



université de bretagne
occidentale

UNIVERSITE
BRETAGNE
LOIRE

THÈSE / UNIVERSITÉ DE BRETAGNE OCCIDENTALE

sous le sceau de l'Université Bretagne Loire

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Mention : Microbiologie

École Doctorale des Sciences de la Mer

présentée par

Coraline Mercier

Préparée à l'Institut Universitaire Européen
de la Mer, au sein du Laboratoire de
Microbiologie des Environnements Extrêmes

De nouveaux systèmes hôtes-virus associés aux sources hydrothermales océaniques profondes

Thèse soutenue le vendredi 16 décembre 2016
devant le jury composé de :

Tamara BASTA-Le BERRE

Maître de conférences, Université Paris-Sud 11 / *Rapporteur*

Télesphore SIME-NGANDO

Professeur, Université Blaise Pascal / *Rapporteur*

Christopher PAYAN

Professeur, Université de Bretagne Occidentale / *Examineur*

Mohamed JEBBAR

Professeur, Université de Bretagne Occidentale / *Directeur de
Thèse*

Claire GESLIN

Maître de conférences, Université de Bretagne Occidentale /
Encadrante scientifique

Remerciements

Tout d'abord je tiens à remercier les membres du jury d'avoir accepté de participer à l'évaluation de ces travaux.

Merci à mon directeur de thèse, le Pr. Mohamed Jebbar de m'avoir accueilli au sein du Laboratoire de Microbiologie des Environnements Extrêmes, de m'avoir suivi durant ces trois années de thèse jusqu'au Japon. Merci de m'avoir permis de participer à Eutrophiles et de nous avoir aidé durant ce périple !

Un grand merci à Claire, mon encadrante scientifique, tes conseils, ton aide et ta gentillesse ont été extrêmement précieux durant ce stage de master et cette thèse. Tu m'as soutenu tout au long de ce chemin, et jusqu'à la fin, cela m'a permis de faire ce que j'ai voulu. Tu as toujours su faire passer ta passion pour la science comme enseignante et encadrante et c'est une des raisons pour lesquelles il est très agréable de travailler avec toi, alors pour tout ça, et tout le reste, merci beaucoup.

Merci aux Dr Tamara Basta-Le Berre et Pr Téphère Sime-Nga d'avoir accepté de prendre de leur temps pour être rapporteurs de ces travaux, et merci au Pr Christopher Payan pour avoir accepté d'être le directeur de ces travaux et contribué à enrichir la discussion autour de ceux-ci.

Merci aux collaborateurs avec qui j'ai pu travailler durant ces trois années de thèse.

Notamment l'équipe du Dr Caillaud-Nesou qui m'a accueilli du début jusqu'à la fin, en particulier au Dr Thoas Havekapp qui m'a beaucoup appris durant ce séjour.

Merci aussi au Dr Anne Claire Baudoux, pour toutes nos journées en cytométrie, merci pour ton accueil au sein du laboratoire, ta gentillesse, ta disponibilité et ton humour, c'est toujours un plaisir de travailler avec toi. Merci aussi pour tes conseils scientifiques tout au long de cette thèse, ta participation à mon projet de thèse et aux discussions de l'atmosphère de caractérisation.

Je tiens également à remercier nos collègues de l'ANSES, Joëlle Cao, Thibault Moiré, L'adrien Louboutin et Laurent Bigarre, qui ont toujours été disponible pour nous sauver lors de nos petits soucis techniques.

Merci aussi à l'équipe du Pr Patrick Fotea, Dr Aurélien Golas, Eveline Maguet, Dr Sukhvinder Madelaine, pour les réunions toujours très enrichissantes et les conseils scientifiques apportés durant ma thèse.

Je tiens également à remercier les membres de mon comité de thèse, Dr Anne Claire Baudoux, Pr Gwenaëlle Le Blay-Laliberté, Dr Claire Le Henaff-Le Marrec pour leurs conseils très enrichissants.

Merci aussi aux Dr Gérard Siquin et Philippe Elies de la plateforme d'analyse et de données de l'UBO pour avoir permis de réaliser ces analyses et toujours avec une bonne humeur.

Un grand merci à tous les membres du laboratoire avec qui j'ai pu travailler.

Nadège Bienvenue Quid pour l'aide et les conseils apportés à la paillasse. Merci aussi pour ces séances de sports avec notre super coach !!

Merci au Dr Samuel Dupont, pour ta bonne humeur, tes blagues et tes coups de mains durant les manip. Ton aide et tes bonnes idées ont été très précieuses durant nos années communes dans la viroteam.

Merci aussi à Valérie Cueff-Gauchard pour ton aide précieuse lors des manip de qPCR.

Un grand merci à mes stagiaires que j'ai eus pendant mon stage : Julia, je suis sûre que ton parcours se poursuivra comme tu le souhaites, je suis certaine que ta gentillesse et ta détermination te mèneront loin ! Lisa, tu vas faire une super thèse au LUBEM et de supers légumes bios dans ton jardin, merci pour ton aide et ton amitié. Et enfin Sarah, merci beaucoup pour ton dévouement, et patience sur les délais, tu t'es donnée à fond et je suis ravie que tu reprennes le flambeau de la viroteam, il ne pouvait pas tomber entre de meilleures mains que les tiennes.

Je tiens aussi à remercier les membres des équipes du LM2E et l'attache sur les équipes de la composante IUEM.

Merci aux enseignants chercheurs, Pr Gwenaëlle Le Blay, Dr Odile Vandenberghe-Trambouze, Dr Karine Alai pour leur disponibilité et l'aide que vous m'avez apportée. Un merci particulier au Dr Frédérique Duthoit, pour la relecture de ce manuscrit. Merci aussi à tous les chercheurs de la composante IFREMER, pour les réunions et discussions toujours passionnantes.

Je tiens aussi à remercier tous les ITA pour leur travail qu'ils font pour le laboratoire, en particulier ceux que j'ai rencontrés : Stéphane gère des pots du labo, Mia biologiste de l'équipe, Alex « Mr Bricolage », merci pour ces moments à papoter autour d'un café et Morgan, tes soutiens et ton énergie nous ont aidés à faire des pas... Un grand merci à Stéphane, le meilleur et le plus sympa de l'univers, toujours disponible, du début à la fin, merci pour ton amitié et ton soutien infatigable.

Merci à mes amis du labo, Julie, Joana, Gégé, Fred... Vous faites une sacrée équipe, votre départ a laissé un vide. Une patience à Julie, tu as beaucoup appris durant mon stage, tu as toujours été de bon conseil et tu as aussi toujours répondu présent aux appels des pots et divers extras ! Merci pour ta bonne humeur, tes conseils et ta bonté.

Merci à mon docteur avec qui j'ai partagé beaucoup d'eau, ou des heures de thèses pour avoir fait sourire tous les jours. Matt et sa dream team girl power, Gaëlle, Sandrine et Aline. Vous avez toujours le sourire, des délires pleins la tête et les moments passés avec vous sont toujours très agréables. Merci Matt pour ton oreille attentive et tes blagues (comique de répétition ...) qui mettaient du fun dans le bureau.

Merci à mes deux camarades de même année, Simon et Cess. Merci Simon pour le fun durant le master et ta grande gentillesse de tous les instants. Merci à toi Cess, nous étions dans le même bateau, tu sais toujours trouver les bons mots au bon moment (UN HOTE o ali e UNE HOTTE est pou le PERE NOEL). Merci pour ton honnêteté et surtout, ton amitié.

Un grand merci aussi aux anciens stagiaires maintenant doctorants, en particulier ma petite Tiph et ton large sourire, tes éclats de rire vont me manquer tout comme nos discussions. Merci aussi à la team bio info, Damien, Flo et Clarisse, toujours prêts à aider, à se marrer ou à manger un burger, merci pour votre bonne humeur et votre gentillesse. Merci à Gwen, toujours haute en couleur, tu as mis de la gaieté dans ce bureau. Merci aux docs IFREMER, Mélanie, Caro, Charlène, Yang, Seb et Vincent, les moments partagés avec vous étaient toujours agréables.

Merci aussi aux Dr Yann Moalic et Axel Thiel, en particulier Yann pour ton humour décalé, tes pulls aux couleurs aléatoires, ta bonne humeur et tes conseils.

Un grand merci à mes amis qui ont rendu ma vie à Brest particulièrement agréable.

Edwige, ma chouette, nous avons commencé ensemble, depuis nos premières années et nous sommes toujours là, nos balades à cheval, nos petits cafés et nos discussions interminables (mais quelles pipelettes ces deux là 'o t appo t eau oup. Mé pour cette belle amitié que nous partageons. Bisous à Manu aussi, qui supporte tout ça depuis toutes ces années, et à Laura qui pour une folle raison est repartie en médecine, courage à toi, je suis certaine que tu vas y arriver !

Me i à es a ie s oll gues, ai te a t ais, Bi ho , e à l'aut eout du o de tu me fais toujours rire, Kevin, tes blagues aléatoires vont me manquer, Rocio, on y est au bout de ette th sela p eue!).

Merci à tous les Diskuizh, ça va être difficile de retrouver un club avec autant de bonne humeur et de fun que celui-ci. Mention spéciale à Val (équipe Une Mr !!! 'est toujou g ial de dis ute s ies, histo et tout leeste ave toi. J'e p ofite pouefai petit coucou à Juliana, je suis certaine que tu as un bel avenir devant toi.

Evidemment, merci à toi Julien, JuC, tu as été et tu restes une pierre fondamentale de ma ussite, sas to aide et to soutiefaillile je 'e se ais pas là aujou d'hui. Me tous les sacrifices que tu as aussi dû concéder durant cette fin de thèse, les corrections, les jobs ingrats et tout leeste. je te souhaite le meilleur pour la suite, je sais que tu y arriveras.

Pou te i e , u g a d e i à es pae ts, vous'avez toujou s soutue'est g âe à vous si j'ai pue fai ue j'ais et avai s au tie ue je fais aujou d'hui. Me aussi à tous mes frères et sœu s et leu s failles, e partie à toi Ol pe, plus u'u e g ade sœu tu as toujou s là pou oi et je saie uoi u'il a ive, je pois compter sur toi.

Table des matières

Remerciements

Table des matières

Liste des figures

Liste des tableaux

Liste des abréviations

Introduction générale 1

I. Les virus 1

- Découverte et Définition 1
- Découvertes récentes 4
- Classification 10
- Les cycles viraux 12
- Course à l'armement : système de défense hôtes – virus 12

II. Les virus de bactéries et d'archées dans les océans 12

- Abondance et diversité 12
- Le rôle des virus dans les océans 12

III. Les environnements thermophiles 5

- Les sources chaudes terrestres 5
- Les sources hydrothermales profondes 81

IV. Bibliographie – Introduction 9

Objectifs des travaux de thèse 9

Chapitre I :	Revue bibliographique « An abyssalmobilome »	
Chapitre II :	Ménage à trois, caractérisation de MPV1	
Chapitre III :	Deux nouveaux virus, MCV1 et MCV2, isolés bactéries du Thermotogae	149
Chapitre IV :	Etude du génome de Thermosipho sp. AT1244-VC14	
Chapitre V :	Virome : validation et normalisation d'un protocole	
Discussion et Perspectives		211
I. Deux nouveaux virus, MCV1 et MCV2, isolés du phylum des Thermotogae et c		
MPV1		214
II. Thermotogae et CRISPR		217
III. Importance de la lysogénie et virome		
IV. Vésicules membranaires et détoxification		2
V. Bibliographie		229

Résumé

Liste des figures

Introduction générale	1
<u>Figure 1</u> : Photographie au microscope électronique à transmission du virus de la mosaïque du tabac	2
<u>Figure 2</u> : Photographie au microscope électronique à transmission d'un virus de Mammiviridae	6
<u>Figure 3</u> : Différents aspects morphologiques du Mammavirus et du virophage Sputnik	7
<u>Figure 4</u> : Comparaison de l'aphélogénétique du vivant de Woese et de la proposition d'une nouvelle classification du monde du vivant par Raoult et Fèrrière.....	9
<u>Figure 5</u> : Schématisation du cycle d'un virus tempéré voie lytique et voie lysogénique.	13
<u>Figure 6</u> : Mécanisme de défense microbienne contre les éléments génétiques mobiles	16
<u>Figure 7</u> : Schématisation du mode de fonctionnement du système CRISPR-cas	17
<u>Figure 8</u> : Illustration du fonctionnement du système R-M de type I	19
<u>Figure 9</u> : Illustration du mode de fonctionnement potentiel du BREX	22
<u>Figure 10</u> : Exemple de stratégie virale pour échapper au système toxine antitoxine.....	26
<u>Figure 11</u> : Relation entre l'abondance virale et l'abondance des micro-organismes (bactéries et archées) dans les sédiments marins profonds de différents océans.	29
<u>Figure 12</u> : Photographie au MET de GTA	32

<u>Figure 13</u> : Différents rôles des vésicules membranaires	33
<u>Figure 14</u> : Observation en MET de vésicules membranaires produites par des Thermococcales	35
<u>Figure 15</u> : Un mobilome abyssal	36
<u>Figure 16</u> : Distribution globale de la diversité virale	39
<u>Figure 17</u> : Illustration humoristique des divers impacts des virus sur les communautés microbiennes	41
<u>Figure 18</u> : Illustration de la stratégie Killing the winner	43
<u>Figure 19</u> : Illustration du viral shunt	46
<u>Figure 20</u> : Photographie de <i>Emiliana huxleyi</i>	47
<u>Figure 21</u> : Rôle des virus dans l'évolution et la pathogénicité de leurs hôtes.....	48
<u>Figure 22</u> : Structures similaires à des queues de caudovirus	51
<u>Figure 23</u> : Illustration de l'utilisation des « virus » de la photosynthèse pour supporter la production valériole d'efficacité de <i>Prochlorococcus</i>	53
<u>Figure 24</u> : (a) Désert du Sahara (b) Antarctique	54
<u>Figure 25</u> : Photographie de geysers et source chaude	56
<u>Figure 26</u> : Diversité des communautés microbiennes retrouvée dans des enrichissements aléatoires à partir d'échantillons de sources chaudes terrestres.....	59
<u>Figure 27</u> : Photographie de MET de bactériovirus isolés de sources chaudes terrestres	63
<u>Figure 28</u> : Alignement de génomes de Fusellovirus	76
<u>Figure 29</u> : Particules de SSV1 analysées en cryo-EM	77

<u>Figure 30</u> : S h a de l'o ga isatio des sois de PFV.....	78
<u>Figure 31</u> : Structures pyramidales	80
<u>Figure 32</u> : Sources hydrothermales profondes sites	81
<u>Figure 33</u> : S h a du fo tio e e t d'u e soue h d othe ale p ofo de.....	83
<u>Figure 34</u> : Rep se tatio de l'a o dae elative des ph la t ie s	85
<u>Figure 35</u> : Rep se tatio de l'a o dae elative des diff ts ph larchéens	86
<u>Figure 36</u> : Pla e e t des The otogales da s l'æ ph log ti ue a t ie pa appo t à l'ARN 6S.....	87
<u>Figure 37</u> : Observation au microscope électronique de virions	92
<u>Figure 38</u> : Photographie en microscopie électronique de D6E	93
<u>Figure 39</u> : Observation en MET de (a) PAV1 et (b) TPV1	94
<u>Figure 40</u> : Relation entre les plasmides de Thermococcales et le virus PAV1	95
<u>Chapitre I</u> : Revue bibliographique « An abyssal mobilome »	111
<u>Figure 1</u> : Electron micrographs of viruses and membrane vesicles isolated from deep-sea hydrothermal vents	116
<u>Figure 2</u> : Multipurpose mobilome within deep-sea hydrothermal vents.	116
<u>Chapitre II</u> : Ménage à trois, caractérisation de MPV1	124
<u>Figure 1</u> : Electron micrograph of two MPV1 particles	128
<u>Figure 2</u> : Virus–host kinetics at atmospheric pressure monitored by flow cytometry	129
<u>Figure 3</u> : Comparison of the MPV1 provirus in M. piezophila and viral DNA.....	129

<u>Figure 4</u> : Comparison of the MPV1 provirus in <i>M. piezophila</i> and viral DNA.....	130
<u>Figure 5</u> : MPV1 infectivity assays performed on <i>Thermosipho</i> sp. AT1244-VC14T and checked by PCR experiments.	130
<u>Figure 6</u> : qPCR assays to assess changes in copy number of <i>M. piezophila</i> chromosomal, proviral and plasmid DNA target genes in the presence or absence of mitomycin C induction	131
<u>Figure S1</u> : Induction assays of MPV1 production by addition of several mitomycin C concentrations	141
<u>Figure S2</u> : RFLP analyses of <i>M. piezophila</i> total DNA, plasmid DNA (pMP1) and viral DNA ...	142
<u>Figure S3</u> : PCR analysis of the total DNA extracted from purified MPV1 viral capsids from an induced culture	143
<u>Figure S4</u> Agarose gel electrophoresis of plasmid DNA extracted by alkaline lysis from <i>Thermosipho</i> sp. strain AT1244-VC14, three subcultures after being infected by MPV1 virions	144
<u>Figure S5</u> : QPCR assays	145
Chapitre III: Caractérisation de deux nouveaux virus, MCV1 et MCV2, isolés de l'océan des Thermotogales	155
<u>Figure 1</u> : Electron micrographs of (a) a MCV1 virion (b) two MCV2 virions	159
<u>Figure 2</u> : Virus-host kinetics of <i>M. camini</i> 97 and <i>M. camini</i> 55	161
<u>Figure 3</u> : Electron micrographs of concentrated MVs produced by <i>M. camini</i> 97	162
<u>Figure 4</u> : Genome comparison of MPV1, MCV1 and MCV2 with GenoPlotR	167

Chapitre V : Virome : validation et optimisation du protocole de purification	203
---	-----

<u>Figure 1</u> : Schéma général du protocole « Filtration » (A) et du protocole « Gradient » (B) ..	207
--	-----

Discussion et Perspectives	211
----------------------------------	-----

<u>Figure 1</u> : Comparaison de 11 génomes de Thermotoga maritima-like.	213
---	-----

<u>Figure 2</u> : Schéma du pipeline de photométrie et génomique d'un système CRISPRas	220
--	-----

<u>Figure 3</u> : Observation en microscopie électronique de MVs produites par M.camini 97	226
--	-----

Liste des tableaux

Introduction générale	1
<u>Tableau 1</u> : Différences entre organismes, virus et constituants cellulaires	3
<u>Tableau 2</u> : Classification selon Baltimore	10
<u>Tableau 3</u> : Bactérovirus des sources chaudes terrestres	61
<u>Tableau 4</u> : Archéovirus des sources chaudes terrestres.	70
Revue bibliographique « An abyssal mobilome »	111
<u>Tableau 1</u> : Bacterial and archaeal viruses isolated from deep-sea hydrothermal vents	114
<u>Tableau 2</u> : Classification selon Baltimore	117
<u>Tableau 3</u> : Archaeal plasmids isolated from deep-sea hydrothermal vents	61
Chapitre IV : Etude du génome de <i>Thermosipho</i> sp. AT1244-VC14	194
<u>Tableau 1</u> : Summary of important genes implicated in the CRISPR-cas system of <i>Thermosipho</i> sp. 1244-VC14	200
Chapitre V : Etude du génome de <i>Thermosipho</i> sp. AT1244-VC14	203
<u>Tableau 1</u> : Concentrations en microorganismes cibles avec lesquelles les échantillons ont été chargés	206

Liste des abréviations

AAF	Aéro-Anaérobie Facultatif
Abi	Abortive infection
ADN	Acide DésoxyriboNucléique
AN	Acide Nucléique
ARN	Acide RiboNucléique
ARNm	Acide RiboNucléique messenger
ARNnc	ARN non codant (ou antisens)
ARNr	Acide RiboNucléique ribosomique
ARNt	Acide RiboNucléique de transfert
ATP	Adénosine TriPhosphate
BREX	BacteRiophage EXclusion
CRISPR-cas	Clustered Regularly Interspaced Short Palindromic Repeats -crispr associated genes
DMS	DiMethyl Sulfur
GTAs	Gene Transfer Agents
GTP	Guanosine TriPhosphate
HGT	Horizontal Gene Transfer
ICTV	International Committee on Taxonomy of Viruses
KtW	Killing the Winner
LUCA	Last Universal Common Ancestor
MEB	Microscope électronique à Balayage
MEF	Microscopie à EpiFluorescence
MET	Microscope électronique à transmission
MOD	Matière Organique Dissoute
MOP	Matière Organique Particulaire
Mtase	MethylTransferase
MVs	Membrane Vesicles
ORFs	Open Reading Frames
pb	paire de bases
Pgl	Phage growth limitation
PICIs	Phage Inducible Chromosomal Islands
PNK	PolyNucleotide Kinase
R-M	Restriction Modification
Rease	Restriction Endonuclease
TA	Toxine Antitoxine
UV	UltraViolet
vAMGs	viruses-encoded Auxiliary Metabolic Genes
VAPs	Virus Associated Pyramids
VLPs	Virus-Like Particles

Introduction générale

I. Les virus

- Découverte et Définition :

Un peu d'histoire

Les virus sont des entités biologiques particulièrement anciennes et, selon certains, ils seraient apparus avant LUCA (Last Universal Common Ancestor) (Abreu Forterre, 2013a, 2016). Les infections dues à des virus telles que la rage existent depuis des millénaires (Lwoff, 1957). Néanmoins les agents responsables de celles-ci ne sont décrits qu'à la fin du XIXème siècle. En effet, même si des traitements de variolisation étaient pratiqués depuis plusieurs siècles, la compréhension et la culture de virus commencent réellement avec l'étude du virus de la mosaïque du tabac. Mayer décrit cette maladie affectant les plants de tabac comme infectieuse, mais une bactérie en est responsable. Par la suite, Dimitri Ivanovski va montrer qu'un agent est capable de franchir les filtres de Chamberland, permettant à l'époque de distinguer des bactéries. Ivanovski maintient pourtant l'explication bactérienne et ce n'est que Beijerinck qui comprendra l'importance de l'observation de son collègue et qui introduit le concept de « contagium vivum fluidum ».

Le virus de la mosaïque du tabac restera un modèle d'étude jusqu'à être cristallisé par Wendell Stanley en 1935. L'année suivante Frederick Bawdon et Rosalind Pirie décriront une structure composée de protéines et d'acides nucléiques de leur confrère. Au début du XXème siècle, Frederick Twort et Félix d'Herelle ont indépendamment l'un de l'autre les virus infectant des bactéries.

C'est Félix D'Hérelle qui, en 1921, les nommera « bactériophages » suite à ses travaux sur *Shigella dysenteriae* durant lesquels il vit apparaître des plages de lyses lors de ses cultures sur boîte de Pétri (Duckworth, 1976).

Il faudra attendre 1939 pour observer pour la première fois un virus en image, grâce à l'invention du microscope électronique à transmission (MET), permettant d'observer des particules grossies jusqu'à 20 000 fois (Klemperer, 1940). A nouveau, le premier virus observé sera celui de la mosaïque du tabac (figure 1).

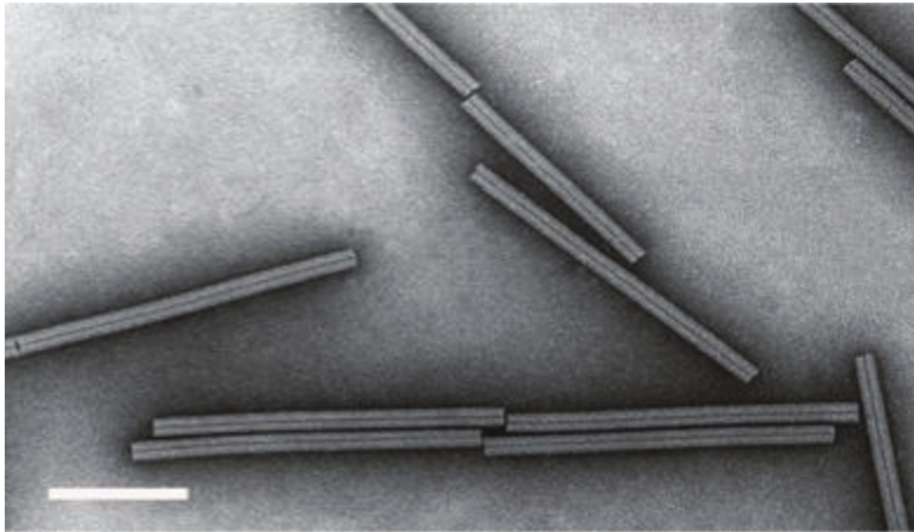


Figure 1 : Photographie au microscope électronique à transmission du virus de la mosaïque du tabac. Echelle 100nm. Source ICTV.

Définition : l'évolution

Les virus ont donc tout d'abord été connus comme un « fluide vivant contagieux » filtrable ayant la capacité de traverser les filtres de Chamberland. Cependant, le premier à émettre une définition complète est André Lwoff dans les années 50. A cette époque, la définition des virus était trouble puisque certains les considéraient comme étant des micro-organismes ou des molécules (Burnet, 1945; Bawden, 1951).

Lwoff propose une définition qui restera celle longtemps admise par la communauté

scientifique : les virus sont « des entités nucléo-protéiques infectieuses de strictement intracellulaires potentiellement pathogènes, elles possèdent un seul nucléique (ADN ou ARN), se multiplient à partir de leur matériel génétique, sont incapables de croissance ou de division et sont dépourvues de machinerie cellulaire de leur hôte afin de se multiplier. Ils possèdent

appelée cascade qui peut ou non être entourée par une enveloppe

only one type, DNA.

Table 3. *Organisms, viruses and cellular constituents*

	Organisms	Viruses	Cellular constituents	
			Genetic material	Organelles endowed with genetic continuity
Types of nucleic acid	2§	1	1	?
Multiplying as nucleic acid and produced from nucleic acid only	0	+	+	0
Growth and division*	+	0	0	+
Presence of Lipmann system	+	0	0	0 (+)†
Infectivity	+	+	0	0

* Concerns micro-organisms only. The replication of a structure by the template mechanism is neither considered as growth nor as division but as replication.

† Some organelles only contain a Lipmann system.

§ 2=DNA+RNA, 1=DNA or RNA.

(2) All typical micro-organisms are reproduced from the integrated sum of their constituent parts, nucleus, cytoplasm, cytoplasmic structures, cortex. Bacteriophage is produced or reproduced from its nucleic acid only. When considering a micro-organism, the *infectum* is the micro-organism itself; when

- Découvertes récentes

La définition de ce qu'est un virus reste néanmoins sujette à débat scientifique. En effet, l'absence de ribosome dans leur matériel génétique du monde du vivant tel qu'il a été décrit par Carl Woese en 1990 (Figure 1.1) certaines découvertes récentes remettent en question la façon dont les virus sont étudiés et classifiés.

Tout d'abord, l'étude de l'abondance virale grâce à des techniques moléculaires (métagénomiques) ou de microscopie à épifluorescence a permis de mettre en évidence que les virus sont les entités biologiques les plus abondantes de la planète. Il est en moyenne, dix fois plus de virus que de cellules dans l'environnement. Cette abondance a longtemps été sous-estimée en raison des limitations techniques de ce qui explique en partie le fait que leur importance ait été négligée.

De plus, les virus ont longtemps été considérés comme des « pickpockets » des cellules, mais il s'avère que les cellules intègrent et utilisent aussi des virus (Cortez et al., 2010, 2013a). En effet, 10 à 15% des génomes de bactéries ou d'eucaryotes contiennent des gènes provenant de virus ou de plasmides ayant été récemment intégrés (Cortez et al., 2009). Ce constat implique un fort impact des virus sur la composition, notamment des communautés microbiennes. Ils peuvent en effet intégrer de nouvelles fonctions, résistances ou propriétés à ces cellules (Weinberg et al., 2009).

La découverte de virus géants a aussi remis en question la façon dont les virus. Historiquement, la petite taille est un critère de sélection des séparer le compartiment bactérien/archéen du compartiment viral, l'utilisation de 0.22 μm est une méthode classique étant donné qu'il était considéré que généralement au travers de cette porosité. Néanmoins, la découverte des appelés « virus géants » qui font la taille de petites bactéries a de ces fondamentaux de la virologie. Ces virus, infectant une amibe de la classe des rhizopodes) de l'espèce *Acanthamoeba polyphaga*, ont confondus avec de petites cellules à Gram négatif (La Scola et al., 2003) les premiers virus observables au microscope optique puisque leur capsid (La Scola et al., 2003). Leur génome de 1,2 Mpb est deux fois plus grand bactéries et code pour plus de 1000 protéines (La Scola et al., 2003; ailleurs, lorsque le virus se réplique à l'intérieur de l'amibe, est aussi grosse que le noyau de l'amibe elle-même. Il peut coder pour enzymatiques précédemment retrouvés uniquement chez les cellules (comme impliqués dans la synthèse protéique) et contient de l'ARN et de la lumière de ces nouvelles données, les approches méthodologiques de nouveaux virus ont évolué et permirent de découvrir de nouveaux d'entre eux sont des virus infectants des amibes (Fischer, 2016). Ils sont en sept familles (*Ascoviridae*, *Asfarviridae*, *Iridoviridae*, *Marseilleviridae*, *Mimiviridae*, *Phycodnaviridae*, and *Poxviridae*) regroupées en un ordre nommé « *Megavirales* » (La Scola et al., 2013). Cet ordre contient des virus ayant des génomes allant d'une centaine de bases pour les plus petits jusqu'à plus de 2,5 Mpb pour les plus grands (Pandey et al., 2016).

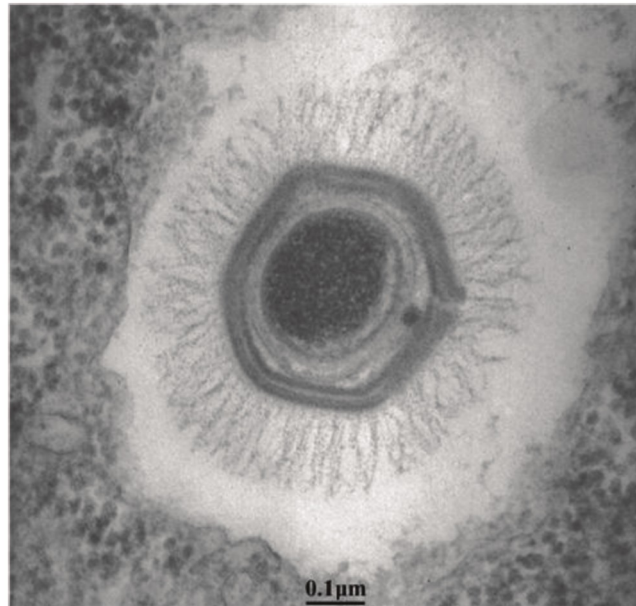


Figure 2 : Photographie au microscope électronique à transmission d'un virion de Mimivirus. Source ICTV.

L'étude de ces virus impressionnants tant par la taille de leurs capsides que par la taille de leurs génomes a aussi permis de mettre en lumière des virus de virus (aussi appelés virophages). C'est ainsi qu'a été caractérisé Sputnik, un virus infectant un Mimivirus qui infecte lui-même une amibe (La Scola *et al.*, 2008). D'autres virophages ont été mis en évidence après Sputnik, comme Mavirus, qui parasite le *Cafeteria roenbergensis* virus (Fischer and Suttle, 2011), ou encore Zamilon décrit en 2014 (Gaia *et al.*, 2014). Ils contiennent des génomes à ADN double brin linéaires ou circulaires de 15 à 30 kb qui sont portés par une capside icosaédrique de 50 à 80 nm de diamètre (Fischer, 2016).

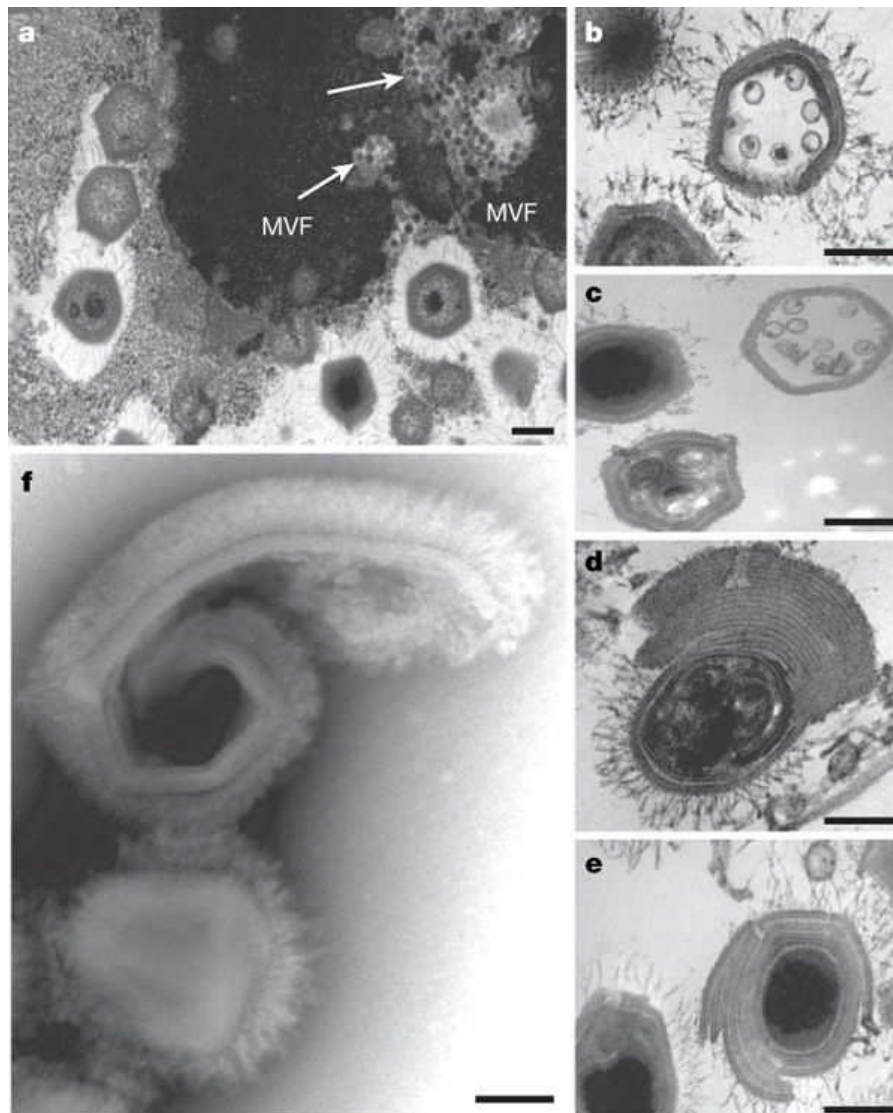


Figure 3 : Différents aspects morphologiques du Mamavirus et du virophage (f) (2008). (a-e) Observations réalisées au microscope électronique à transmission (a) « Mamavirus virus factory » (b) Sputnik dans des capsides de Mamavirus. (c) Production de petits virus de Mamavirus avec des morphologies anormales dues à l'infection.

L'ensemble de ces découvertes remet donc fondamentalement en question la définition de ce qu'est un virus mais aussi la façon dont ils sont considérés et classifiés. L'élaboration de nouveaux concepts tels que celui de la « Virocell » a été proposée par le Pr Patrick Forterre dans les années 2010, suite aux nombreuses découvertes en question la définition des virus (Forterre, 2011, 2013b).

L'une des bases de cette théorie repose sur le fait que, dans la commens large, les virus et les virions sont souvent confondus, les virions étant extracellulaire du virus (capside et acide nucléique). Par exemple, dans la il est écrit que celui-ci ne contient qu'un seul type d'acide nuclé réplique il passe par une étape d'ARN messenger et certains virus comm capables de passer de l'ARN à l'ADN (puis de nouveau à l'ARN messa

Or, d'autres auteurs, à l'instar du Pr Jean Michel Claverie séquencé le génome du premier virus géant (Raoult et al., 2004)), pense virus et les virions revient à confondre l'homme et ses spermatozoïdes. alors une théorie selon laquelle le virus serait l'usine virale lors de l'infec cette usine virale n'est pas présente partout et est notamment absente exemple.

C'est ainsi qu'a été émis le concept de Virocell du Pr Patrick Forter cellulaire de la réplication du virus (ou organisme codant pour une capsid un processus biologique intégrant tous les aspects du cycle de repro virocell, et éventuellement génome intégré) (Forterre, 2011, 2016). Ce la proposition d'une nouvelle classification des organismes, intégrant ainsi les du vivant (par opposition à la proposition communément admise de Carl al., 1990; Woese, 1994; Raoult and Forterre, 2008) (Figure 4).

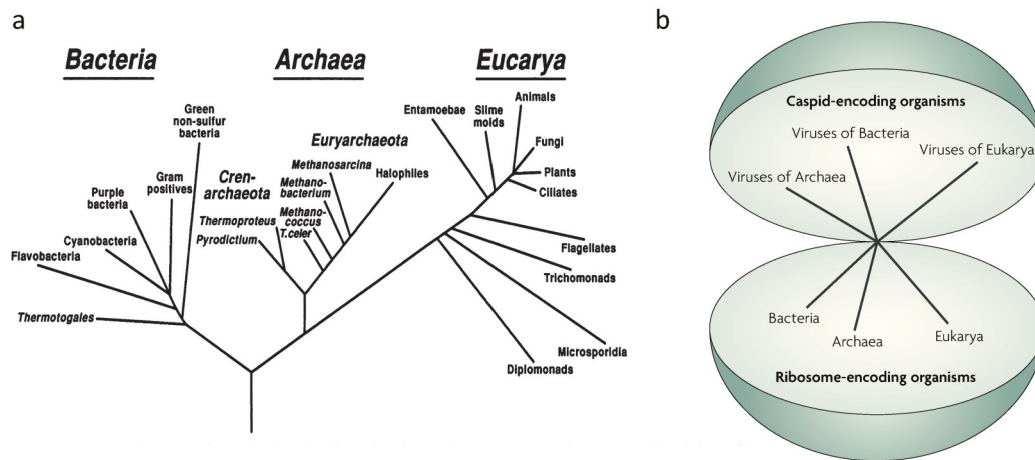


Figure 4 : Comparaison de l'arbre phylogénétique du vivant de Woese (Woese, 1994) et de la proposition d'une nouvelle classification du monde du vivant par Raoult et Forterre (Raoult and Forterre, 2008). (a) Arbre phylogénétique du vivant montrant les trois domaines Archaea, Bacteria et Eucarya. (b) Représentation de virus avec leurs capsides et des trois domaines du vivant ayant évolués à partir de LUCA. Les trois domaines ont des ribosomes mais n'ont pas de capside. Illustration de la nouvelle définition des virus basée sur le fait qu'ils aient des capsides mais pas de ribosome, l'incluant dans un nouveau type d'arbre du vivant.

Dans ce cas, le monde du vivant serait séparé en deux grands types d'organismes, ceux codant pour des ribosomes (Ribosome encoding organisms) et ceux codant pour des capsides (Caspid encoding organisms) (Raoult and Forterre, 2008; Forterre, 2013a, 2016). En effet, la capacité à coder pour des capsides ou des ribosomes permet de bien séparer les deux types d'organismes tout en rappelant leurs caractéristiques fondamentales (Abrescia *et al.*, 2012). Sur la base de cette séparation bipartite, les auteurs proposent une évolution de la définition d'un virus, plus cohérente avec les récentes découvertes, comme étant « un organisme codant pour une capside composée de protéines et d'acides nucléiques, assemblés en une nucléocapside et utilisant un organisme codant pour des ribosomes pour compléter son cycle de vie » (Raoult and Forterre, 2008). Il semble évident que la définition de Lwoff décrivant les virus comme de petits agents (moins de 0,2 μm) infectieux mais non autonomes et contenant

un seul type d'acide nucléique paraît obsolète. Néanmoins, l'intégration des virus du vivant reste controversé (López-García, 2012; López-García and Moreira, 2012).

D'autres modèles ont aussi été proposés et notamment celui d'un quatrième vivant, en plus des bactéries, archées et eucaryotes, nommé « Akamara » (López-García et al., 2011).

