirus 1: demonstration of oyster spat susceptibility. Vet. Res. 42, 27. https://doi.org/10.1186/1297-9716-42-27

Scientific Opinion on the increased mortality events in Pacific oysters, Crassostrea gigas, 2010. . EFSA J. 8, 1894. https://doi.org/10.2903/j.efsa.2010.1894

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

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, 580. 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 Res. 153, 92–99. 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 Pathol. 47, 129–136. https://doi.org/10.3147/jsfp.47.129

Simões, E.A.F., Forleo-Neto, E., Geba, G.P., Kamal, M., Yang, F., Cicirello, H., Houghton, M.R., Rideman, R., Zhao, Q., Benvin, S.L., Hawes, A., Fuller, E.D., Wloga, E., Pizarro, J.M.N., Munoz, F.M., Rush, S.A., McLellan, J.S., Lipsich, L., Stahl, N., Yancopoulos, G.D., Weinreich, D.M., Kyratsous, C.A., Sivapalasingam, S., 2020. Suptavumab for the Prevention of Medically Attended Respiratory Syncytial Virus Infection in Preterm Infants. Clin. Infect. Dis. https://doi.org/10.1093/cid/ciaa951

Snieszko, S.F., 1974. The effects of environmental stress on outbreaks of infectious diseases of fishes. J. Fish Biol. 6, 197–208. https://doi.org/10.1111/j.1095-8649.1974.tb04537.x*

Stahl, E.A., Dwyer, G., Mauricio, R., Kreitman, M., Bergelson, J., 1999. Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. Nature 400, 667–671. https://doi.org/10.1038/23260

Takeuchi, T., Takemoto, Y., Matsubara, T., 1960. Haematological study of bacteria affected oysters. Hiroshima-ken Suisan Shikenjo Hokoku 22, 1–7.

Tellier, A., Moreno-Gámez, S., Stephan, W., 2014. Speed of Adaptation and Genomic Footprints of Host–Parasite Coevolution Under Arms Race and Trench Warfare Dynamics. Evolution 68, 2211–2224. https://doi.org/10.1111/evo.12427

Thézé, J., Li, T., du Plessis, L., Bouquet, J., Kraemer, M.U.G., Somasekar, S., Yu, G., de Cesare, M., Balmaseda, A., Kuan, G., Harris, E., Wu, C., Ansari, M.A., Bowden, R., Faria, N.R., Yagi, S., Messenger, S., Brooks, T., Stone, M., Bloch, E.M., Busch, M., Muñoz-Medina, J.E., González-Bonilla, C.R., Wolinsky, S., López, S., Arias, C.F., Bonsall, D., Chiu, C.Y., Pybus, O.G., 2018. Genomic Epidemiology Reconstructs the Introduction and Spread of Zika Virus in Central America and Mexico. Cell Host Microbe 23, 855-864.e7. https://doi.org/10.1016/j.chom.2018.04.017

Thomas, J.C., Godfrey, P.A., Feldgarden, M., Robinson, D.A., 2012. Candidate targets of balancing selection in the genome of Staphylococcus aureus. Mol. Biol. Evol. 29, 1175–1186. https://doi.org/10.1093/molbev/msr286

Thrall, P.H., Laine, A.-L., Ravensdale, M., Nemri, A., Dodds, P.N., Barrett, L.G., Burdon, J.J., 2012. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. Ecol. Lett. 15, 425–435. https://doi.org/10.1111/j.1461-0248.2012.01749.x

Thunberg, C.P., 1793. ekning och Beskrifning på en stor Ostronsort ifrån Japan. Kongliga Vetenskaps Academiens Nya Handlingar., Academiens nya handlingar. ed.

Tonegawa, S., 1983. Somatic generation of antibody diversity. Nature 302, 575–581. https://doi.org/10.1038/302575a0

Torre, J.C. de la, Holland, J.J., 1990. RNA virus quasispecies populations can suppress vastly superior mutant progeny. J. Virol. 64, 6278–6281.

Trimpert, J., Groenke, N., Kunec, D., Eschke, K., He, S., McMahon, D.P., Osterrieder, N., 2019. A proofreading-impaired herpesvirus generates populations with quasispecies-like structure. Nat. Microbiol. https://doi.org/10.1038/s41564-019-0547-x

Vahlenkamp, T.W., Verschoor, E.J., Schuurman, N.N., Vliet, A.L. van, Horzinek, M.C., Egberink, H.F., Ronde, A. de, 1997. A single amino acid substitution in the transmembrane envelope glycoprotein of feline immunodeficiency virus alters cellular tropism. J. Virol. 71, 7132–7135.

Van Valen, L., 1973. A New Evolutionary Law. Evol. Theory 1 1–30.

Vásquez-Yeomans, R., García-Ortega, M., Cáceres-Martínez, J., 2010. Gill erosion and herpesvirus in Crassostrea gigas cultured in Baja California, Mexico. Dis. Aquat. Organ. 89, 137–144. https://doi.org/10.3354/dao02189

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., Andino, R., 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439, 344–348. https://doi.org/10.1038/nature04388

Wainberg, M.A., Drosopoulos, W.C., Salomon, H., Hsu, M., Borkow, G., Parniak, M.A., Gu, Z., Song, Q., Manne, J., Islam, S., Castriota, G., Prasad, V.R., 1996. Enhanced Fidelity of 3TC-Selected Mutant HIV-1 Reverse Transcriptase. Science 271, 1282–1285. https://doi.org/10.1126/science.271.5253.1282

Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164–e164. https://doi.org/10.1093/nar/gkq603

Weigert, M.G., Cesari, I.M., Yonkovich, S.J., Cohn, M., 1970. Variability in the lambda light chain sequences of mouse antibody. Nature 228, 1045–1047. https://doi.org/10.1038/2281045a0

Whitney, J.B., Oliveira, M., Detorio, M., Guan, Y., Wainberg, M.A., 2002. The M184V Mutation in Reverse Transcriptase Can Delay Reversion of Attenuated Variants of Simian Immunodeficiency Virus. J. Virol. 76, 8958–8962. https://doi.org/10.1128/JVI.76.17.8958-8962.2002

Whittington, R.J., Paul-Pont, I., Evans, O., Hick, P., Dhand, N.K., 2018. Counting the dead to determine the source and transmission of the marine herpesvirus OsHV-1 in Crassostrea gigas. Vet. Res. 49, 34. https://doi.org/10.1186/s13567-018-0529-7

Wilfert, L., Jiggins, F.M., 2012. The dynamics of reciprocal selective sweeps of host resistance and a parasite counter-adaptation in Drosophila. Evol. Int. J. Org. Evol. 67, 761–773. https://doi.org/10.1111/j.1558-5646.2012.01832.x