Celui-ci regrouperait les agents infectieux acellulaires possédant un génome d'acide nucléique (López-García et al., 2011). Ce nouveau domaine serait séparé en deux règnes, celui des Euryviriota (virus satellites (codant pour une capsid) et celui des Viroidia (composés de ARN codant pas pour leur capsid) (Hurst, 2011).

D'autres équipes ont néanmoins repris le concept de Virocell (López-García et al., 2016) en relevant que se placer du point de vue des virus permet de mieux comprendre leur implication et leur importance dans les écosystèmes, notamment océaniques (cf Partie « Rôle dans les océans » p41).

- **Classification**

Actuellement, il existe deux grands types de classifications.

Le premier est la classification par la méthode de Baltimore, définie en 1971. Celle-ci classe les virus en 7 groupes en fonction de leur type d'ARN, du nombre de brins (simple ou double brin), de la polarité et du mode de réplication (Tableau 2).

Réplication ↓	Classe						
	I	II	III	IV	V	VI	VII
	ADN double brin	ADN simple brin	ARN double brin	ARN simple brin sens positif	ARN simple brin sens négatif	ARN simple brin (avec rétro transcription)	ADN double brin (avec rétro transcription)
		ADN double brin		ARN simple brin sens négatif		ADN double brin	
	ARN messager						

Tableau 2 : Classification selon Baltimore. Les classes sont définies par rapport au type d'acide nucléique porté par le virion. La première ligne indique le numéro de la classe, la seconde le type d'AN porté dans la capsidie et les deux suivantes les états par lesquels passent les AN viraux lors de la réplication.

Ce tableau montre la classification selon Baltimore, la première ligne indique le type d'acide nucléique lorsque le génome viral est encapsidé sous forme de virion. Lors de sa réplication dans l'hôte, le virus est alors transcrit en ARN messager (ARNm) mais peut aussi, dans certains cas (ex. classe II), passer par un autre type d'acide nucléique ne correspondant ni à un ARNm ni à celui encapsidé. C'est par exemple le cas très connu du VIH. Cette séparation en 7 classes a l'avantage de répartir efficacement les virus, même s'il est reste complexe de faire une séparation pertinente entre les différents virus en se basant sur ces seuls critères.

Le second type de classification est celui de l'**ICTV** (International Committee on Taxonomy of Viruses) communément adopté par la majorité de la communauté scientifique. La définition d'une espèce virale selon l'ICTV est la suivante « classe polythétique de virus qui constitue une lignée répliquative et occupe une niche écologique particulière ». Les espèces virales forment donc des classes polythétiques c'est à dire des groupes dont les membres possèdent toujours plusieurs propriétés en commun sans qu'aucune propriété particulière ne soit obligatoirement présente chez tous les membres de l'espèce. La différenciation en espèces virales n'est pas aisée. Plus globalement, les virus sont classés selon la nature de leur acide

nucléique (ADN, ARN, simple ou double brin), la présence ou non d'une enveloppe, la capsid, le type de symétrie de celle-ci et la taille du virion. Cette classification périodiquement, comprend une base de données internationale qui recense la majorité des espèces de virus décrites à ce jour. A l'heure actuelle, les virus sont réparties en 7 ordres et 111 familles. Les 7 ordres sont Caudovirales, Herpesvirales, Mononegavirales, Nidovirales, Picornavirales, Tymovirales et le dernier créé les Ligamenvirales. Pour être acceptée, une nouvelle espèce de virus doit contenir au moins un virus isolé parmi ses membres.

Quasiment toutes les espèces de virus sont classées au sein d'un genre. Cependant, certaines restent non assignées car leur caractérisation est inadéquate. Ces classifications ne sont pas incompatibles puisque l'ICTV reprend en partie la classification de Baltimore tout en incluant notamment le type d'hôte infecté et la morphologie (Abrescia et al., 2012).

- Les cycles viraux

Les virus entretiennent trois principaux types de relations avec leurs hôtes (Weinberg, Madigan and Martinko, 2007) :

Les virus virulents réalisent un cycle lytique en détournant à leur profit les ressources de l'hôte et les virions néo synthétisés sont libérés suite à la lyse provoquée.

Les virus tempérés, qui effectuent un cycle lysogénique, persistent sous forme de virus nucléiques, le plus souvent intégrés au sein du génome de l'hôte (ou sous forme extra-chromosomique), à l'état de provirus. Ils se répliquent avec le génome de la cellule hôte jusqu'à ce qu'un stress (comme les antibiotiques ou UV) ne conduise les virus à déclencher un cycle lytique.

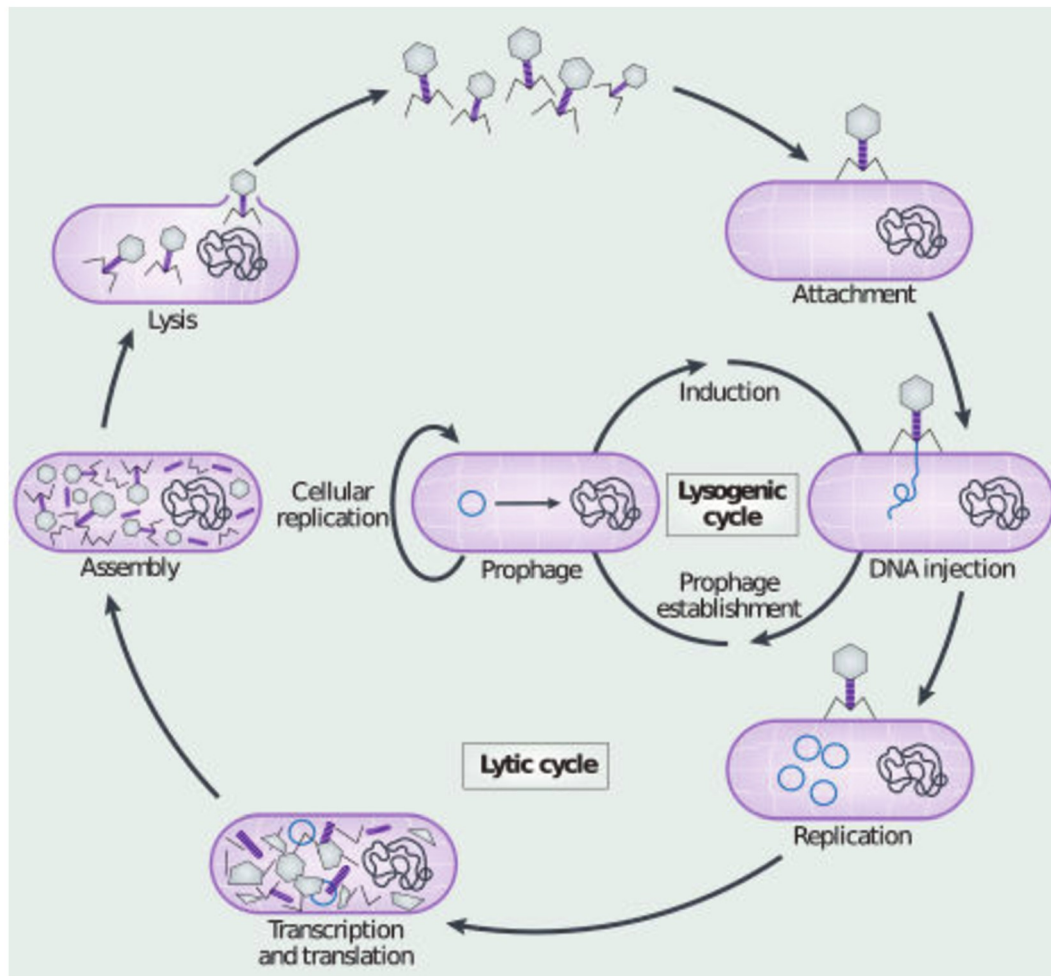


Figure 5 : Schématisation du cycle d'un virus tempéré voie lytique et voie lysogénique (Salmond and Fineran, 2015). Le cycle du virus lytique est représenté sur le cercle extérieur. Les virions libres s'adsorbent sur la cellule puis y injectent leur AN, avant de se répliquer, de s'assembler puis de provoquer la lyse cellulaire menant au relargage de virions néoformés dans l'environnement. Le cycle lysogénique, représenté au centre débute comme le cycle lytique par l'attachement des virions sur les cellules mais après injection de l'AN dans l'hôte, celui-ci s'intègre au génome de l'hôte sous forme de provirus se répliquant ainsi avec la cellule. Si une induction intervient, le provirus s'excise du génome de l'hôte entrant ainsi en cycle lytique.

L'état porteur ou chronique, correspond à un équilibre entre la production virale et la croissance cellulaire. Il y a une production constante de virions qui sortent de la cellule hôte par bourgeonnement. Contrairement au cycle lytique, ce relargage ne s'accompagne pas forcément de la lyse de l'hôte. De plus, la production virale ne peut être induite (Weinbauer, 2004; Madigan and Martinko, 2007).

Des étapes communes lors de l'infection sont partagées par les différents bactéries ou des archées quel que soit le type de cycle qu'ils mettent en œuvre (Muller, 2004).

Tout d'abord, le virion s'adsorbe sur une surface cellulaire. Cette étape est un temps réversible puis les fibres caudales virales se lient de façon définitive aux récepteurs présents sur la membrane de la cellule. Ensuite, les enzymes de la membrane cellulaire permettant l'injection du génome viral dans la cellule. Dans le cytoplasme, le génome peut être intégré au chromosome hôte ou rester libre. Pour permettre la néo-synthèse de virions, le génome doit être sous forme de réplicon. Les gènes viraux sont alors exprimés, conduisant à la réplication du génome et la synthèse de protéines virales qui formeront notamment les capsides. Les génomes sont ensuite empaquetés dans les capsides virales néo-synthétisées, puis libérés dans l'environnement par bourgeonnement, extrusion ou suite à la lyse de la cellule hôte. La vie d'un virion se termine soit par l'infection d'un nouvel hôte soit par la destruction de l'environnement.

- Course à l'armement : système de défense hôtes – virus

Au cours des milliers d'années d'évolution entre les cellules et les techniques de résistances ont été mises au point par les cellules pour la survie virale. Il est d'ailleurs supposé que l'une d'entre elles ait permis le passage du virus à l'homme (Forterre, 2016).

Il existe divers mécanismes de défense contre les infections virales (Figure 6) classés en deux groupes :

Le premier regroupe les systèmes basés sur la reconnaissance de l'élément exogène et sa différenciation avec les AN cellulaires. C'est, par exemple, le système CRISPR-cas, du système de restriction modification (R-M), du système de restriction endonucléase (système DND) et probablement d'un système récemment décrit, le BREX (Makarova et al., 2013).

Le second groupe rassemble des systèmes de défenses basés sur la mort cellulaire programmée ou la dormance des cellules infectées, comme le système Abi ou le système toxine-antitoxine (TA) (Makarova et al., 2013).

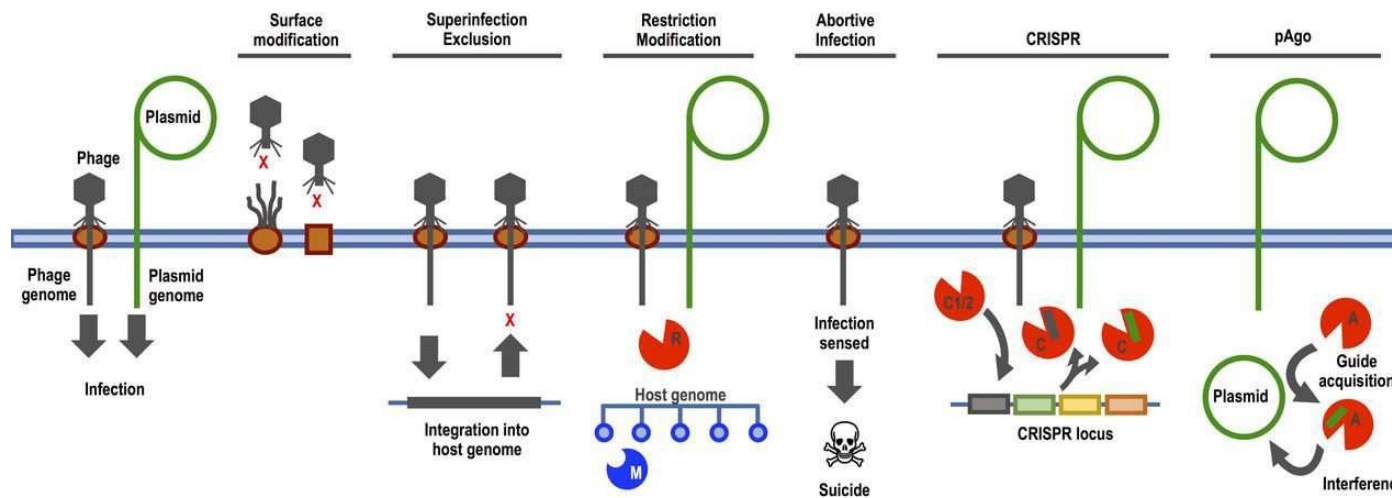


Figure 6 : Mécanisme de défense microbienne contre les éléments génétiques (2016). Les bactériovirus sont représentés en gris et les plasmides en vert. La modification de la surface cellulaire empêchant la fixation des phages, la présence d'un autre provirus, le système de restriction modification empêchant la fixation des plasmides, la présence d'un autre plasmide, l'infection abortive menant au « suicide » de la cellule, le système CRISPR-cas comprenant une « mémoire » des infections passées. Ces mécanismes empêchent la fixation des éléments génétiques.

Système CRISPR-cas

L'un des mécanismes de défenses contre les invasions d'ADN exogènes est le système CRISPR-cas. Cette famille de séquences d'ADN répétées a été découverte chez *E. coli* puis ce système fut nommé CRISPR-cas (Clustered Regularly Interspaced Short Palindromic Repeats – associated with cas genes) après la découverte de ce système chez les archées et les bactéries. Ce système permet de contrôler l'entrée de nucléiques exogènes tels que les virus ou les plasmides. Le locus de CRISPR est composé de plusieurs répétitions de séquences courtes partiellement palindromiques non codantes (spacers) séparées par des spacers correspondant à des séquences nucléiques exogènes comme par exemple des morceaux de génome bactérien très conservés dans un locus CRISPR donné et ont toujours une structure palindromique afin d'avoir une conformation secondaire stable, les séquences spacer

elles hypervariables. Les gènes cas sont la partie effective du système protéines portant des domaines fonctionnels typiques tels que nucléase, et protéines de liaisons aux polynucléotides. Les CRISPR permettent à l'acide nucléique exogène alors que les protéines produites par les gènes l'éliminer (Horvath and Barrangou, 2010).

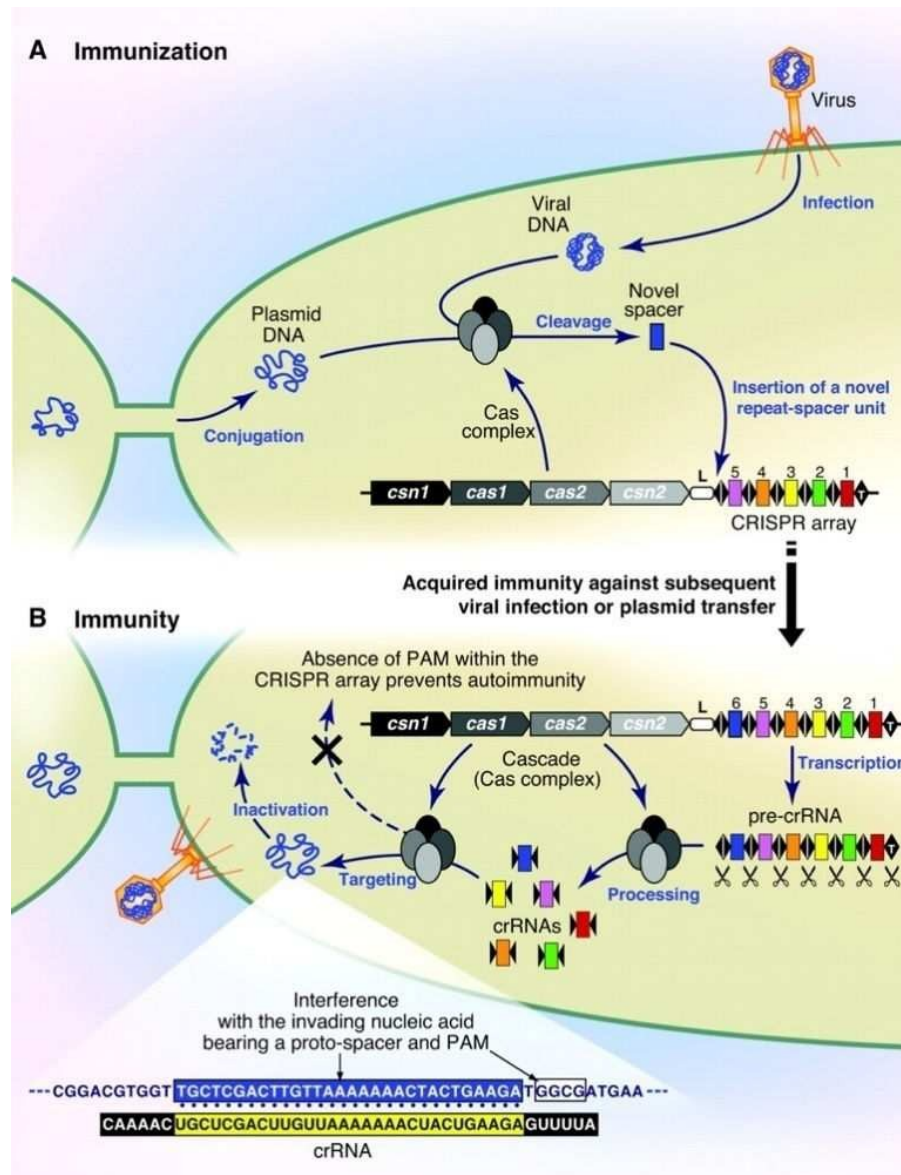


Figure 7 Schématisation du mode de fonctionnement du système CRISPR-cas (Horvath and Barrangou, 2010). (a) Processus d'immunisation : entrée de l'AN exogène, production de crRNA, clivage de l'AN exogène, création d'un nouveau spacer au sein du locus CRISPR (b) Processus d'immunité : entrée d'un AN exogène, transcription du crRNA précédemment transcrit et inactivation de la cible.

Le mécanisme général d'action est le suivant(Figure 7) :

Processus d'immunisation

- Entrée d'acide nucléique exogène
- Reconnaissance
- Création d'une nouvelle séquence repeat
- Sélection et création du spacer
- Intégration dans le génome

Mise en place de l'immunité

- Transcription du loci CRISPR en pre-crRNA
- Transformation des pre-crRNA en crRNA
- crRNA guide le complexe cas vers la cible
- Dégradation de l'acide nucléique cible

C'est un système très répandu car il est présent dans 87% des génomes bactériens (Sorek et al., 2008; Makarova et al., 2015). Ces systèmes sont particulièrement sujets aux transferts latéraux de gènes, même entre espèces apportant un avantage évolutif certain notamment face aux infections (Sorek et al., 2008).

Système Restriction – Modification (R-M) (Tock and Dryden, 2005; Makarova

Un autre mécanisme particulièrement bien étudié est le système de restriction (R-M). Tous les systèmes R-M fonctionnent sur le même principe : discrimination endogènes des AN exogènes afin de cliver ces derniers. Cela met en

- Une méthyltransferase (MTase), qui méthyle des bases adénine de la cellule pour qu'il puisse être reconnu et protégé du clivage.

- Une enzyme de restriction de type endonucléase (REase) qui clive des AN étrangers non méthylés (viraux ou plasmidiques).

Les système R-M sont classés en 4 grands types selon la composition des sous-unités, les cofacteurs requis (ATP(GTP)), le site de reconnaissance et le mécanisme de clivage. Le système de **type I** inclut différents gènes : 2 sous unités REase pour le clivage (R), 2 sous unités MTase (M) pour la méthylation et une sous unité S pour la spécificité, formant ainsi le complexe hétéro-oligomérique R-MA (Figure 8). Le système de **type II** est le plus simple et le plus commun (plus de 3500 systèmes différents). Il repose sur l'association basique d'une MTase et d'une REase, généralement codées par le même opéron. Le **type III** ressemble au type II puisqu'il ne contient que 2 sous unités : R et M. Celui-ci ressemble aussi au type I car leurs sous unités R partagent des similarités : elles contiennent toutes les 2 un domaine hélicase et leur réaction est ATP-dépendante. Enfin, le type IV contient 2 sous unités distinctes, une GTPase de la famille AAA+ et une endonucléase qui clive l'ADN cible de façon non spécifique.

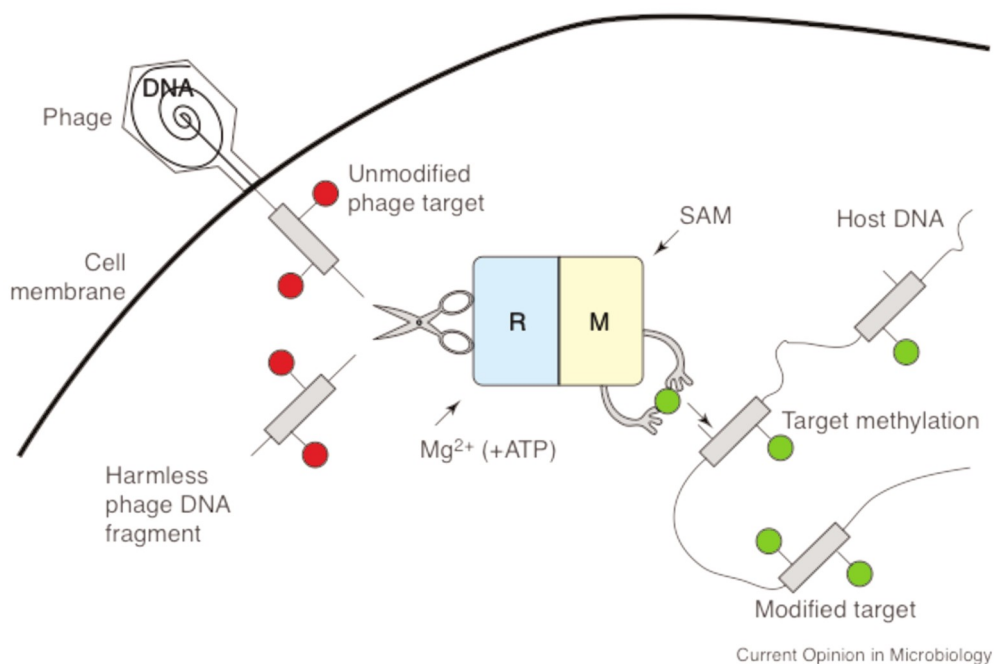


Figure 8 : Illustration du fonctionnement du système R-M de type I (Tock and Dryden, 2005). Les enzymes de restriction (ciseaux) reconnaissent l'ADN exogène non méthylé (rouge) alors que l'ADN cellulaire est, lui, méthylé (vert) et donc protégé.

Système de phosphorothioation de l'ADN (DND) (Xu et al., 2010; Makarova

Le principe du système DND, un peu comme le système R-M via une S-modification de l'ADN (phosphorothioation : un atome de soufre d'oxygène non-liant du phosphate de la liaison phospho-diester), et de non marqués. Comme souvent pour les îlots de gènes de défenses, gènes arrangés en cluster. L'activité de restriction comprend au moins autres gènes contigus sont nécessaires pour l'activité de modification de l'ADN que ce système DND soit proche du système R-M de type I. En de gènes et le mode de fonctionnement (destruction des AN exogènes AN cellulaires sont protégés par le marquage) semble proche du système

Système BREX (BacteRiophage EXclusion) (Barrangou and van der Oost, 2014; 2015; Nunes-Alves, 2015).

Le système BREX contient des similarités avec un système plus ancien (growth limitation). Celui-ci est basé sur l'utilisation du gène *plgZ* qui est système de résistance de *Streptomyces coelicolor*. Dans une étude récente, environ 1500 génomes de bactéries et d'archées, des gènes contenant été retrouvés dans 10% des génomes. Or dans plus de la moitié des retrouvés dans un cluster composé de 6 gènes dont 2 gènes ont une système Pgl. Cette observation est cohérente avec des données précédemment soulignaient le fait que les mécanismes de défenses contre les infections généralement groupés en clusters (Makarova et al., 2011). C'est ce cluster qui a été appelé système BREX. Il s'agit donc de 6 gènes permettant gamme de virus, contrairement au système Pgl qui a une gamme

Cette cassette inclus une Lon-like protéase putative (serine peptidase ATP-d protéine contenant un domaine phosphatase alcaline, une protéine putative l'ARN, une ADN méthylase, une protéine à domaine ATPase et une inconnue. BREX semble être un système complexe contenant des protéines ayant de nombreuses activités biochimiques (protéase, phosphatase et méthylase). Même si ce système semble proche de système Pgl puisqu'il y a partage d'une quelques différences notables. En effet, le système Pgl permet une libération de virions et permet aussi la lysogénie. Le système BREX infection ni la lysogénie, ce qui semble prouver que les deux mode de fonctionnement (Figure 9). D'autre part, le système BREX l'adsorption du virus mais bloque la réplication du génome viral. Néanmoins, ce qu'il se passe pour le système restriction modification, l'ADN viral ne clivé par le BREX, ce qui suggère un nouveau mécanisme de défense. Par contre, le chromosome de la cellule est méthylé quand les cellules contiennent le suggérant que celui-ci est capable de promouvoir une méthylation site-spécifique du chromosome cellulaire. Cette méthylation, probablement médiée par le gène pglX, permettre une sorte de reconnaissance entre les acides nucléiques endogènes. Aussi, si le gène pglX est inactivé, la résistance aux infections n'est forte importance de celui-ci dans le mécanisme de défense. Le système du BREX n'est pas encore connu néanmoins il semblerait qu'il soit actif réplication de l'ADN viral. Il semble aussi que, comme pour le système BREX soit une cible de choix pour les transferts horizontaux de gènes. Pour virus testés lors de l'étude ont montré une résistance au système BREX, la course à l'armement entre virus et cellules est soumise à une

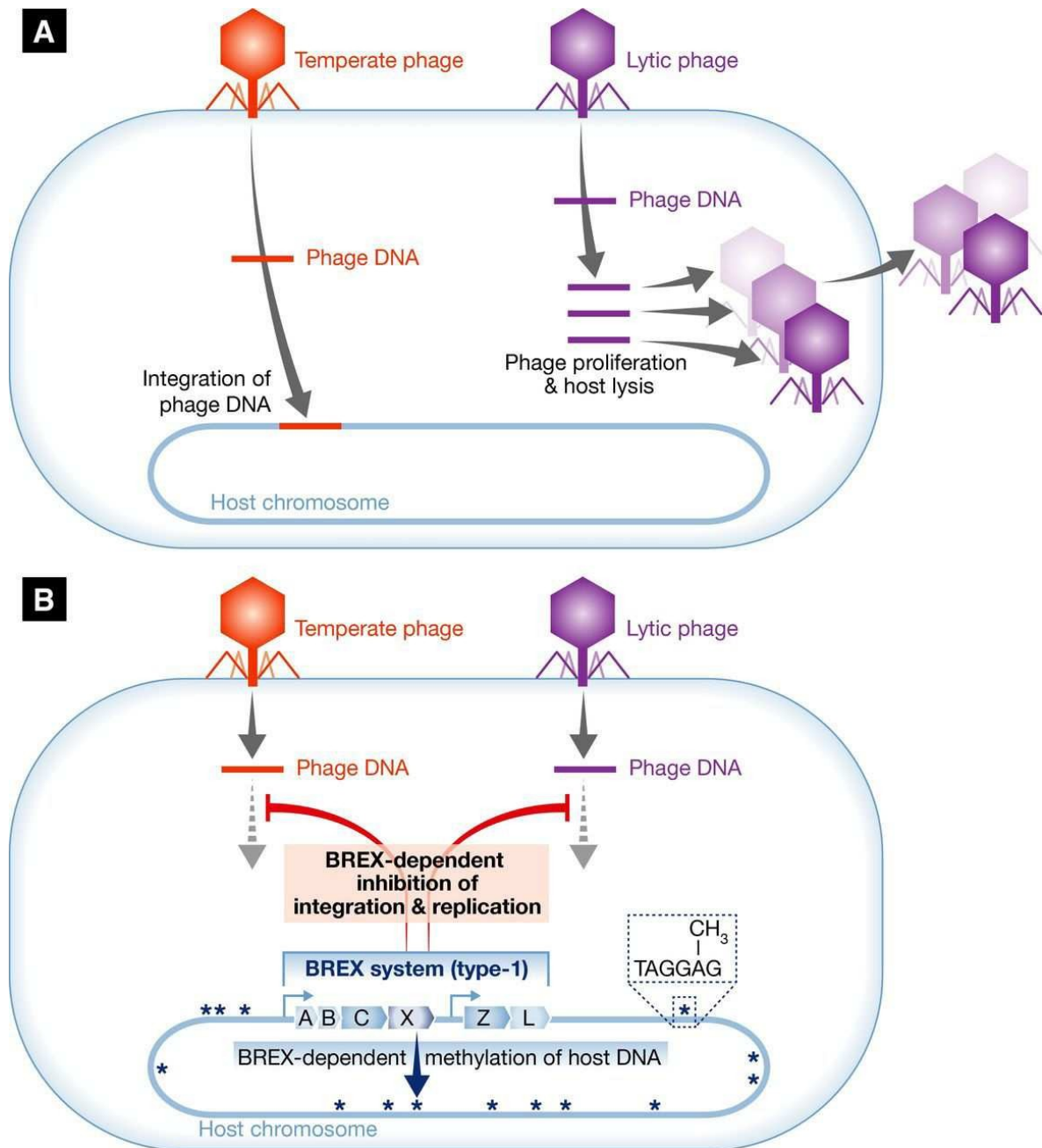


Figure 9 : Illustration du mode de fonctionnement hypothétique du BREX (Barra 2014). (a) Cycles classiques : lysogénique ou lytique (gauche et droite) (b) Cas le système de défense BREX entraînant l'inhibition de la production virale, voire encore réellement connu à ce jour. (chez *Bacillus cereus*). Le mécanisme

Système toxine – antitoxine (TA) (Hayes and Van Melder, 2011; Makarova Triki, 2014)

Il existe divers types de système TA classés selon les toxines et l'organisation des gènes impliqués et leurs modes d'actions. Il y a 6 Le type I met en œuvre un ARN non codant ou antisens (ARNnc) protéine (généralement une holin-like protéine) comme toxine. Le type les archées et les bactéries. Il utilise des protéines comme toxine généralement transcrites dans le même opéron. Pour le système de type sous forme d'ARNnc qui peut directement se fixer sur la toxine. Trois autres récemment proposés (IV, V et VI) basés sur des mécanismes d'actions antitoxines respectives.

Quel que soit le type de système, ils sont tous basés sur une toxine si elle est exprimée au-dessus d'un certain niveau, va entraîner antitoxine instable qui peut inactiver la toxine et/ou réguler son expression mort cellulaire. Les toxines sont des protéines dirigées contre des cibles spécifiques alors que les antitoxines sont des protéines ou de petits ARN pour cible la toxine. La plupart des toxines sont des endoribonucléotides qui freinent la synthèse protéique en réponse à divers stress. Leur effet est réversible ce qui leur donne une capacité d'adaptation face aux différentes rencontrées. Les autres toxines peuvent avoir des effets sur l'action d'une topoisomérase ou l'assemblage de la membrane cellulaire ou sur les structures du cytosquelette.

Globalement, le système TA a pour objectif d'induire une entrée en la cellule, grâce à une production contrôlée de toxine, pour répondre à l'invasion par des AN exogènes.

Le système TA joue un rôle majeur dans la lutte contre l'invasion dans la survie cellulaire, la formation de biofilm et la résistance aux

Système Abi (Chopin et al., 2005; Fineran et al., 2009; Makarova

Le système Abi (Abortive infection), aussi appelé « phage exclusion » du virus en entraînant la mort prématurée de l'hôte lors de l'infection caractérisé par un début d'infection virale classique, avec adsorption du virion de l'AN viral dans la cellule. L'infection est ensuite stoppée menant quantité de nouveaux virions (voir aucun) dans l'environnement mais aussi prématurée de la cellule infectée. Ce système permet de prévenir environnantes et est ainsi considéré comme un « suicide altruiste » mécanisme moléculaire qui est mis en jeu reste assez méconnu. Il n'y trentaine de systèmes Abi décrits, provenant majoritairement de organismes modèles (*Lactococcus lactis* et *Escherichia coli*) ce qui suggère assez peu répandu, ne représentant qu'une fraction mineure de la modules de défenses microbiens. Néanmoins il partage des similarités (comme par exemple la formation de pores dans la membrane) et au système R-M, pouvant donc potentiellement interagir avec celui-ci.

La réponse des virus

Ces mécanismes de défenses cellulaires partagent des similarités comme des cibles prioritaires pour les transferts latéraux de gènes, ou qu'ils ciblent les virus que les plasmides.

Un autre point commun partagé par nombre de ces systèmes est des mécanismes pour les éviter d'où le nom de « course à l'armement ».

L'un des mécanismes les plus couramment utilisé par les virus pour échapper de la défense est de créer une mutation dans leur propre génome. Par exemple, CRISPR-cas, une seule mutation dans les proto-spacers (spacers encore dans le génome) permet de contrecarrer le système de défense cellulaire (Horvath, 2005; Horvath and Barrangou, 2010; Samson et al., 2013). La délétion d'une partie du génome viral peut aussi servir de mécanisme de défense bactérien. Les virus peuvent ainsi la cible connue par le système de défense bactérien. Les virus interagissent directement avec le système de défense en inhibant l'activité enzymatique du système R-M ou en interférant avec les protéines codées par les gènes du système R-M. Ils peuvent aussi modifier le système à leur avantage en codant pour l'antitoxine du système TA, ou en activant le système R-M pour que celui-ci méthyle aussi les AN viraux (Samson et al., 2013).

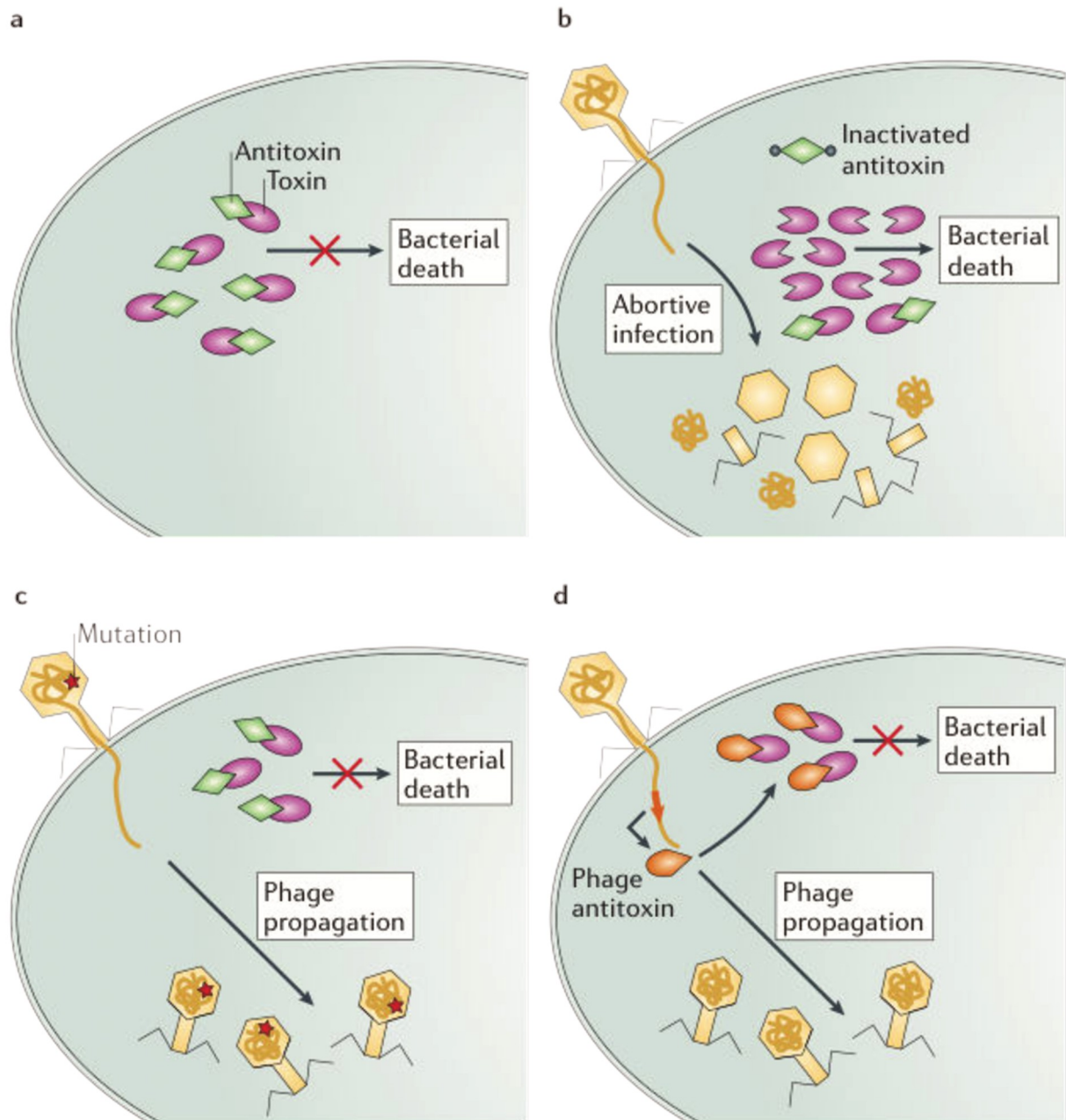


Figure 10 : Exemple de stratégie virale pour échapper au système toxine antitoxine (Samson et al., 2013). (a) Croissance cellulaire normale, les antitoxines (vert) neutralisent les toxines (violet) évitant ainsi la mort cellulaire. (b) Lors de l'infection par un virus, les antitoxines sont inactivées ou en nombre inférieur et les toxines libres entraînent la mort cellulaire prématurée. (c) Cas d'une infection par un virus possédant une mutation permettant d'empêcher l'activation du système TA. (d) Cas de virus capable de coder pour une protéine similaire à l'antitoxine cellulaire permettant d'éviter l'activation du système.

II. Les virus de bactéries et d'archées dans les océans

Alors que les bactériovirus ont été découverts au début du XX^{ème} siècle, il a fallu attendre les années 70 pour que le premier archéovirus soit mis en évidence et identifié, le concept même d'archée n'étant pas conçu (Torsvik and Dundas, 1988). Les virus associés à des bactéries ou archées sont de loin les plus nombreux dans les océans. Cependant leur diversité morphologique est assez faible puisque 96% d'entre eux ont une symétrie binaire (Ackermann and Prangishvili, 2012). Il s'agit de virus appartenant à des Caudovirales, divisé en 3 familles : les Myoviridae, les Podoviridae et les Siphoviridae. Les 4% restants sont cubiques comme les Tectiviridae, filamenteux comme les Inoviridae ou pléomorphes comme les Fuselloviridae ou les Ampullaviridae. Ces virions contiennent généralement un génome à double brin dans la majeure partie des cas, mais ils peuvent aussi porter un génome à simple brin ou de l'ARN simple ou double brin, même si ce dernier cas est très rare. Ces données s'appliquent à l'ensemble des bactériovirus et archéovirus, compris ceux présents dans les océans dont les caractéristiques vont être développées dans le prochain

• Abondance et diversité

L'océan viral, estimations de l'abondance et « fausses » particules. Les virus sont les entités biologiques les plus abondantes de la biosphère. Dans les océans, on estime qu'il y a en effet 10³⁰ à 10³¹ virus par litre d'eau, dont la plupart d'entre eux sont des virus de bactéries ou d'archées (Breitbart and Rohwer, 2005; Suttle, 2005). La nouvelle ère de la virologie océanique a débuté durant les années 90 (Torrella and Morita, 1979). Ces dernières années ont vu un développement des environnements où la quantité de données sur les virus océaniques est la plus importante (Breitbart and Rohwer, 2005; Brum and Sullivan, 2015; Chow

2015). On considère qu'il y a environ 10 à 15 fois plus de dans les océans et cette forte abondance a été confirmée par les données disponibles (Suttle, 2005; Ackermann, 2007; Breitbart et al., 2007; Rohwer 2009). A l'image des hôtes qu'ils infectent, l'abondance des divers virus est l'abondance virale est estimée (VLP/ml ou virus-like particles) en ce qui concerne côtières, celle-ci diminue avec la profondeur et l'éloignement par rapport 2007; Danovaro et al., 2008). Ces résultats sont cohérents si on considère que l'abondance virale est corrélée à celle des micro-organismes (Figure 11) (Danovaro d'ailleurs être considéré que les virus utilisent les stratégies « *r* » et « *K* » décrivant habituellement des populations bactériennes. Il est alors possible de les classer en 2 groupes (Chow and Suttle, 2015) :

- Les virus les plus abondants, sont caractérisés par une forte virulence, rapide, un burst size important et de petits génomes et suivent la stratégie « *r* ».
- Les virus présents en minorité sont, eux, caractérisés par une faible virulence, un burst size, de grands génomes et mettent plutôt en œuvre des cycles de latence des états porteurs et suivent une stratégie « *K* ». Cependant, la plus grande diversité génotypique (Breitbart and Rohwer, 2005).