Wolf, J.B.W., Ellegren, H., 2017. Making sense of genomic islands of differentiation in light of speciation. Nat. Rev. Genet. 18, 87–100. https://doi.org/10.1038/nrg.2016.133

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. Nat. Genet. 32, 569–577. https://doi.org/10.1038/ng1202-569

Wu, Z., Asokan, A., Grieger, J.C., Govindasamy, L., Agbandje-McKenna, M., Samulski, R.J., 2006. Single Amino Acid Changes Can Influence Titer, Heparin Binding, and Tissue Tropism in Different Adeno-Associated Virus Serotypes. J. Virol. 80, 11393–11397. https://doi.org/10.1128/JVI.01288-06

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. Virol. J. 12, 110. https://doi.org/10.1186/s12985-015-0334-0

Yao, T., Zhang, Y., Yan, X., Wang, Z., Li, D., Su, J., Yu, R., 2015. Interspecific hybridization between Crassostrea angulata and C. ariakensis. J. Ocean Univ. China 14, 710–716. https://doi.org/10.1007/s11802-015-2546-8

Young, K.-C., Lindsay, K.L., Lee, K.-J., Liu, W.-C., He, J.-W., Milstein, S.L., Lai, M.M.C., 2003. Identification of a ribavirin-resistant NS5B mutation of hepatitis C virus during ribavirin monotherapy. Hepatol. Baltim. Md 38, 869–878. https://doi.org/10.1053/jhep.2003.50445

Zhang, Guofan, Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P.W.H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, Jiafeng, Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng, Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Lošo, T., Du, Y., Sun, X., Zhang, Shoudu, Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Wang, W., Fu, W., Wang, T., Wang, B., Zhang, J., Peng, Z., Li, Yingxiang, Li, Na, Wang, Jinpeng, Chen, M., He, Y., Tan, F., Song, X., Zheng, Q., Huang, R., Yang, Hailong, Du, X., Chen, L., Yang, M., Gaffney, P.M., Wang, S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, Shu, Huang, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, Junyi, Wang, Q., Steinberg, C.E.W., Wang, H., Li, Ning, Qian, L., Zhang, Guojie, Li, Yingrui, Yang, Huanming, Liu, X., Wang, Jian, Yin, Y., Wang, Jun, 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490, 49–54. https://doi.org/10.1038/nature11413

Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., Zhang, G., 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. 5, 8693. https://doi.org/10.1038/srep08693

Zhang, S.-M., Adema, C.M., Kepler, T.B., Loker, E.S., 2004. Diversification of Ig Superfamily Genes in an Invertebrate. Science 305, 251–254. https://doi.org/10.1126/science.1088069

Zhang, S.M., Léonard, P.M., Adema, C.M., Loker, E.S., 2001. Parasite-responsive IgSF members in the snail Biomphalaria glabrata: characterization of novel genes with tandemly arranged IgSF domains and a fibrinogen domain. Immunogenetics 53, 684–694. https://doi.org/10.1007/s00251-001-0386-8

Annexe

Annexe 1. ***Hôtes bivalves signalés de virus de type herpès / herpèsvirus.*** *143*

Annexe 2. ***Pipeline bioinformatique d'appel de variant.*** *145*

Annexe 3 ***Pipeline bioinformatique d'assemblage de novo*** *146*

Annexe 4. ***Pipeline d'assemblage utilisé pour l'article II.*** *156*

Annexe 5 ***Récapitulatif des biais à prendre en compte dans l'analyse et la comparaison de la diversité génétique.*** *157*

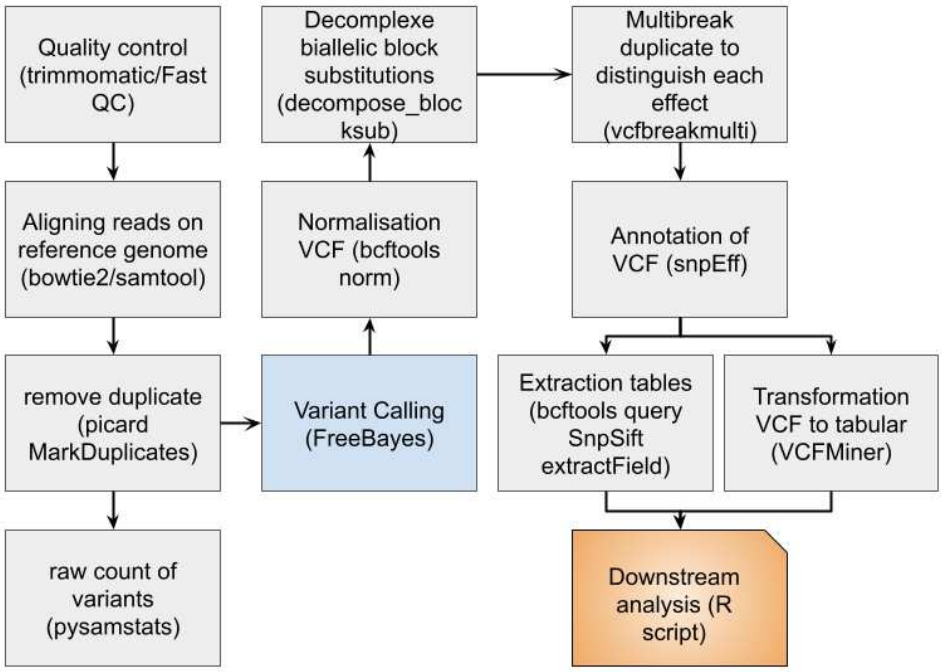
Annexe 1. **Hôtes bivalves signalés de virus de type herpès / herpèsvirus**.

Les hôtes du virus peuvent avoir été retrouvés plusieurs fois

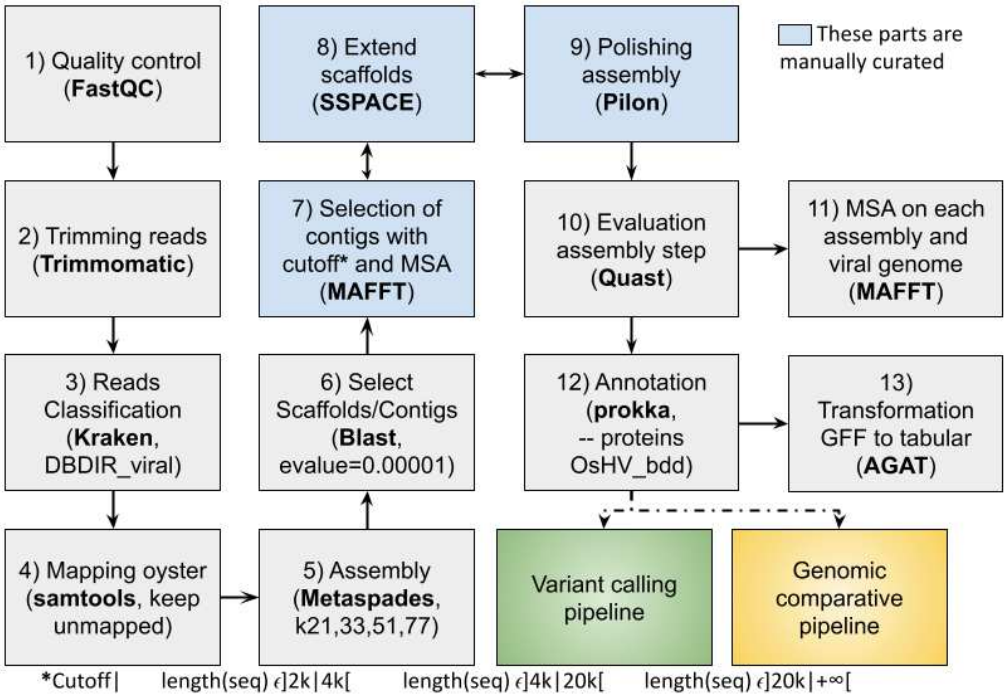
Host	**publication**	**1er description**	**note**	**taxonomie**	**classe**	**présence**	**nom commun**
Crassostrea Virginica	Farley et al. 1972	1972-11-17	histology/MET	oyster	bivalve	+	Eastern oyster
Crassostrea gigas	Hine et al. 1992, Nicolas et al. 1992	1992-02-10	histology/MET	oyster	bivalve	+	Pacific oyster
Ostrea edulis	Comps et al. 1993	1993-02-17	histology/MET	oyster	bivalve	+	Flat oyster
Ostrea angasi	Hine et al. 1997	1997-03-11	histology/MET, adults	oyster	bivalve	+	Southern mud oyster
Tiostrea chilensis	Hine et al. 1998	1998-04-03	histology/MET	oyster	bivalve	+	
Ruditapes philippinarum	Renault et al. 1998	1998-02-01	histology/MET	clams	bivalve	+	
Ruditapes decussatus	Renault et al. 1998	1998-02-01	histology/MET	clam	bivalve	+	
Chlamys farreri	Chongming et al. 2002			scallops	bivalve	+	Chinese scallop
Ruditapes philippinarum	Arzul et al. 2001a	2001-01-05	MET,PCR	clams	bivalve	+	
Ruditapes decussatus	Arzul et al. 2001a	2001-01-05	MET,PCR	clam	bivalve	+	
Crassostrea angulata	Arzul et al. 2001b	2001-08-22	MET,PCR	oyster	bivalve	+	Portuguese oyster
Crassostrea rivularis	Arzul et al. 2001b	2001-08-22	MET,PCR	oyster	bivalve	+	
Pecten maximus	Arzul et al. 2001c	2001-01-05	MET,PCR	scallops	bivalve	+	
Haliotis diversicolor	Chang et al. 2005	2005-06-14	histology/MET, experimental infection	abalone	Gastero pode	+	

Crassostrea hongkongensis	Moss et al. 2007	2007-10-15	MET,PCR	oyster	bivalve	+	Hong Kong oyster
Crassostrea ariakensis	Moss et al. 2007	2007-10-15	MET,PCR	oyster	bivalve	+	
Crassostrea sikamea	Moss et al. 2007	2007-10-15	MET,PCR	oyster	bivalve	+	Kumamoto oyster
Crassostrea sikamea	Burge et al. 2011	2011-04-05	qPCR	oyster	bivalve	+	Kumamoto oyster
Mytilus galloprovincialis	Burge et al. 2011	2011-04-05	qPCR	mussel	bivalve	+	Mediterranean mussel
Venerupis philippinarum	Burge et al. 2011	2011-04-05	qPCR	scallops	bivalve	-	Manila clam
Ostrea lurida	Burge et al. 2011	2011-04-05	qPCR	oyster	Bivalve	-	Olympia oyster
Ostrea edulis	Burge et al. 2011	2011-04-05	qPCR	oyster	bivalve	+	Flat oyster
Saccostrea glomerata	Jenkins et al. 2013	2010-11	PCR/Sequencing /histology	oyster	Bivalve	-	Sydney rock oyster
Crassostrea angulata	Batista et al 2014	2001-08-22	PCR/Sequencing	oyster	bivalve	+	Portuguese oyster
Scapharca broughtonii	Bai et al. 2014	2014-12-05	PCR/Sequencing, doit être trouvé avant	scallops	bivalve	+	blood ark shell
Chlamys farreri	Bai et al. 2014		PCR/Sequencing	scallops	bivalve	+	
Patinopecten yessoensis	Bai et al. 2014	2014-12-05	PCR/Sequencing	scallops	bivalve	+	
Meretrix meretrix	Bai et al. 2014	2014-12-05	PCR/Sequencing	clams	bivalve	+	
Ruditapes philippinarum	Bai et al. 2014	1998-02-01	PCR/Sequencing	clams	bivalve	+	
Crassostrea hongkongensis	Bai et al. 2014	2007-10-15	PCR/Sequencing	oyster	bivalve	+	
Argopecten irradians	Bai et al. 2014	2014-12-05	PCR	scallops	Bivalve	-	bay scallop
Crassostrea angulata	López Sanmartín et al. 2016	2001-08-22	PCR, vertical transmission	oyster	bivalve	+	Portuguese oyster
Crassostrea corteziensis	Martinez et al. 2017	2017-01-10	PCR	oyster	bivalve	+	Cortez oyster
Crassostrea brasiliana	Mello et al. 2017	2017-10-13	PCR	oyster	bivalve	+	mangrove oyster
Saccostrea	Evans et al. 2017	2017-05-08	qPCR	oyster	Bivalve	+	Sydney rock

glomerata							oyster
Anadara trapezia	Evans et al. 2017	2017-05-08	qPCR	clams	Bivalve	+	Sydney cockle



Annexe 2. **Pipeline bioinformatique d'appel de variant**.
Les outils sont indiqué entre parenthèse, la version avec pysamstats a été utilisé dans l'article I et la version avec Freebayes a été utiliser dans l'article II.