Ce qui va caractériser le mode de vie d'un virus, son évolution et son cycle de vie peut donc être résumé en 4 grands critères : le type de porteur (bactériophage, eucaryote, etc.) ; l'étendue de sa gamme d'hôte (virus spécialiste ou généraliste) ; enfin les caractéristiques propres des hôtes qu'il infecte (abondance, état physiologique, etc.). Certaines familles virales peuvent être caractérisées par ces stratégies (Suttle

Par exemple, les Myovirus sont la plupart du temps lytique, avec une gamme d'hôtes étendue et prolifèrent rapidement, ils représentent plutôt la stratégie « r ». A l'inverse, les Podovirus ont une gamme d'hôtes restreinte, sont peu compétitifs, moins représentés au niveau du viroplancton et sont donc plus généralement considérés comme « K ». Enfin, les Siphovirus, avec une gamme d'hôtes intermédiaire, sont connus pour être en majorité tempérés (mettant en œuvre un cycle lysogénique) ce qui les classe plutôt comme utilisant une stratégie « K » (Suttle, 2007; Sime-Ngando, 2014).

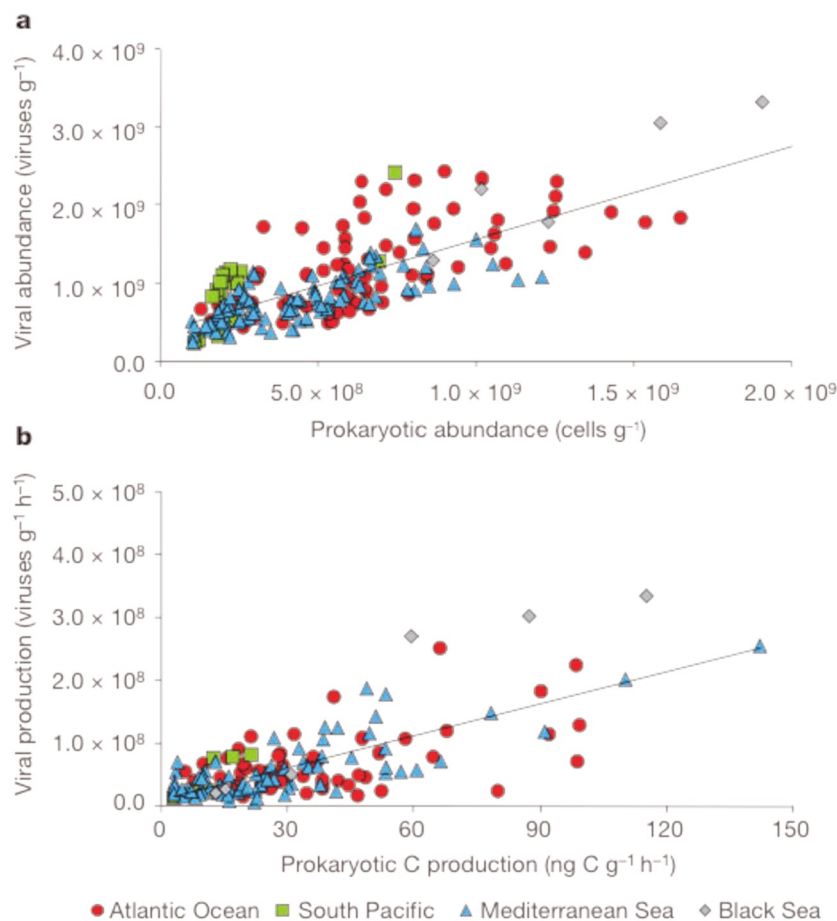


Figure 11 : Relation entre la concentration virale et l'abondance des micro-organismes (bactéries et archées) dans les sédiments marins profonds de différents océans (Danovaro et al., 2008). Le premier graphique représente l'abondance virale en fonction de l'abondance procaryotique et le second représente l'abondance virale en fonction de la production de carbone procaryotique. Les différentes couleurs représentent l'origine géographique des échantillons. Une forte corrélation entre l'abondance virale et l'abondance procaryotique.

Les estimations d'abondances virales sont ou devraient être exprimées en « en « Virus-like particles ».

Ces comptages étaient initialement réalisés grâce à la MET, ce qui permet de compter des particules simplement en ayant des indications sur la taille et la morphologie. Néanmoins, cette méthode sous-estime le nombre de particules car le nombre est souvent assez élevé et aucune information sur l'infectiosité ne peut être obtenue. La MET est particulièrement chronophage et assez coûteuse (Weinbauer, 2004).

Aujourd'hui les techniques les plus utilisées pour les estimations sur échantillons environnementaux sont la microscopie à épifluorescence et la cytométrie en flux. Ces techniques sont plus rapides et plus précises que la précédente même si la morphologie des VLP est perdue. L'infectiosité ne peut pas non plus être déterminée. Ces techniques sont basées sur le même principe, c'est à dire la coloration des acides nucléiques, ce qui entraîne divers biais. D'une part la coloration au bromure de méthyle fait globalement que sur des ADN double brin, sous-estimant par conséquent la présence des virus à ARN et d'autre part, cette coloration ne discrimine pas les autres acides nucléiques non infectieux (Brum et al., 2015; Chow and Suttle, 2015). La coloration se fait en effet indifféremment sur tout type de VLPs, qu'il s'agisse d'un Agent de transfert de gène (Gene Transfert Agent : Agent de transfert de gène), des vésicules membranaires ou même des acides nucléiques libres agrégés sur d'autres particules. Ces colorations généralistes, peuvent donc mener à une surestimation de l'abondance virale océanique. Elles sont encore aujourd'hui sujettes à débat (Forterre et al., 2013). De ces dernières années, l'importance et l'abondance des MV et des GTA a été mise en évidence dans les océans (Lang et al., 2012; Biller et al., 2014; Scanlan, 2014).

En effet, à la lumière de récentes études, il apparaît que
abondantes (ex : 5. 6. 10⁶ à 10¹⁰ par ml d'eau de mer) et peuvent être
comme VLP puisque leur taille et le type d'acide nucléique qu'elles
différents de certains virus (Forterre et al., 2013; Biller et al., 2014;

Les GTA sont des particules ressemblant morphologiquement à de petits bactéries
capables de transférer de gènes (Figure 12). Néanmoins, même s'ils ressemblent
leur manque la caractéristique principale de ceux-ci : la capacité
de capsidation. En effet, ils empaquettent aléatoirement des fragments du chromosome
sans pour autant contenir le matériel nécessaire à la production des
entités caudées (Lang et al., 2012). Ils n'ont donc jamais l'opportunité de transcrire
dans la cellule réceptrice la capacité de produire à son tour d'autres GTAs.
issus de transduction généralisée car, dans ce cas, le virus encapside partiellement
de chromosome de l'hôte à la place du génome viral (faible pourcentage de
transductrices dans un lysat). Les GTAs sont néanmoins produits par un système
par la cellule et sont probablement relargués dans l'environnement par la cellule
hôte. Ces particules sont retrouvées chez diverses bactéries et archées comme des
Proteobactéries, des Spirochètes ou encore des Methanococcales (Paul, 2008;
2012) et sont sans doute particulièrement abondantes dans les océans (McDaniel

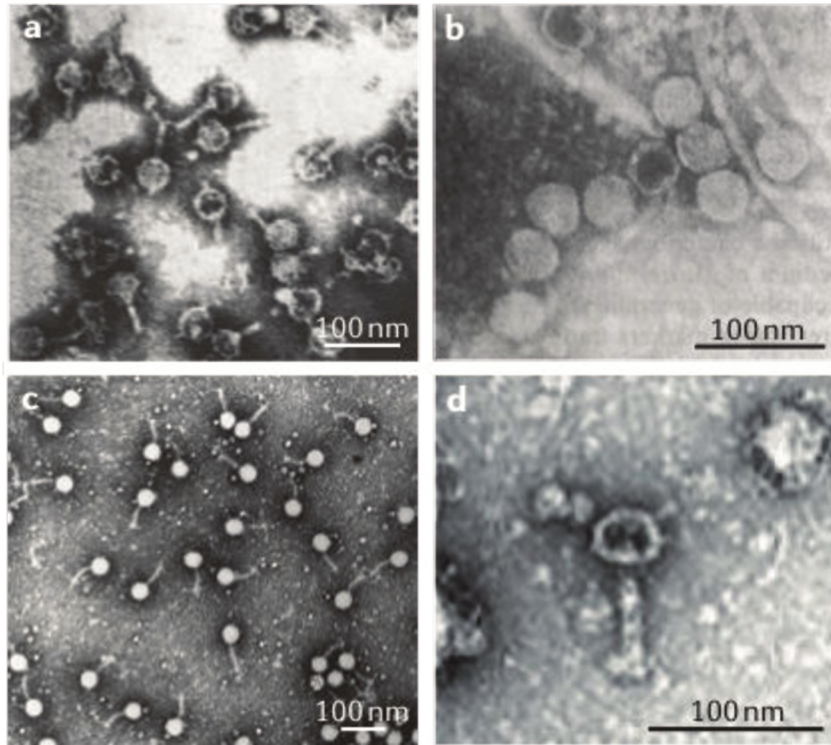


Figure 12 : Photographie au MET de GTA (Lang et al., 2012). (a) GTA de *Rhodobacter capsulatus*. (b) Particules de Dd1 issues de *Desulfovibrio desulfuricans*. (c) Particules virales de *Serpulina hyodysenteriae* (VSH-1) dans *Brachispira hyodysenteriae*. (d) Particules de « Voltae transfer agent » issues de *Methanococcus voltae*.

Les vésicules membranaires sont, quant à elles, de petits compartiments sphériques de 50 à plus de 250 nm de diamètre produites par bourgeonnement à partir de la membrane cellulaire et pouvant être sécrétées dans le milieu extracellulaire (Kulp and Kuehn, 2010). Elles sont représentées dans les trois domaines du vivant et peuvent avoir de nombreuses formes, tailles, compositions et fonctions (Gaudin *et al.*, 2014). Ces MVs ont été très étudiées chez les Eucaryotes mais aussi chez les bactéries à Gram négatif. Il est admis que le processus de formation de celles-ci soit contrôlé, et probablement bien conservé puisqu'elles sont présentes dans les trois domaines du vivant. On sait qu'elles sont d'origine membranaire, que leur composition est très proche de celle de la membrane cellulaire mais que certaines protéines y sont plus ou moins abondantes voire totalement absentes.

Néanmoins, il n'y a pas de consensus quant à la compréhension du mécanisme de formation des MVs (Kulp and Kuehn, 2010). On sait toutefois qu'elles peuvent endosser de nombreux rôles dont voici quelques exemples (Figure 13) (Kulp and Kuehn, 2010; Manning and Kuehn, 2011; Gaudin, 2012):

- Excrétion de divers composés : insolubles ou toxiques (détoxification de déchets biologiques et autres composés nocifs)
- Transport de différents éléments pour les transférer à une cellule receveuse : acide nucléique (transfert de gènes), toxine (compétition avec d'autres cellules), molécules impliquées dans le Quorum Sensing (communication intercellulaire)
- Leurre cellulaire contre les attaques virales.

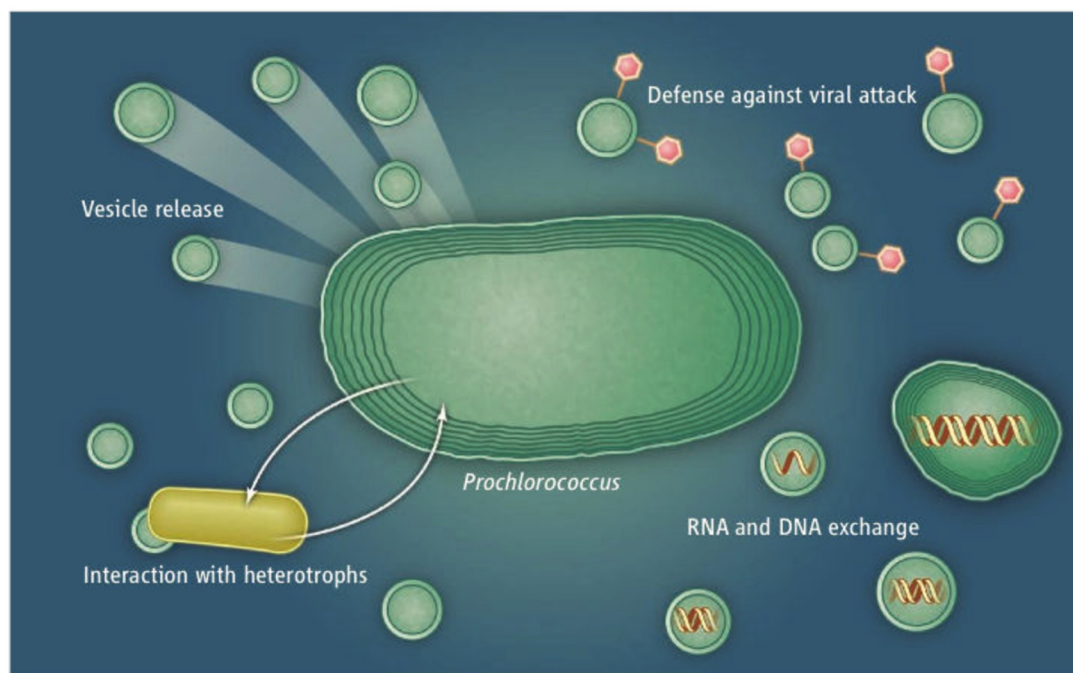


Figure 13 : Différents rôles des vésicules membranaires (Scanlan, 2014). Les MVs peuvent prévenir les infections virales en jouant le rôle de leurre cellulaire, intervenir dans les transferts horizontaux de gènes et interagir avec des hétérotrophes.

Récemment, la production de vésicules membranaires a aussi été décrite chez les archées. Le premier modèle à avoir été décrit est celui de Sulfobacillus acidophilus (la sulfobacilline) associée à des MVs (Prangishvili et al., 2000). D'autres MVs ont été caractérisées chez les archées, toujours dans le phylum des Crenarchaeota. Elles sont visiblement pour la plupart entourées de la S-layer, membrane caractéristique des archées (Ellen et al., 2009). Des MVs ont aussi été découvertes associées à des bactéries, notamment des Thermococcales (Figure 14) (Soler et al., 2008, 2011; Gorlas et al., 2013). Les MVs ont des caractéristiques particulières. En effet, elles peuvent produire des filaments filamenteuses en forme de colliers de perles, mais elles peuvent aussi ressembler à des virions (notamment les Globulaviridae) (Soler et al., 2008). Outre leur aspect morphologique, certaines de ces MVs peuvent aussi transporter du matériel génétique (ADN plasmidique ou viral), la résistance de celui-ci aux traitements thermiques devient alors très importante. Ce type de transport pourrait donc être une stratégie de protection des acides nucléiques facilitant les transferts latéraux de gènes entre cellules dans des conditions environnementales dites extrêmes (Soler et al., 2008, 2011, 2015, Gaudin et al., 2013, 2014).

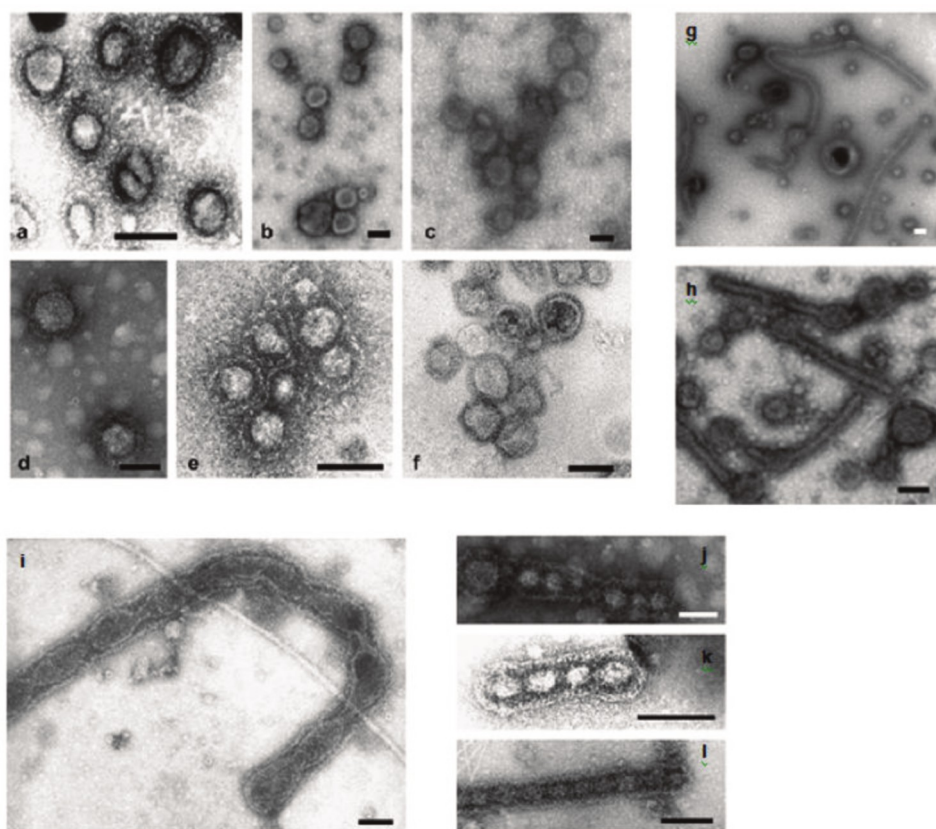


Figure 14 : Observation en MET de vésicules membranaires produites par des *Thermococcales* (Soler et al., 2008). (a-f) « Virus-like vesicles » (g, h) structures filamenteuses en forme de bâtonnets (i-l) ou de colliers de perles produites par des *Thermococcales*. **Echelle 100nm.**

Il est admis que les virus sont de véritables moteurs de l'évolution et qu'ils sont des acteurs déterminants dans l'échange de matériel génétique. Cependant les GTA, les MVs, mais aussi les plasmides et autres éléments génétiques mobiles ne doivent pas être négligés. Etant donné leur abondance, en particulier dans les océans, ils ont aussi un rôle majeur dans les transferts latéraux de gènes et donc dans l'évolution des communautés microbiennes (Figure 15, pour plus de détails, voir la revue constituant le chapitre I de ce manuscrit (Lossouarn, Dupont, et al., 2015)). La difficulté rencontrée lors des comptages pour obtenir des estimations de l'abondance virale dans les océans (Forterre et al., 2013) réside dans la discrimination des MVs et des GTA par rapport aux virions. Il faut donc garder à l'esprit qu'il en résulte probablement une surestimation de l'abondance virale.

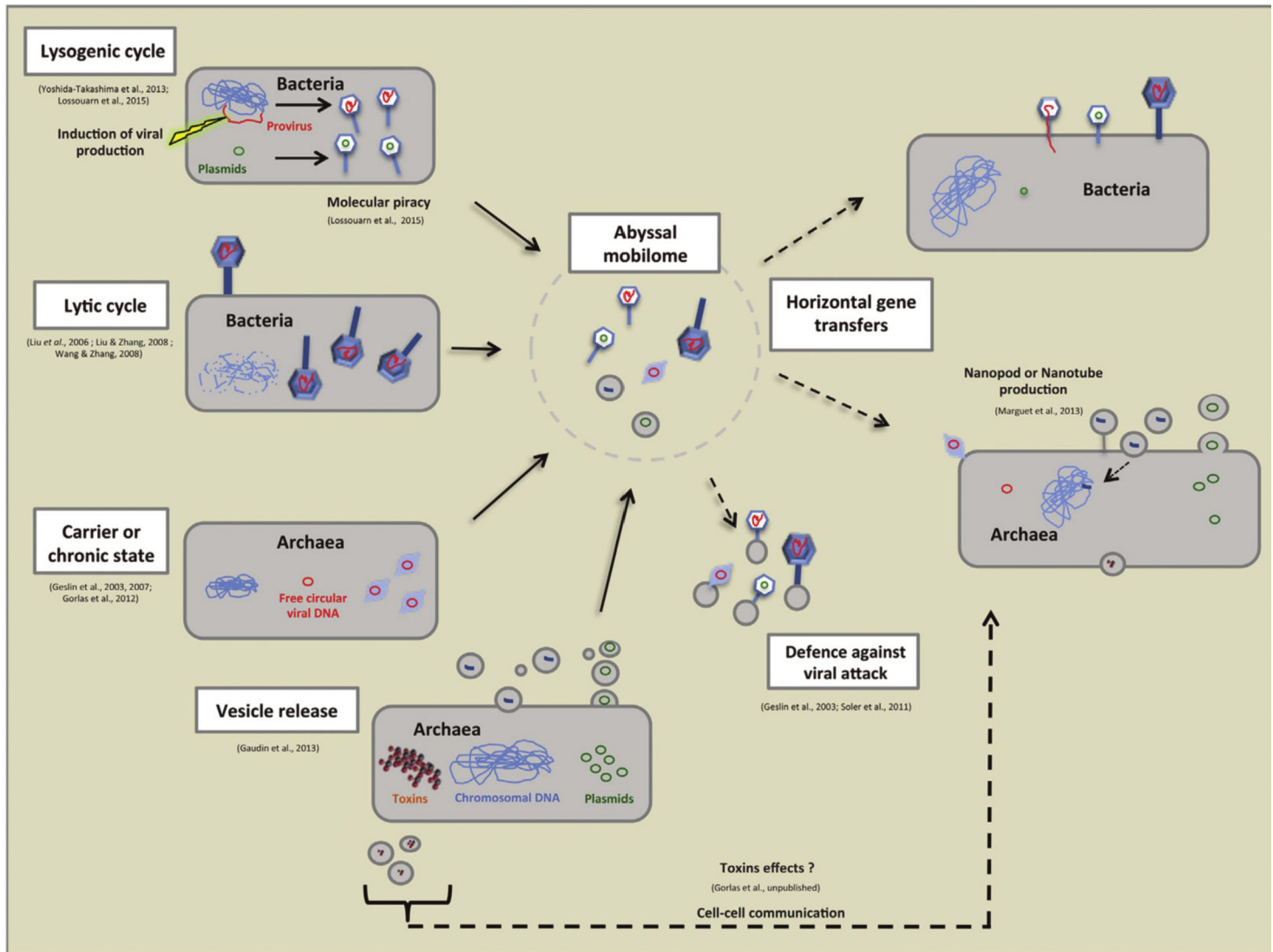


Figure 15 : Un mobilome abyssal (Lossouarn, Dupont, et al., 2015). Description schématique des divers rôles possibles des virus, MVs et plasmides au niveau des sources hydrothermales océaniques profondes. Implication notamment dans les transferts latéraux de gènes (ADN cellulaire (bleu), viral (rouge) ou plasmidique (vert)). Rôle des MVs dans la défense contre les attaques virales, le transport de toxines et autres types de communication entre cellules. Les flèches en traits pleins signifient que les résultats ont été démontrés alors que les flèches en pointillés caractérisent des rôles potentiels. Pas de notion d'échelle sur le schéma.

Diversité du viroplancton

La diversité génomique et morphologique virale en milieu marin est très elle reste encore largement méconnue (Breitbart and Rohwer, 2005). La répertoriés sont associés à des bactéries, ils sont caudés ou non et Ngando, 2014). Selon diverses études, les virus caudés, de l'ordre des Caudovirales représenteraient 10 à 40% de l'abondance virale en milieu aquatique (Sime Ces estimations paraissent assez cohérentes sachant que 96% des virus infect ou des archées sont des virus caudés (Ackermann and Prangishvili, 2012, l'importance des virus de type « phagique » mais non caudés a estimée mais reste sujette à débat. Certaines études estiment qu'ils représentent du viroplancton (Brum et al., 2013). Les virus d'archées semblent avoir beau diversité morphologique que les virus infectant les bactéries, en partic les virus infectant des archées thermophiles ou hyperthermophiles (Prangishvili 2016; Lossouarn, Dupont, et al., 2015).

La plupart des populations virales retrouvées dans les océans sont assez ubiqu des données obtenues lors de l'expédition Tara océan, montrent que les sont rarement endémiques à une station (15%) mais plutôt communément obser de 4 stations (47%) et certaines jusque dans 24 des 26 stations (Brum et al., 2015). Ces données sur la distribution des populations virales une autre étude menée sur plus de 3000 échantillons (provenant de divers notamment marins) qui montre que 86% des géotypes viraux sont retrouv échantillon et 73% dans plus de 5 échantillons différents (Figure 16) (Paez- Des séquences proches ou identiques ont été retrouvées à des endro éloignés ce qui implique une forte dispersion des virus dans ces environ

Même si la plupart des populations virales sont largement distribuées, la communautés virales réside dans la variation de l'abondance de chacun des s'adaptant aux différentes conditions environnementales (Brum et al., 2015; Chow 2015). Il est aussi intéressant de remarquer que la diversité génomique virale des environnements côtiers semble plus grande que celle des environnements bathypélagiques (Paez-Espino et al., 2016). D'autre part, la diversité virale locale est toujours dans les océans mais pour autant la diversité globale reste proportionnellement. Ce constat peut être expliqué par un modèle appelé le « Bank 2005). Les virus seraient séparés en deux grandes catégories les virus très actifs, numériquement majoritaires mais ne représentant qu'une faible partie de la diversité autres virus, inactifs, fonctionnant comme une réserve de gènes. Lorsque les conditions environnementales deviennent favorables à leurs hôtes, des virus issus de la population précédente prolifèrent et leur population devient alors dominante. La population précédemment majoritaire, dont les hôtes sont en déclin, voit alors son abondance diminuer, dans la fraction inactive constituant ainsi la « banque virale »

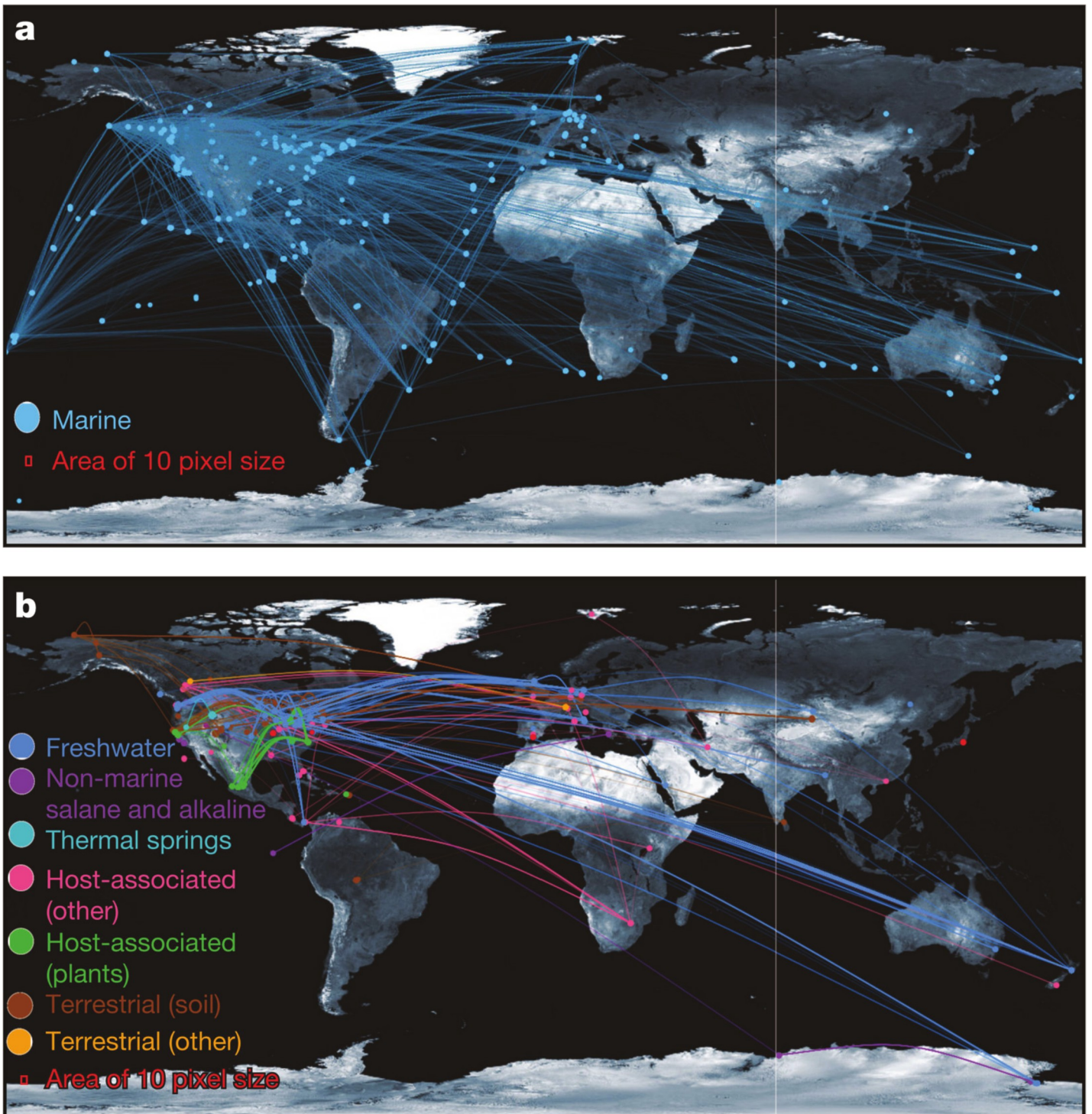
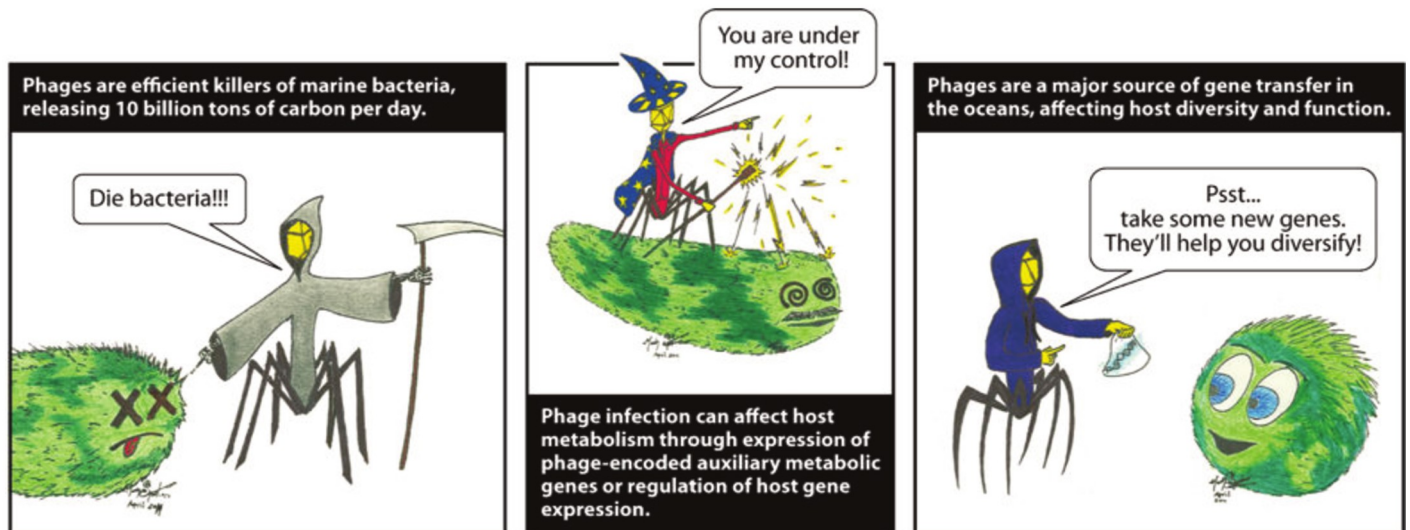


Figure 16 : Distribution globale de la diversité virale (Paez-Espino et al., 2016). La répartition de mêmes populations virales ou « singletons » (cercles) entre les échantillons est représentée par des lignes de différentes couleurs selon la nature de l'environnement, (a) pour les échantillons marins (b) pour les échantillons non marins. La dispersion des virus est très importante, en particulier au niveau marins où de nombreux virus sont retrouvés dans à différents lieux géographiquement très éloignés.

Le challenge réside donc aujourd'hui en la détection, la compréhension nouveaux virus (Brum and Sullivan, 2015). En effet, moins de 2200 génomes double brin et rétrovirus) sont disponibles via NCBI comparé aux plus de bactériens déposés (Reddy et al., 2014). Par ailleurs, seul un faible infectant des populations dominantes peuvent être affiliés à des virus cultivés. Une étude sur le Tara océan, moins de 1%) reflétant la pauvreté des bases de données en terme de quantité et de diversité de génomes séquencés. Les virus sont liés aux populations dominantes bien connues dans les océans tels que SAR11, SAR16 ou Roseobacter ou les autotrophes tels que Prochlorococcus et (Brum et al., 2015). Une étude réalisée sur des viromes issus des océans et des séquences présentes dans ceux-ci n'ont pas de similarités avec des séquences dans les bases de données (Hurwitz and Sullivan, 2013). Ces données permettent de faire un bond en avant concernant la quantité de données de génomes viraux. Néanmoins, un des challenges restant particulièrement complexe est de connecter les virus à leurs hôtes, d'où l'importance de continuer à développer des systèmes hôte-virus en laboratoire.

- **Le rôle des virus dans les océans**

Les virus sont aujourd'hui reconnus comme des acteurs essentiels des environnements marins. Ils sont impliqués notamment dans l'évolution des hôtes qu'ils infectent et plus globalement dans les cycles biogéochimiques (Suttle, 2005, 2007; Breitbart, 2012). Ils ont en effet un rôle important dans le recyclage du carbone et autres éléments ; ils sont d'importants agents de transfert de gènes (Figure 17). De ce fait ce sont des acteurs déterminant dans la variation et la sélection des espèces microbiennes, se positionnant ainsi comme des véritables moteurs de l'évolution (Forterre, 2013a).



From death dealers to DNA dealers, phage-host interactions are extremely diverse...

Figure 17 : Illustration humoristique des divers impacts des virus sur les communautés microbiennes (Breitbart, 2012). De gauche à droite sont illustrés les impacts sur la mortalité microbienne, sur la manipulation du métabolisme cellulaire et sur les transferts latéraux de gènes.

Impact sur la mortalité microbienne et cycles biogéochimiques

Les virus sont un facteur de mortalité très important pour les communautés océaniques. En effet, en se basant sur des observations directes de cellules, ils seraient responsables de 10 à 50% de la mortalité quotidienne de la communauté hétérotrophe (Sime-Ngando, 2014). Cette observation fait des virus une cause de mortalité équivalente à celles des organismes brouteurs tel que les flagellés ou les ciliés (Sime-Ngando and Colombet, 2009).

L'une des conséquences de cette forte mortalité induite par les virus est la modification de la composition des communautés microbiennes présentes, contrôlant ainsi la diversité. La diversification de ces communautés (Personnic et al., 2006; Sime-Ngando, 2014). Cela a mené à l'élaboration d'une théorie nommée « Killing the winner » (Thingstad and Lignell, 1997). Celle-ci repose sur divers constats. D'une part, plus un organisme est abondant plus il a de chances de rencontrer un virus. D'autre part, les organismes actifs à division rapide sont des proies prioritaires pour les infections virales (Sime-Ngando, 2010; Ankrah et al., 2014; Sime-Ngando, 2014). La théorie KtW suppose que dans les communautés les plus actives, et souvent les plus abondantes, sont des organismes qui sont les plus touchés par les infections virales. Les virus peuvent donc contrôler les populations présentes en massivement les populations dominantes, laissant la place à des communautés moins compétitives et minoritaires. Les virus supportent ainsi le maintien d'une diversité en permettant la coexistence de diverses communautés bactériennes plus ou moins compétitives.

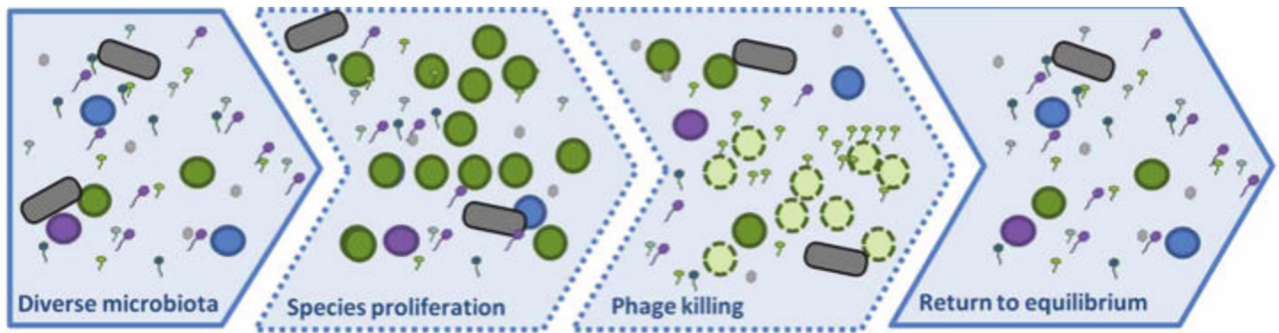


Figure 18 : Illustration de la stratégie Killing the winner (De Paepe et al., 2014). De gauche à droite, illustration de la diversité microbienne, puis prolifération d'une espèce, production importante de virus infectant cette espèce conduisant enfin à la décimation de celle-ci et donc au retour à une diversité plus grande.

Cette théorie semble admise par une grande partie de la communauté scientifique mais elle reste tout de même sujette à débat. En effet, les virus comme tous les parasites et prédateurs spécialisés sont face à un dilemme: un trop grand succès dans l'infection peut entraîner un désastre. Si les virus tuent toutes leurs proies, ils sont condamnés à disparaître puisqu'ils n'auront plus d'hôtes (Knowles *et al.*, 2016; Thingstad and Bratbak, 2016). Une étude récente a montré que le ratio VLP/micro-organismes avait tendance à décroître lorsqu'il y avait une forte population microbienne (Knowles *et al.*, 2016). Néanmoins le fait qu'il y ait une corrélation entre l'abondance virale et l'abondance des micro-organismes se révèle plutôt exact quand il s'agit de virus tempérés. Pour l'heure la relation entre densité de population microbienne et lysogénie reste floue. L'induction des provirus tend à montrer que la lysogénie est plus présente lorsqu'il y a une faible densité de population (et que le shift vers le cycle lytique se fait lors de l'augmentation de la population). Différentes études montrent au contraire que dans les récifs coralliens l'augmentation de la population microbienne se traduit, au niveau viral, par un passage du cycle lytique au cycle lysogénique. L'une de ces études, réalisée sur des viromes issus de récifs coralliens, montre que la présence de gènes marqueurs de la lysogénie augmente avec l'abondance microbienne (Knowles *et al.*, 2016).

Cette observation entraîne la naissance d'une nouvelle théorie « Piggyback the Winner » sur les épaules des gagnants. Pour les auteurs de cette théorie, lorsque l'abondance microbienne devient très importante, les virus tempérés sont privilégiés par rapport aux virus virulents (Knowles et al., 2016). Néanmoins, ils considèrent que cette théorie est une simplification de Killing the winner. En effet, la stratégie Killing the winner fonctionne à certaines densités de population où « Piggyback the Winner » est la stratégie dominante. À haute densité cellulaire, la dynamique lytique serait supprimée pour laisser la place aux virus tempérés (Knowles et al., 2016; Thingstad and Bratbak, 2016), ce qui leur permet de conférer un avantage non négligeable aux cellules. En effet lorsqu'un virus s'intègre au sein d'un chromosome, l'hôte peut acquérir une immunité face à ce virus. Le virus peut aussi améliorer le fitness de la cellule lui apportant de nouveaux gènes de son intégration dans le génome. Ces travaux mettent en avant l'importance et de l'impact de la lysogénie dans la dynamique des communautés microbiennes (Bratbak, 2016).

Quel que soit le modèle proposé, il est indéniable que les virus jouent un rôle important dans les océans. Lors de ces lyses de communautés dominantes, les virus libèrent un relargage de grandes quantités de matière organique, la rendant disponible pour d'autres organismes. Ils ont donc un impact important sur la chaîne alimentaire de l'écosystème océanique et de la boucle microbienne (Weinbauer, 2004). Il s'agit de l'implication des micro-organismes dans le transfert du carbone via leur matière organique dissoute (MOD) issue notamment de l'excrétion phytoplanctonique (Personnic et al., 2006). Les communautés microbiennes sont aussi intégrées dans la chaîne trophique puisqu'elles sont broutées par les eucaryotes unicellulaires planctoniques et certains zooplanctons.

Les virus interviennent donc dans la boucle microbienne en lysant une partie des organismes, modulant ainsi les flux de carbone de l'écosystème aquatique (Weinbauer, 2004; Weinbauer et al., 2011). Il apparaît donc logique que si l'abondance virale varie, cela entraîne des modifications sur l'abondance des micro-organismes. Diverses expériences ont montré que la concentration et l'activité enzymatique cellulaire, la concentration et la dynamique des matières organiques dissoutes ou particulaires, la production primaire (Weinbauer, 2004). Par ailleurs, il est estimé que les virus marins sont responsables de plus d'un quart du recyclage du carbone fixé dans les océans, alimentant ainsi les réseaux trophiques en rendant disponible le carbone pour les cellules lysées (Suttle, 2007; Weinbauer et al., 2011). Cela entraîne par conséquent un court-circuitage des réseaux trophiques supérieurs au profit d'une exportation vers les océans profonds (Weinbauer, 2004; Weinbauer et al., 2011). Ce phénomène est appelé « shunt », il s'agit du relargage des produits lysés des différents niveaux de la chaîne alimentaire organique dissoute (MOD) ou particulaire (MOP) (Figure 19). Ce relargage de matière organique engendré par les virus permet notamment de promouvoir le développement des populations microbiennes hétérotrophes (Weinbauer et al., 2011; Shelford et al., 2011). Cela entraîne aussi un transfert de carbone vers les océans profonds, jouant un rôle clé dans la pompe biologique.

L'influence des virus se manifeste sous différents aspects (Weinbauer *et al.*, 2011):

- La lyse du matériel cellulaire transformé en matière dissoute ou colloïdale augmente le temps de rétention à la surface des océans, ce qui laisse plus de temps aux organismes pour consommer le carbone et les nutriments.
- La formation d'agrégats issus des produits de lyse augmente l'exportation du carbone sauf si celle-ci entraîne une meilleure flottabilité des particules augmentant alors leur temps de rétention dans la zone euphotique
- Les nutriments rendus disponibles peuvent stimuler la production primaire entraînant diverses conséquences sur la pompe biologique

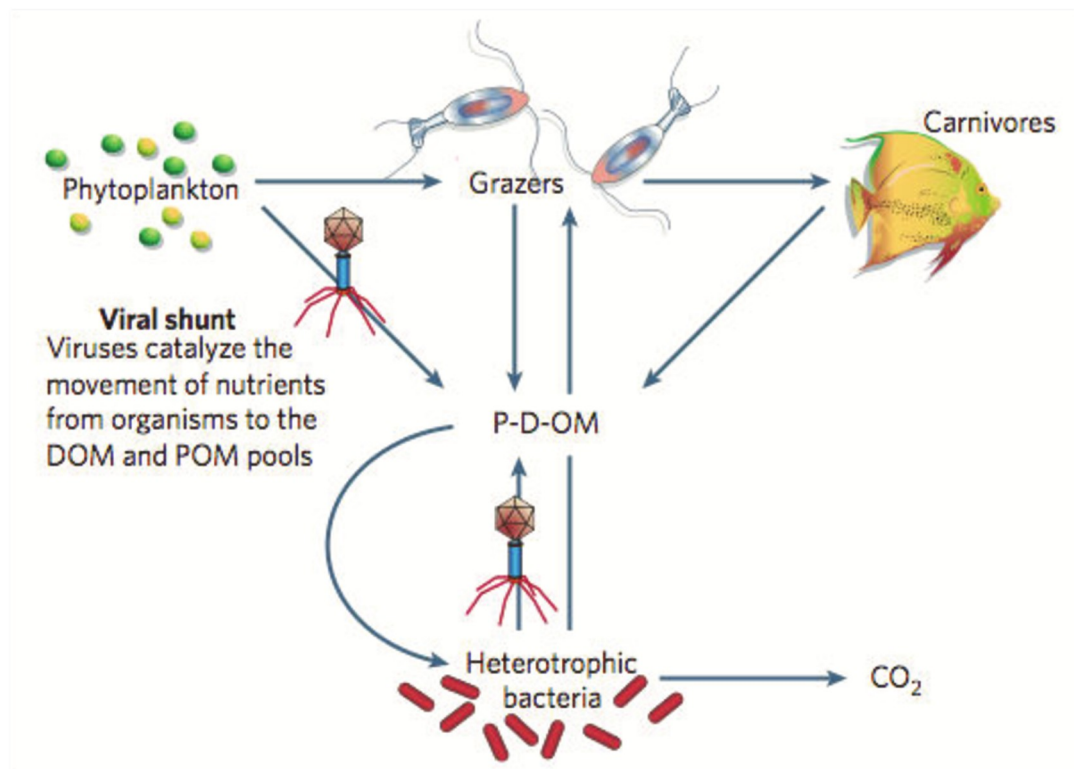


Figure 19: Illustration du viral shunt (Suttle, 2005). Les virus en lysant divers organismes comme le plancton ou les « brouteurs » provoquent un court-circuitage de la chaîne alimentaire et catalysent le mouvement des nutriments des organismes vers la matière organique dissoute ou particulaire.

Ainsi, l'impact des virus dans les océans ne se limite pas aux modulations des populations microbiennes mais peut aller jusqu'à une incidence sur le climat, notamment via l'arrêt des blooms phytoplanctoniques (aussi appelés efflorescences algales).

Exemple de *Emiliania huxleyi*-EhV : *E. huxleyi* est une coccolithophore, algue unicellulaire calcifiée responsable de vastes blooms dans les océans (notamment au nord de l'Atlantique) pouvant recouvrir jusqu'à plusieurs centaines de kilomètres carrés d'océans (Holligan *et al.*, 1993). Malgré sa petite taille, elle représente environ 1/3 de la production de CaCO_3 dans les océans, et est aussi un producteur majeur de DMS (Dimethyl sulfur) lors de sa dissolution dans l'eau de mer (généralement provoquée par la lyse virale). Le DMS se répand ainsi dans l'atmosphère et, une fois oxydé, est partiellement responsable de l'effet de serre impactant notamment la formation des nuages (Alcolombri *et al.*, 2015). En limitant la propagation de ces blooms et en lysant ces cellules, les virus jouent ainsi un rôle déterminant dans la régulation des flux de carbone et de sulfure dans les océans (Personnic *et al.*, 2006).

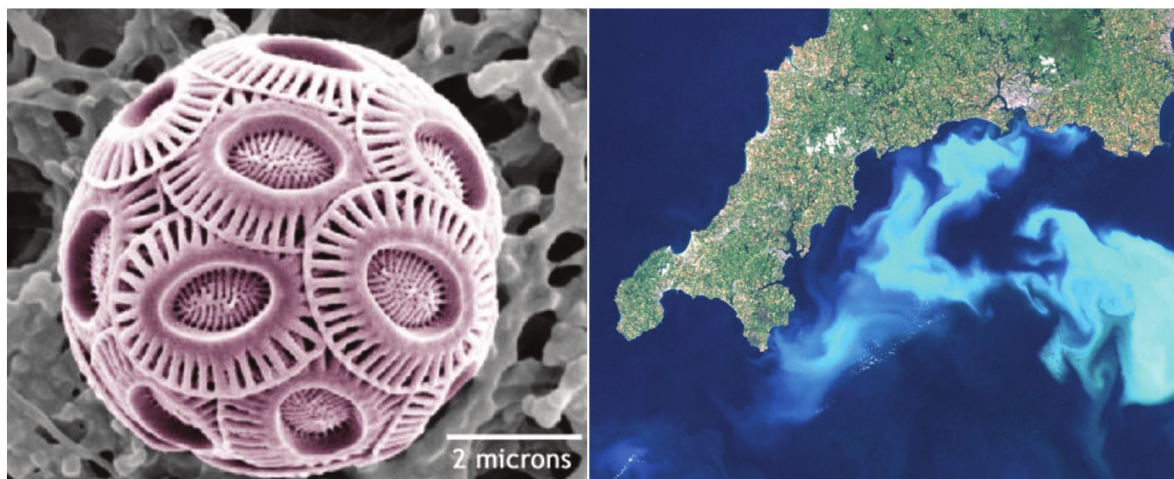


Figure 20 : Photographie de *Emiliania huxleyi*(a) au microscope électronique à balayage (Earthguide.image library, Coccolithophore – *Emiliania huxleyi*) (b) vue de l'espace en phase de bloom (*Emiliania huxleyi*| Collection | Expedition Arctic).

Transfert de gènes et manipulations des populations microbiennes

Les virus sont capables de moduler la composition des communautés microbiennes, notamment via le modèle « Killing the Winner », comme vu précédemment. Néanmoins, un autre impact majeur des virus découle de leur capacité à réaliser de nombreux transferts horizontaux ou latéraux (HGT pour Horizontal gene transfer) de gènes, apportant ainsi de nouvelles fonctions à leurs hôtes. En effet, on estime à 2.10^{16} le nombre de HGT médiés par les virus de micro-organismes chaque seconde dans les océans (Brüssow *et al.*, 2004).

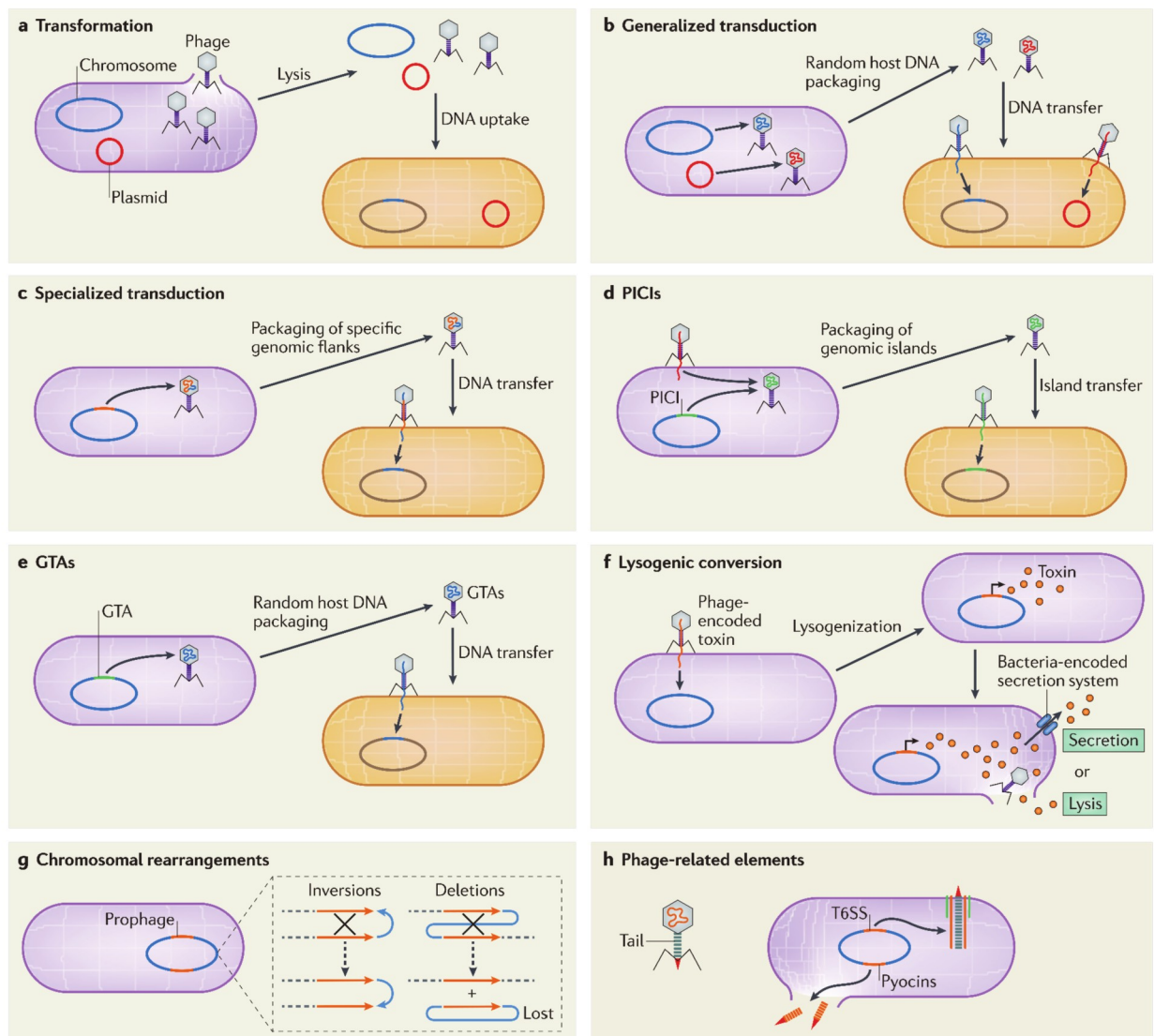


Figure 21 : Rôle des virus dans l'évolution et la pathogénicité de leurs hôtes (Salmond and Fineran, 2015).

Les virus peuvent être l'intermédiaire des HGTs par différentes voies (Fineran, 2015).

Tout d'abord, lors de la lyse virale, il y a relargage des acides nucléiques dans l'environnement. Les cellules présentes autour de la cellule lysée peuvent intégrer des morceaux d'acides nucléiques via le phénomène de transformation. (Figure 1.1)

Les virus sont aussi capables de transduction, qu'elle soit généralisée ou spécialisée (Figure 1.2a et c). La transduction généralisée peut être réalisée par des virus à ADN.

Lors de l'empaquetage des acides nucléiques dans la capsid, le virus englobe accidentellement de l'acide nucléique cellulaire au lieu du génome viral.

Après avoir infecté une nouvelle cellule, il injecte alors l'acide nucléique cellulaire. Ce nouvel acide nucléique peut alors rester sous forme libre (et être perdu lors de la génération suivante) ou être intégré au chromosome hôte par recombinaison.

La transduction spécialisée est, quant à elle, le fait des virus tempérés capables d'intégrer leur génome au chromosome de l'hôte. Lorsque le génome viral s'excise de façon imprévisible de l'hôte, il peut prendre une partie du chromosome de l'hôte avec son matériel génétique.

Après avoir infecté une nouvelle cellule, le virion parvient à intégrer un nouvel hôte. Il peut alors transmettre les gènes de son ancien hôte vers le nouveau.

Ces deux types de transductions sont capables de faciliter la mobilité des gènes de virulence ou de résistance aux antibiotiques.

Dans les environnements marins (Jiang and Paul, 1998; Penadés et al., 2015), la grande abondance de cellules lysogénisées dans l'environnement, les phénomènes de transduction sont probablement des acteurs majeurs des transferts horizontaux.

Dans la biosphère marine (Jiang and Paul, 1998).

En effet, 10 à 15% des gènes de bactéries et archées sont des plasmidique récemment intégrés au chromosome hôte (Cortezet al., 2009).

Le cas des PICIs (Phage inducible chromosomal islands, ou îlots chromosomiques inducibles) est un peu particulier car il met en jeu un virus dit chromosomiques peuvent « pirater » un virus pour promouvoir leur transfert drastiquement l'efficacité de leur transduction. Ce mécanisme est aussi appelé piratage moléculaire et est notamment connu chez *Staphylococcus aureus* avec les îlots de pathogénicité SaPIs (Novick et al., 2010; Christie and Dokland, 2012). Cette piratage moléculaire est maintenant aussi connue dans les océans profonds ce manuscrit (Lossouarn, Nesbø, et al., 2015)).

Concernant les GTAs (déjà évoqués précédemment, cf partie I de l'introduction) aussi médiés des HGTs. Il est possible que ces GTAs soient une évolution devenus défectifs (Figure 21 e).

L'un des exemples les plus connus de virus apportant une nouvelle infectée, est celui de *Vibrio cholerae* découvert en 1996 (Waldor and virus CTX permet à son hôte de coder pour un facteur de virulence toxine) rendant *Vibrio cholerae* particulièrement virulent et pathogène (Waldor and Mekalanos, 1996). Ce phénomène est appelé conversion lysogénique (Figure 21 f), une situation où le virus apporte de nouveaux gènes bénéfiques pour

Dans le cas où l'hôte porte deux provirus, si ceux-ci partagent des similitudes des réarrangements chromosomiques chez l'hôte en entraînant des inversions ou des délétions par exemple (Figure 21 g).

Ces réarrangements peuvent fortement moduler le fitness des hôtes entraînant une modification de leur compétitivité et par extension une modulation de la communautés microbiennes.

Les éléments apparentés aux virus (phage-related elements) font partie du cellulaire (Figure 21 h). Il s'agit notamment des pyocines de type R de type VI qui sont des systèmes cellulaires dérivés de structures initiales. Ces systèmes sont aussi présents dans certaines bactéries marines symbiotiques (*Pseudoalteromonas luteoviolacea*) qui libèrent des ensembles de structures (Figure 22) qui induisent une métamorphose chez les vers tubiformes Hydroides ulmaria, avec lesquelles elles vivent en symbiose. Ces données indiquent que l'utilisation de ces structures peut intervenir dans la communication avec l'hôte eucaryote (Shikuma et al., 2014).

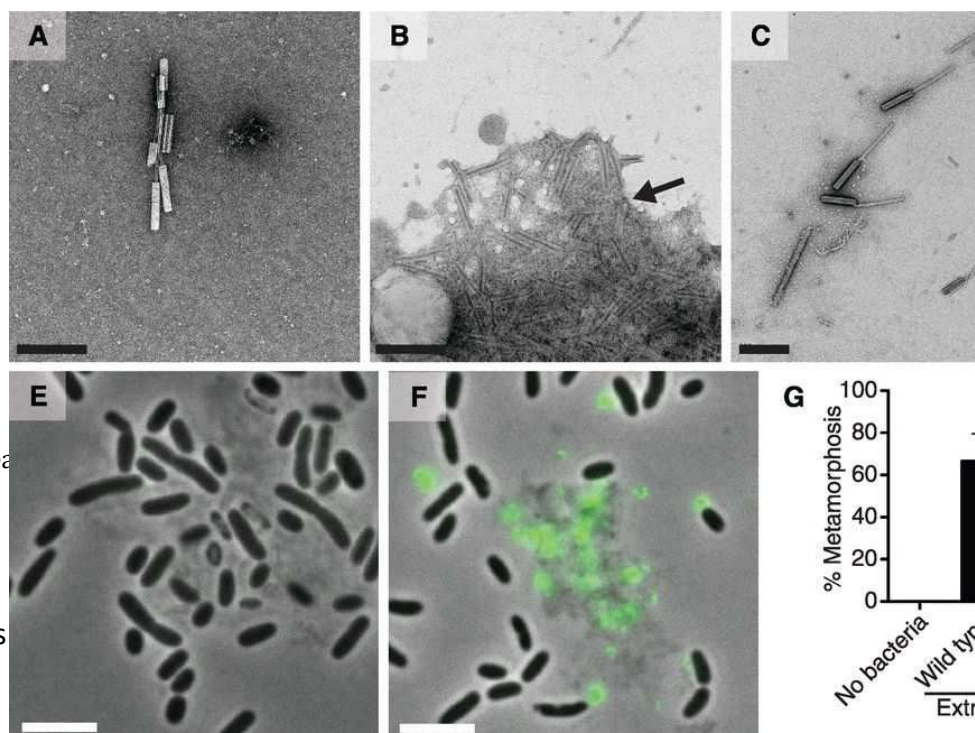


Figure 22 : Structures (*Pseudoalteromonas luteoviolacea* et al., 2014).

Outre ces divers effets des virus sont aussi capables

ils altèrent le métabolisme de l'hôte de façon à favoriser leur réplication allant même jusqu'à exprimer des gènes (de leur génome) supportant le l'hôte.

Ces gènes codant pour des protéines impliquées dans le métabolisme de l'hôte appelés « gènes auxiliaires » (ou « Virus-encoded auxiliary meta Ces vAMGs peuvent parfois être déjà présents dans le chromosome jusqu'alors inhibés ou peu produits. Ces gènes sont divers, et peuvent être nombreuses fonctions de l'hôte comme la photosynthèse, la voie des la régulation du phosphate, le métabolisme du soufre, la synthèse métabolisme des sphingolipides ou le métabolisme (processus) l'ADN et de l'ARN (Rosenwasser et al., 2016).

Exemple : Gènes de la photosynthèse chez les virus marins

L'un des exemples les plus connus de vAMGs au sein du viroplancton est codant pour des gènes impliqués dans la photosynthèse (Figure 23). Découvert système montre que S-PM2 (virus infectant des cyanobactéries) contient deux gènes majeurs de la photosynthèse. Il s'agit de *psbA* et *psbD*, deux protéines (D1 et D2) ayant un rôle central dans le photosystème II. cyanobactéries est généralement limitée par la photo-inhibition, en particulier océans. Les protéines D1 et D2 jouent un rôle important dans la l'excès de protéine D1 photo-inactivée et endommagée qui l'entraîne, amenant diminution de l'efficacité du photosystème. D'autre part, durant l'infection D1 cellulaire chute fortement entraînant une baisse de l'apport d'énergie. Néanmoins, l'expression des protéines D1 et D2 codées par les virus permet le photosystème efficace, passant outre la photo-inactivation, donnant ainsi plus virus pour qu'ils puissent se répliquer massivement et libérer un grand nombre dans l'environnement (Mann et al., 2003; Lindell et al., 2004).

Suite à cette étude, ces gènes codés par des virus et impliqués dans la photosynthèse se sont révélés particulièrement répandus chez les virus infectant des cyanobactéries (Sullivan *et al.*, 2006; Breitbart *et al.*, 2007).

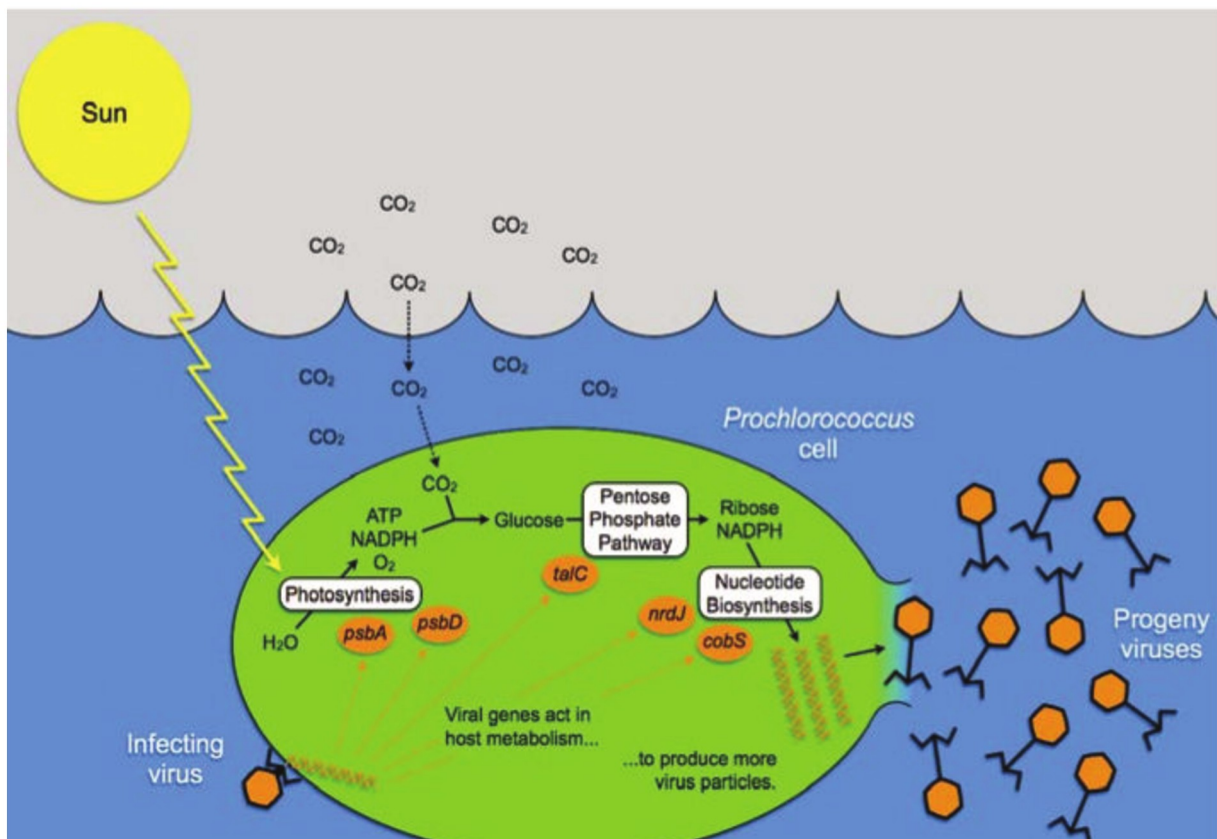


Figure 23 : Illustration de l'utilisation des gènes « viraux » de la photosynthèse pour supporter la production virale lors d'une infection de *Prochlorococcus* (Breitbart *et al.*, 2007). Trois voies métaboliques supportées par des gènes d'origine virale (orange) sont illustrées ici, la photosynthèse, la voie des pentoses-phosphates et la synthèse des nucléotides. L'expression de ces gènes durant l'infection permet de soutenir le métabolisme de l'hôte augmentant *in fine* la production virale.

III. Les environnements thermophiles

Les thermophiles, organismes vivants à des températures optimales au-dessus de 45°C font partie d'un groupe plus vaste appelé *extrophiles*. Cette notion est profondément anthropomorphique puisque nous appelons « environnements extrêmes » les biotopes qui nous seraient hostiles. Or, pour ces micro-organismes, ce sont nos conditions de vies qui sont extrêmes. Ces organismes très divers, vivent sous au moins une condition physico-chimique (température, pression osmotique, pression hydrostatique, pH, concentration en oxygène... paraissant être en hors des limites compatibles de la vie non extrémophiles. Les environnements qui sont considérés comme extrêmes peuvent par exemple être les zones polaires, les étendues désertiques (chaudes ou froides), les fonds marins, les lacs de soude, lacs acides ou hypersalés (Figure 24).



Figure 24 : (gauche) Désert du Sahara (source : commons.wikimedia.org) (droite) Antarctique (source : laterredufutur.com).

Il a longtemps été considéré que la vie à haute température était impossible car dénaturant les protéines. La chaleur est d'ailleurs l'un des paramètres de sélection les plus importants et aussi le plus pénible. Pourtant, c'est la condition de vie même que nous ne connaissons pas des micro-organismes thermophiles et hyperthermophiles.

La limite de croissance de *Bacillus stearothermophilus*, la plus haute température (Stark and Tetrault, 1951; Prieur, 2014). Néanmoins, les travaux de Thomas Brock et son équipe dans les années 70 ont révolutionné nos connaissances sur ces micro-organismes. En effet, ils ont isolé les premières souches microbiennes ayant une température optimale de croissance au-dessus de 60°C après l'habitillage de sources chaudes terrestres du Parc national de Yellowstone aux Etats-Unis (Madigan and Martinko, 1995; Prieur, 2014). Aujourd'hui, les thermophiles se divisent en trois sous-catégories : les thermophiles modérés (45 à 60°C), les thermophiles extrêmes (60 à 85°C) et les hyperthermophiles (au-dessus de 85°C) (Mehta and Madigan, 2014). A l'heure actuelle, la limite de la vie est fixée à 122°C, température à laquelle une méthanogène piezophile décrite en 2008 peut encore proliférer sous 20 MPa de pression (Takai et al., 2008). Depuis les travaux fondateurs de Brock et Woese, l'intérêt de la communauté scientifique pour les thermophiles a considérablement augmenté. En effet, outre leurs divers intérêts biotechnologiques (le plus connu étant la Taq Polymérase de *Thermus aquaticus* (Chien et al., 1976)), ces micro-organismes sont aussi étudiés afin de permettre une meilleure compréhension des cycles biogéochimiques globaux, du rôle des micro-organismes à l'échelle planétaire, dans le cadre des recherches sur les origines de la vie... Globalement, l'étude de leur écologie et de leur physiologie est nécessaire car encore trop peu de données sont disponibles sur ces organismes (Forterre, 2007; Prieur, 2014).

x Les sources chaudes terrestres

Les sources chaudes terrestres sont un grand réservoir de thermophiles. Elles se trouvent généralement dans les zones à activité volcanique, où les environnements chauds affleurent à la surface. En effet, dans ces zones, l'eau s'infiltre dans la croûte terrestre, se chauffe sous l'effet d'un chauffage magmatique mais elle n'est pas soumise à la pression lithostatique. Sous l'effet de la chaleur, la pression augmente faisant jaillir à la surface l'eau chargée de minéraux et autres éléments (Mehta and Satyanarayana, 2013; Prieur, 2014).

Ces manifestations hydrothermales sont présentes dans divers endroits du globe mais les plus connues restent le Parc National de Yellowstone aux USA (Figure 25), la péninsule de Kamchatka en Russie, le Chili, le Japon, la Nouvelle-Zélande, l'Italie et l'Islande. Elles peuvent se manifester sous forme de solfatares, de geysers ou de sources chaudes (Mehta and Satyanarayana, 2013).

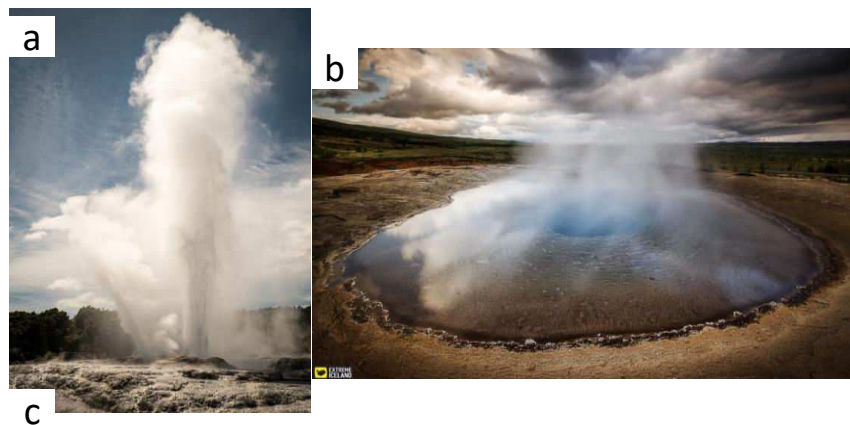


Figure 25 : Photographie de geysers (a) en Nouvelle-Zélande (source : eugene.kaspersky.fr) (b) en Islande (source : extremeiceland.is). (c) Photographie de source chaude à Yellowstone (USA, source : yellowstoneparknet.com).

Les micro-organismes isolés des sources chaudes

Les sources chaudes terrestres sont un réservoir très important de biodiversité bactérienne et archéenne. L'apport de nutriments et les gradients physico-chimiques présents laissent l'espace à une grande diversité d'espèces microbiennes. Nous avons cette diversité grâce à l'usage de méthodes d'amplification à partir de l'ARN (López et al., 2013). Les groupes de micro-organismes majoritaires de cet écosystème sont le phylum des Crenarchaeota pour les archées et l'ordre des Aquificales pour les bactéries. Deux phyla plus récents ont aussi été découverts dans ces sources, appartenant au domaine Archaea: les Korarchaeota et les Nanoarchaeota. Les membres de ces deux phyla, tout comme les Aquificales font partie des embranchements les plus profondes de l'arbre phylogénétique (López et al., 2013). La composition des communautés de ces écosystèmes est très variable et dépend notamment du type de source chaude étudiée.

Exemple des champs de solfatares

Les solfatares sont des fumerolles qui émettent des gaz soufrés. Ils sont retrouvés dans des « champs de solfatares » qui comprennent des sols, boues et eaux de surfaces (riches en soufre), chauffés par la chambre magmatique se situant en dessous, pouvant atteindre des températures de l'ordre de 100 °C. Leurs paramètres physico-chimiques peuvent varier d'un champ de solfatares à un autre, avec des pH allant de 2 à 10. L'acidité est due à la présence d'acide sulfurique formé suite à l'oxydation biologique du sulfure d'hydrogène et du soufre. Les solfatares sont généralement très riches en minéraux de fer, émettent du dioxyde de carbone et ont des concentrations variables en dihydrogène, méthane, azote et monoxyde de carbone (Mehta and Satyanarayana, 2013).

Les communautés de bactéries et d'archées qui colonisent ces environnements sont variées, pouvant être acidophiles, thermophiles extrêmes ou modérées. La plupart de ces populations sont hétérotrophes (comme *Desulfurococcus mobilis* ou *Thermoproteus uzoniensis*) ou chimioautolithotrophes (comme *Pyrobaculum islandicum* ou *Thermoproteus neutrophilus*), utilisant souvent le soufre dans leur métabolisme énergétique. Ces solfatares peuvent permettre la croissance de souches aérobies comme *Sulfolobus solfataricus*, anaérobies comme *Thermoproteus tenax* ou encore des souches AAF (aéro-anaérobie facultatif) comme *Acidianus infernus* (Seeger et al., 1986; Mehta and Satyanarayana, 2013). Des bactéries comme *Fervidobacterium islandicum* ou *Thermotoga thermarum* appartiennent à l'ordre des Thermotogales colonisent aussi ces environnements mais sont moins souvent isolées de ceux-ci que les archées (Windberger et al., 1989; Huber et al., 1990; Huber and Hannig, 2006).

Exemple : **Soufre chauds de l'Okidani** au Japon

Aujourd'hui encore, la recherche des micro-organismes est incultivée. Les études en microbiologie sont toujours en grande partie basées sur des échantillons de soufre qui n'ont pas encore pu être étudiés ultérieurement. Une approche expérimentale a été mise en œuvre dans une étude récente. Les auteurs utilisent des échantillons naturels congelés et les soumettent à diverses conditions de culture. Cette méthode a permis de cultiver et d'isoler de nouvelles espèces issues de sources chaudes acides (68°C, pH 1,4) (Sakai and Kurosawa, 2016). Ces espèces appartiennent au genre *Sulfolobus* (HS-1 et HS-3) qui est bien représenté dans les sources chaudes. Une dizaine de phylotypes d'archées ont été obtenus dans les enrichissements réalisés dans cette étude, majoritairement des *Sulfolobus*, mais aussi des *Thermoplasma*, des *Metallosphaera* et des *Sulfurisphaera*... (Figure 26).

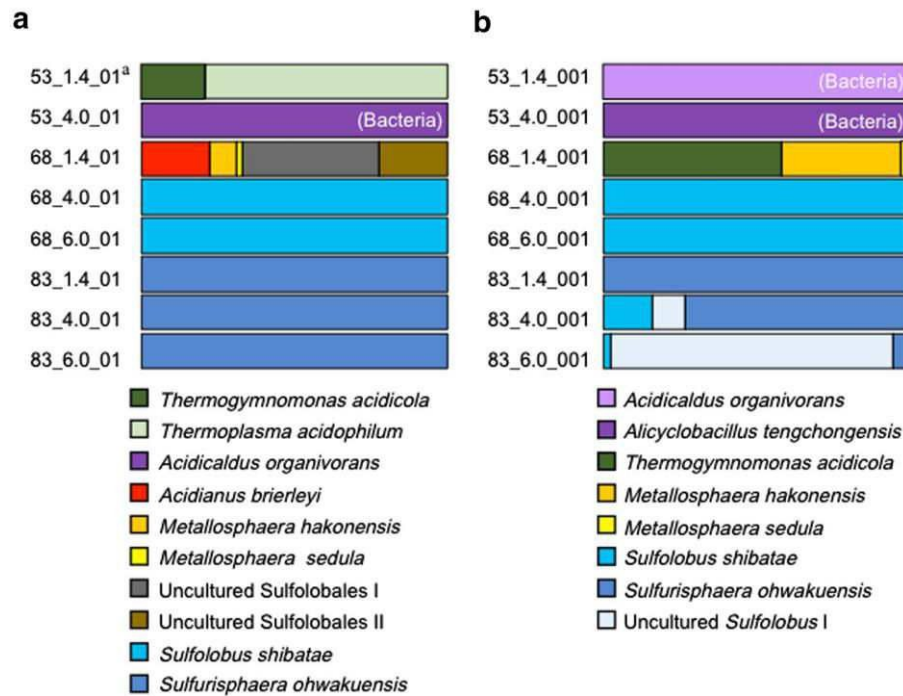


Figure 26 : Diversité des communautés microbiennes retrouvée dans des enrichissements réalisés à partir d'échantillons de sources chaudes terrestres (Saito and Kurosawa, 2016). Les chiffres à gauche des figures représentent la température de culture, le pH et le pourcentage en extrait de levure (a) 0, % b , %. Cette figure révèle notamment qu'il y aurait une prépondérance de bactéries parmi les micro-organismes cultivés ici.

Virus associés aux bactéries et archées des sources chaudes terrestres

Dans les écosystèmes où la pression d'eutrophisation se fait sentir, les virus sont le facteur principal de mortalité des communautés microbiennes (Schoenfeld et al., 2008; López-López et al., 2010). Ils influencent donc la structure, la diversité et la composition de ces populations. Pour autant, il y a encore trop peu de systèmes hôtes-virus caractérisés dans ces environnements. En effet, une vingtaine d'archéovirus et une quinzaine de bactériovirus ont été bien décrits à ce jour. Néanmoins, leur diversité morphologique et génomique est impressionnante en particulier concernant les archéovirus (Rachel et al., 2002; Basta et al., 2009; Prangishvili et al., 2016).

Bactériovirus

Il existe aujourd'hui sept espèces de virus thermophiles (entre 55 et 75°C) issus de sources chaudes bien décrits dans la littérature (Uldahl and Peng, 2013; Nagayoshi et al., 2016).

Néanmoins, lors d'une étude récente, nous avons isolé et partiellement décrits, mais leur caractérisation n'a pas été possible, ils sont passés inaperçus et caractérisés. Ces virus infectent des bactéries appartenant au genre *Thermus*, mais aussi aux genres *Geobacillus*, *Meiothermus* et *Rhodothermus* (Yu et al., 2006; Uldahl and Peng, 2013)(Tableau 3).

Parmi les 14 systèmes décrits seuls certains seront détaillés et pour permettre une meilleure lisibilité, les autres sont décrits dans le tableau ci dessous.

Tableau 3 : Résumé des bactériovirus des sources chaudes terrestres. Les étoiles indiquent que ces virus font l’objet d’un développement dans le corps du texte.