Annexe 3 **Pipeline bioinformatique d'assemblage *de novo***.
Ce pipeline a été utilisé uniquement dans l'article II. Le dépôt de code Git correspondant est en cours de nettoyage et sera accessible lors de la publication de l'article.

```
################################## DiVir ###################################


## DESCRIPTION ##

# Metagenomics analysis targeted viral reads


## Related program ##

#print_comment.py (HemoVir)


#############################################################################

# FONCTION USED

#############################################################################


function show_help {

    echo "Usage: $0 [options] -i input.fastq.gz -o output_dir -f host_genome.fa -l 200 -r
virus_ref.fa -p 2 -t 4 -m 32

    Options:

    -h              Show this help message and exit

    -i input        Input fastq.gz file from Minion sequencing

    -o dir          Path of the output directory

    -f file         Fasta file form host (to sort against the host)

    -l #            min_length of long reads (default=200nt)

    -r file         Fasta file known virus

    -p #            Nombre de polishing à faire (default = 1)

    -t #            Number of threads

    -m #            Memory used (Gb)


Example:

$0 -i input.fastq.gz -o output_dir -f host_genome.fa -r oshv.fa -l 200 -p 2 -t 4 -m 32"

}

# Fonction to align reads with contigs

Alignment_bowtie()

{

    if [ $# -eq 4 ]; then

        echo "Alignment_bowtie function with $NCPUS CPUs"

        echo -e "genome=$1"

        echo -e "R1=$2"

        echo -e "R2=$3"
```

```
# Structure of chunk announcement - (check_start_chunk - code - check_end_chunk)

# announcement see l181

announcement () { python2 ${SOFT}/print_comment.py "$1" "#" >> ${logfile}; echo -e "\n## ${1}\n" >>
$report ; }


check_start_chunk ()

{

    echo -e "$1 start at: $(date +%Y-%m-%d_%Hh%Mm%Ss)" >> $report

    echo -e "$1 start at: $(date +%Y-%m-%d_%Hh%Mm%Ss)" >> $logfile

    cd $output

    # env

    source activate ${path_conda_env}/$1

}


check_end_chunk ()

{

    echo -e "$1 End at: $(date +%Y-%m-%d_%Hh%Mm%Ss)" >> $report

    echo -e "$1 End at: $(date +%Y-%m-%d_%Hh%Mm%Ss)" >> $logfile

    # env

    conda deactivate

    cd $output

}


################################################################################

#

#SET UP CONFIGURATION VARIABLES & VALIDATE INPUTS

#

################################################################################


export PATH=/appli/anaconda/3.7/bin/:$PATH


if [ ! -d "$outdir" ]; then

    echo -e "Error 'OUTDIR' dosen't exist, please make the output path first\n"

    show_help

    exit 0

else

    echo "Analysis using DiVir pipeline !"

    cd $outdir

fi
```

```
# variable depend of HPC

if [ -z ${max_memory} ]; then echo "max_memory defaut parameter : 110" ;

    max_memory=110 ; fi

if [ -z ${NCPUS} ]; then echo "max_memory defaut parameter: 28" ;

    NCPUS=28 ; fi

if [ -z ${path_conda_env} ]; then echo "path_conda_env defaut parameter:
/home1/datahome/jdelmott/conda-env" ;

    path_conda_env="/home1/datahome/jdelmott/conda-env" ; fi

# Pretty modified from METAWRAP pipeline

if [ -z ${SOFT} ]; then echo "SOFT defaut parameter: /home1/datahome/jdelmott" ;

    SOFT=/home1/datahome/jdelmott ; fi

if [ -z $adapter_trimming ]; then echo "adapter_trimming defaut parameter :
/home1/datahome/jdelmott/all_adapter.fa" \

        ; adapter_trimming=/home1/datawork/jdelmott/raw_Hemovir/adapter_Hemovir.fa ; fi


# Analysis selection

if [ -z ${skip_assembly} ]; then echo "skip_assembly defaut parameter : no" \

        ; skip_assembly="no" ; fi


# Output directory

if [ ! -d "$name" ];

then

    mkdir $name

    cd $name

    echo "Start DiVir pipeline analysis on $name"

else

    cd $name

fi


################################################################################

#

#SET UP GLOBAL VARIABLES

#

################################################################################


#output=${outdir}/${name}

basefile=$(basename $reads1 .fastq.gz)

logfile=$output/${name}_logfile.txt

report=$output/${name}_$(date +%Y-%m-%d_%Hh%Mm)_report.md
```

```
#############################################################################
#
# Start Pipeline
#
#############################################################################

echo -e "Commande: ${0}\nStart at: $(date +%Y-%m-%d_%Hh%Mm%Ss)\nInput:\n" >> $logfile

echo -e "output=$output
basefile=$basefile
reads1=$reads1
reads2=$reads2
name=$name
host_genome=$host_genome
viral_fasta=$viral_fasta
viral_kraken_db=$viral_kraken_db
insersize=$insersize
viral_target_proteins=$viral_target_proteins
viral_genome=$viral_genome
viral_gff=$viral_gff
max_memory=$max_memory
NCPUS=$NCPUS
path_conda_env=$path_conda_env
adapter_trimming=$adapter_trimming
skip_assembly=$skip_assembly
viral_input=$viral_input" >> $logfile

echo -e "# Report of DiVir analysis\n\nStart at: $(date +%Y-%m-%d_%Hh%Mm%Ss)\n" > $report

if [ ! -f "$fastQC_out1" ]; then
    announcement "FastQC 1 Control quality of raw reads"
    check_start_chunk "fastqc"
    if [ $reads2 == "" ]; then
        echo -e "\nFirst quality controle on ${reads1} (single ends)... \n" >> $logfile
        fastqc $reads1
    else
        mkdir $fastQC
        echo -e "### First quality controle" >> $report
```

```
if [ ! -f "$readsTrim1" ]; then
    announcement "Trimming reads with trimomatic"
    check_start_chunk "trimmomatic"
    if [ $reads2 == "" ]; then
        #mkdir $trimmomatic
        echo -e "\n Trimming ${reads1}..." >> $logfile
        #trimmomatic PE -phred33 -threads $NCPUS ${reads1} \
        #${reads2} \
        #${readsTrim1} \
        #$(basename $reads1 .fastq.gz)_1unpaired.fastq.gz \
        #ILLUMINACLIP:/home1/datahome/jdelmott/all_adapter.fa:2:30:10 \
        #LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50
        #echo -e "Reads trim in : ${reads1%.fastq.gz}_trim.fastq.gz \n " >> $logfile
    else
        mkdir $trimmomatic
        cd 02-trimmomatic
        echo -e "\n### Perform Triming of reads" >> $report
        echo -e "\nTrimmomatic on ${reads1} and ${reads2} (paired ends)" >> $report
        echo -e "Parameters are :
        Remove leading low quality or N bases (below quality 3) (LEADING:3)
        Remove trailing low quality or N bases (below quality 3) (TRAILING:3)
        Scan the read with a 4-base wide sliding window, cutting when the average quality per base
drops below 15 (SLIDINGWINDOW:4:15)
        Drop reads below the 50 bases long (MINLEN:50)
        source: [trimmomatic](http://www.usadellab.org/cms/?page=trimmomatic)" >> $report
        if [ ! -f $adapter_trimming ]; then
            echo "Error : ${adapter_trimming} dosen't exit,
            Verifiy if the file exist or specified using 'adapter_trimming' parameter" \
                >> $logfile
            exit 0
        fi
        trimmomatic PE -phred33 -threads $NCPUS \
            ${reads1} \
            ${reads2} \
            ${readsTrim1} \
            $(basename $reads1 .fastq.gz)_1unpaired.fastq.gz \
            ${readsTrim2} \
            $(basename $reads2 .fastq.gz)_2unpaired.fastq.gz \
            ILLUMINACLIP:$adapter_trimming:2:30:10 \
```

```
# Reads_classification


### This parts is from the metaviromics pipeline, both of this pipeline worked together ###


if [ ! -f "03-Reads_classification/${ID}_unclassified.fasta" ]
then
    announcement "Reads classification with Kraken" >> $logfile
    cd ${outdir}/${ID}
    if [ ! -d 03-Reads_classification ]; then
        mkdir 03-Reads_classification
    fi
    cd 03-Reads_classification
    # env
    source activate /home1/datahome/jdelmott/conda-env/krakenuniq
    # Classification
    if [ $reads2 == "" ] ;
        then
            echo -e "\nError reads should be paired read (not developp yet) \n" >> $logfile
        else
            echo -e "\nStarting Reads classification... \n" >> ${outdir}/${ID}/$logfile
            krakenuniq --report-file ${ID}_kraken.txt \
                --db ${KRAKEN_DB} \
                --gzip-compressed \
                --unclassified-out ${ID}_unclassified.fasta \
                --classified-out ${ID}_classified.fasta \
                --threads $NCPUS \
                --fastq-input \
                --paired ${outdir}/${ID}/${readsTrim1} ${outdir}/${ID}/${readsTrim2} \
                    2>> ${outdir}/${ID}/$logfile \
                    &> krakenuniq_logFile.txt
            source deactivate
            # env
            source activate /home1/datahome/jdelmott/conda-env/krona
            echo -e "Create krona report..." >> ${outdir}/${ID}/$logfile
            ktImportTaxonomy \
                -o ${ID}_kraken_krona.html \
                -t 7 \
                -s 2 \
```

```
# Aligning reads on Host genome and keep others reads
if [ ! -f "${outdir}/${ID}/04-Bowtie2_alignment_HOST/${Host_bam%.bam}_unmapped_1.fastq" ] ;
then
    announcement "Filtering reads against Host genome" >> $logfile
    cd ${outdir}/${ID}
    if [ ! -d 04-Bowtie2_alignment_HOST ]; then
        mkdir 04-Bowtie2_alignment_HOST
    fi
    # env
    source activate /home1/datawork/jdelmott/virus-discovery-pipeline
    ## Create bowtie2 index from genome file.
    if [ ! -f ${genomefile}.1.bt2 ]
    then
        echo -e '\n Building genome index... \n'
        bowtie2-build --quiet $genomefile $genomefile
    fi

    ## Aligning reads on Host genome
    if [ ! -f "${outdir}/${ID}/04-Bowtie2_alignment_HOST/$Host_bam" ] ;
    then
        if [ "$reads2" == "" ] ;
        then
            echo -e '\nAligning single reads with genome... \n'
            #bowtie2 \
            #    -p $NCPUS \
            #    -x $genomefile \
            #    -U $readsTrim1 | \
            #    samtools view -b > $Host_bam 2>>${logfile}
            #samtools flagstat $Host_bam >> ${logfile}
        else
            echo -e '\nAligning paired reads with genome...\n'
            bowtie2 \
                -p $NCPUS \
                -x $genomefile \
                -1 ${outdir}/${ID}/${ID}_viral_read_1.fastq.gz \
                -2 ${outdir}/${ID}/${ID}_viral_read_2.fastq.gz | \
                samtools view -b > ${outdir}/${ID}/04-Bowtie2_alignment_HOST/$Host_bam \
                    2>>${logfile}
```

```
    ## Convert unaligned reads to fasta format

    if [ ! -f "${outdir}/${ID}/04-Bowtie2_alignment_HOST/${Host_bam%.bam}_mapped_singleton.fastq" ]
;

    then

        echo -e '\n converting mapped-reads.bam to fastq format... \n' >> ${outdir}/${ID}/$logfile

        if [ "$reads2" == "" ] ;

        then

            echo -e "Error it should be paired reads"

            #samtools fastq $getUnmappedBam > $unmappedFastqS

        else

            samtools fastq \

            -@ $NCPUS \

            -1 ${outdir}/${ID}/04-Bowtie2_alignment_HOST/${Host_bam%.bam}_unmapped_1.fastq \

            -2 ${outdir}/${ID}/04-Bowtie2_alignment_HOST/${Host_bam%.bam}_unmapped_2.fastq \

            -s ${outdir}/${ID}/04-Bowtie2_alignment_HOST/${Host_bam%.bam}_unmapped_singleton.fastq \

            ${outdir}/${ID}/04-Bowtie2_alignment_HOST/${getUnmappedBam} \

            2>> ${outdir}/${ID}/$logfile

            echo -e "Unaligned pair 1\t"`grep -c "^@" ${outdir}/${ID}/04-
Bowtie2_alignment_HOST/${Host_bam%.bam}_mapped_1.fastq` \

            >> ${outdir}/${ID}/$logfile

            echo -e "Unaligned pair 2\t"`grep -c "^@" ${outdir}/${ID}/04-
Bowtie2_alignment_HOST/${Host_bam%.bam}_mapped_2.fastq` \

            >> ${outdir}/${ID}/$logfile

            echo -e "Unaligned single\t"`grep -c "^@" ${outdir}/${ID}/04-
Bowtie2_alignment_HOST/${Host_bam%.bam}_mapped_singleton.fastq` \

            >> ${outdir}/${ID}/$logfile

        fi

    else

        echo -e "\nSkipped bam to fastq" >> ${outdir}/${ID}/$logfile

    fi

    source deactivate

    cd ${outdir}/${ID}

else

    echo -e "\nSkipped alignment" >> $logfile

fi

# Assembling the contigs from the mapped reads. Not Filter contigs by size.

if [ ! -f "${outdir}/${ID}/05-MetaSPAdes_assembly/scaffolds.fasta" ] ;

then

    announcement "Assembly contigs with MetaSpades" >> $logfile

    cd ${outdir}/${ID}
```

```
# Fonction to polish assembly

Polishing_assembly()

{

    # Controle parameter

    if [ $# -eq 3 ]; then

        echo "Polishing assembly function with $NCPUS CPUs"

        echo -e "genome=$1"

        echo -e "Frags=$2"

        echo -e "out=$3"

        source activate /home1/datahome/jdelmott/conda-env/pilon

    else

        echo -e "There are $# arguments instead of 3"

    fi

    local genome=$(basename $1)

    local Frags=$(basename $2)

    local out=$3

    #outfile=$4

    ln -s $1 .

    ln -s $2 .

    #pilon \

    /home1/datahome/jdelmott/conda-env/pilon/bin/java -Xms512m -Xmx150g -jar \

    /home1/datahome/jdelmott/conda-env/pilon/share/pilon-1.23-2/pilon-1.23.jar \

        --genome ${genome} \

        --frags ${Frags} \

        --threads $NCPUS \

        --verbose \

        --vcf \

        --output ${out}_contigs \

        --outdir $out \

        &> ${basefile}_pilon_logfile.log \

        2>> ${outdir}/${ID}/$logfile

    source deactivate

}

### Bining

#if [ ! -f "${outdir}/${ID}/07-Binning_contigs/" ] ;

#then

#    announcement "Binning contigs with Concoct" >> $logfile

#    cd ${outdir}/${ID}

#    mkdir 07-Binning_contigs
```

```
ln -f -s ${outdir}/${ID}/06-Blast_oshv/${ID}_oshv_contigs.fasta \

        ${outdir}/${ID}/${ID}_oshv_contigs.fasta

    source deactivate

    cd ${outdir}/${ID}

else

    echo -e "\nSkipped Blast ${ID}_oshv_contigs.fasta with evalue=${evalue}" >>
${outdir}/${ID}/$logfile

fi



# Polishing assembly 01

if [ ! -d "${outdir}/${ID}/09-Pilon_polishing/" ] ;

then

    announcement "Polishing contigs N°01" >> $logfile

    cd ${outdir}/${ID}

    if [ ! -d 09-Pilon_polishing ]; then

        mkdir 09-Pilon_polishing

    fi

    cd 09-Pilon_polishing

    if [ ! -f "${outdir}/${ID}/09-Pilon_polishing/ReAligneContigs.bam" ] ;

    then

        # Aligning reads with de novo assembly genome of Oshv1

        ln -f -s ${outdir}/${ID}/07-
Extend_contigs/${ID}_oshv_contigs_extend/${ID}_oshv_contigs_extend.final.scaffolds.fasta . # modif -
cap3 pb

        # modif - cap3 pb

        Alignment_bowtie "${ID}_oshv_contigs_extend.final.scaffolds.fasta" \

            "${outdir}/${ID}/${readsTrim1}" \

            "${outdir}/${ID}/${readsTrim2}" \

            "Polishing_01.bam"

        Polishing_01=Polishing_01.bam

        # Aligning reads with de novo assembly genome of Oshv1  # modif arg1 - cap3 pb

        Polishing_assembly "${ID}_oshv_contigs_extend.final.scaffolds.fasta" \

            $Polishing_01 \

            "Polishing_assembly_01"

        ln -f -s ${outdir}/${ID}/09-
```