Famille / Morphologie	Virus	Hôte	Origine	Génome	Caractéristiques	Références
Myoviridae Capside icosédrique, queue contractile	F YS40*	Thermus thermophilus	Source chaude, Japon	Linéaire 152,7 kb	Cycle lytique Partage 155 ORFs avec F TMA	(Sakaki and Oshima, 1975).
	F TMA*	Thermus thermophilus HB27	Source chaude, Japon	Linéaire 151,5 kb	Cycle lytique Partage 155 ORFs avec F YS40	(Tamakoshi et al., 2011)
	TS2126*	Thermus scotoductus	Source chaude, Islande	Linéaire 90 kb	RNA ligase 1 similaire à celle du bactériovirus T4	(Blondal et al., 2005)
	RM378*	Rhodothermus marinus	Source chaude, Islande	Linéaire 130 kb	RNA ligase 1, plynucleotide kinase et protéine de capside gp23 similaires à celles du bactériovirus T4	(Hjorleifsdottir et al., 2014)
MMP17 (Methanothermobacter phage 17)	MMP17 (Methanothermobacter phage 17)	Methanothermobacter	Source chaude, Chine	33,5 à 39,5 kb	Cycle lytique	(Lin et al., 2010)
	GBSV1*	Geobacillus sp. 6K51	Source chaude, Chine	Linéaire 34,6 kb		(Liu et al., 2009)
	BV1*	Geobacillus sp. 6K512	Source chaude, Chine	Linéaire 35 kb	Génomes ayant des similarités avec un pathogène humain	(Liu et al., 2010)

Tableau 3 : Résumé des bactériovirus des sources chaudes terrestres. Les étoiles indiquent que ces virus font l’objet d’un développement dans le corps du texte (suite).

Famille / Morphologie	Virus	Hôte	Origine	Génome	Caractéristiques	Références
<i>Siphoviridae</i> , Capside icosaoédrique, très longue queue flexible	TSP4 (<i>Thermus</i>	<i>Thermus</i>	Source chaude,	80 kb	Cycle lytique	(Lin <i>et al.</i> , 2010)
	<i>siphoviridae</i> phage 4)		Chine		1 ORF codant pour une protéine similaire à DnaB des bactériovirus P74-26 et P24-46	
	P23-45*	<i>Thermus thermophilus</i> HB8	Source chaude acide, Kamchatka,	84 kb	Cycle lytique	(Minakhin <i>et al.</i> , 2008)
	P74-26*	<i>Thermus thermophilus</i> HB8	Source chaude acide, Kamchatka, Russie	83 kb		
<i>Tectiviridae</i> , capsides icosaoédriques contenant une membrane interne lipidique	P23-77*	<i>Thermus thermophilus</i> ATCC 33923	Source chaude Nouvelle-Zélande	Linéaire 17 kb	Cycle lytique	(Jaatinen <i>et al.</i> , 2008)
	F IN93	<i>Thermus thermophilus</i> TZ2	Source chaude, Japon	Linéaire 20 kb	Protéine de capside particulière	
					Cycle lysogénique	(Matsushita and Yanase, 2009)
<i>Innoviridae</i> , longs filaments	PH75	<i>Thermus thermophilus</i> HB8	Source chaude, Nouvelle-Zélande	Linéaire 6,5 kb	Non virulent	(Pederson <i>et al.</i> , 2001)
					ADN simple brin	
	F OH3	<i>Thermus thermophilus</i> HB8	Source chaude, Japon	Linéaire 5,7 kb	Capside composée d'une protéine répétée plusieurs milliers de fois	(Nagayoshi <i>et al.</i> , 2016)

Contrairement à la diversité morphologique impressionnante retrouvée pour les archéovirus, les bactériovirus issus des sources chaudes terrestres se répartissent en 4 morphologies, correspondant à 4 familles de virus : les *Myoviridae* et *Siphoviridae*, de l'ordre des *Caudovirales*, les *Tectiviridae* et les *Inoviridae*. Les *Myoviridae* et les *Siphoviridae* ont la morphologie classique dite « tête-queue » comme la majorité des bactériovirus alors que les *Inoviridae* sont filamenteux et que les *Tectiviridae* sont des virus non caudés à capside icosaédrique (Yu *et al.*, 2006; Ackermann and Prangishvili, 2012; Uldahl and Peng, 2013). Néanmoins, il est intéressant de noter que, dans l'étude portant sur les 115 virus isolés de sources chaudes alcalines (Islande, Japon, USA, Russie), plus de 50% des virus retrouvés étaient non caudés (capside icosaédrique ou filamenteux) (figure 27) (Yu *et al.*, 2006). Les *Caudovirales* semblent proportionnellement moins représentés dans ces sources chaudes alcalines que dans les autres écosystèmes. Les auteurs émettent l'hypothèse que cela est dû au grand nombre de souches de *Thermus* préférentiellement infectées par des *Inoviridae* ou des *Tectiviridae* (Yu *et al.*, 2006; Ackermann and Prangishvili, 2012).

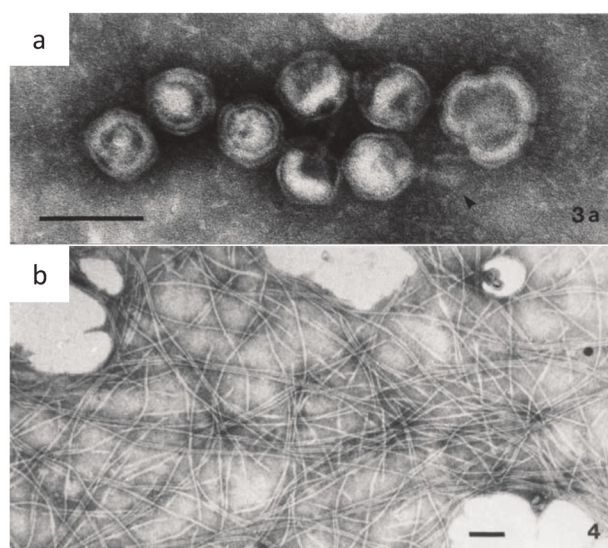


Figure 27 : Photographie de MET de bactériovirus isolés de sources chaudes terrestres (a) Tectiviridae (b) Inoviridae. Echelle 100 nm (Yu *et al.*, 2006).

Les Myoviridae se caractérisent par une capsidie icosaédrique à laquelle est attachée une queue contractile, ils sont généralement icosahédriques, et ont une queue contractile.

Le premier virus thermophile infectant une souche issue de source chaude fut le myovirus

) YS40. Isolé au Japon, il infecte via un cycle lytique la souche *Thermus thermophilus* HB8 (Sakaki and Oshima, 1975).

) TMA est aussi un Myoviridae isolé d'une souche chaude au Japon. Il est associé à *Thermus thermophilus* HB27, qui comme HB8 croît de façon optimale à 65°C.

) TMA et) YS40 partagent plusieurs similitudes. Tout d'abord, ils possèdent des morphologies très similaires avec une capsidie icosaédrique de 125 nm de diamètre et une queue contractile de 178 nm de long pour 27 nm de large. Ensuite, ils possèdent de grands génomes. En effet, leurs ADN doublement hélicoïdaux sont respectivement de 180 et 185 kbp. Leur taille n'est pas leur seule particularité puisqu'ils partagent ORFs dont les séquences ne présentent pas de similarités dans les bases de données) et leur organisation génomique est similaire. Il semble aussi que ces deux virus nécessitent la présence de pili sur la souche HB8 pour s'adsorber sur cette cellule et injecter leur génome. Cette caractéristique est aussi utilisée par les Inoviridae (Stassen et al., 1994; Naryshkina et al., 2006; Tamakoshi et al., 2011).

Les bactériovirus de thermophiles sont un réservoir de nouvelles protéines et notamment d'enzymes thermostables dont de nombreuses applications biotechnologiques (Uldahl and Peng, 2013).

C'est dans le cadre de l'étude des myovirus TS2126 et RM378 ont été découverts (Hjörleifsdottir et al., 2002; Blondal, Thorisdottir, et al., 2005; Hjörleifsdottir et al., 2014).

Pour ces deux virus, une RNA ligase 1 montrant des similarités avec celle du phage T4 a été mise en évidence. Cette RNA ligase est chez le phage T4, impliquée dans le contournement d'un cis-actin-ribose plaqué par l'hôte. Ce cis-actin est assujé à la ligation d'ARN de transferts (ARNt) viraux. La ligase a pour objectif de réparer ces ARNt. Elle est associée à une polynucléotidylotransférase PNK qui effectue des modifications sur l'ARNt pour le transformer en substrat disponible pour la ligation. Néanmoins, la RNA ligase de T4 n'est pas associée à une PNK connue. En effet, aucun ORF du génome à ADN double brin linéaire de 90 kb de T4 n'a été identifié à la PNK de T4. Nous avons donc recherché si une PNK différente soit présente dans son génome. D'autre part, l'activité de cette ligase s'est révélée particulièrement efficace pour effectuer des ligations inter ou intra moléculaires, sur des ARN, ADN simple et double brin (Blondal, Thorisdottir, et al., 2005). En effet, la présence d'une PNK et d'une RNA ligase similaires à celles de T4 ont été identifiées dans le génome à ADN double brin de 130 kb de RM378 (Blondal et al., 2003; Blondal, Hjörleifsdottir, et al., 2005; Hjörleifsdottir et al., 2014).

GBSV1 et BV1 ont eux aussi été isolés de souches provenant de sources chaudes en Chine. Ce sont deux myovirus infectant deux souches de *Geobacillus* sp. 6K51 pour GBSV1 et *Geobacillus* sp. 6K512 pour BV1. Leurs hôtes ont une température de croissance optimale de 65°C. Les virions sont dotés d'une capside icosaédrique de diamètre de 25 nm et d'une queue contractile de 10 nm de long. Leurs génomes sont composés d'ADN double brin linéaire de 34,6kb (GBSV1) et 35kb (BV1).

Ils contiennent tous les deux 54 ORFs qui présentent de fortes similarités avec des *Geobacillus* et *Bacillus* ainsi que les virus qui leurs sont associés. Ces souches de *Geobacillus* et *Bacillus* sont majoritairement issues des sols mais il est à noter que plusieurs ORFs de ces virus ont des similarités avec des gènes de *Geobacillus kaustophilus*, bactérie thermophile notamment retrouvée au niveau de la fosse des Mariannes (Takami et al., 2004). Fait intéressant, 7 à 8 ORFs de ces génomes partagent de fortes similarités avec des protéines produites par *Listeria innocua*. Cette bactérie est un pathogène intracellulaire opportuniste humain. Ces similarités laissent à penser qu'il existe un lien évolutif potentiel entre ces virus et le pathogène (Wang and Zhang, 2008; Liu et al., 2009, 2010).

Parmi les Siphoviridae, 3 systèmes ont été décrits. Cette famille de virus se caractérise par la présence d'une capsid hexaédrique à laquelle est attachée une queue flexible contractile. Les trois siphovirus infectent des souches de *Thermus*.

Les deux siphovirus, P23-45 et P74-26 infectent via un cycle lytique la même souche *Thermus thermophilus* HB8. Ils ont été isolés de sources chaudes acides dans la péninsule de Kamchatka en Russie. Les virions sont similaires et morphologiquement proches de TSP4 (*Thermus siphoviridae* phage) puis qu'ils ont une capsid hexaédrique de diamètre de 25 nm et une queue flexible de 100 nm de long. Les capsides contiennent des gènes d'ADN double brin de 84 kb pour P23-45 et 83 kb pour P74-26. Ils comprennent respectivement 117 et 116 ORFs dont 95% sont similaires. Ces génomes présentent une caractéristique inhabituelle. En effet ils contiennent, dans leurs régions intergéniques, de grandes séquences répétées polypurine-polypyrimidine. Ces séquences sont fréquemment retrouvées chez les eucaryotes mais rarement chez les bactéries ou archées. Elles forment des triples hélices stables qui jouent le rôle de la pléiotomie et la transposition, permettant l'insertion du génome viral.

Les expériences réalisées *in vitro* montrent que ces séquences forment bien un ADN triple brin ce qui bloque la réplication des virus. Cette conformation particulière pourrait donc être impliquée dans la régulation de l'expression des virus. Autre fait intéressant, une protéine potentielle de liaison à l'ARN polymérase a été identifiée dans le génome de P23-45. Celle-ci se situe à proximité de l'origine de l'hostilité de l'hostilité de l'hostilité des gènes au (Yu et al., 2006; Minakhin et al., 2008; Berdyuglova et al., 2011).

La famille des Tectiviridae est caractérisée par des virions aux capsides icosaédriques contenant une membrane interne lipidique. Les deux virus connus sont P23-77 et IN93. P23-77 infecte *Thermus thermophilus* isolé d'une source chaude en Nouvelle-Zélande et IN93 infecte *Thermus thermophilus* isolé d'une source chaude au Japon. Ces hosties se développent respectivement à 70 et 65°C. P23-77 est virulent alors que IN93 est tempéré comme l'indique la présence de gènes impliqués dans la lysogénie des *S. typhimurium* (Jaatinen et al., 2008; Matsushita and Yanase, 2009). Les protéines de capsides de P23-77 ont été particulièrement étudiées (Jaatinen et al., 2008; Jalasvuori et al., 2010; Rissanen et al., 2013). En effet, les capsides de ces deux virus présentent des similarités avec celles du virus SH1 infectant l'archaéophile *Haloarcula hispanica*. Trois gènes communs ont été mis en évidence dans leurs génomes, deux codant pour les protéines de capsides mineurs et majeurs et un codant pour une ATPase potentielle impliquée dans l'empaquetage du génome viral. D'autre part, la formation des capsides des virus est particulière. En effet, chez "H. o" pour les deux virus présentés ici, la formation de la capside se base sur une symétrie T=28 et des capsomères de décoration en forme de tours sont présents à la surface des virions. La conformation des protéines de capside serait basée sur de multiples feuillets bêta.

Cette conformation semble particulièrement bien conservée et est présente dans différents virus P23-77 like. Elle serait ancienne, révélant un possible ancêtre commun lointain pour ces virus. C'est sûrement parce qu'il a été proposé d'illustrer les deux Tectiviridae dans les Sphaerolipoviridae ou peut-être d'autres virus d'archéobactéries halophiles dont "H" (Jaatinen et al., 2008; Jalasvuori et al., 2010; Rissanen et al., 2012, 2013; Pawlowski et al., 2014).

Pour terminer, la famille Inoviridae se caractérise par des virions ayant une structure en forme de longs filaments sortant de leurs hôtes par extrusion sans provoquer la lyse cellulaire. Cette famille comprend deux virus : PH75 et) OH3 qui portent un ORF commun.

Archéovirus

La diversité morphologique et génomique des virus infectant les archées est impressionnante et particulièrement chez les hyperthermophiles. “eizeo photpes d’a h o i usont été décrits à ce jour (dont 4 sont uniques) alors que parmi les plus de 6000 bactériovirus connus, seuls 9 o photpes o t t d its et au u ou eau o photpe ’a t d ou e t depuis les années 70 (Ackermann and Prangishvili, 2012; Pietilä et al., 2014; Uldahl, 2015). Par ailleurs, la vaste majorité des gènes composant leurs génomes restent sans similarités avec les séquences présentes dans les bases de données et/ou ont des fonctions inconnues.

Un peu plus de 80 archéovirus ont été décrits à ce jour. La majorité est représentée par des Crenarchaeota hyperthermophiles appartenant aux genres *Sulfolobus* et *Acidianus* ou aux Euryarchaeota halophiles des genres *Haloarcula* et *Halorubrum* (Uldahl, 2015). Parmi les archaeovirus des sources chaudes, 11 morphologies différentes sont retrouvées et réparties en différentes familles (Tableau 4). Comme pour les bactériovirus, pour une meilleure lisibilité, seule est présentée la liste des noms des archéovirus caractérisés à ce jour est présenté ci-dessous.

Tableau 4 : Résumé des archéovirus des sources chaudes terrestres. Les étoiles indiquent que ces virus font l’objet d’un développement dans le corps du texte.

Famille / Morphologie	Virus	Hôte	Origine	Génome	Caractéristiques	Références
<i>Ampullaviridae</i> Bouteille	ABV (<i>Acidianus</i> bottle-shaped virus)*	<i>Acidianus convivator</i>	Source chaude acide, Italie	Linéaire 23,9 kb	Non virulent	(Håring et al., 2005)
					Homologue DNA pol B	
<i>Bicaudaviridae</i> 2 queues (ou broche)	ATV (<i>Acidianus</i> two-tailed virus)	<i>Acidianus convivator</i>	Source chaude acide, Italie	Circulaire 62,7 kb	Cycle lysogénique Développement extracellulaire	(Prangishvili et al., 2006)
					Intégrase	
<i>Clavaviridae</i> Bâtonnets Une extrémité arrondie et l'autre pointue	APBV1 (<i>Aeropyrum</i> permix bacilliform virus 1)	<i>Aeropyrum permix</i>	Source chaude côtière, Japon	Circulaire 5,2 kb	Etat porteur	(Mochizuki et al., 2010)
<i>Globuloviridae</i> Sphérique, enveloppé	PSV (<i>Pyrobaculum</i> spherical virus)	<i>Pyrobaculum</i>	Source chaude Yellowstone, USA	Linéaire 28,3 kb	Non virulent 15 ORFs communs	(Håring et al., 2004)
	TTSV1 (<i>Thermoproteus</i> tenax spherical virus 1)	<i>Thermoproteus tenax</i>	Source chaude, Indonésie	Linéaire 21,6 kb	Les autres ORFs montrent peu de similarités avec les séquences disponibles dans les bases de données	(Ahn et al., 2006)
<i>Guttaviridae</i> Goutte, enveloppé, fibres caudales	SNDV (<i>Sulfolobus</i> <i>neozelandicus</i> droplet-shaped virus)	<i>Sulfolobus neozelandicus</i>	Champ de solfatares, Nouvelle-Zélande	Circulaire 20 kb	Etat porteur Méthylation dam Virus non disponible en laboratoire	(Arnold, Ziese, et al., 2000)

Tableau 4 : Résumé des archéovirus des sources chaudes terrestres. Les étoiles indiquent que ces virus font l’objet d’un développement dans le corps du texte (suite).

Famille / Morphologie	Virus	Hôte	Origine	Génome	Caractéristiques	Références
Fuselloviridae	ASV1 (<i>Acidianus</i> spindle-shaped virus 1)	<i>Acidianus brierleyi</i>	Source chaude acide, Islande	Circulaire 24,2 kb		(Redder <i>et al.</i> , 2009)
	SMF1 (<i>Sulfolobales</i> Mexican fusellovirus 1)	<i>Sulfolobus</i>	Source chaude acide, Mexique	Circulaire 14,8 kb		(Servín-Garcidueñas <i>et al.</i> , 2013a)
	SSV1 (<i>Sulfolobus shibatae</i> virus 1)*	<i>Sulfolobus shibatae</i>	Source chaude acide, Beppu, Japon	Circulaire 15,5 kb		(Palm <i>et al.</i> , 1991)
	SSV2 (<i>Sulfolobus</i> spindle shaped virus 2)	<i>Sulfolobus solfataricus</i>		Circulaire 14,8 kb	Non virulents Génomes similaires	(Stedman <i>et al.</i> , 2003)
	SSV4 (<i>Sulfolobus</i> spindle shaped virus 4)	<i>Sulfolobus islandicus</i>		Circulaire 15,1 kb	13 ORFs communs Intégrase	(Peng, 2008)
	SSV5 (<i>Sulfolobus</i> spindle shaped virus 5)	<i>Sulfolobus solfataricus</i>	Source chaude acide, Islande	Circulaire 15,3 kb	Protéines de capsides VP1 et 3 Régulateurs de transcriptions	(Redder <i>et al.</i> , 2009)
	SSV6 (<i>Sulfolobus</i> spindle shaped virus 6)	<i>Sulfolobus solfataricus</i>		Circulaire 15,6 kb	Homologue à DnaA	(Redder <i>et al.</i> , 2009)
	SSV7 (<i>Sulfolobus</i> spindle shaped virus 7)	<i>Sulfolobus solfataricus</i>		Circulaire 17,6 kb		(Redder <i>et al.</i> , 2009)
	SSVK1 (<i>Sulfolobus</i> spindle shaped virus K1)	<i>Sulfolobus</i>	Source chaude acide,	Circulaire 17,4 kb		(Wiedenheft <i>et al.</i> , 2004)
	SSVRH (<i>Sulfolobus</i> spindle shaped virus RH)	<i>Sulfolobus</i>	Kamchatka, Russie	Circulaire 16,4 kb		(Wiedenheft <i>et al.</i> , 2004)
Fusifforme (ou Citron), fibres caudales						

Tableau 4 : Résumé des archéovirus des sources chaudes terrestres. Les étoiles indiquent que ces virus font l’objet d’un développement dans le corps du texte (suite).

Famille / Morphologie	Virus	Hôte	Origine	Génome	Caractéristiques	Références
Lipothirixiridae Filamenteux, enveloppe lipidique, structures particulières aux extrémités	AFV1 (<i>Acidianus</i> filamentous virus 1)	<i>Acidianus hospitalis</i>	Source chaude acide, Yellowstone, USA	Linéaire 21,1 kb		(Bettstetter <i>et al.</i> , 2003)
	AFV2 (<i>Acidianus</i> filamentous virus 2)	<i>Acidianus</i>	Source chaude acide, Italie	Linéaire 31,8 kb		
	AFV3 (<i>Acidianus</i> filamentous virus 3)	<i>Acidianus</i>	Source chaude acide, Italie	Linéaire 40,5 kb		
	AFV6 (<i>Acidianus</i> filamentous virus 6)	<i>Acidianus convivator</i>	Source chaude acide, Italie	Linéaire 39,6 kb		(Vestergaard, Aramayo, <i>et al.</i> , 2008)
	AFV7 (<i>Acidianus</i> filamentous virus 7)	<i>Acidianus convivator</i>	Source chaude acide, Italie	Linéaire 36,9 kb	Pas d'intégration dans le chromosome de l'hôte	
	AFV8 (<i>Acidianus</i> filamentous virus 8)	<i>Acidianus convivator</i>	Source chaude acide, Italie	Linéaire 38,2 kb	Séquences répétés inversés aux extrémités	
	AFV9 (<i>Acidianus</i> filamentous virus 9)	<i>Acidianus convivator</i>	Lac chaud acide, Kamchatka, Russie	Linéaire 41,2 kb		(Bize <i>et al.</i> , 2008)
	SIFV (<i>Sulfolobus islandicus</i> filamentous virus)	<i>Sulfolobus islandicus</i>	Champ de solfatares, Islande	Linéaire 40,9 kb		(Arnold, Zillig, <i>et al.</i> , 2000)
	TTV1 (<i>Thermoproteus</i> <i>tenax</i> virus 1)			Linéaire 15,9 kb		
	TTV2 (<i>Thermoproteus</i> <i>tenax</i> virus 2)	<i>Thermoproteus tenax</i>	Mare de boue, Islande	Linéaire 15,9 kb		(Janekovic <i>et al.</i> , 1983)
	TTV3 (<i>Thermoproteus</i> <i>tenax</i> virus 3)			Linéaire 27 kb		
	PfV1 (<i>Pyrobaculum</i> filamentous virus 1)*	<i>Pyrobaculum arsenaticum</i>	Source chaude, Puzzioli solfatara, Italie	Linéaire 17,7 kb	Lytique	(Rensen <i>et al.</i> , 2016)

Tableau 4 : Résumé des archéovirus des sources chaudes terrestres. Les étoiles indiquent que ces virus font l’objet d’un développement dans le corps du texte (suite).

Famille / Morphologie	Virus	Hôte	Origine	Génome	Caractéristiques communes	Références
Rudiviridae Bâtonnet rigide, fibres caudales	ARV1 (<i>Acidianus</i> rod-shaped virus 1)	<i>Acidianus</i>	Source chaude acide, Italie	Linéaire 24,7 kb	Séquences inversées répétées aux extrémités	(Vestergaard et al., 2005)
	SIRV1 (<i>Sulfolobus islandicus</i> rod-shaped virus 1)		Champ de solfatares, Islande	Linéaire 32,3 kb		
	SIRV2 (<i>Sulfolobus islandicus</i> rod-shaped virus 1)*	<i>Sulfolobus islandicus</i>		Linéaire 35,5 kb	17 ORFs communs (dont gènes codant pour des protéines de structures et des protéines impliquées dans la réplication de l'ADN)	(Prangishvili et al., 1999) (Vestergaard, Shah, et al., 2008) (Servin-Garcidueñas et al., 2013b)
	SRV (<i>Stygiolobus</i> rod-shaped virus)	<i>Stygiolobus</i>	Source chaude acide, Açores, Portugal	Linéaire 28,1 kb		
	SMR1 (<i>Sulfolobus</i> Mexican rudivirus 1)	<i>Sulfolobus</i>	Source chaude acide, Açores, Mexique	Linéaire 27,4 kb		
<i>Spiraviridae</i> Cylindre	ACV (<i>Aeropyrum</i> coil-shaped Virus)	<i>Aeropyrum permix</i>	Source chaude, Japon	Circulaire 24,9 kb	ADN simple brin	(Mochizuki et al., 2012)
<i>Turnviridae</i> Icosahédrique, enveloppé, structure bourgeonnante à la surface de la capside	STIV1 (<i>Sulfolobus turreted</i> icosahedral virus 1)*	<i>Sulfolobus solfataricus</i>	Source chaude acide Yellowstone, USA	Circulaire 17,7kb	Organisation génomique similaire 25 ORFs communs	(Rice et al., 2004)
	STIV2 (<i>Sulfolobus turreted</i> icosahedral virus 2)*	<i>Sulfolobus islandicus</i>	Source chaude acide, Islande	Circulaire 16,6 kb	Mécanisme de sortie de l'hôte particulier (structure pyramidales)	(Happonen et al., 2010)

Tableau 4 : Résumé des archéovirus des sources chaudes terrestres. Les étoiles indiquent que ces virus font l’objet d’un développement dans le corps du texte (suite).

Famille / Morphologie	Virus	Hôte	Origine	Génome	Relation à l’hôte / Caractéristiques	Références
Fuseau – non classé (Monocaudavirus)	ATSV (<i>Acidianus</i> tailed spindle-shaped virus)	<i>Acidianus</i>	Source chaude acide, Yellowstone, USA	Circulaire 70,8 kb		(Hochstein <i>et al.</i> , 2015)
	STSV1 (<i>Sulfolobus tengchongensis</i> spindle-shaped virus 1)	<i>Sulfolobus tengchongensis</i>	Source chaude acide et mare de boue, Tengchong, Chine	Circulaire 75,3 kb	25% de gènes communs Similarité avec ATV notamment protéine de capside et Intégrase	(Xiang <i>et al.</i> , 2005)
	STSV2 (<i>Sulfolobus tengchongensis</i> spindle-shaped virus 2)	<i>Sulfolobus tengchongensis</i>	Source chaude acide et mare de boue, Tengchong, Chine	Circulaire 76 kb		(Erdmann <i>et al.</i> , 2014)
Co-infection avec un plasmide conjugatif						
					pMGB1	
	SMV1 (<i>Sulfolobus monocaudavirus</i> 1)	<i>Sulfolobus solfataricus</i>	Source chaude acide, Yellowstone, USA	Circulaire 48,8 kb	Similarité avec ATV notamment protéine de capside et Intégrase	(Erdmann <i>et al.</i> , 2013)

Ampulaviridae

C'est une des familles les plus petites d'ultraprotéomorphologie. Elle comprend un unique virus : ABV (*Acidianus bottle shaped virus*). ABV a été isolé à partir d'une souche chaude-acide en Italie et infecte *Acidianus convivator* qui croît à une température optimale de 75°C. Les virions font 230 nm de long pour 4 nm de large à la pointe la plus fine et 75 nm de large à la base du pied. La partie externe du virion pourrait être constituée de l'adsorption et l'insertion du génome viral dans la cellule hôte. L'autorégulation du virion comporte une vingtaine de filaments, insérés dans un disque ou un anneau, qui ne semblent pas être liés à l'adsorption à l'hôte, ce qui pourrait avoir un rôle dans la stabilité du virion. La libération des virions ne causant pas la lyse cellulaire mais un retard de croissance, ce virus pourrait être une œuvre utile de biotechnologie. Le génome viral est un ADN double brin linéaire d'environ 3,5 kb. Celui-ci est composé de 57 ORFs dont seulement 3 ont des similarités avec des séquences présentes dans les bases de données incluant une glycosyltransferase, une thymidylate kinase et ADN polymérase de la famille B. Cette dernière pourrait être d'intérêt si elle était similaire à des protéines eucaryotes mais aussi à des polynéses d'Archaea. Enfin, la procapside de ce virus montre une homologie de structure avec celle de phi29, impliquant un potentiel ancêtre commun entre ces deux virus (Häring et al., 2005; Peng et al., 2007).

Fuselloviridae

Cette famille de virus, dont les virions sont caractérisés par une forme de citron, infecte des archées des genres *Sulfolobus* (SMF1, SSV1, 2, 4, 5, 6, 7 K1, RH) et *Acidianus* (ASV1) (Tableau 1). Ils ont été isolés pour la première fois dans des sources chaudes de l'Islande, mais certains proviennent du Japon, de Russie, du Mexique ou des USA. Ils contiennent de petits génomes linéaires d'environ 3,5 kb (Pi a et al., 2011; Uldahl and Peng, 2013).

Ces Fusellovirus partagent un « core genome » composé de 13 ORFs (Figure 28). Ces ORFs

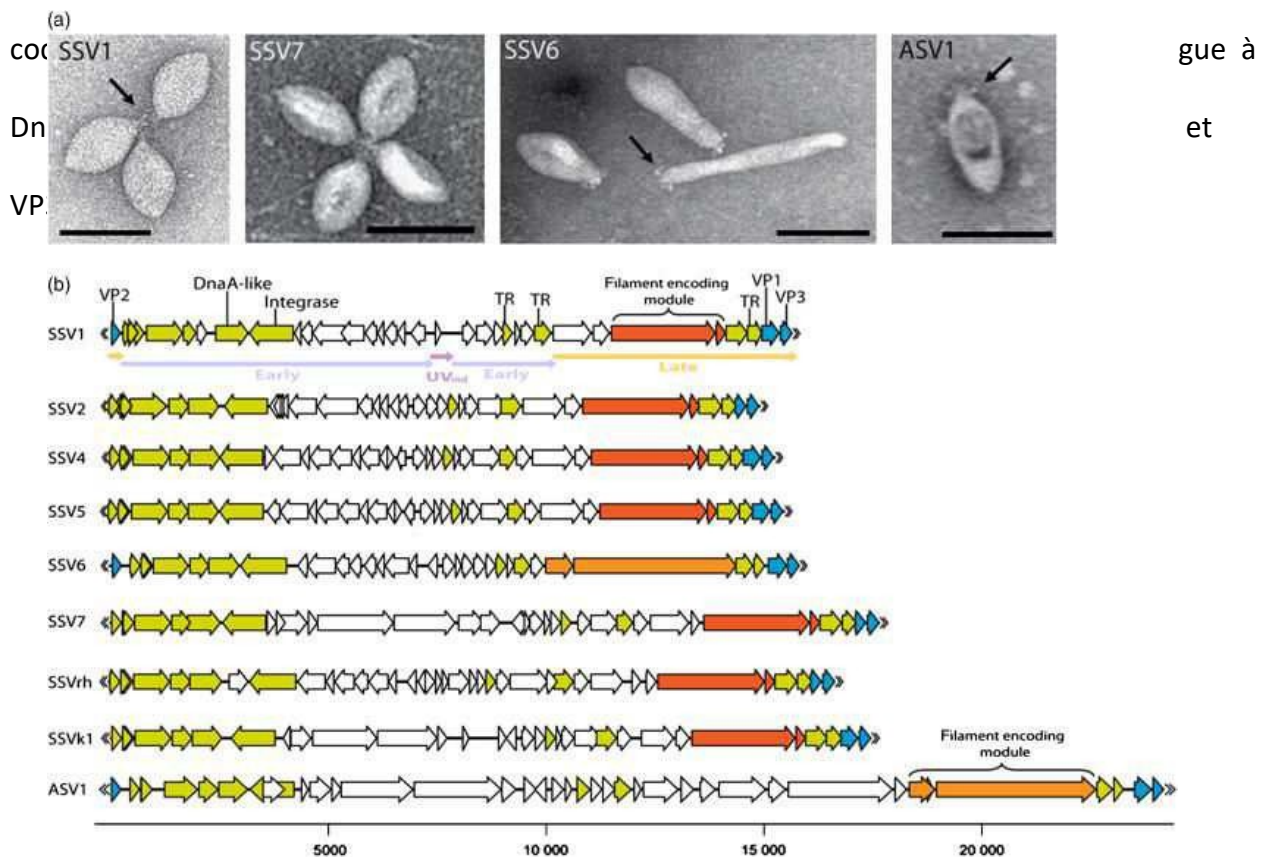


Figure 28 : Alignement de génomes de Fusellovirus (Pina et al., 2011). Les génomes ont été linéarisés pour la figure. En vert est représenté le core genome, en bleu les protéines de structure, en rouge (SSV1-like) ou orange (SSV6 like) le module codant pour les fibres caudales. Les ORFs codant pour des protéines aux fonctions connues sont annotées (VP2, DnaA-like, integrase, transcriptional regulator (TR), filament encoding module, VP1 et VP3).

Un Fusellovirus a été particulièrement étudié et est devenu un modèle : SSV1 pour *Sulfolobus shibatae*. Il fut découvert dans des hottilons poe et d'une souche chaude acide au Japon (Martin et al., 1984; Palm et al., 1991). Les virions de 100 X 60nm portent des fibres caudales caractéristiques des Fuselloviridae. Son ADN double brin circulaire de 22,5 kb a la particularité d'être superenroulé, il peut être sous forme libre dans le cytoplasme ou intégré au chromosome de son hôte (Nadal et al., 1986). Les virions de SSV1 sont notamment composés des protéines mineures et majeures de capside et d'une protéine

de liaison à l'ADN (Neite et al., ; P agish ili et al., ; Me o et al., ; Que i et al., 2015). Une analyse en cryomicroscopie a permis de révéler en détail la structure du virion, montrant notamment une symétrie sur 6 axes de la queue et la présence de nombreuses aspérités à la surface de la capsid (Figure 29)(Stedman et al., 2015).

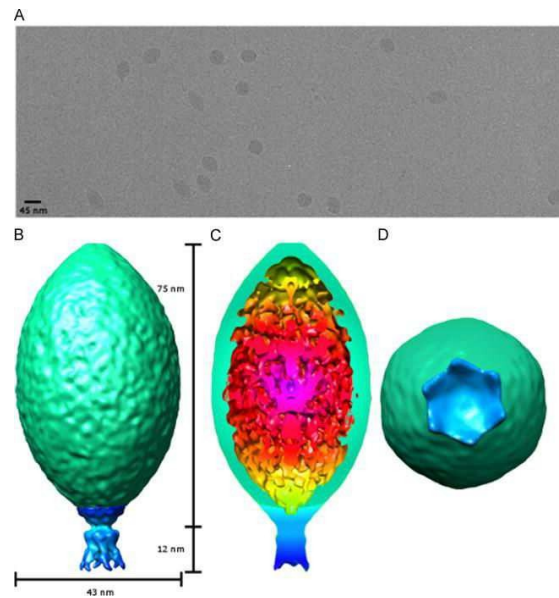


Figure 29 : Particules de SSV1 analysées en cryo-EM (Stedman et al., 2015). (A) Particules de SSV1 en cryo-EM. Reconstruction tridimensionnelle de SSV1 (B) vue d'ensemble (C) coupe transversale (D) vue de dessous.

Lipothrixiviridae

La famille Lipothrixiviridae est celle qui contient le plus de membres caractérisés provenant des sources chaudes terrestres. Les virions de cette famille ont une forme de longs filaments (| 900 à 2000nm de long) entourés par une enveloppe lipidique provenant de leur hôte. La

ajoute de partie d'été et de l'été des a h es appate at au gée Acidianus (AFV1, 2, 3, 6, 7 8 et 9) mais sont aussi retrouvés chez trois virus infectant Thermoproteustenax (TTV1, 2 et 3) et un virus infectant Sulfolobus islandicus (SIFV). Ils proviennent de sources chaudes en Italie, en Russie, en Islande et aux U.S.A. Leurs gènes sont composés d'ADN à double hélice linéaires, généralement d'une taille de kilobases et si les plus petits SIFV, TTV ne font que 15,9kb (Janekovic et al., 1983; Uldahl and Peng, 2013).

Virus filamenteux non classé : PFV1

PFV1 (*Pyrobaculum* filamentous virus 1) infecte *via* un cycle lytique une souche de *Pyrobaculum* isolée d'une source chaude italienne. Les virions de PFV1 font environ 400 nm de long pour 3 nm de large. Ils sont composés d'une nucléocapside hélicoïdale en forme de bâtonnet, constituée de répétitions de deux protéines de liaison à l'ADN (VP1+2). Celle-ci est surmontée d'une gaine protéique formée par de multiples répétitions d'une seule protéine VP3 (Figure 30). L'ensemble (génome + VP1-3) est contenu dans une enveloppe. Cette capside contient un ADN double brin linéaire de 17,7 kb qui inclut 39 ORFs dont la majeure partie n'ont pas de similarités avec les séquences présentes dans les bases de données (Rensen *et al.*, 2016).

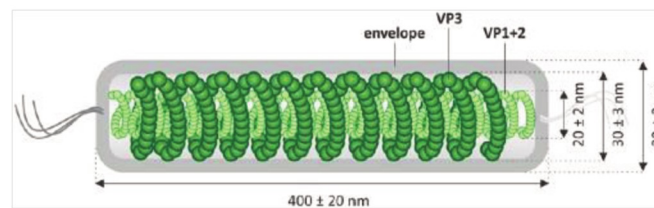


Figure 30 : Schéma de l'organisation des virions de PFV1 (Rensen *et al.*, 2016). En vert sont représentés les protéines VP 1 à 3, entourées par une enveloppe (en gris).

Rudiviridae

Les virus de cette famille se caractérisent par des virions en forme de longs bâtonnets rigides se terminant par trois fibres caudales et ne présentant pas d'enveloppe. Ils infectent des archées des genres *Sulfolobus* (**SIRV1 et 2, SMR1**), *Acidianus* (**ARV1**) et *Stygiolobus* (**SRV**). Les virions, dont la taille dépend de celle du génome font entre 610 et 900 nm de long pour environ 23 nm de large. Ces génomes à ADN double brin linéaires font entre 24,7 et 35,5 kb et montrent plusieurs similarités. Tout d'abord, les extrémités des génomes de ces virus sont toutes dotées de longues séquences inversées répétées.

Ensuite, 17 ORFs sont partagés par ces virus, incluant des ORFs codant pour des protéines de structure et d'attachement à la capsule de l'ADN. Ces derniers sont retrouvés dans un génome viral récemment séquencé, celui de SMR1 (*Sulfolobus Mexicanus*). Ce génome a été identifié à partir d'un tag obtenu sur une souche isolée. Il est possible qu'il infecte le genre *Sulfolobus* car les séquences de ces protéines sont présentes dans le tag obtenu. D'autre part, il a aussi été identifié dans la famille des *Rudiviridae* car il possède, non seulement les séquences inversées répétées terminales, mais aussi des ORFs homologues d'autres *Rudivirus*. Ces ORFs communs codent notamment pour des protéines de structure, des glycosyl-transférases et des protéines impliquées dans la capsule de l'ADN (Magish et al., 1999; Vestergaard et al., 2005, 2008; Servín-Garcidueñas et al., 2013; Uldahl and Peng, 2013).

Famille des *Turriviridae*

Les deux membres composant cette famille STIV1 et 2 (*Sulfolobus turreted icosahedral virus* 1 et 2) sont caractérisés par des virions de forme icosaédrique d'environ 25 nm de diamètre. Ils sont constitués d'une enveloppe et possèdent de petites structures géométriques à la surface de la capsid. Ces structures sont potentiellement impliquées dans la reconnaissance et/ou l'adhésion à l'hôte. STIV1 infecte *Sulfolobus solfataricus* via un cycle lytique. Sa capsid contient un ADN double brin circulaire de 17,6 kb. STIV2 infecte lui *Sulfolobus* G4ST-2 et est composé d'un génome à ADN double brin circulaire de 18,5 kb (Rie et al., 2004; Happonen et al., 2010).

Ces deux virus semblent similaires au niveau génomique et structurel. En effet, leurs génomes contiennent une organisation similaire par rapport aux ORFs. Ils possèdent aussi une série de protéines de l'hôte assez particulières et une série de structures protéiques, aussi appelées VAP (Virus Associated Proteins), jouant un rôle dans l'enveloppe cellulaire.