Annexe 4. **Pipeline d'assemblage utilisé pour l'article II**.
Le code présenté ici est un mixe des pipeline DiVir et Metaviromics. Ce pipeline est globalement similaire à l'outils Maeva qui est train d'être développé par le SeBimer

Paramètres	**Description**
Temporelle	Différence entre les dates d'échantillonnages
Spatiale	Différente localisation
Nombre de génome	La quantité de génomes utilisés lors de la comparaison
Matériel biologique	Séquençage en individuel ou non
Echantillon	Le type de séquence utilisé : ADN ou ARN
Analyse bioinformatique	Assemblage *de novo*, appel de variants

Annexe 5 **Récapitulatif des biais à prendre en compte dans l'analyse et la comparaison de la diversité génétique**.

Remerciements

Tout au long de ces années de thèse, et de toutes les histoires qui l'ont précédé, j'ai été encouragé par bon nombre de cœurs bienveillants. Il serait négligent de ne pas le faire savoir, alors voilà.

Tout d'abord, je souhaite remercier **Elisabeth Herniou** et **Maryline Houssin**. Merci d'avoir accepté d'être les rapporteurs de ma thèse et de prendre le temps d'évaluer mon travail. Je voudrais également adresser mes sincères remerciements à **Mathieu Sicard**, à **Stéphane Blanc** à **Germain Chevignon** et enfin à **Jean-Baptiste Lamy**. Je souhaite exprimer toute ma reconnaissance aux membres de mes comités de suivis de thèse: à **Christelle Desnue**, à **Benjamin Morga**, et à **Anne-Sophie Gosselin-Grenet** pour avoir pris le temps de me conseiller dans la construction de ce projet de thèse.

Je remercie **Guillaume Mitta**, directeur du Laboratoire *Interactions Hôtes Pathogènes Environnements* (IHPE) ainsi que **Delphine Destoumieux-Garzon**, responsable du site de Montpellier pour m'avoir accueilli au sein du laboratoire. Je remercie le directeur de cette thèse **Jean-Michel Escoubas**, avec qui la communication n'a pas toujours été simple, mais qui a été présent jusqu'à la fin. Un grand merci également à la co-encadrante de cette thèse Caroline Montagnani, pour s'être impliquée dans mon travail.

Un grand merci tout particulier à **Jean-Christophe** et **Richard** pour leur bienveillance, vous avez toujours su trouver les mots justes pour m'aider et je vous en suis infiniment reconnaissant. Muchas gracias à **Sarahi**, Somos el equipo de Rudo Team y permaneceremos en el egotrip de la ciencia.

Je n'oublie pas non plus toute l'équipe du laboratoire IHPE. Merci à tous les membres du laboratoire de Perpignan avec lequels j'ai eu l'opportunité de travailler, merci à **Christophe**, **Benjamin**, **Céline**, **Hélène**, **Arnaud**, **Eve**, **Jérôme**, **Ingo**, **Olivier**. Je remercie chaleureusement également tous les membres du site de Montpellier **Emmanuel**, **Jérémie**, **Julien**, **Yannick**, **Guillaume**. Merci également à **Juliette** et **Agnès** ! Merci à **Marco,** c'est toujours un plaisir de blaguer avec toi sur les dernières séries qui sortent. Je salue aussi le travail de **Anne**, **Diane** et **Sylvie** qui ont toujours été là pour m'aider à m'orienter dans la jungle administrative !

Un énorme merci à mes camarades de Perpignan **Stephen**, **Boris**, **Manon**, **Camille** vous êtes des gens géniaux et je vous souhaite pleins de réussite, ce fut un réel plaisir de passer du temps à vos côtés lors de mes séjours à Perpi ! Merci infiniment à mes frères et sœurs de batailles durant cette thèse ! Merci à **Erika**, **Noémie**, **Daniel**, d'avoir partagé des apéros à l'époque où les bars existaient. Merci **Aurélie** pour nos échanges, j'espère t'avoir transmis un peu de ma passion pour R, à **Maurine** (dit Momorito !!) j'ai appris trop de trucs en discutant d'évolution avec toi, ça m'a un peu trop plu alors j'en ai fait un chapitre entier dans la thèse. Merci à **Janan**, ton soutien et ta bonne humeur sont une grande force, vivement qu'on puisse à nouveau sortir pour aller prendre des bières et acquérir pleins de nouvelles identités ! Un grand merci à **Etienne**, les pauses clopes/café me manquent, j'espère que tu t'épanouis dans ta nouvelle vie à Bruxelles. Un grand merci à **Pierre-louis**, tu étais pendant longtemps la seule personne à comprendre mon jargon ce fut un plaisir de coder ensemble! À **Elyne** qui est depuis peu la jeune ambassadrice de l'Atlantique, ce fut un réel plaisir de partager du temps ensemble, je te souhaite plein de succès pour cette nouvelle aventure. Aux cinq jeunes thésards, à **Erwan** (dit le jeune), toujours à fond

sur la fitness et **Angélique** (dit la jeune) je suis toujours choqué de tes questions pertinentes en conf ! À **Valérie** je te souhaite pleins de succès dans ta thèse. Un grand merci à **Hajar** pour son caractère incroyable qui illuminait les journées du laboratoire, je te souhaite plein de réussite dans ta thèse. À **Sandra** avec qui j'ai bien tapé des barres, je te souhaite du succès pour la suite !

J'ai également eu la chance de rencontrer des gens incroyables au cours de ces 3 années. Je tiens donc à les remercier pour leurs soutiens et leur bonne humeur. Merci à **Dominique,** le gardien de mes nuits. À **Antonin** qui m'aide à décompresser. Merci infiniment à la **communauté R**, je suis venu pour le code et je suis resté pour cet état d'esprit bienveillant et incroyable. Quel plaisir de venir au meet-up, merci à **Criscely** pour sa bonne humeur et tous nos moments partagés. Un merci infinie (et plus) à **Cécile** pour m'avoir tant apporté et soutenue j'ai tellement appris grâce à toi ! J'ai pris énormément confiance en moi grâce à nos travaux sur le "super" projet BibliographeR. Je comprends quand tu disais que ce qui t'as fait tenir c'est tous les petits projets sous-jacent à la thèse. Merci aussi à ta petite **Lexie**, maintenant j'ai plus peur des gosses. On a vécu l'aventure *UseR2019* tous les quatres et c'était incroyable !!!! Merci à mes amis de la glisse **Léo**, **Maoui** et **Kév** qui m'ont fait découvrir le *surf skate* ! Merci à mes amis **Marvin**, **Marc**, **Alex** et **Lucas** pour tous ces bons moments passés à jeter des dés. Merci infiniment à **Malou** et **Romain** d'être des personnes aussi formidables, grâce à vous j'ai mangé mes premières lacrimo ! Un grand merci à mes gars sûr **Mathious**, **GuiGui** (dit Chinois), **Iris**, **Léo**, **Rudy**, **Dadidou**, **Jeremy**, **Erwanoux**, **Camélia** et **Kev** pour nos soirées. Merci également à **Hugo** d'avoir partagé des moments uniques ! Merci à mon ptit keum **FloFlo** pour les sorties annuelles (on est en 2021 va falloir en faire une) ! Merci aussi à **laMeufDeFIP** qui m'a accompagné avec des sons ouf!

Passons ensuite aux joyeux drilles, mes amis de la première heure, je ne vous ai pas assez vus durant ces 3 ans. Vous êtes des gens incroyables et vous m'avez toujours soutenu, je serais bref car je pourrais écrire l'équivalent de 1000x la BnF en éloge à votre égard, mais après vous allez prendre la confiance... Merci infiniment à **Marie** et **Luc** pour ces moments incroyables passés ensemble ! Merci à **Sarah TT77**, ma binôme de l'espace. Merci mon confrère **Cyril**, t'inquiètes le projet cancer avance ;) . Merci à **Papa Nico**, **Camille** et le petit **Gaspard** d'avoir toujours été là pour moi. Merci aux futurs maîtres du monde **Chaimoule** et **Bubar** d'être dans ma vie. Merci ma chère **Coline**, même si ça c'est ta faute tout ça, "tu verras STL ça va te plaire" moui moui voila où j'en suis maintenant.

Je voudrais remercier particulièrement deux femmes exceptionnelles, je n'oublierai jamais votre soutien indéfectible. La première, ma **Caro**, tu es mon frère de thèse, ensemble on a partagé bien des choses, des barres de rire, des joies, des pleurs, des folies, des défaites, des victoires… je n'aurais su avoir meilleur compagnon durant ces années difficiles. Sache que je serai toujours là pour toi frère ! Merci d'avoir été là pour moi et dans ma vie. La seconde ma **Tiphaine**, tu m'as fait rentrer dans ta vie, et j'espère qu'ensemble nous irons jusqu'aux bouts de nos rêves.

Pour finir j'aimerai remercier ma famille, mon **père**, ma **mère** et à ma **sœur** dont le soutien indéfectible m'a accompagné tout au long de ma vie. Merci également à **Edouard**, **Anaïs**, **Julia**, et mon p'tit filleul **Thomas** j'espère être bientôt à vos côtés. Merci à tous ces gens qui ont crue en moi !!