Figure 1. Microscopie électronique (indiquée par STIV) observée chez une souche (Echelle : 1Pm).

x Les sources hydrothermales profondes

Les océans recouvrent plus de 70% de la surface de la planète. Jusqu'à dans les années 1970, il était admis par la communauté scientifique que les océans, en dessous de la couche euphotique, étaient oligotrophes. Néanmoins en 1977, au large des îles Galápagos, ont été découvertes les premières sources hydrothermales (Corliss & Ballard, 1977). Les organismes de ces véritables oasis de vie au fond des océans vivent sans lumière et se développent grâce à l'énergie chimique issue du fluide hydrothermal. Depuis leur découverte, elles sont le sujet de nombreuses études. En effet, la macrofaune et la microfaune sont riches et diverses, composées de bivalves, vers et nombreux micro-organismes.

Les sources hydrothermales profondes sont parmi les environnements les plus extrêmes de la planète. Elles se situent au niveau des dorsales océaniques, généralement entre 800 et 3500m de profondeur. De ce fait elles sont caractérisées par de fortes pressions hydrostatiques et une absence de lumière. Ce sont néanmoins de véritables oasis de vie grâce à la présence du fluide hydrothermal qui apporte l'énergie nécessaire au développement de divers organismes colonisant cet environnement (Figure 32)(Prieur, 1997; Flores and Reysenbach, 2011).

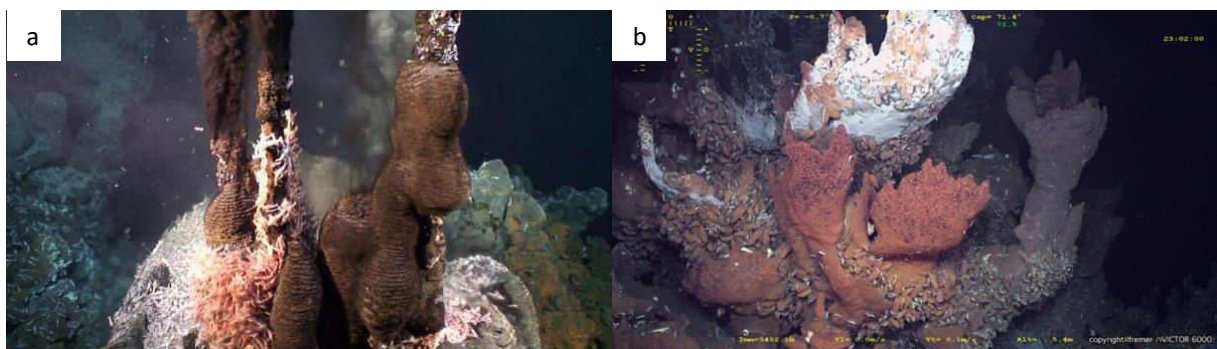


Figure 32: Sources hydrothermales profondes sites « les ruches » (a) et « l'Élan » (b). Photographie issues de la campagne Ifremer BICOSE (Biodiversité, Interactions, Connectivité et Symbioses en Milieux Extrêmes).

Leur formation résulte de l'altération de l'eau de mer dans la croûte terrestre fracturée océanique. L'eau de mer se chauffe progressivement et au contact de la roche magmatique, elle s'enrichit également en divers gaz dissous, comme le méthane, le sulfure d'hydrogène, le dioxyde de carbone, le dioxyde de soufre et le diiode de soufre. Sous l'effet de la chaleur, l'eau monte jusqu'à la surface du plancha où elle se lessive les éléments qu'elle transporte. Elle s'altère alors en éléments dissous, oxygène et des ions métalliques comme le zinc ou le fer puis rejaillit sous forme de fluide hydrothermal (Von Damm et al., 1995; Flores and Reysenbach, 2011).

Ces fluides sont très chauds (> 350°C), acides, anoxiques et enrichis en gaz dissous et en divers éléments (cités ci-dessus). Les éléments contenus dans ces fluides précipitent au contact de l'eau de mer à 25°C, pour former des hydrothermales (Figure 33). Au niveau de celles-ci se créent des gradients de température, de pH et de concentration en oxygène du fait de la différence entre l'eau de mer et le fluide hydrothermal. La confrontation de ces paramètres physico-chimiques très différents entraîne une ambiance chaotique qui ajoute au caractère extrême de cet environnement (Prieur, 1997; Desbruyères et al., 2000; Miroshnichenko, 2004; Flores and Reysenbach, 2011)

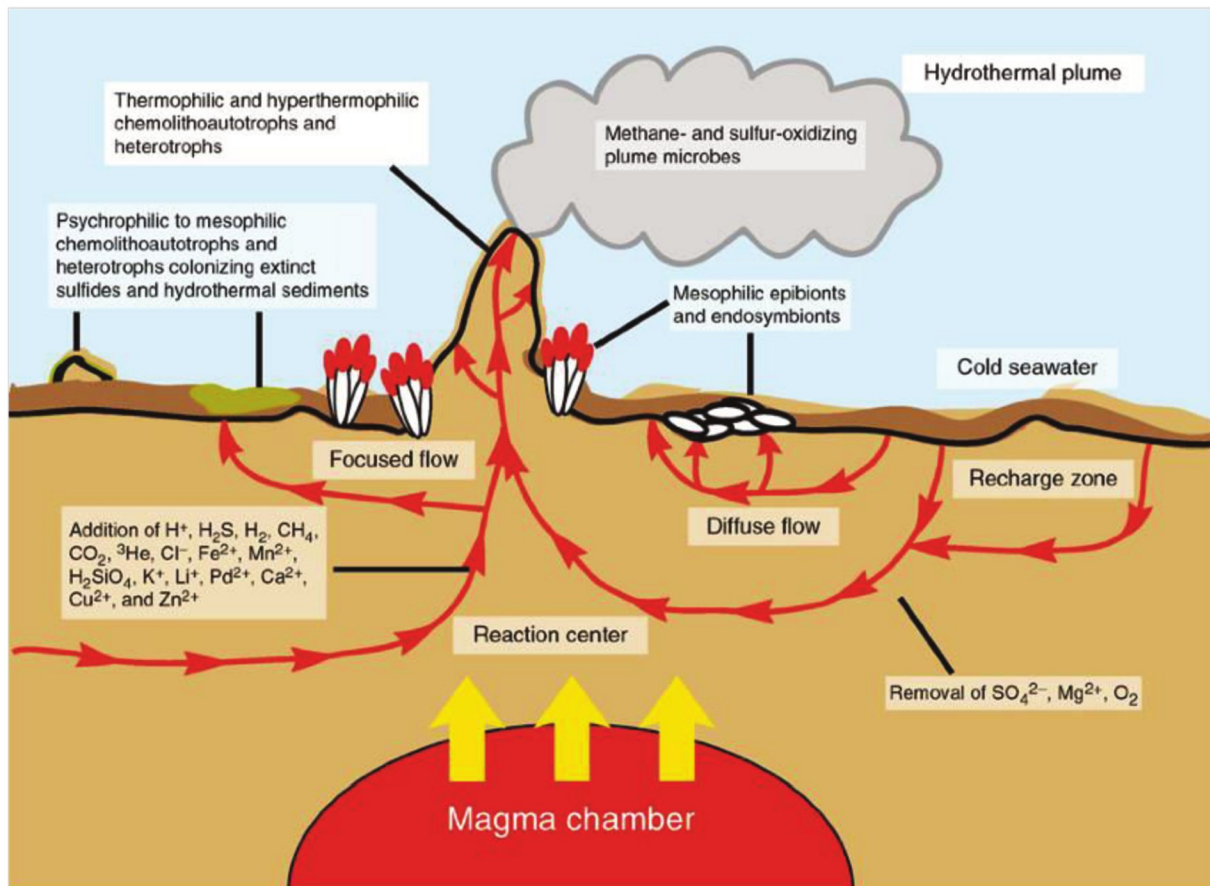


Figure 33 : Schéma du fonctionnement d'une source hydrothermale profonde (Flores and Reysenbach, 2011). Les flèches rouges indiquent le flux de fluide et les flèches jaunes indiquent les transferts de chaleurs issus de la chambre magmatique. Après s'être infiltré dans la croûte océanique, le fluide se réchauffe et s'enrichit et peut ressortir sous forme de flux diffus, ciblé ou sous forme de cheminée hydrothermale. Selon le type de sortie de fluide, diverses communautés colonisent l'environnement.

Communautés microbiennes des sources hydrothermales profondes

Malgré toutes ces conditions qui paraissent hostiles, de nombreuses espèces d'eucaryotes peuplent ces biotopes. En effet, plus de 500 espèces différentes ont été décrites réparties sur environ 50 sites. La plupart de celles-ci sont caractéristiques de ces écosystèmes comme le ver tubiforme *Riftia pachyptila* ou les crevettes *Rimicaris exoculata*. Il s'agit en majorité d'arthropodes, de mollusques et d'annélides (Prieur, 1977; Tarasov *et al.*, 2005; Thornburg *et al.*, 2010).

A la base de ces écosystèmes sont retrouvés des micro-organismes chimiosynthétiques qui sont à l'origine de la production primaire. Il s'agit aussi d'archées, la plupart sont anaérobies et peuvent être thermophiles ou hyperthermophiles. La concentration en H_2S dans ces environnements étant particulièrement haute (de 3 à 110 Mol/kg de sédiments l'oxidation des sulfures est la source principale d'énergie de ces micro-organismes chimiosynthétiques (Thornburg et al., 2010). Les Protéobactéries représentent une grande partie de la population bactérienne colonisant ces écosystèmes (figure 34). Elles peuvent être sulfo-oxydantes ou ferro-oxydantes, mésophiles à modérément thermophiles. Il s'agit souvent de symbiotes associés à leurs hôtes et parfois participant à la détoxification des composés soufrés (Kelley et al., 2002; Pflugfelder et al., 2005; Thornburg et al., 2010; Eichinger et al., 2014). Des Aquificales thermophiles hydrogéné-oxydantes et anaérobies sont aussi bien représentées. Elles sont autotrophes et retrouvées dans de nombreuses sources et sont utiles pour l'étude de la vie primitive et de la phylogénétique du vivant (Miroshnichenko, 2004; Flores et al., 2011). De nombreuses archées hyperthermophiles peuplent le géosphère. Il s'agit notamment des thermophiles aérobies des genres *Bacillus*, *Geobacillus* et *Thermus* appartenant au phylum des Firmicutes et des *Deinococcus-Thermus* (Thornburg et al., 2010; Flores and Reysenbach, 2011). Enfin, les bactéries thermophiles du phylum des Thermotogae sont souvent isolées à partir de ces écosystèmes et seront décrites plus en détail par la suite.

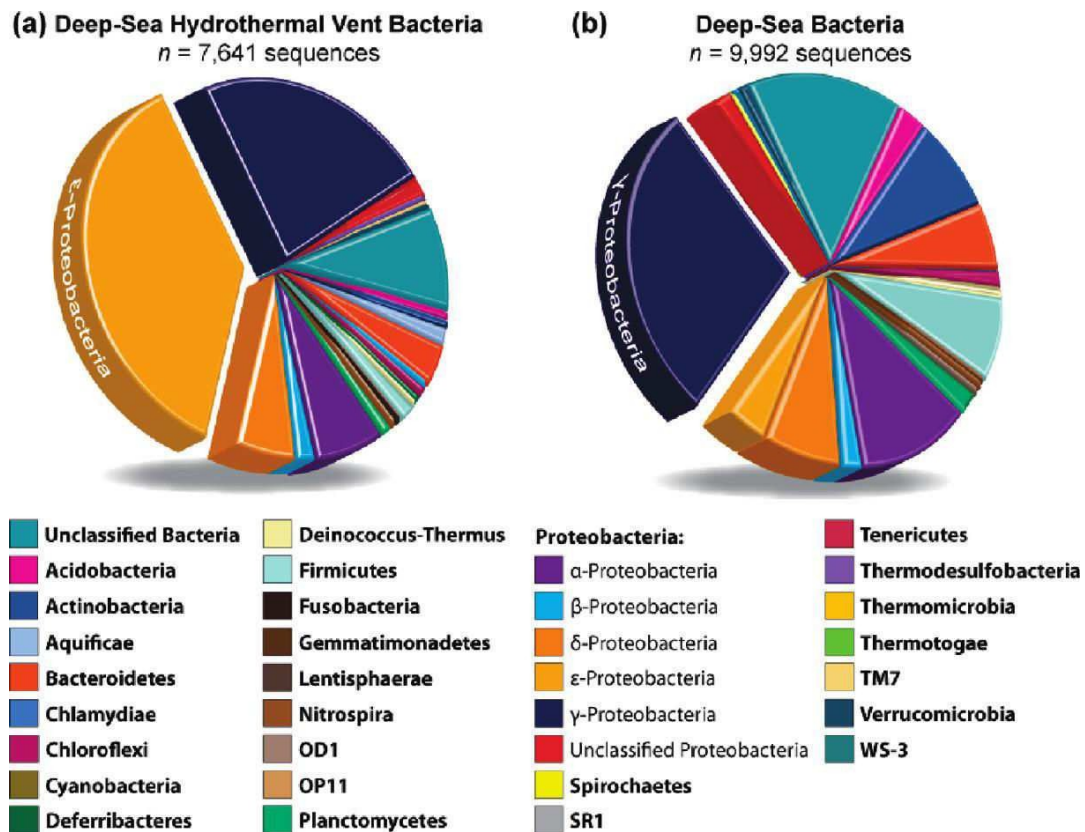


Figure 34: Abondance relative des phyla bactériens (a) au niveau des sources hydrothermales profondes, forte prédominance des epsilon protéobactéries, puis des alpha protéobactéries (b) dans les océans profonds, prédominance des alpha protéobactéries (Thornburg et al., 2010).

Concernant le domaine Archaea, des archées méthanogènes appartenant notamment à l'ordre des Methanococcales ou des archées organotrophes fermentaires appartenant à l'ordre des Thermococcales sont omniprésentes. D'autres hyperthermophiles sont aussi présentes comme les Desulfurococcales qui réduisent le soufre ou les Archaeoglobales pouvant réduire des composés soufrés ou ferreux (Figure 35) (Thornburg et al., 2010; Flores et al., 2011).

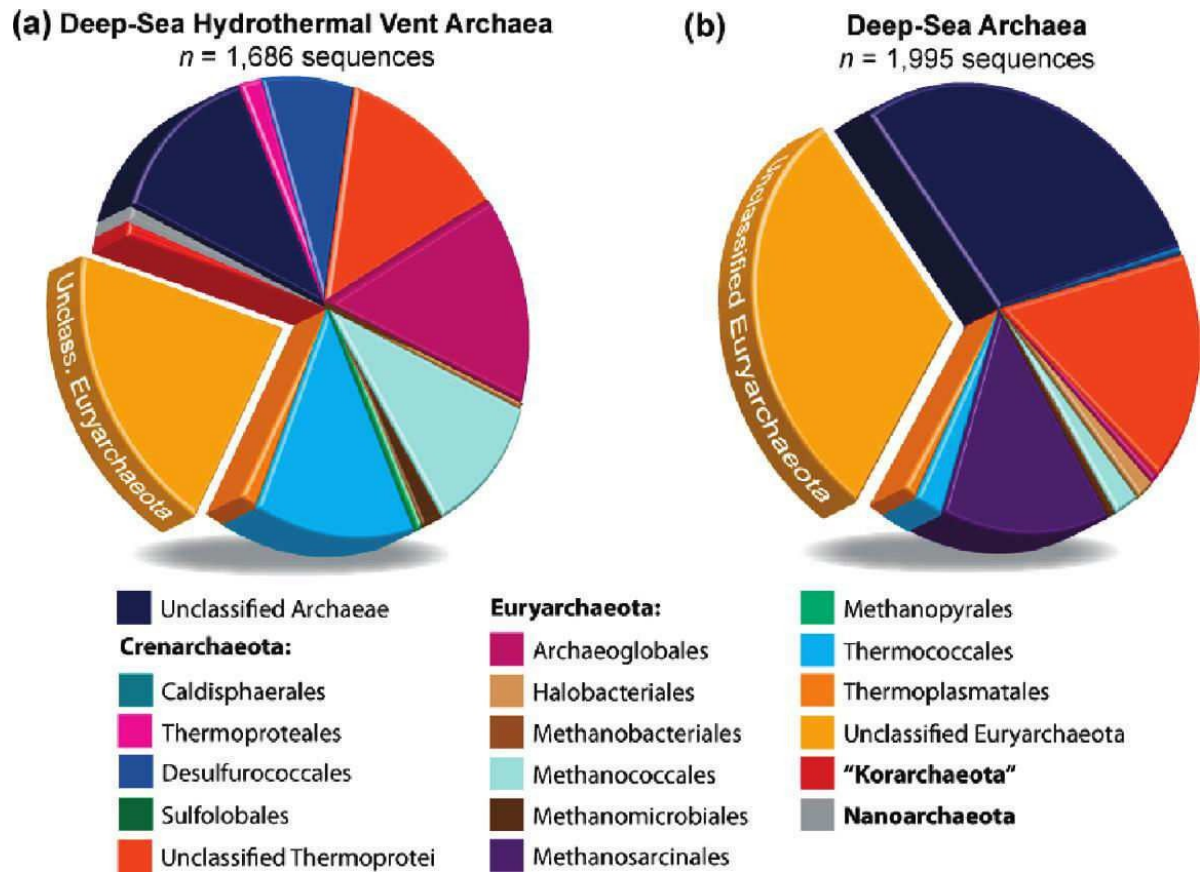


Figure 35 : Représentation de l'abondance relative des différents phyla archéens des sources hydrothermales profondes, (b) dans les océans profonds, prédominance des euryarchées non classées dans les deux environnements (Thornburg et al., 2010).

Les Thermotogae

Le phylum des Thermotogae regroupe plus de 70 espèces décrites, réparties au sein de 3 ordres, 4 familles et 12 genres différents (Pollo et al., 2015). Les Thermotogae sont des bactéries très majoritairement thermophiles ou hyperthermophiles (un genre non thermophile : les Mesotoga), anaérobies et à Gram négatif. Elles se présentent sous forme de bâtonnets portant une « toge » caractéristique. La composition de cette toge reste peu connue, elle contient deux protéines identifiées : la protéine d'attachement OmpA1 et la porine OmpB. Ces bactéries vivent dans des environnements chauds géothermiques ou volcaniques comme les sources hydrothermales profondes, les sources chaudes terrestres ou les puits de pétrole. Sur le plan métabolique, les Thermotogae peuvent fermenter de nombreux sucres et substrats protéiques simples. Cette faculté s'accompagne, pour la plupart, d'une capacité de dégradation de composés soufrés (Bonch-Osmolovskaya, 2008).

La position phylogénétique des Thermotogae, est sujette à débat (Bonch-Osmolovskaya, 2008; Zhaxybayeva et al., 2009; Gupta and Bhandari, 2011). En comparant les séquences des gènes codant pour les ARNr 16S et les protéines ribosomales des Thermotogae, celles-ci sont proches des Aquificales. Cette proximité contribue au fait que les Thermotogae occupent la seconde place la plus profonde dans l'arbre phylogénétique des bactéries (Figure 6).

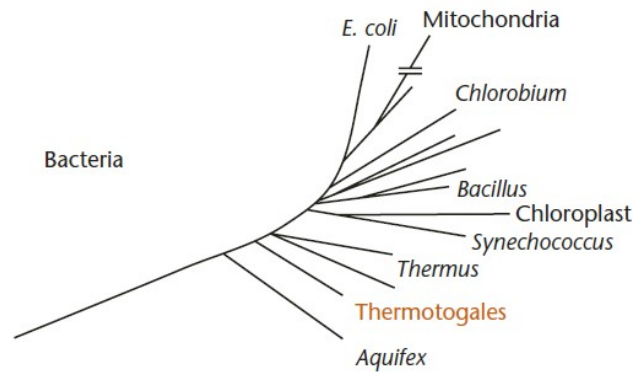


Figure 36: Placement des Thermotogales (par extension des Thermotogae) dans l'arbre phylogénétique bactérien par rapport à l'ASNBonchOsmolovskaya, 2008). Elles sont situées profonds dans cette arbre, proches des Aquificales.

N a o i s, l o s d' a a s e s p h l o g t i u e s s u l a e s t e d u g o e, l e a p p o h e e t a e l e s A q u i f i c a l e s e s t b e a u c o u p m o i n s é v i d e n t. L e s T h e r m o t o g a e s' a e t a i f e s t e e t p l u s p r o c h e s d e s F i r m i c u t e s, n o t a m m e n t d e s C l o s t r i d i a l e s, m a i s é g a l e m e n t d e s A r c h a e a . L' t u d e d e s g é n o m e s d e d i f f é r e n t e s s o u c h e s d e T h e r m o t o g a e (T . m a r i t i m a, T . p e t r o p h i l a, T . l e t t i n g a e, T . e l a e s i e s i s...) a m o n t r é q u e c e u x - c i p a r t a g e n t 4 5 % d e s i m i l a r i t é a v e c d e s F i r m i c u t e s (n o t a m m e n t T h e r m o a n a e r o b a c t e r a l e s, o r d r e f a i s a n t p a r t i e d e s C l o s t r i d i a l e s) e t 1 0 % d e s i m i l a r i t é a v e c d e s a r c h é e s (n o t a m m e n t l e s T h e r m o c o c c a l e s) (Z h a x y b a y e v a e t a l ., 2 0 0 9 ; G u p t a a d B h a d a i, . L' h i s t o i r e e v o l u t i v e d e s g o e s d e T h e r m o t o g a e p o u r r a i t é t r e l e r e f l e t d' a o d a t s t a s f e t s l a t a u d e g e s. P a e e p l e, u o p o p u t a t i f o d a t p o u u e f e r r é d o x i n e o x y d o r é d u c t a s e a é t é m i s e n é v i d e n c e c h e z l a m a j e u r e p a r t i e d e s e s p è c e s d é c r i t e s d e T h e r m o t o g a e. C e t o p é r o n e s t p a r t a g é a v e c P y r o c o c c u s f u r i o s u s q u i e s t u n e a r c h é e d e l' o d e d e T h e r m o c o c c a l e s, o c c u p a n t l a m ê m e n i c h e é c o l o g i q u e q u e c e r t a i n e s T h e r m o t o g a e, n o t a m m e n t a u n i v e a u d e s s o u r c e s h y d r o t h e r m a l e s p r o f o n d e s. I l a é t é d é m o n t r é q u e c e t o p é r o n a é t é o b t e n u p a r t r a n s f e r t l a t é r a l d e g è n e s d e s a r c h é e s v e r s l e s b a c t é r i e s, c e q u i e s s a i t l' i p l i a t i o d' l e t s g t i u e s o i l e s a u o u s d e l' o l u t i o d e e s e s p e s (Z h a x y b a y e v a e t a l ., 2 0 0 9).

Ces t asfe ts latau de g es a e t do à e isage l'i pli atio pote tielle de us via le ph o e de t a sdu tio et d'ig atio s ue es p o iales . Cette h poth se se trouve renforcée par la description de CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) au sein de nombreux génomes de Thermotogae ; ces séquences t oig e t de la ise e pla e d'i u it is- is d'ife tio s i ales et p ou e tue de telles ife tio s o t au oi s d jà eu lieu au s de l'hist o luti e des Thermotogae (Zhaxybayeva et al., 2009; Pollo et al., 2015). Pour autant, à ce jour, un seul virus a été décrit chez les Thermotogae. Il s'agit de MPV isole Marinitoga piezophila et récemment décrit (Lossouarn, Nesbø, et al., 2015). A celui-ci se rajoutent deux mini-plasmides cryptiques, très similaires, pQR7 et pMC24, retrouvés respectivement chez Thermotoga sp. RQ7 et Thermotoga maritima MC24, qui sont les seuls éléments génétiques mobiles décrits au sein de cet ordre bactérien (Bonch-Osmolovskaya, 2008).

Virus associés aux micro-organismes thermophiles

D'i po ta ts effo ts o t o sa s, ota e t au lao atoi e, afi de ieu o p e d e la diversité génétique, les capacités métaboliques et les adaptations physiologiques des communautés microbiennes hydrothermales. Notre connaissance de la diversité virale associée à ces micro-organismes demeure, en revanche, encore assez limitée. Alors u'e i o esp es de at ies et a h es so t o ues pou olo ise es environnements, seuls quelques virus ont été décrits (Williamson et al., 2008; Jebbar et al., 2015; Lossouarn, Dupont, et al., 2015).

Tableau 5 : Récapitulatif des virus de thermophiles décrits issus des sources hydrothermales océaniques profondes (Lossouarn, Dupont, *et al.*, 2015).

Bacterial and archaeal viruses isolated from deep-sea hydrothermal vents.					
	Host strains (growth temperature)	Family	Virions morphology	Virus-host relationship	Genomes
Bacterioviruses references					
BVW1 [29]	<i>Bacillus</i> sp. w13 (65 °C)	Unclassified	Long flexible tail, 300 nm	Lytic	Double-stranded linear DNA 18 kb
GVE1 [29]	<i>Geobacillus</i> sp. E26323 (65 °C)	<i>Siphoviridae</i>	Hexagonal head, 70 nm diameter Flexible tail, 180 nm	Lytic	Double-stranded linear DNA 41 kb
GVE2 [30]	<i>Geobacillus</i> sp. E263 (65 °C)	<i>Siphoviridae</i>	Hexagonal head, 130 nm diameter Not described	Lytic, potentially lysogenic	Double-stranded linear DNA 40.9 kb
D6E [31]	<i>Geobacillus</i> sp. E263 (65 °C)	<i>Myoviridae</i>	Probably similar to GVE1 Contractile tail 60 nm	Lytic	Double-stranded linear DNA 49.3 kb
Nsr-1 [32]	<i>Nitratiruptor</i> sp. SB155-2 (55 °C)	<i>Siphoviridae</i>	Hexagonal head, 60 nm diameter Flexible tail, 210 nm	Lysogenic	Double-stranded linear DNA 37.1 kb
MPV1 [34]	<i>Marinitoga piezophila</i> KA3 (65 °C)	<i>Siphoviridae</i>	Hexagonal head, 64 nm diameter Flexible tail, 200 nm Hexagonal head, 50 nm	Lysogenic	Double-stranded linear/circular DNA 43.715 kb
Archeoviruses references					
PAV1 [23,24]	<i>Pyrococcus abyssi</i> GE23 (85 °C)	<i>Fuselloviridae</i>	Lemon-shaped 120 nm length, 80 nm width	Carrier state	Double-stranded linear DNA 18 kb
TPV1 [25]	<i>Thermococcus pieurii</i> (80 °C)	<i>Fuselloviridae</i>	Lemon-shaped 140 nm length, 80 nm width	Carrier state	Double-stranded linear DNA 21.5 kb

Chez les Bacteria, 6 virus ont été caractérisés; d'e t e eu so t asso à des sou hes aérobie de *Geobacillus* et de *Bacillus* dont les températures optimales de croissance se situent entre 65 et 70°C. Le 5^{ème} est associé à une epsilon-protéobactérie dont la température de croissance optimale est de 55°C. Le dernier est associé à une Thermotogae dont la température optimale de croissance est de 65°C.

Les deux premiers virus à avoir été caractérisés sont BVW1 (*Bacillus virus W1*) et GVE1 (*Geobacillus virus E1*) (Figure 37 a et . BVW esto pos d'u e apside osa d i ue de de dia t e à la uelle est fie u e ueue de . Le g o e i al esto situ d'u ADN double brin linéaire de 18 kb. Il infecte la souche *Bacillus W13* via un cycle lytique et reste à l'heu e atuell lass Liu et al., 2006).

GVE1 appartient à la famille Siphoviridae. Les i io s so t o pos s d'u e apside i osa d i ue de de dia t e et d'u e ueue fle ile de de lo g. Il infecte la souche de *Geobacillus E26323* via un cycle lytique. Il apparaît proche de GVE2 (*Geobacillus virus E2*) qui infecte *Geobacillus E263* supposément via un cycle lysogénique. Ces hôtes ont une température optimale de croissance de 65°C. Ces deux virus contiennent un ADN double i li aie d'e i o k . Le g o e et les aut est isti ues de GVE o t fait l'o jet de nombreuses études (Liu et al., 2006; Liu and Zhang, 2008; Wang and Zhang, 2008a, 2008b; Wei and Zhang, 2008; Wu et al., 2009; Song et al., 2011; Jin et al., 2013, 2014, 2015). Une l si e the ostale de à °C a ota e t t ide tifie Ye ad)hag, . D'aut es pot i es telles u e pot i e de liaiso à l'ADN l si i , u e th id late s thase thermostable, ou une protéine portale ont été caractérisées à partir du génome de ce virus (Wang and Zhang, 2008a, 2008b; Wei and Zhang, 2008). Les protéines portales sont très i po tates pou les i us puis u elles so t i pli u es das l'e pa uetage du g o e i al à l'it i eur de la capsid.

Comme pour beaucoup de protéines de bactériovirus thermophiles, celle-ci montre peu de similitude avec les protéines des virus thermophiles. Elle a donc dû évoluer indépendamment. Cette évidence grâce à sa structure secondaire prédite qui montrait de fortes similarités avec les autres protéines portales connues. Sa fonction a ensuite été confirmée notamment grâce à un usage à l'opéetta t de o i sa positio das la st u tu e d u i o (Figure 7 c) (Wang and Zhang, 2008a).

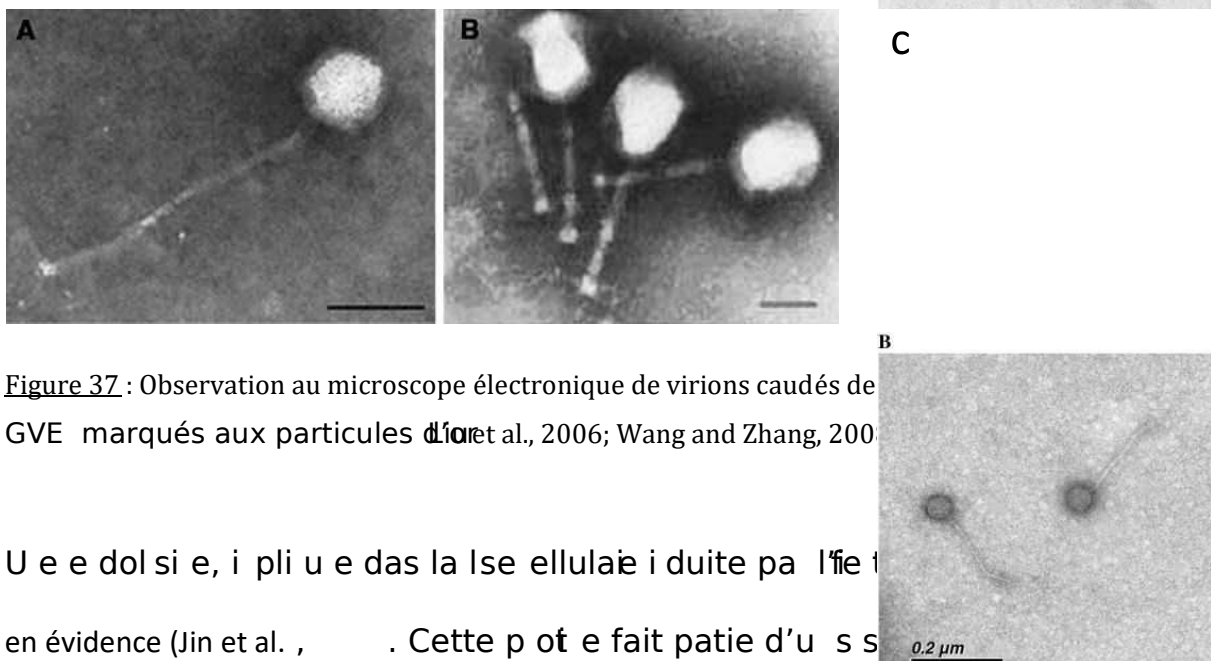


Figure 37 : Observation au microscope électronique de virions caudés de GVE marqués aux particules d'or (Liu et al., 2006; Wang and Zhang, 2008).

Une étude plus récente, réalisée par Jin et al. (2013), a mis en évidence (Jin et al., 2013). Cette propriété fait partie d'un système qui crée des pores dans la membrane cellulaire et entraîne une perte de perméabilité menant à la lyse cellulaire et donc à la libération des ions dans l'environnement (Jin et al., 2013). (Liu and Zhang, 2008; Ye and Zhang, 2008)

D6E infecte une souche de *Geobacillus* (Figure 38). C'est un virus plus virulent dont les virions sont composés d'une capside hexaédrique de diamètre de 20 nm et d'une queue contractile de 100 nm. Il est capable de pénétrer à l'ADN du hôte et de libérer l'ADN qui provoque la lyse cellulaire. Il est très similaire à d'autres virus thermophiles et notamment GVE2. En effet, leurs modules de transcription et de réplication sont très similaires (Wang and Zhang, 2010).

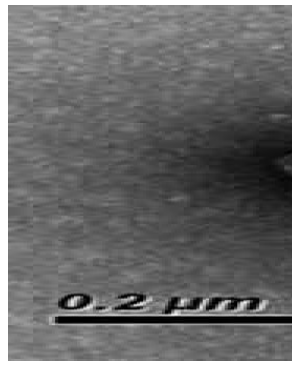


Figure 38 : Photographie en microscopie électronique de D6E (Wang and Zhang, 2010).

NrS-1 (*Nitratiruptor* sp. Siphoviridae 1) infecte une epsilon-proteobactérie du genre *Nitratiruptor* qui croît à une température optimale de 55°C. Sa capside de 64 nm de diamètre à laquelle est attachée une queue de 210 nm de long, contient un ADN double brin linéaire de 37kb. Ce virus tempéré a été classé dans la famille Siphoviridae (Yoshida-Takashima et al., 2013).

Enfin, MPV1 (*Marinitoga piezophila* Virus 1) est le premier et le seul virus infectant une *Thermotoga* décrite à ce jour. C'est un siphoviriste qui infecte *Marinitoga piezophila* KA3. Il sera décrit plus longuement dans la partie résultats (Lossouarn, Nesbø, et al., 2015).

Chez les Archaea, les deux seuls virus hyperthermophiles marins connus ont été isolés et caractérisés au laboratoire LM2E; ils sont associés à deux souches de Thermococcales. PAV1 (*Pyrococcus abyssi* virus 1) et TPV1 (*Thermococcus piezophilus* virus 1) ont des virions en forme de citron (Figure 39). Pour ces deux virus, la relation à l'hôte est du type latente, au cours de laquelle les virus sont libérés par bourgeonnement. Cependant pour TPV1, il est à noter que la production peut être induite au UV sans s'accompagner de lyse de l'hôte, son cycle de reproduction est alors plutôt défini comme lysogénique (Geslin, Le Romancer, Erauso, et al., 2003; Geslin et al., 2007; Gorlas et al., 2012).

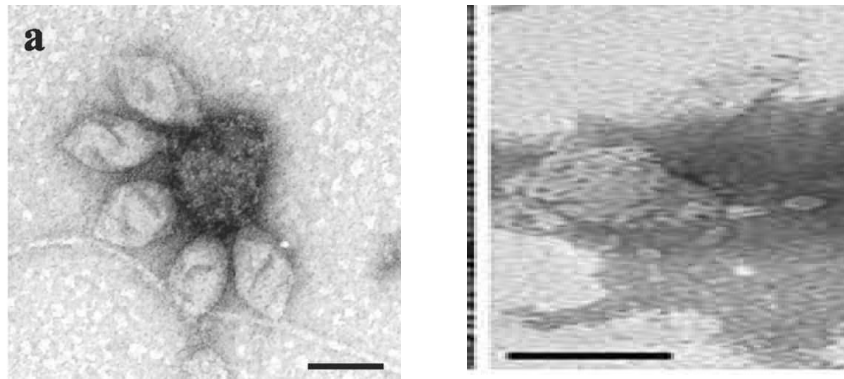


Figure 39 : Observation en MET de deux virus en forme de citrons (a) PAV1 et (b) TPV1 (Geslin, Le Romancer, Gaillard, et al., 2003; Gorlas et al., 2012) Echelle (a) 100 nm (b) 120 nm.

Leurs génomes sont à ADN double brin circulaire de 18 kb pour PAV1 et 21 kb pour TPV1. Ces génomes ne présentent que peu de similarités avec les séquences virales disponibles. Néanmoins, deux de leurs gènes partagent une forte similarité. Ces gènes codent pour des protéines potentiellement impliquées dans l'adhésion et la croissance cellulaire. D'autre part, PAV partage plusieurs gènes avec des plasmides d'autres thermococcales, ils sont tous regroupés et occupent la moitié du génome viral (Figure 40). Il est intéressant de noter que les protéines de structure du virus, dont les gènes avec TPV, se situent dans l'autre moitié du génome viral supposant que celui-ci est le fruit d'une co-localisation de deux éléments génétiques mobiles différents (Geslin et al., 2007; Gorlas et al., 2012; Krupovic et al., 2013).

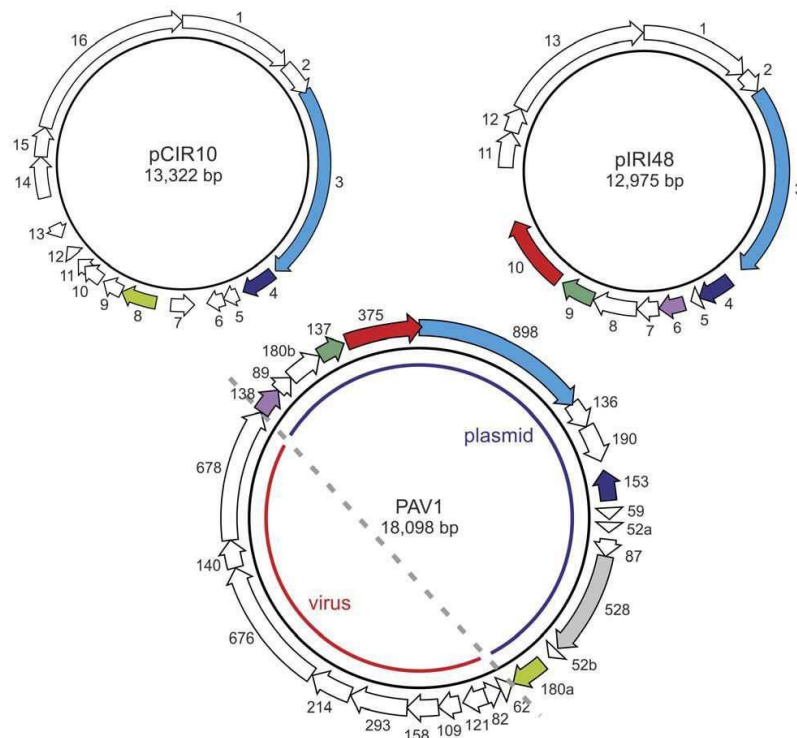


Figure 40 : Relation entre les plasmides de Thermococcales pCIR10 et pIRI48 et le virus PAV1 (Krupovic et al., 2013). Les gènes homologues portent la même couleur. PAV1 (en bas) est séparé en deux parties l'une d'origine plasmidique (partie haute) et l'autre d'origine virale (partie basse), en accord avec l'hypothèse selon laquelle il serait issu de la fusion de deux éléments génétiques mobiles.

IV. Bibliographie – Introduction

-A-

- Abrescia, N.G.A., Bamford, D.H., Grimes, J.M., and Stuart, D.I. (2012) Structure unifies the viral universe. *Annu. Rev. Biochem.* 81: 795–822.
- Ackermann, H. (2009) Phage classification and characterization. In, Clokie, M. and Kropinski, A. (eds), *Bacteriophages: Methods and protocols.*, pp. 127–140.
- Ackermann, H.W. (2007) 5500 Phages examined in the electron microscope. *Arch. Virol.* 152: 227–243.
- Ackermann, H.W. and Prangishvili, D. (2012) Prokaryote viruses studied by electron microscopy. *Arch. Virol.* 157: 1843–1849.
- Alcolombri, U., Ben-Dor, S., Feldmesser, E., Levin, Y., Tawfik, D.S., and Vardi, A. (2015) Identification of the algal dimethyl sulfide-releasing enzyme: A missing link in the marine sulfur cycle. *Science* (80-.). 348: 1466–1469.
- Ankrah, N.Y.D., May, A.L., Middleton, J.L., Jones, D.R., Hadden, M.K., Gooding, J.R., et al. (2014) Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J.* 8: 1089–1100.