Index

adaptation à la variabilité antigénique, *15*

balayages sélectifs, *7*

captage, *28*

coévolution, *3*

coévolution antagoniste, *4*

coévolution diffuse, *4*

coévolution spécifique, *4*

course à l'armement, *5*

dérive génétique, *1*

déterminants de la virulence, *17*

espace de séquence, *16*

FREP, *14*

Generation Gap, *9*

guerre des tranchées, *5*

hôtes, *3*

hypermutation somatique, *14*

interactions hôte-pathogène, *2*

L'huître, *25*

micro-variants

µVar, *22*

mutabilité catastrophique, *16*

naissain, *26*

ostréiculture, *28*

Ostreid Herpèsvirus 1, *19*

parasites, *3*

pathogènes, *3*

polymorphisme de compatibilité, *14*

quasi-espèces virales, *15*

recombinaison V(D)J, *13*

sélection fréquence dépendant, *6*

sélection directionnelle, *7*

séquençage, *11*

signatures génomiques, *8*

stade juvénile, *25*

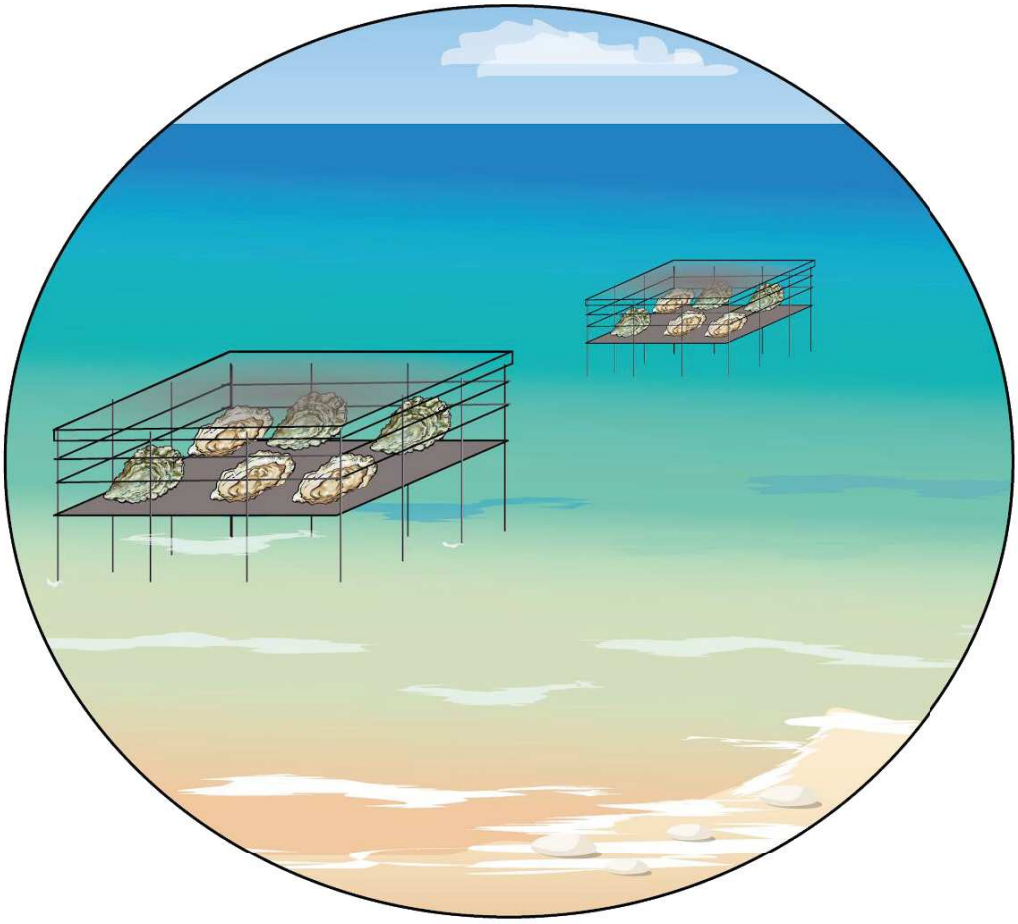
taux de mutation, *15*

théorie synthétique de l'évolution, *1*

triade épidémiologique, *31*

virus, *15*

virus à ADN, *17*



A la mémoire de Aaron Swartz…

Résumé

La diversité génétique contemporaine des virus est le résultat de l'interaction continue et dynamique des processus écologiques et évolutifs passés avec son hôte. Depuis plus de 50 ans des épisodes de mortalités impactent l'industrie ostréicole en France. Celles-ci ont des étiologies complexes, mais sont souvent associées à un Herpèsvirus : l'Ostreid Herpèsvirus 1 (OsHV-1). Depuis 2008 l'apparition d'événements de mortalité sans précédent chez l'espèce *Crassostrea gigas* ont été signalés à l'échelle mondiale. L'apparition de ces surmortalités coïncident avec l'émergence d'un nouveau génotype du virus : l'OsHV-1 µVar. Après plusieurs années de recherche pour comprendre ce pathosystème, il a été démontré récemment qu'un Herpèsvirus : l'Ostreid herpesvirus 1 (OsHV-1) était l'agent initiant la pathogenèse et aboutissant à la mort de l'hôte. En parallèle, l'observation d'autres variants apparentés du virus OsHV-1 ont été associés à d'autres épidémies de mortalité pour d'autres espèces de bivalves. Or jusqu'à récemment, les études sur la diversité génétique du virus utilisent un court fragment du génome du virus. Il est donc urgent de caractériser cette diversité génétique d'OsHV-1 à l'échelle du génome entier. Dans ce manuscrit de thèse nous apportons les outils nécessaire à la caractérisation complète du génome du virus. Nous confirmons que les populations virales d'OsHV-1 sont différentes selon leurs localisations géographiques avec de potentiels signatures génomiques entre les sites. Enfin, grâce à l'assemblage complet de vingt et un génomes nous proposons un schéma de dissémination du virus au travers d'analyses d'épidémiologie moléculaire à l'aide d'outils bioinformatique innovants. Nos résultats semblent indiquer que la diversité génétique du virus au sein de l'espèce *C. gigas* est relativement restreinte. Cette observation semble contraster avec les travaux antérieurs suggérant une diversité génétique importante du virus OsHV-1. Cependant, nous montrons qu'il est nécessaire de mieux définir la diversité au sein d'une espèce en particulier et dans un temps donné pour interpréter correctement celle-ci. Cette diversité n'est cependant pas inexistante et nous discutons quelle pourrait être son moteur. Ce travail de thèse est donc un point de départ notable pour les recherches futures caractérisant la diversité génétique pour appréhender ses conséquences biologiques. Ainsi il fournit le cadre théorique et méthodologique pour mieux comprendre les interactions hôtes pathogènes.

Abstract

The contemporary genetic diversity of viruses is the result of the continuous and dynamic interaction of past ecological and evolutionary processes with its host. For more than 30 years, mortality episodes have been impacting the oyster industry. These have complex etiologies but are often associated with a Herpesvirus: Ostreid Herpesvirus 1 (OsHV-1). Since 2008, unprecedented mortality events in *Crassostra gigas* species have been reported worldwide. The appearance of these excess mortalities coincides with the emergence of a new genotype of the virus: the OsHV-1 µVar. After several years of research to understand this pathosystem, it was recently demonstrated that a Herpesvirus: Ostreid herpesvirus 1 (OsHV-1) was the agent initiating pathogenesis and leading to host death. In parallel, the observation of other related variants of OsHV-1 virus have been associated with other mortality epidemics in other bivalve species. However, until recently, studies on the genetic diversity of the virus used a short fragment of the virus genome. There is therefore an urgent need to characterize this OsHV-1 genetic diversity at the whole genome level. This thesis manuscript provides us with the tools necessary for a complete characterization of the virus genome. We confirm that OsHV-1 viral populations are different according to their geographical location with potential genomic signatures between sites. Finally, thanks to the complete assembly of twenty-one genomes, we propose a scheme of virus dissemination through molecular epidemiology analysis using innovative bioinformatics tools. Our results suggest that the genetic diversity of the virus within the C. gigas species is relatively small. This observation seems to contrast with previous work suggesting a significant genetic diversity of the OsHV-1 virus. However, we show that it is necessary to better define the diversity within a particular species and within a given time frame in order to correctly interpret it. However, this diversity is not non-existent and we discuss what could be the driving force behind the observed diversity. This thesis work is thus a notable starting point for future research characterizing genetic diversity in order to understand its biological consequences. Thus it provides the theoretical and methodological framework for a better understanding of pathogenic host interactions.