-B-

- Barrangou, R. and van der Oost, J. (2014) Bacteriophage exclusion, a new defense system. *EMBO J.* e201490620.
- Basta, T., Garrett, R.A., and Prangishvili, D. (2009) Viruses of the Archaea. *Encycl. Life Sci.*
- Bawden, F.C. (1951) The multiplication of viruses. *Sci. Prog.* 39: 1–12.
- Berdygulova, Z., Westblade, L.F., Florens, L., Koonin, E. V, Chait, B.T., Ramanculov, E., et al. (2011) Temporal regulation of gene expression of the *Thermus thermophilus* bacteriophage P23-45. *J. Mol. Biol.* 405: 125–142.
- Bergh, O., Børshheim, K.Y., Bratbak, G., and Heldal, M. (1989) High abundance of viruses found in aquatic environments. *Nature* 340: 467–8.
- Bibi-Triki, S. (2014) Identification et caractérisation fonctionnelle et structurale du système toxine-antitoxine HicA3-HicB3 de *Yersinia pestis*.
- Biller, S.J., Schubotz, F., Roggensack, S.E., Thompson, A.W., Summons, R.E., and Chisholm, S.W. (2014) Bacterial vesicles in marine ecosystems. *Science* (80-.). 343: 183–186.
- Blondal, T., Hjorleifsdottir, S., Aevarsson, A., Fridjonsson, O.H., Skirnisdottir, S., Wheat, J.O., et al. (2005) Chaate izatio of α -Pol u leotide Kase/-Phosphatase from Bacteriophage RM378. *J. Biol. Chem.* 280: 5188–5194.
- Blondal, T., Hjorleifsdottir, S.H., Fridjonsson, O.F., Aevarsson, A., Skirnisdottir, S., Hermannsdottir, A.G., et al. (2003) Discovery and characterization of a thermostable bacteriophage RNA ligase homologous to T4 RNA ligase 1. *Nucleic Acids Res.* 31: 7247–7254.
- Blondal, T., Thorisdottir, A., Unnsteinsdottir, U., Hjorleifsdottir, S., Aevarsson, A., Ernstsson, S., et al. (2005) Isolation and characterization of a thermostable RNA ligase 1 from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties. *Nucleic Acids Res.* 33: 135–142.
- Bonch-Osmolovskaya, E. (2008) Thermotogales. In, eLS. John Wiley & Sons, Ltd.
- Breitbart, M. (2012) Marine Viruses: Truth or Dare. *Ann. Rev. Mar. Sci.* 4: 425–448.
- Breitbart, M. and Rohwer, F. (2005) Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13: 278–284.
- Breitbart, M., Thompson, L.R., Suttle, C.A., and Sullivan, M.B. (2007) Exploring the Vast Diversity of Marine Viruses. *Oceanography* 135–139.

- Brum, J.R., Ignacio-Espinoza, J.C., Roux, S., Doulcier, G., Acinas, S.G., Alberti, A., et al. (2015) Patterns and ecological drivers of ocean viral communities. *Science* (80-.). 348:.
- Brum, J.R., Schenck, R.O., and Sullivan, M.B. (2013) Global morphological analysis of marine viruses shows minimal regional variation and dominance of non-tailed viruses. *ISME J.* 7: 1738–51.
- Brum, J.R. and Sullivan, M.B. (2015) Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat. Rev. Microbiol.* 13: 147–159.
- Brumfield, S.K., Ortmann, A.C., Ruigrok, V., Suci, P., Douglas, T., and Young, M.J. (2009) Particle assembly and ultrastructural features associated with replication of the lytic archaeal virus *Sulfolobus turreted* icosahedral virus. *J. Virol.* 83: 5964–5970.
- Brüssow, H., Canchaya, C., and Hardt, W.-D. (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68: 560–602.

Burnet, F.M. (1945) *Virus as Organism. Evolutionary and Ecological Aspects of some Human Virus Diseases.*
Virus as Org. Evol. Ecol. Asp. some Hum. Virus Dis.

-C-

- Chien, A., Edgar, D.B., and Trela, J.M. (1976) Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* 127: 1550–1557.
- Chopin, M.-C., Chopin, A., and Bidnenko, E. (2005) Phage abortive infection in lactococci: variations on a theme. *Curr. Opin. Microbiol.* 8: 473–479.
- Chow, C.-E.T. and Suttle, C.A. (2015) Biogeography of Viruses in the Sea. *Annu. Rev. Virol.* 2: 41–66.
- Christie, G.E. and Dokland, T. (2012) Pirates of the Caudovirales. *Virology* 434: 210–21.
- Colso , P., De La alle ie, X., Yuřin., Asga i, “., Bigot, Y., Bideshi, D.K., et al. Megai ales , a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch. Virol.* 158: 2517–21.
- Cortez, D., Forterre, P., and Gribaldo, S. (2009) A hidden reservoir of integrative elements is the major source of recently acquired foreign genes and ORFans in archaeal and bacterial genomes. *Genome Biol.* 10: 1.

-D-

- Da o ao, R., Dell’A o, A., Co ialdesi, C., Magag, iM., No le, R., Ta u i i, C, and Weinbauer, M. (2008) Major viral impact on the functioning of benthic deep-sea ecosystems. *Nature* 454: 1084–1087.
- De Paepe, M., Leclerc, M., Tinsley, C.R., and Petit, M.-A. (2014) Bacteriophages: an underestimated role in human and animal health? *Front. Cell. Infect. Microbiol.* 4: 39.
- Desbruyères, D., Almeida, A., Biscoito, M., Comtet, T., Khripounoff, A., Le Bris, N., et al. (2000) A review of the distribution of hydrothermal vent communities along the northern Mid-Atlantic Ridge: dispersal vs. environmental controls. *Hydrobiologia* 440: 201–216.
- Du k o th, D.H. Who Dis o e ed Bate iophage? *Bacteriol. Rev.* 40: 739–802.

-E-

- Earthguide.image library, Coccolithophore – *Emiliana huxleyi*.
- Eichinger, I., Schmitz-Esser, S., Schmid, M., Fisher, C.R., and Bright, M. (2014) Symbiont-driven sulfur crystal formation in a thiotrophic symbiosis from deep -sea hydrocarbon seeps. *Environ. Microbiol. Rep.* 6: 364–372.
- Ellen, A.F., Albers, S.-V., Huibers, W., Pitcher, A., Hobel, C.F. V, Schwarz, H., et al. (2009) Proteomic analysis of secreted membrane vesicles of archaeal *Sulfolobus* species reveals the presence of endosome sorting complex components. *Extremophiles* 13: 67–79.
- Emiliana huxleyi* | Collection | Expedition Arctic.

-F-

- Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S., and Salmond, G.P.C. (2009) The phage abortive infection system, ToxIN, functions as a protein –RNA toxin–antitoxin pair. *Proc. Natl. Acad. Sci.* 106: 894–899.
- Fischer, M.G. (2016) Giant viruses come of age. *Curr. Opin. Microbiol.* 31: 50–57.
- Fischer, M.G. and Suttle, C.A. (2011) A Virophage at the Origin of Large DNA Transposons. *Science*. 332.
- Flores, G.E., Campbell, J.H., Kirshtein, J.D., Meneghin, J., Podar, M., Steinberg, J.I., et al. (2011) Microbial community structure of hydrothermal deposits from geochemically different vent fields along the Mid-Atlantic Ridge. *Environ. Microbiol.* 13: 2158–2171.
- Flores, G.E. and Reysenbach, A.-L. (2011) Hydrothermal environments, marine. In, *Encyclopedia of Geobiology*. Springer, pp. 456–467.
- Forterre, P. (2010) Defining life: the virus viewpoint. *Orig. Life Evol. Biosph.* 40: 151–60.
- Forterre, P. (2013a) Les virus à nouveau sur le devant de la scène. *Biol. Aujourd'hui*. 207: 153–168.
- Forterre, P. (2011) Manipulation of cellular syntheses and the nature of viruses: The virocell concept. *Comptes Rendus Chim.* 14: 392–399.
- Forterre, P. (2013b) The virocell concept and environmental microbiology. *ISME J.* 7: 233–236.
- Forterre, P. (2016) To be or not to be alive: how recent discoveries challenge the traditional definitions of viruses and life. *Stud. Hist. Philos. Sci. Part C Stud. Hist. Philos. Biol. Biomed. Sci.*
- Forterre, P., Soler, N., Krupovic, M., Marguet, E., and Ackermann, H.-W. (2013) Fake virus particles generated by fluorescence microscopy. *Trends Microbiol.* 21: 1–5.

-G-

- Gaia, M., Benamar, S., Boughalmi, M., Pagnier, I., Croce, O., Colson, P., et al. (2014) Zamilon, a Novel Virophage with Mimiviridae Host Specificity. *PLoS One* 9: e94923.
- Gaudin, M. (2012) Etude des vésicules membranaires produites par les Archées hyperthermophiles marines de l'ordre des Thermococcales.
- Gaudin, M., Gauliard, E., Schouten, S., Houel-Renault, L., Lenormand, P., Marguet, E., and Forterre, P. (2013) Hyperthermophilic archaea produce membrane vesicles that can transfer DNA. *Environ. Microbiol. Rep.* 5: 109–16.
- Gaudi, M., Kupo, J. M., Marguet, E., Gauliard, E., Kupo, J. V., Le Ca, E., et al. (2013) Extracellular membrane vesicles harbouring viral genomes. *Environ. Microbiol.* 16: 1167–1175.
- Geslin, C., Gaillard, M., Flament, D., Rouault, K., Le Romancer, M., Prieur, D., and Erauso, G. (2007) Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from *Pyrococcus abyssi*. *J. Bacteriol.* 189: 4510–4519.
- Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003) PAV1, the first virus-like particle isolated from the hyperthermophilic archaeon *Pyrococcus abyssi*. *J. Bacteriol.* 185: 3888–3894.
- Geslin, C., Le Romancer, M., Gaillard, M., Erauso, G., and Prieur, D. (2003) Observation of virus-like particles in high temperature enrichment cultures from deep-sea hydrothermal vents. *Res. Microbiol.* 154: 303–307.
- Geslin, C., Le Romancer, M., Gaillard, M., and Prieur, D. (2005) Diversité virale associée aux écosystèmes hydrothermaux océaniques profonds et aux sources chaudes terrestres. *Virologie* 9: 357–366.
- Glansdorff, N. (2009) Extremophiles - volume 1 UNESCO. Gerday, C. and Glansdorff, N. (eds) Eolss, Oxford.
- Goldfarb, T., Sberro, H., Weisstock, E., Cohen, O., Doolittle, W. F., Chakrabarty, A., et al. (2015) BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* 34: 169–183.
- Gorlas, A., Koonin, E. V., Bienvenu, N., Prieur, D., and Geslin, C. (2012) TPV1, the first virus isolated from the hyperthermophilic genus *Thermococcus*. *Environ. Microbiol.* 14: 503–516.

- Gorlas, A., Marguet, E., Gill, S., Geslin, C., Guigner, J.M., Guyot, F., and Forterre, P. (2015) Sulfur vesicles from Thermococcales: A possible role in sulfur detoxifying mechanisms. *Biochimie* 118: 356–364.
- Gupta, R.S. and Bhandari, V. (2011) Phylogeny and molecular signatures for the phylum Thermotogae and its subgroups. *Antonie Van Leeuwenhoek* 100: 1–34.
- H-
- Happonen, L.J., Redder, P., Peng, X., Reigstad, L.J., Prangishvili, D., and Butcher, S.J. (2010) Familial relationships in hyperthermo-and acidophilic archaeal viruses. *J. Virol.* 84: 4747–4754.
- Häring, M., Rachel, R., Peng, X., Garrett, R.A., and Prangishvili, D. (2005) Viral Diversity in Hot Springs of Pozzuoli, Italy, and Characterization of a Unique Archaeal Virus, Acidianus Bottle-Shaped Virus, from a New Family, the Ampullaviridae. *J. Virol.* 79: 9904–9911.
- Hayes, F. and Van Melderren, L. (2011) Toxins-antitoxins: diversity, evolution and function. *Crit. Rev. Biochem. Mol. Biol.* 46: 386–408.
- Hjorleifsdottir, S., Aevansson, A., Hreggvidsson, G.O., Fridjonsson, O.H., and Kristjansson, J.K. (2014) Isolation, growth and genome of the Rhodothermus RM378 thermophilic bacteriophage. *Extremophiles* 18: 261–270.
- Hjörleifsdottir, S., Hreggvidsson, G.O., Fridjonsson, O.H., Aevansson, A., and Kristjansson, J.K. (2002) Bacteriophage RM 378 of a thermophilic host organism.
- Holligan, P.M., Fernández, E., Aiken, J., Balch, W.M., Boyd, P., Burkill, P.H., et al. (1993) A biogeochemical study of the coccolithophore, *Emiliana huxleyi*, in the North Atlantic. *Global Biogeochem. Cycles* 7: 879–900.
- Horvath, P. and Barrangou, R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science* (80- .). 327: 167–170.
- van Houte, S., Buckling, A., and Westra, E.R. (2016) Evolutionary Ecology of Prokaryotic Immune Mechanisms. *Microbiol. Mol. Biol. Rev.* 80: 745–763.
- Huber, R. and Hannig, M. (2006) Thermotogales. In, Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H., and Stackebrandt, E. (eds), *The Prokaryotes: Volume 7: Proteobacteria: Delta, Epsilon Subclass*. Springer New York, New York, NY, pp. 899–922.
- Huber, R., Woese, C.R., Langworthy, T.A., Kristjansson, J.K., and Stetter, K.O. (1990) *Fervidobacterium islandicum* sp. nov., a member of the thermophilic Thermotogales. *Arch. Microbiol.* 154: 105–111.
- Hurst, C.J. (2011) An introduction to viral taxonomy with emphasis on microbial and botanical hosts and the proposal of Akamara, a potential domain for the genomic acellular agents. *Stud. viral Ecol. Microb. Bot. host Syst.* 1: 41–65.
- Hurwitz, B.L. and Sullivan, M.B. (2013) The Pacific Ocean Virome (POV): A Marine Viral Metagenomic Dataset and Associated Protein Clusters for Quantitative Viral Ecology. *PLoS One* 8: .
- J-
- Jaatinen, S.T., Happonen, L.J., Laurinmäki, P., Butcher, S.J., and Bamford, D.H. (2008) Biochemical and structural characterisation of membrane-containing icosahedral dsDNA bacteriophages infecting thermophilic *Thermus thermophilus*. *Virology* 379: 10–19.
- Jalasvuori, M., Pawlowski, A., and Bamford, J.K.H. (2010) A unique group of virus-related, genome-integrating elements found solely in the bacterial family Thermaceae and the archaeal family Halobacteriaceae. *J. Bacteriol.* 192: 3231–3234.
- Janekovic, D., Wunderl, S., Holz, I., Zillig, W., Gierl, A., and Neumann, H. (1983) TTV1, TTV2 and TTV3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaeobacterium *Thermoproteus tenax*. *Mol. Gen. Genet. MGG* 192: 39–45.
- Jebbar, M., Franzetti, B., Girard, E., and Oger, P. (2015) Microbial diversity and adaptation to high hydrostatic pressure in deep-sea hydrothermal vents prokaryotes. *Extremophiles* 19: 721–740.

- Jiang, S.C. and Paul, J.H. (1998) Gene transfer by transduction in the marine environment. *Appl. Environ. Microbiol.* 64: 2780–2787.
- Jin, M., Chen, Y., Xu, C., and Zhang, X. (2014) The effect of inhibition of host MreB on the infection of thermophilic phage GVE2 in high temperature environment. *Sci. Rep.* 4:.
- Jin, M., Xu, C., and Zhang, X. (2015) The effect of tryptophol on the bacteriophage infection in high-temperature environment. *Appl. Microbiol. Biotechnol.* 99: 8101–8111.
- Jin, M., Ye, T., and Zhang, X. (2013) Roles of bacteriophage GVE2 endolysin in host lysis at high temperatures. *Microbiology* 159: 1597–1605.
- K-
- Kelley, D.S., Baross, J.A., and Delaney, J.R. (2002) Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu. Rev. Earth Planet. Sci.* 30: 385–491.
- Klemperer, O. (1940) Electron microscopes of high magnification. *Reports Prog. Phys.* 7: 107.
- Knowles, B., Silveira, C.B., Bailey, B.A., Barott, K., Cantu, V.A., Cobián-Güemes, A.G., et al. (2016) Lytic to temperate switching of viral communities. *Nature* 531: 466–470.
- Krupovic, M., Gonnet, M., Hania, W. Ben, Forterre, P., and Erauso, G. (2013) Insights into dynamics of mobile genetic elements in hyperthermophilic environments from five new *Thermococcus* plasmids. *PLoS One* 8: e49044.
- Krupovic, M., Quemin, E.R.J., Bamford, D.H., Forterre, P., and Prangishvili, D. (2014) Unification of the globally distributed spindle-shaped viruses of the Archaea. *J. Virol.* 88: 2354–2358.
- Kulp, A. and Kuehn, M.J. (2010) Biological Functions and Biogenesis of Secreted Bacterial Outer Membrane Vesicles. *Annu. Rev. Microbiol.* 163–184.
- L-
- Lang, A.S., Zhaxybayeva, O., and Beatty, J.T. (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* 10: 472.
- La Scola, B., Audic, S., Robert, C., Jungang, L., Lamballerie, X. de, Drancourt, M., et al. (2003) A Giant Virus in Amoeboe. *Science* (80-). 299: 2033.
- La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., et al. (2008) The virophage as a unique parasite of the giant mimivirus. *Nature* 455: 100–104.
- Lindell, D., Sullivan, M.B., Johnson, Z.I., Tolonen, A.C., Rohwer, F., and Chisholm, S.W. (2004) Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc. Natl. Acad. Sci. U. S. A.* 101: 11013–11018.
- Liu, B., Wu, S., Song, Q., Zhang, X., and Xie, L. (2006) Two Novel Bacteriophages of Thermophilic Bacteria Isolated from Deep-Sea Hydrothermal Fields. *Curr. Microbiol.* 53: 163–166.
- Liu, B., Wu, S., and Xie, L. (2010) Complete genome sequence and proteomic analysis of a thermophilic bacteriophage BV1. *Acta Oceanol. Sin.* 29: 84–89.
- Liu, B. and Zhang, X. (2008) Deep-sea thermophilic *Geobacillus* bacteriophage GVE2 transcriptional profile and proteomic characterization of virions. *Appl. Microbiol. Biotechnol.* 80: 697–707.
- Liu, B., Zhou, F., Wu, S., Xu, Y., and Zhang, X. (2009) Genomic and proteomic characterization of a thermophilic *Geobacillus* bacteriophage GBSV1. *Res. Microbiol.* 160: 166–71.
- López-García, P. (2012) The place of viruses in biology in light of the metabolism-versus-replication-first debate. *Hist. Philos. Life Sci* 34: 391–406.
- López-García, P. and Moreira, D. (2012) Viruses in Biology. *Evol. Educ. Outreach* 5: 389–398.
- López-López, O., Cerdán, E.M., and González-Siso, I.M. (2013) Hot Spring Metagenomics. *Life* 3: 308–320.
- Lossouarn, J., Dupont, S., Gorlas, A., Mercier, C., Bienvenu, N., Marguet, E., et al. (2015) An abyssal mobilome: Viruses, plasmids and vesicles from deep-sea hydrothermal vents. *Res. Microbiol.* 166: 742–752.

Lossouarn, J., Nesbø, C.L., Mercier, C., Jha, A., Johansson, M., Chahk, R., et al. (2015) A selfish gene that uses its host to propagate. *The ISME Journal* 17: 3278–3288.

Lwoff, A. (1957) The concept of virus. *Microbiology* 17: 239–253.

-M-

Madigan, M. and Martinko, J. (2006) *L'essentiel de la biologie*. I, F. A. E. P. E. ed. Pearson Education France, pp. 232–256.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., et al. (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13: 722–736.

Makarova, K.S., Wolf, Y.I., and Koonin, E. V (2013) Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* gkt157.

Makarova, K.S., Wolf, Y.I., Van der Oost, J., and Koonin, E. V (2009) Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct* 4: 1.

Makarova, K.S., Wolf, Y.I., Snir, S., and Koonin, E. V (2011) Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J. Bacteriol.* 193: 6039–6056.

Mann, N.H., Cook, A., Millard, A., Bailey, S., and Clokie, M. (2003) Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* 424: 741.

Manning, A.J. and Kuehn, M.J. (2011) Contribution of bacterial outer membrane vesicles to innate bacterial defense. *BMC Microbiol.* 11: 258.

Martin, A., Yeats, S., Janekovic, D., Reiter, W.-D., Aicher, W., and Zillig, W. (1984) SAV 1, a temperate uv-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12. *EMBO J.* 3: 2165.

Matsushita, I. and Yanase, H. (2009) The genomic structure of thermus bacteriophage {phi}IN93. *J. Biochem.* 146: 775–85.

McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., and Paul, J.H. (2010) High frequency of horizontal gene transfer in the oceans. *Science* (80-.). 330: 50.

Mehta, D. and Satyanarayana, T. (2013) Diversity of Hot Environments and Thermophilic Microbes. In, *Thermophilic Microbes in Environmental and Industrial Biotechnology*. Springer Netherlands, Dordrecht, pp. 3–60.

Menon, S.K., Maaty, W.S., Corn, G.J., Kwok, S.C., Eilers, B.J., Kraft, P., et al. (2008) Cysteine usage in *Sulfolobus* spindle-shaped virus 1 and extension to hyperthermophilic viruses in general. *Virology* 376: 270–278.

Minakhin, L., Goel, M., Berdyuglova, Z., Ramanculov, E., Florens, L., Glazko, G., et al. (2008) Genome comparison and proteomic characterization of *Thermus thermophilus* bacteriophages P23-45 and P74-26: siphoviruses with triplex-forming sequences and the longest known tails. *J. Mol. Biol.* 378: 468–480.

Miroshnichenko, M.L. (2004) Thermophilic microbial communities of deep-sea hydrothermal vents. *Microbiology* 73: 1–13.

-N-

Nadal, M., Mirambeau, G., Forterre, P., Reiter, W.-D., and Duguet, M. (1986) Positively supercoiled DNA in a virus-like particle of an archaebacterium. *Nature* 321: 256–258.

Nagayoshi, Y., Kumagai, K., Mori, K., Tashiro, K., Nakamura, A., Fujino, Y., et al. (2016) Physiological Properties and Genome of the Hyperthermophilic Filamentous Phage ϕ OH. *With Life* *Thermus thermophilus* HB8. *Front. Microbiol.* 7: 50.

- Naryshkina, T., Liu, J., Florens, L., Swanson, S.K., Pavlov, A.R., Pavlova, N. V, et al. (2006) Thermus theophilus Bacteriophage ϕ YGe genome and Proteomic Characterization of Virus. J. Mol. Biol. 364: 667–677.
- Novick, R.P., Christie, G.E., and Penadés, J.R. (2010) The phage-related chromosomal islands of Gram-positive bacteria. Nat. Rev. Microbiol. 8: 541–551.
- Nunes-Alves, C. Bacteriophage biology: Putting the BREX on phage replication. Nat. Rev. Microbiol. 13: 129.
- P-
- Paez-Espino, D., Eloe-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., et al. Uncovering Earth's diversity. Nature advance on:
- Palm, P., Schleper, C., Grappin, B., Yeats, S., McWilliam, P., Reiter, W.-D., and Zillig, W. (1991) Complete nucleotide sequence of the virus SSV1 of the archaeobacterium Sulfolobus shibatae. Virology 185: 242–250.
- Paul, J.H. (2008) Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? ISME J. 2: 579–589.
- Pawlowski, A., Rissanen, I., Bamford, J.K.H., Krupovic, M., and Jalasvuori, M. (2014) Gammasphaerolipovirus, a newly proposed bacteriophage genus, unifies viruses of halophilic archaea and thermophilic bacteria within the novel family Sphaerolipoviridae. Arch. Virol. 159: 1541–1554.
- Penadés, J.R., Chen, J., Quiles-Puchalt, N., Carpena, N., and Novick, R.P. (2015) Bacteriophage-mediated spread of bacterial virulence genes. Curr. Opin. Microbiol. 23: 171–178.
- Peng, X., Basta, T., Häring, M., Garrett, R.A., and Prangishvili, D. (2007) Genome of the Acidianus bottle-shaped virus and insights into the replication and packaging mechanisms. Virology 364: 237–243.
- Personnic, S., Duhamel, S., Bettarel, Y., Sime-Ngando, T., and Jacquet, S. (2006) Les virus planctoniques : un compartiment biologique clé des milieux aquatiques. *Courr. l'évolution et l'INRA* 19–34.
- Pflugfelder, B., Fisher, C.R., and Bright, M. (2005) The color of the trophosome: elemental sulfur distribution in the endosymbionts of Riftia pachyptila (Vestimentifera; Siboglinidae). Mar. Biol. 146: 895–901.
- Pietilä, M.K., Demina, T.A., Atanasova, N.S., Oksanen, H.M., and Bamford, D.H. (2014) Archaeal viruses and bacteriophages: comparisons and contrasts. Trends Microbiol. 22: 334–344.
- Pina, M., Bize, A., Forterre, P., and Prangishvili, D. (2011) The archaeoviruses. FEMS Microbiol. Rev. 35: 1035–1054.
- Pollo, S.M.J., Zhaxybayeva, O., and Nesbø, C.L. (2015) Insights into thermoadaptation and the evolution of mesophily from the bacterial phylum Thermotogae. Can. J. Microbiol. 1–16.
- Prangishvili, D., Arnold, H.P., Götz, D., Ziese, U., Holz, I., Kristjansson, J.K., and Zillig, W. (1999) A novel virus family, the Rudiviridae: structure, virus-host interactions and genome variability of the Sulfolobus viruses SIRV1 and SIRV2. Genetics 152: 1387–1396.
- Prangishvili, D., Basta, T., Garrett, R.A., and Krupovic, M. (2016) Viruses of the Archaea. In, Encyclopedia of Life Sciences. John Wiley & Sons, Ltd.
- Prangishvili, D., Holz, I., Stieger, E., Nickell, S., Kristjansson, J.K., and Zillig, W. (2000) Sulfolobins, specific proteinaceous toxins produced by strains of the extremely thermophilic archaeal genus Sulfolobus. J. Bacteriol. 182: 2985–8.
- Prangishvili, D., Stedman, K., and Zillig, W. (2001) Viruses of the extremely thermophilic archaeon Sulfolobus. Trends Microbiol. 9: 39–43.
- Prieur, D. Les activités de l'éte, des assés à l'espace de. De Boek "Upu s.a.
- Prieur, D. (1997) Microbiology of deep-sea hydrothermal vents. Trends Biotechnol. 15: 242–244.
- Prieur, D. (1977) Microbiology of deep-sea hydrothermal vents. Mar. Biotechnol. 7799: 1–3.

-Q,R-

- Quemin, E.R.J., Pietilä, M.K., Oksanen, H.M., Forterre, P., Rijpstra, W.I.C., Schouten, S., et al. (2015) Sulfolobus Spindle-Shaped Virus 1 Contains Glycosylated Capsid Proteins, a Cellular Chromatin Protein, and Host-Derived Lipids. *J. Virol.* 89: 11681–11691.
- Rachel, R., Bettstetter, M., Hedlund, B.P., Häring, M., Kessler, a, Stetter, K.O., and Prangishvili, D. (2002) Remarkable morphological diversity of viruses and virus-like particles in hot terrestrial environments. *Arch. Virol.* 147: 2419–29.
- Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., et al. (2004) The 1.2-Megabase Genome Sequence of Mimivirus. *Science.* 306: 1344–1350.
- Raoult, D. and Forterre, P. (2008) Redefining viruses: lessons from Mimivirus. *Nat. Rev. Microbiol.* 6: 315 – 319.
- Reddy, T.B.K., Thomas, A.D., Stamatis, D., Bertsch, J., Isbandi, M., Jansson, J., et al. (2014) The Genomes OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta)genome project classification. *Nucleic Acids Res.*
- Reiter, W.-D., Palm, P., Yeats, S., and Zillig, W. (1987) Gene expression in archaeobacteria: physical mapping of constitutive and UV-inducible transcripts from the Sulfolobus virus-like particle SSV1. *Mol. Gen. Genet.* MGG 209: 270–275.
- Rensen, E., Krupovic, M., and Prangishvili, D. (2015) Mysterious hexagonal pyramids on the surface of Pyrobaculum cells. *Biochimie* 118: 365–367.
- Rensen, E.I., Mochizuki, T., Quemin, E., Schouten, S., Krupovic, M., and Prangishvili, D. (2016) A virus of hyperthermophilic archaea with a unique architecture among DNA viruses. *Proc. Natl. Acad. Sci.* 113: 2478–2483.
- Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., et al. (2004) The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc. Natl. Acad. Sci. U. S. A.* 101: 7716–7720.
- Rissanen, I., Grimes, J.M., Pawlowski, A., Mäntynen, S., Harlos, K., Bamford, J.K.H., and Stuart, D.I. (2013) Bacteriophage P23-77 capsid protein structures reveal the archetype of an ancient branch from a major virus lineage. *Structure* 21: 718–726.
- Rissanen, I., Pawlowski, A., Harlos, K., Grimes, J.M., Stuart, D.I., and Bamford, J.K.H. (2012) Crystallization and preliminary crystallographic analysis of the major capsid proteins VP16 and VP17 of bacteriophage P23-77. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 68: 580–583.
- Rohwer, F. and Thurber, R.V. (2009) Viruses manipulate the marine environment. *Nature* 459: 207–12.
- Rosenwasser, S., Ziv, C., Creveld, S.G. van, and Vardi, A. (2016) Virocell Metabolism: Metabolic Innovations During Host-Virus Interactions in the Ocean. *Trends Microbiol.*

-S-

- Sakai, H.D. and Kurosawa, N. (2016) Exploration and isolation of novel thermophiles in frozen enrichment cultures derived from a terrestrial acidic hot spring. *Extremophiles* 20: 207–214.
- Sakaki, Y. and Oshima, T. (1975) Isolation and characterization of a bacteriophage infectious to an extreme thermophile, *Thermus thermophilus* HB8. *J. Virol.* 15: 1449–1453.
- Salmond, G.P.C. and Fineran, P.C. (2015) A century of the phage: past, present and future. *Nat. Rev. Microbiol.* 13: 777–786.
- Samson, J.E., Magadán, A.H., Sabri, M., and Moineau, S. (2013) Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* 11: 675–687.
- Scanlan, D. (2014) Bacterial vesicles in the ocean. *Science* (80-.). 343: 143–144.
- Schoenfeld, T., Patterson, M., Richardson, P.M., Wommack, K.E., Young, M., and Mead, D. (2008) Assembly of Viral Metagenomes from Yellowstone Hot Springs. *Appl. Environ. Microbiol.* 74: 4164–4174.

- Seeger, A., Neuner, A., Kristjansson, J.K., and Stetter, K.O. (1986) *Acidianus infernus* gen. nov., sp. nov., and *Acidianus brierleyi* comb. nov.: facultatively aerobic, extremely acidophilic thermophilic sulfur-metabolizing archaeobacteria. *Int. J. Syst. Evol. Microbiol.* 36: 559–564.
- Servín-Garcidueñas, L.E., Peng, X., Garrett, R.A., and Martínez-Romero, E. (2013) Genome sequence of a novel archaeal rudivirus recovered from a mexican hot spring. *Genome Announc.* 1: e00040-12.
- Shelford, E.J., Middelboe, M., Møller, E.F., and Suttle, C.A. (2012) Virus-driven nitrogen cycling enhances phytoplankton growth. *Aquat. Microb. Ecol.* 66: 41–46.
- Shikuma, N.J., Pilhofer, M., Weiss, G.L., Hadfield, M.G., Jensen, G.J., and Newman, D.K. (2014) Marine tubeworm metamorphosis induced by arrays of bacterial phage tail –like structures. *Science* (80-.). 343: 529–533.
- Sime-Ngando, T. (2014) Environmental bacteriophages: viruses of microbes in aquatic ecosystems. *Front. Microbiol.* 5: 355.
- Sime-Ngando, T. and Colombet, J. (2009) Virus et prophages dans les écosystèmes aquatiques. *Can. J. Microbiol.* 55: 95–109.
- Snyder, J.C., Brumfield, S.K., Peng, N., She, Q., and Young, M.J. (2011) *Sulfolobus* turreted icosahedral virus c92 protein responsible for the formation of pyramid-like cellular lysis structures. *J. Virol.* 85: 6287 – 6292.
- Soler, N., Gaudin, M., Marguet, E., and Forterre, P. (2011) Plasmids, viruses and virus-like membrane vesicles from Thermococcales. *Biochem. Soc. Trans.* 39: 36–44.
- “ole , N., K upo, iM., Ma guet, E., ad Fo tter P. Me ae esiles i atu al e io e ts : a major challenge in viral ecology. *ISME J.* 9: 793–796.
- Soler, N., Marguet, E., Verbavatz, J.M., and Forterre, P. (2008) Virus-like vesicles and extracellular DNA produced by hyperthermophilic archaea of the order Thermococcales. *Res. Microbiol.* 159: 390–399.
- Song, Q., Ye, T., and Zhang, X. (2011) Proteins responsible for lysogeny of deep-sea thermophilic bacteriophage GVE2 at high temperature. *Gene* 479: 1–9.
- Song, Q. and Zhang, X. (2008) Characterization of a novel non-specific nuclease from thermophilic bacteriophage GBSV1. *BMC Biotechnol.* 8: 1.
- Sorek, R., Kunin, V., and Hugenholtz, P. (2008) CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6: 181–186.
- Stark, E. and Tetrault, P.A. (1951) Isolation of bacterial, cell-free, starch saccharifying enzymes from the medium at 70 C. *J. Bacteriol.* 62: 247.
- Stassen, A.P.M., Folmer, R.H.A., Hilbers, C.W., and Konings, R.N.H. (1994) Single-stranded DNA binding protein encoded by the filamentous bacteriophage M13: structural and functional characteristics. *Mol. Biol. Rep.* 20: 109–127.
- Stedman, K.M., DeYoung, M., Saha, M., Sherman, M.B., and Morais, M.C. (2015) Structural insights into the architecture of the hyperthermophilic Fusellovirus SSV1. *Virology* 474: 105–109.
- Sullivan, M.B., Lindell, D., Lee, J.A., Thompson, L.R., Bielawski, J.P., and Chisholm, S.W. (2006) Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol.* 4: e234.
- Suttle, C.A. (2007) Marine viruses - major players in the global ecosystem. *Nat. Rev. Microbiol.* 5: 801–812.
- Suttle, C.A. (2005) Viruses in the sea. *Nature* 437: 356–361.

-T-

- Takai, K., Nakamura, K., Toki, T., Tsunogai, U., Miyazaki, M., Miyazaki, J., et al. (2008) Cell proliferation at 122°C and isotopically heavy CH₄ production by a hyperthermophilic methanogen under high-pressure cultivation. *Proc. Natl. Acad. Sci.* 105: 10949–10954.

- Takami, H., Nishi, S., Lu, J., Shimamura, S., and Takaki, Y. (2004) Genomic characterization of thermophilic *Geobacillus* species isolated from the deepest sea mud of the Mariana Trench. *Extremophiles* 8: 351–356.
- Tamakoshi, M., Murakami, A., Sugisawa, M., Tsuneizumi, K., Takeda, S., Saheki, T., et al. (2011) Genomic analysis of the late stage of the life cycle of the thermophilic bacteriophage ϕ TMA. *Bacteriophage* 1: 152–164.
- Tarasov, V.G., Gebruk, A. V., Mironov, A.N., and Moskalev, L.I. (2005) Deep-sea and shallow-water hydrothermal vent communities: Two different phenomena? *Chem. Geol.* 224: 5–39.
- Thingstad, T.F. and Bratbak, G. (2016) Microbial oceanography: Viral strategies at sea. *Nature*.
- Thingstad, T.F. and Lignell, R. (1997) Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat. Microb. Ecol.* 13: 19–27.
- Thornburg, C.C., Zabriskie, T.M., and McPhail, K.L. (2010) Deep-Sea Hydrothermal Vents: Potential Hot Spots for Natural Products Discovery? *J. Nat. Prod.* 73: 489–499.
- Tock, M.R. and Dryden, D.T.F. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.* 8: 466–472.
- Torrella, F. and Morita, R.Y. (1979) Evidence by electron micrographs for a high incidence of bacteriophage particles in the waters of Yaquina Bay, Oregon: ecological and taxonomical implications. *Appl. Environ. Microbiol.* 37: 774–778.
- Torsvik, T. and Dundas, I.D. (1974) Bacteriophage of *Halobacterium salinarum*. *Nature* 248: 680–681.
- U,V-
- Uldahl, K. and Peng, X. (2013) Biology, biodiversity and application of thermophilic viruses. In, *Thermophilic Microbes in Environmental and Industrial Biotechnology*. Springer, pp. 271–304.
- Uldahl, K.B. (2015) Hyperthermophilic archaeal viruses as novel nanoplateforms.
- Vestergaard, G., Häring, M., Peng, X., Rachel, R., Garrett, R.A., and Prangishvili, D. (2005) A novel rudivirus, ARV1, of the hyperthermophilic archaeal genus *Acidianus*. *Virology* 336: 83–92.
- Vestergaard, G., Shah, S.A., Bize, A., Reitberger, W., Reuter, M., Phan, H., et al. (2008) SRV, a new rudiviral isolate from *Stygiolobus* and the interplay of crenarchaeal rudiviruses with the host viral-defence CRISPR system. *J. Bacteriol.*
- Von Damm, K.L., Oosting, S.E., Kozlowski, R., Buttermore, L.G., Colodner, D.C., Edmonds, H.N., et al. (1995) Evolution of East Pacific Rise hydrothermal vent fluids following a volcanic eruption. *Nature* 375: 47–50.
- W,X,Y,Z-
- Waldor, M.K. and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* (80-). 272: 1910.
- Wang, Y. and Zhang, X. (2008a) Characterization of a novel portal protein from deep-sea thermophilic bacteriophage GVE2. *Gene* 421: 61–66.
- Wang, Y. and Zhang, X. (2010) Genome Analysis of Deep-Sea Thermophilic Phage D6E. *Appl. Environ. Microbiol.* 76: 7861–7866.
- Wang, Y. and Zhang, X. (2008b) Identification and characterization of a novel thymidylate synthase from deep-sea thermophilic bacteriophage *Geobacillus* virus E2. *Virus Genes* 37: 218–224.
- Wei, D. and Zhang, X. (2008) Identification and characterization of a single-stranded DNA-binding protein from thermophilic bacteriophage GVE2. *Virus Genes* 36: 273–278.
- Weinbauer, M.G. (2004) Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28: 127–181.
- Weinbauer, M.G., Chen, F., and Wilhelm, S.W. (2011) Virus-mediated redistribution and partitioning of carbon in the global oceans. In, *Microbial carbon pump in the ocean*, pp. 54–56.

- Williamson, S.J., Cary, S.C., Williamson, K.E., Helton, R.R., Bench, S.R., Winget, D., and Wommack, K.E. (2008) Lysogenic virus host interactions predominate at deep-sea diffuse-flow hydrothermal vents. *ISME J.* 2: 1112–1121.
- Windberger, E., Huber, R., Trincone, A., Fricke, H., and Stetter, K.O. (1989) *Thermotoga thermarum* sp. nov. and *Thermotoga neapolitana* occurring in African continental solfataric springs. *Arch. Microbiol.* 151: 506–512.
- Winter, C., Bouvier, T., Weinbauer, M.G., and Thingstad, T.F. (2010) Trade-offs between competition and defense specialists among *Ullula* planktonic organisms: the killing the eater hypothesis revisited. *Microbiol. Mol. Biol. Rev.* 74: 42–57.
- Woese, C.R. (1994) There must be a prokaryote somewhere in the biologist's search for itself. *Microbiol. Rev.* 58: 1–9.
- Woese, C.R., Kandler, O., and Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci.* 87: 4576–4579.
- Wu, S., Liu, B., and Zhang, X. (2009) Identification of a tail assembly gene cluster from deep-sea thermophilic bacteriophage GVE2. *Virus Genes* 38: 507–514.
- Xu, T., Yao, F., Zhou, X., Deng, Z., and You, D. (2010) A novel host-specific restriction system associated with DNA backbone S-modification in *Salmonella*. *Nucleic Acids Res.* 38: 7133–7141.
- Ye, T. and Zhang, X. (2008) Characterization of a lysin from deep-sea thermophilic bacteriophage GVE2. *Appl. Microbiol. Biotechnol.* 78: 635–641.
- Yoshida-Takashima, Y., Takaki, Y., Shimamura, S., Nunoura, T., and Takai, K. (2013) Genome sequence of a novel deep-sea vent epsilonproteobacterial phage provides new insight into the co-evolution of Epsilonproteobacteria and their phages. *Extremophiles* 17: 405–419.
- Yu, M.X., Slater, M.R., and Ackermann, H.W. (2006) Isolation and characterization of *Thermus* bacteriophages. *Arch. Virol.* 151: 663–679.
- Zhaxybayeva, O., Swithers, K.S., Lapierre, P., Fournier, G.P., Bickhart, D.M., DeBoy, R.T., et al. (2009) On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. U. S. A.* 106: 5865–5870.

Objectifs des travaux de thèse

Nous savons, aujourd'hui, les virus sont omniprésents dans tous les océans, compris les plus extrêmes, où ils représentent le plus grand réservoir de biodiversité non caractérisée. Néanmoins, notre connaissance de la diversité virale associée aux bactéries et archées issues des sources hydrothermales océaniques profondes demeure encore limitée. Peu de données sont disponibles sur l'abondance virale ou sur l'impact des virus sur les micro-organismes colonisant ces écosystèmes. En effet, seuls 8 virus infectant des archées et 14 autres isolés de bactéries ont été identifiés. Deux d'entre eux sont associés à des alphaproteobactéries appartenant à l'ordre des Thermococcales. Il s'agit de PAV et TPV, isolés au LME. Les six autres virus sont associés à des bactéries appartenant aux Bacillales, Epsilon-proteobacteria et Thermotogales. Une revue présentant les éléments génétiques mobiles, virus, plasmides et vésicules, isolés de sources hydrothermales océaniques vient d'être publiée : An abyssal mobilome: virus, plasmid and vesicle from deep sea hydrothermal vents par J. Lossouarn, S. Dupont, A. Gorlas, C. Mercier, N. Bienvenu, P. Forterre and C. Geslin in *Research in microbiology*. Celle-ci constitue le premier chapitre du manuscrit.

Afin de poursuivre l'effort d'investigation de cette diversité virale au sein des sources hydrothermales océaniques profondes, un criblage des bactéries du phylum des Thermotogae avait été initié au laboratoire afin d'identifier de nouveaux éléments génétiques mobiles. Cet ordre bactérien, composé de bactéries chimio-organotrophes, anaérobies et majoritairement thermophiles, est largement représenté au niveau de la biosphère profonde. Partageant la même niche écologique que les Thermococcales, les Thermotogae ont une histoire évolutive complexe et particulière qui suggère la présence de

nombreux transferts latéraux de gènes. Ces transferts sont souvent dus à différents types d'éléments génétiques mobiles.

D'autre part, la présence de CRISPR et de *pus* a été mise en évidence dans de nombreux génomes de *Thermotogae*. Malgré toutes ces preuves laissant penser à des infections virales présentes ou passées, jusqu'à aujourd'hui, seul un élément infectant une bactérie de cette famille a été identifié. C'est d'ailleurs dans notre thèse que nous avons découvert l'élément d'asie d'origine latérale de MPV1 (*Marinitoga piezophila* virus 1), premier virus associé à une *Thermotoga*. La caractérisation de ce système très original, menée dans le cadre de la thèse de J. Lossouarn, a ainsi pu être finalisée et publiée (Lossouarn J, Nesbø CL, Mercier C, Zhaxybayeva O, Johnson MS, Charchuck R, Farasin J, Bienvenu N, Baudoux AC, Michoud G, Jebbar M, and Geslin C. (2015) "Ménage à trois": a selfish genetic element uses a virus to propagate within *Thermotogales*. In *Environ Microbiol.* doi: 10.1111/1462-2920.12783). Il constitue le second chapitre de ce manuscrit.

Lors du criblage des *Thermotogae*, d'autres virus ont été mis en évidence et notamment deux virus infectant deux souches de *Marinitoga camini*. J'ai eu l'opportunité de caractérisation fonctionnelle et génomique de ces nouveaux virus (MCV1 et MCV2) qui fera l'objet d'un prochain article. D'autre part, les génomes des deux souches hôtes de ces virus ont été séquencés et nous avons fait l'objet d'un « genome announcement » (publication acceptée). Ces travaux constituent le troisième chapitre du manuscrit.

Enfin, lors de l'étude des trois virus, une souche appartenant au genre *Thermosiphon* a particulièrement attiré notre attention. En effet, pMP1 (plasmide porté par *M. piezophila* KA3, hôte du virus MPV1) peut se propager au sein de cette souche de *Thermosiphon* sp. AT1244-VC14. Néanmoins, MCV1 et MCV2 ne sont pas capables d'infecter cette souche. Ces résultats ont été publiés au sujet de cette souche ainsi qu'à l'étude des CRISPR

présents dans son génome. **Ces travaux feront l'objet d'une partie de type «Genome report »** et sont présentés dans le quatrième chapitre de ce manuscrit.

En parallèle, j'ai travaillé à un projet destiné à étendre le plateau protocolaire vu de produire un volume de séquences à plusieurs profonds. En effet, il y a une peu d'analyses de métagénomique virale publiées sur des échantillons de sources hydrothermales océaniques profondes. La plus complète, portant sur des échantillons de flux hydrothermaux diffus, révèle que les virus hydrothermaux tempérés seraient moins diversifiés que ceux composant la communauté virioplanctonique globale mais qu'ils présenteraient un très grand nombre de gènes aux fonctions inconnues (Williamson et al., 2008). Dans ce contexte, le LM2E souhaite analyser par une approche métagénomique des échantillons de sédiments hydrothermaux afin d'accéder à une diversité génomique de nouveaux virus infectant des bactéries et archées. Les expériences débutées lors de ce projet, la validation et la **normalisation du protocole** pour obtenir un virome (étude préliminaire), sont poursuivies dans le cadre d'une thèse réalisée au LUBEM de Brest par Lisa Billard sous la direction des Pr. C. Payan et Dr S. Vallet.

Chapitre I : Revue bibliographique

« An abyssal mobilome »

Présentation des travaux

Depuis leur découverte en 1977, les sources hydrothermales profondes font l'objet de nombreuses études. Elles abritent une riche diversité de micro-organismes aux métabolismes variés et au potentiel biotechnologique notable. Pourtant, les éléments génétiques mobiles (EGMs) qui y sont associés restent encore peu décrits (Lossouarn, Dupont, et al., 2015).

Les EGMs regroupent les virus, les vésicules membranaires (MVs), les plasmides, les agents de transferts de gènes (GTAs, gene transfer agents), les transposons et les transpovirions. Ils jouent un rôle important dans l'évolution des communautés microbiennes notamment de part leur implication dans les transferts latéraux de gènes. Ils sont aussi très présents, particulièrement abondants dans les environnements extrêmes comme les sources hydrothermales profondes (Anderson et al., 2014).

L'objectif de cette revue est de présenter les connaissances actuelles jusqu'à présent sur les virus, vésicules membranaires, et plasmides associés aux bactéries et archées thermophiles et hyperthermophiles colonisant les sources hydrothermales océaniques profondes.



ELSEVIER



CrossMark



Institut Pasteur

Research in Microbiology 166 (2015) 742e752



www.elsevier.com/locate/resmic

An abyssal mobilome: viruses, plasmids and vesicles from deep-sea hydrothermal vents

Julien Lossouarn^a, Samuel Dupont^a, Aurore Gorlas^b, Coraline Mercier^a, Nadege Bienvenu^a,
Evelyne Marguet^a, Patrick Forterre^c, Claire Geslin^{a,*}

^a Laboratory of Microbiology of Extreme Environments (LMEE), UMR 6197/UBO/Ifremer/CNRS, IUEM, Place Nicolas Copernic, Technopôle Brest Iroise, 29280, Plouzané, France

^b Institute of Integrative Cellular Biology, CEA, CNRS, Université Paris sud, 91405 Orsay, France

^c Institut Pasteur, Laboratoire de Biologie Moléculaire du Gène chez les Extrémophiles, 25 Rue du Docteur Roux, Paris Cedex 15, France

Received 30 January 2015; accepted 9 April 2015

Available online 22 April 2015

Abstract

Mobile genetic elements (MGEs) such as viruses, plasmids, vesicles, gene transfer agents (GTAs), transposons and transpovirons, which collectively represent the mobilome, interact with cellular organisms from all three domains of life, including those thriving in extreme environments. While efforts have been made to better understand deep-sea vent microbial ecology, our knowledge of the mobilome with prokaryotes inhabiting deep-sea hydrothermal vents remains limited. Here we focus on the abyssal mobilome by reviewing data on viruses, plasmids and vesicles associated with thermophilic and hyperthermophilic Bacteria and Archaea present in deep-sea hydrothermal vents.

© 2015 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

Keywords: Deep-sea hydrothermal vent; Bacteria; Archaea; (Hyper-)thermophiles; Mobilome

1. Introduction

Deep-sea hydrothermal vents represent one of the most extreme environments on Earth. These ecosystems are characterized by steep physicochemical gradients, high hydrostatic pressure, high temperatures, obscurity and the prevalence of chemosynthesis. These extreme environments are home to a vast diversity of mesophilic and (hyper-)thermophilic prokaryotes belonging to the Bacteria and Archaea [1e3]. Although our knowledge of deep-sea hydrothermal vent microbial communities is progressing, the impact of MGEs on microbial ecology and evolution remains largely overlooked in these abyssal ecosystems [2]. MGEs such as viruses, plasmids, membrane vesicles, gene transfer agents (GTAs), transposons and transpovirons, which collectively represent the mobilome, interact with cellular organisms from all three domains of life, including those thriving in extreme environments [4,5]. Many reviews have highlighted how MGEs, especially viruses, are powerful agents that affect not only the diversity and evolution of microbial communities but also global biochemical cycles in marine environments [6e12]. Evidence was recently reviewed supporting the hypothesis that MGEs could also play a key role in deep-sea hydrothermal vental, notably by facilitating horizontal gene transfer (HGT) [2]. MGEs have potential as powerful drivers of cellular host adaptations to extreme marine environments [2]. Here we present the first review focused on deep-sea hydrothermal MGEs.

* Corresponding author.

E-mail addresses: julienlossouarn@yahoo.fr (J. Lossouarn), samueldupont92@gmail.com (S. Dupont), aurore.gorlas@gmail.com (A. Gorlas), coraline.mercier@univ-brest.fr (C. Mercier), nadege.bienvenu@univ-brest.fr (N. Bienvenu), evelyne.marguet@igmors.u-psud.fr (E. Marguet), patrick.forterre@pasteur.fr (P. Forterre), claire.geslin@univ-brest.fr (C. Geslin).

<http://dx.doi.org/10.1016/j.resmic.2015.04.001>

0923-2508/© 2015 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

Table 1
Bacterial and archaeal viruses isolated from deep-sea hydrothermal vents.

Host strains (growth temperature)	Family	Virions morphology	Virus-host relationship	Genomes
Bacteriophages references				
BVW1 [29] <i>Bacillus</i> sp.w13 (65°C)	Unclassified	Long flexible tail 300 nm Hexagonal head 70 nm diameter	Lytic	Double-stranded linear DNA 18 kb
GVE1 [29] <i>Geobacillus</i> sp.E26323 (65°C)	<i>Siphoviridae</i>	Flexible tail 180 nm Hexagonal head 130 nm diameter	Lytic	Double-stranded linear DNA 41 kb
GVE2 [30] <i>Geobacillus</i> sp.E263 (65°C)	<i>Siphoviridae</i>	Not described	Lytic, potentially lysogenic	Double-stranded linear DNA 40.9 kb
D6E [31] <i>Geobacillus</i> sp.E263 (65°C)	<i>Myoviridae</i>	Contractile tail 60 nm Hexagonal head 60 nm diameter	Lytic	Double-stranded linear DNA 49.3 kb
Nsr-1 [32] <i>Nitratiruptor</i> sp.SB155-2 (55°C)	<i>Siphoviridae</i>	Flexible tail 210 nm Hexagonal head 64 nm diameter	Lysogenic	Double-stranded linear DNA 37.1 kb
MPV1 [34] <i>Marinitoga piezophila</i> KA3 (65 °C)	<i>Siphoviridae</i>	Flexible tail 200 nm Hexagonal head 50 nm	Lysogenic	Double-stranded linear/circular DNA 43.715 kb
Archeoviruses references				
PAV1 [23,24] <i>Pyrococcus abyssi</i> GE23 (85°C)	<i>Fuselloviridae</i>	Lemon-shaped 120 nm length 80 nm width	Carrier state	Double-stranded linear DNA 18 kb
TPV1 [25] <i>Thermococcus pieurii</i> (80°C)	<i>Fuselloviridae</i>	Lemon-shaped 140 nm length 80 nm width	Carrier state	Double-stranded linear DNA 21.5 kb

associated with (hyper-)thermophilic prokaryotes, collectively denoted as the abyssal mobilome.

2. Viruses in deep-sea hydrothermal vents

2.1. Evidence for viral activity

Only a few viral ecological studies have been performed in deep-sea hydrothermal vents [13e17]. Viral abundance and viral production were notably investigated in diffuse flow hydrothermal vent fluids. In these samples collected from vents within the Endeavour Ridge system [13] and the Pacific Rise [14], average VLP (virus-like particles) abundances were estimated at 10 VLPs per milliliter and were ~10-fold higher than prokaryote abundances. In comparison, VLP abundances in productive coastal waters were estimated at ~10 VLPs/mL and exceeded those of prokaryotes by 10-fold [6,7]. These VLPs might represent bona fide viral particles (virions), but also membrane vesicles containing plasmid or viral DNA (viral membrane vesicles) (see section 3).

Viral activities occurring in these extreme ecosystems are also highlighted by genome sequence analysis of deep-sea bacteria and Archaea [2]. The presence of clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated (Cas) systems has been reported in many thermophilic bacterial and archaeal genomes [18,19]. These systems provide acquired, yet heritable, sequence-specific “adaptive” immunity against and other horizontally acquired elements, such as conjugative plasmids [19]. CRISPR loci consist of several non-contiguous direct repeats separated by stretches of variable sequences called spacers, which correspond to fragments derived from invading DNAs such as viruses and plasmids. CRISPR regions therefore act as a record of past infections that occurred in the history of the prokaryotes [2,19,20].

Interestingly, it was reported that thermophilic strains harbored a higher number of CRISPRs in their genomes than mesophilic and psychrophilic strains [21]. This may indicate that viral infections play a major role in the ecology and evolution of thermophilic communities, notably those inhabiting deep-sea hydrothermal vents [2].

A viral metagenomic study using CRISPRs, indicated that a diffuse flow sample collected from Hulk vent on the Juan Fernandez Ridge in the Pacific ocean contained a range of viruses infecting mesophilic and thermophilic hosts from both the archaeal and bacterial domains [22]. Most of the viral reads belonged to the *Myoviridae*. Other tailed viruses frequently identified were the *Podoviridae* and the *Siphoviridae*, which also recovered. Archaeoviruses belonging to the *Rudoviridae*, *Fuselloviridae* and *Lipothirixviridae*, which are frequently found in hot spring virus assemblages, were largely absent from this marine ventome [22]. The abundance of Archaea in deep-sea hydrothermal vents is strongly suggests that archaeoviruses were present in the marine vent virome. The problem is that little is known about the virosphere of hydrothermal environments, which is more the consequence of insufficient screening.

Low virus abundance led to date only two viruses have been isolated from described marine hyperthermophilic Archaea [23e25], in addition to proviruses and plasmids (see next sections). Therefore, deep-sea hydrothermal systems play host to novel archaeoviruses.

Lysogeny is actually presumed to be a more common cycle in deep-sea hydrothermal vents than those listed in Table 1. Comparative analysis of cellular and viral metagenomes obtained from a Hulk diffuse flow sample revealed higher enrichment of phages in the cellular fraction than in a range of other aquatic and terrestrial metagenomes [26]. This result complements the high proportion of inducible lysogenic microorganisms previously

reported in deep-sea hydrothermal vents [14]. Furthermore, the genome composed of 51 ORFs and its sequence organization are distinct from those of any other previously characterized siphoviruses. Homologues of *NrS-1* genes were found to be relatively high abundance of auxiliary metabolic genes were found in the vent viral fraction compared to the cellular fraction. All these results suggest the prevalence of temperate distributed among the bacterial genomes of *Epsilonproteobacteria* in deep-sea vents. The wide distribution of these genes may illustrate early co-evolution between *Epsilonproteobacteria* and their temperate bacteriophages prior to the divergence of these habitats. The integrated state represents a means to avoid the physiological adaptations [32].

2.2. Virus host systems characterized

Despite obvious viral activity in deep-sea hydrothermal environments, few viruses have been isolated and characterized to date.

Within the Bacteria, 6 bacteriophages (i.e. bacteriophages) [35]. MPV1 is a temperate *Siphoviridae*-like virus with a 43.7 kb genome (Fig. 1 A). This double-stranded viral DNA has head and tail morphotypes reveals a connection to genomes of *Srificutes* and bacteriophages known to infect them. Surprisingly, MPV1 virions carry not only the viral DNA, but preferentially package a plasmid of 13.3 kb (pMP1). This plasmid is the second mobile genetic element carried by *M. piezophila*. The system described in *M. piezophila* is the first virus-mediated plasmid exchange reported in deep-sea hydrothermal vents and could correspond to a new example of molecular piracy (Fig. 2) [36].

Four of these deep-sea hydrothermal bacteriophages are exchange reported in deep-sea hydrothermal vents and could correspond to a new example of molecular piracy (Fig. 2) [36]. To date, only two viruses have been isolated and described from Archaea living in deep-sea hydrothermal vents, both infecting *Euryarchaea* of the order of *Thermococcales* (Table 1). These bacteriophages infect *Thermococcales* have been frequently isolated from marine environments, notably from deep-sea vents [37]. The order is represented by 3 genera: *Thermococcus*, *Pyrococcus*, and *Palaeococcus*, and includes obligate anaerobes, sulfur-metabolizing hyperthermophiles. *Thermococcales* are widely studied model microorganisms in various fields of investigation, including microbial adaptation to extreme temperature and ionizing radiation. Replication mechanisms, metabolism, phylogeny and genome evolution have been extensively studied. Analyses revealed extensive mosaicism in the D6E viral genome with other mesophilic and thermophilic bacteriophages [38]. Accordingly, intensive effort has been made to characterize D6E replication and transcription sequence diverse genomes of members of this order, resulting in the availability of 14 genomes of *Thermococcus*, 6 genomes of *Pyrococcus*, and 1 of *Palaeococcus*. The availability of all these data makes *Thermococcales* a good model for exploring the dispersion of DNA sequences in genomes of deep-sea hyperthermophilic and marine virosphere. Currently, remains largely uncharacterized. Phages have been detected in several genomes of *Thermococcales* (TKV1 to TKV4 in *Thermococcus kodakaraensis* [38], TGV1 and TGV2 in *Thermococcus gammatolerans* [39], PHV1 in *Pyrococcus horikoshii* genome [40]).

NrS-1 is a temperate siphovirus associated with a member of the *Epsilonproteobacteria* [32]. Members of this bacterial phylum have frequently been found to dominate microbial communities inhabiting deep-sea hydrothermal environments [3]. *NrS-1* infects the chemolithoautotrophic, anaerobic and microaerobic moderately thermophilic strain *Nitratiruptor* SB155-2 cultivated at 55 °C [33]. This lysogenic strain was isolated from a hydrothermal chimney sample collected at the Ibeya North field (27°N, 126°33'E, 1000 m depth). *NrS-1* virions contain a double-stranded linear DNA genome of 37.1 kb which is circularly permuted and terminally redundant. *NrS-1* is affiliated with the *Siphoviridae* family, its growth and led to the detection of a vast morphological

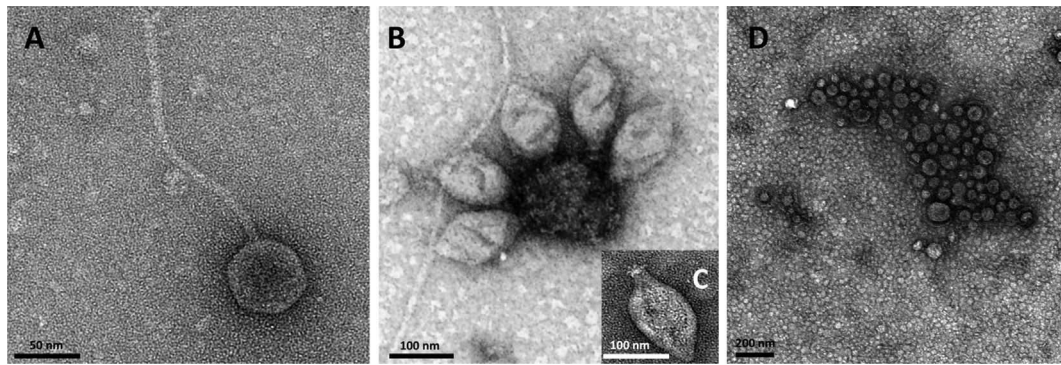


Fig. 1. Electron micrographs of viruses and membrane vesicles isolated from deep-sea hydrothermal vents, negatively stained with 2% uranyl acetate. A: PAV1, a virus isolated from *Marinitoga piezophila* KA3. B: PAV1, a virus isolated from *Pyrococcus abyssi* GE23. C: TPV1, a virus isolated from *Thermococcus prius*. D: Membrane vesicles from *Thermococcus prius*.

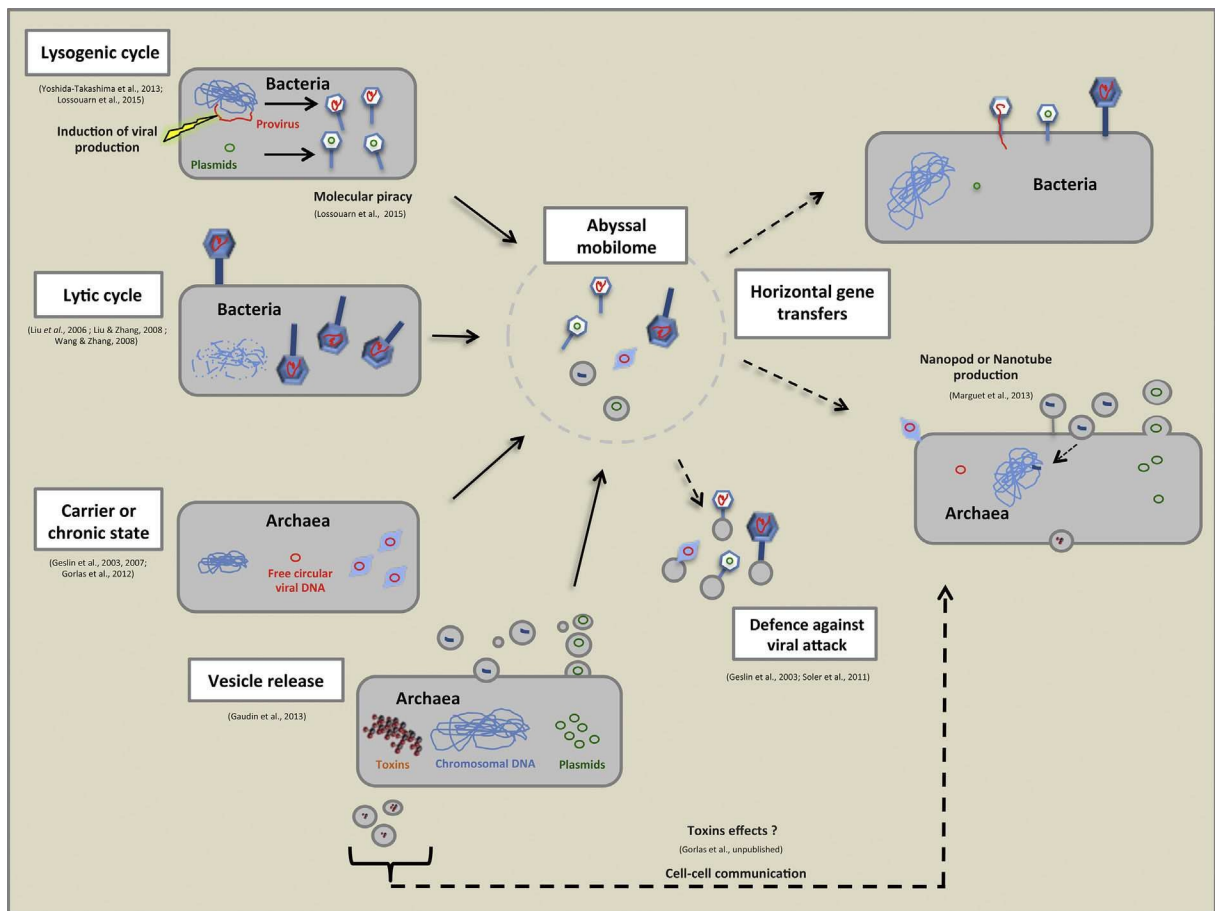


Fig. 2. Multipurpose mobilome within deep-sea hydrothermal vents. Possible roles of viruses (via their viral cycles, lysogenic, carrier state or chronic), vesicles and plasmids in HGT of cellular, viral or plasmid DNA (in blue, red or green, respectively); a specific role for vesicles in defense against viral attack. Many other roles in cell-cell communication. This is a schematic presentation and is not to scale. Arrows with solid lines indicate demonstrated roles, dotted lines show more speculative ones. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

diversity of VLPs by transmission electron microscopy [41]. Among the different morphotypes observed, the lemon-shaped type prevailed. Outcrops and new pleomorphic morphologies were also reported. Additionally, this study highlighted the striking similarity between VLP morphologies from deep-sea hydrothermal vents and those from terrestrial environments [41]. Thereafter, two viruses associated with *Thermococcales* strains were isolated and characterized. The virus PAV1 was discovered in *Pyrococcus abyssi* strain GE23, isolated from a deep-sea sample collected in the North Fidi Basin (White Lady Site 16°59'S, 173°55'W, 2000 m depth) and cultivated at 85°C [42]. PAV1 is lemon-

Table 2
 Bacterial plasmids (associated with thermophilic strains) isolated from deep-sea hydrothermal vents.

Plasmids	Size (kp)	Host stains (growth temperatures)	Replication	Informations	References
Unnamed	53.6	<i>Persephonella marina</i> Ex-H1 (73°C)	unknown	DNA replication protein DnaC DNA topoisomerase IA	[49]
pDF308	308.54	<i>Deferribacter desulfuricans</i> SSM1 (60e65°C)	unknown	CRISPR/cas system Transposases	[52]
pOCEPR01	135.3	<i>Oceanithermus profundus</i> 506 (60°C)	unknown	DNA helicase DNA polymerase DNA primase polymerase bacteriolytic enzymes CRISPR/cas system	[50]
pTHEAM01	76.6	<i>Thermovibrio ammonificans</i> HB-1 (75°C)	unknown	DNA topoisomerase IA	[51]
pMP1	13.3	<i>Marinitoga piezophila</i> KA3 (65°C)	unknown	DNA primase Integrase/recombinase	[34,53]

shaped (120 " 80 nm) with a short tail terminated by fibers with two plasmids pTP1 (3.1 kb) and pTP2 (2.0 kb). These 3 (Fig. 1 B). PAV1 virions contain a double-stranded circular DNA of 18 kb which is also present in high copy number in the host cell, without any apparent plasmid form in the host genome size and antagonistic effect on each other [25,44].

virion morphology, PAV1 strikingly resembles the SSV fuselloviruses isolated from aerobic hot spring *Crenarchaeota* belonging to both *Sulfolobus* and *Acidianus* genera [23]. However, none of its 25 predicted proteins, except one (MCP), exhibits similarity to SSV proteins. Instead, PAV1 is evolutionarily related to plasmids of *Thermococcales*, with three genes having homologues in plasmid pTN22 in plasmid pP12-1 and one in pP11. In addition, PAV1 shares one gene with provirus TKV4 of *T. kodakaraensis* [24,36,44]. PAV1 persists in the host strain in a stable carrier state. Indeed, PAV1 virions are spontaneously and continuously released from the host cell without causing its lysis or growth delay [23]. The infectivity of PAV1 has been demonstrated by the inverted spot test method, which indicates that PAV1 can infect *Pyrococcus glycovorans* leading to growth retardation but does not infect any of the *Thermococcus* strains tested [42].

Virus TPV1 was discovered in *Thermococcus prieurii* isolated from a hydrothermal chimney sample collected from the East Pacific Rise ("Sarah Spring" area, 7°25'24" S, 107°47'66" W, 2700 m depth) and cultivated at 80°C [43]. TPV1 is lemon-shaped (140 " 80 nm) with a short tail terminated by fibers (Fig. 1 C), similarly to PAV1 [25]. TPV1 contains a double-stranded circular DNA of 21.5 kb which is also present in high copy number in a free form in the host cell. The TPV1 genome encompasses 28 predicted genes. The proteins and transcriptional regulators predicted to be involved in genome replication were identified. TPV1 encodes also a predicted integrase (SSV-type tyrosine recombinase) and a family. The only two genes that are homologous between TPV1 and PAV1 encode proteins containing a concanavalin A-like lectin/glucanase domain that might be involved in host recognition. TPV1 shares more homologous genes with proviruses, (TKV2, TKV3, TKV4, TGV1) and plasmid (pT26: two cryptic miniplasmids from *Thermococcus* species, than with PAV1 [25]. Infection of TPV1 with TPV1 does not cause host lysis, and viral replication can be induced by UV irradiation. TPV1 can infect several reference species belonging to the *Thermococcus* genus leading to growth retardation [42]. Interestingly, TPV1 shares its host

3. Plasmids in deep-sea hydrothermal vents

Plasmids, including those of (hyper-)thermophiles, are abundant in the biosphere. They actively participate in HGT, which plays a major role in microbial plasticity, adaptation and evolution [45e47]. To date, only 6 bacterial plasmids and 24 archaeal plasmids have been isolated and described from deep-sea hydrothermal vents.

Bacterial domain

Among the few plasmids isolated from deep-sea vent bacteria, pDF308 is associated with a mesophilic strain [48] while the other five are carried by thermophilic strains [49e53]. These five bacterial plasmids are given in Table 2. The latter have generally not been described in detail, and most have only been annotated and briefly mentioned in the genome description of their hosts. We will briefly focus on two of them.

pDF308 is a 308.54 kb megaplasmid associated with the thermophilic, anaerobic and heterotrophic strain *Deferribacter desulfuricans* SSM1 cultivated at 65°C [52,54]. This megaplasmid, whose two-thirds of the CDSs have no apparent similarity with any CDSs in the database, interestingly encodes 17 copies of a gene cluster encoding two transposases, the IS200 and IS605 family, whereas none were found in the host genome [52]. Very similar CRISPR/Cas systems in genome replication were identified. pDF308 and *D. desulfuricans* genomes [52].

pMP1 is a plasmid recently described within the *Thermomargarita* order [34,53]. This 13.3 kb genetic element, carried by *Marinitoga piezophila* (previously introduced in 1.2 section), is unrelated to the only known plasmids among the *Thermotoga* family: two cryptic miniplasmids of the genus *Thermotoga* (pT26: a 1265 bp miniplasmid in *Thermotoga prima* [56] and a 1724 bp miniplasmid in *Mesotoga prima* [57]). pMP1 encodes a site-specific recombinase XerD (primase family) and a protein with both a DnaB (helicase) and a DnaG (primase) domain. A large fraction of the encoded proteins might actually be involved in DNA binding, replication and

Table 3
Archaeal plasmids isolated from deep-sea hydrothermal vents.

Plasmids	Family	Subfamily	Size (kb)	Host strains (growth temperatures)	Replication Informations	References
pTN1	pTN1		3.4	<i>Thermococcus nautili</i> 30-1 (85°C)	Rolling-circle type Rep74	[61]
pGT5	pTN1		3.4	<i>Pyrococcus abyssi</i> GE5 (95°C)	Rolling-circle type Rep75	[62]
pTP1	pTN1		3.1	<i>Thermococcus prieurii</i> (80°C)	Rolling-circle type RepTP1	[44]
pTP2	pTP2		2	<i>Thermococcus prieurii</i> (80°C)	Rolling-circle type RepTP2	[44]
pRT1	pRT1		3.4	<i>Pyrococcus</i> sp.JT1 (95°C)	Theta type Rep63	[70]
pAMT11	pRT1		20.5	<i>Thermococcus</i> sp.AMT11 (85°C)	Theta type Rep72 integrase SSV-type	[71]
pT26-2	pT26-2		21.6	<i>Thermococcus</i> sp.26/2 (85°C)	Theta type ATPase (replicative helicase) Integrase SSV-type	[40]
pTBMP1	pTBMP1		54.2	<i>Thermococcus barophilus</i> MP (85°C)	Theta type	[73]
pTN2	pTN2	pTN2-like	13	<i>Thermococcus nautili</i> 30-1 (85°C)	Theta type Helicase protein (SFI) DNA primase-polymerase	[40]
pP12-1	pTN2	pTN2-like	12.2	<i>Pyrococcus</i> sp.12/1 (95°C)	Theta type Helicase protein (SFI) DNA primase-polymerase	[40]
pCIR10	pTN2	pTN2-like	13.3	<i>Thermococcus</i> sp.CIR10 (85°C)	Theta type Helicase protein (SFI) DNA primase-polymerase	[40]
pIRI48	pTN2	pTN2-like	12.9	<i>Thermococcus</i> sp.IRI48 (85°C)	Theta type Helicase protein (SFI) Primase-polymerase domain	[60]
pEXT9a	pTN2	pEXT9a-like	10.5	<i>Thermococcus</i> sp.EXT9 (85°C)	Theta type Helicase protein (SFI)	[60]
pIRI33	pTN2	pEXT9a-like	11	<i>Thermococcus</i> sp.IRI33 (85°C)	Theta type Helicase protein (SFI)	[60]
pAMT7	pTN2	pEXT9a-like	8.5	<i>Thermococcus</i> sp.AMT7 (85°C)	Theta type Helicase protein (SFI)	[60]
pMETVU01	pTN2	pEXT9a-like	10.7	<i>Methanocaldococcus vulcanius</i> M7 (80°C)	Theta type Helicase protein (SFI) Proteins homologous to proteins of pEXT9a-like plasmids	NC_013408.1 [60]
pTN3	pTN3		18.3	<i>Thermococcus nautili</i> 30-1 (85°C)	Theta type DNA replicative helicase (MCM family) Integrase SSV-type	[74]
pGS5			2.8	<i>Archaeoglobus profundus</i> AV18 (80°C)	Rolling-circle type RC-Rep	[68]
pMETVU02			4.7	<i>Methanocaldococcus vulcanius</i> M7 (80°C)	Theta type Putative MCM family protein	NC_013409.1
pFS01			12.2	<i>Methanocaldococcus</i> sp.FS406-22 (90°C)	Theta type Putative MCM family protein DNA polymerase domain	NC_013888.1
ECE1			58.4	<i>Methanocaldococcus jannaschii</i> (85°C)	Theta type Putative MCM protein and Pa	[7]
ECE2			16.5	<i>Methanocaldococcus jannaschii</i> (85°C)	Theta type Restriction/modification system	[7]
pMEFER01			22.2	<i>Methanocaldococcus fervens</i> AG86 (80°C)	Theta type Putative MCM family protein	NC_013157.1
pMETOK01			14.9	<i>Methanothermococcus okinawensis</i> IH1 ⁺ (60e65°C)	Theta type Helicase and nuclease domain	NC_015632

propagation of pMP1. The plasmid, which shares its host with the pTN1 family comprises 3 small plasmids of almost the temperate siphovirus MPV1 (mentioned in section 3.1.2); pTN1 isolated from *Thermococcus nautili* [61], pGT5 induced by mitomycin along with the provirus and re-isolated from *P. abyssi* GE5 [62] and pTP1 recently isolated responds by replicating and escaping from the bacterial hosts by *T. prieurii* [44]. They encode a new family of rolling-circle replication initiator proteins (Rep74, Rep75 and RepTP1) which contain the three characteristic motifs of RC-Rep superfamily II, with a single tyrosine in motif 3 [6] as well as a fourth motif conserved with RC transposases, pGT5 and pTN1

3.2. Archaeal domain

The vast majority of archaeal plasmids isolated from deep-sea hydrothermal vents are associated with hyperthermophilic strains with optimal growth temperatures between 80 and 95 °C, belonging to *Thermococcales* and *Methanococcales* orders; these are given in Table 3. The pTP2 family could be represented by the 2.0 kb pTP2 which propagates in *T. prieurii* cells with pTP1 element) [1,36,58e60]. However, only 16 of them have been mentioned above. pTP2 is the smallest known plasmid of the sequenced and described. They have recently been classified by hyperthermophilic archaea and is unrelated to other *Thermococcales* plasmids [44]. It encodes an RC-Rep containing ty

Plasmids are common in *Thermococcales* strains (approx. 40% were found to harbor at least one extrachromosomal plasmid pTP2 which propagates in *T. prieurii* cells with pTP1 element) [1,36,58e60]. However, only 16 of them have been mentioned above. pTP2 is the smallest known plasmid of the sequenced and described. They have recently been classified by hyperthermophilic archaea and is unrelated to other *Thermococcales* plasmids [44]. It encodes an RC-Rep containing ty

tyrosines in motif 3 (superfamily which shares significant plasmids also encode several proteins that have homologues sequence similarity with RC-Rep. plasmid pGS5 isolated with the virus PAV1 [23,24]. from *Archaeoglobus profundus* AV18 [67]. Furthermore, it has the pTN3 family includes pTN3 (18.3 kb), isolated from *T. nautili* sp. 30-1 [74]. pTN3 encodes an SSV-type integrase and has been shown that pGS5 is negatively supercoiled due to the presence of a gyrase introducing negative supercoiling [68], closely related to the virus-like integrated element TKV4. In contrast with other plasmids isolated from hyperthermophilic *Kodakaraensis* [74], pTN3 and TKV4 actually represent relaxed or slightly positively supercoiled [69].

The pRT1 family contains pRT1, a 1.4 kb plasmid common to viruses belonging to PRD1-adenovirus viral lineage, isolated from *Pyrococcus* sp. JT1 [70] and a much larger the capsid protein which contains the double-jelly roll fold [71]. plasmid of 20.5 kb, pAMT11, isolated from *Thermococcus* sp. AMT11 [71]. They encode the homologous replication proteins Rep63 and Rep72, respectively, unrelated to RC-Rep proteins and may replicate via a theta mechanism [72]. Interestingly, pAMT11 genome revealed homology and synteny with the genome of TKV1 a virus-like integrated element of *T. kodakaraensis* [38].

The pT26-2 family constitutes a new large family of archaeal plasmids and integrated elements that probably date the separation of *Thermococcales* and *Methanococcales* [40]. The plasmid pT26-2 from *Thermococcus* sp. 26-2, encodes several proteins including an SSV-type integrase, a new DNA replicative helicase related to the superfamily AAA β ATPases probably involved in theta-type replication. Interestingly, the plasmid pMETVU01 also isolated from *M. thermophilus* pT26-2 shares homologues with virus-like integrated elements of *Thermococcales* (TKV2 and TKV3 present in *Kodakaraensis* genome, TGV1 present in *T. gammatolerans* genome and PHV1 present in *P. horikoshii* genome) and of *Methanococcales* (MMPV1, MMC7V1, MMC7V2, MMC6V1 in *Methanococcus maripaludis* strains, MVV1 in *Methanococcus voltae* genome), and with plasmid pMEFER1 isolated from *Methanocaldococcus fervens*. The pT26-2 family is identified by the presence of nine core genes [40,60] and phylogenetic analyses have shown that these “pT26-2 core genes” have co-evolved with cellular hosts [40].

The pTBMP1 family includes only pTBMP1, which is the largest plasmid of *Thermococcales* known to date. This 54.2 kb element sequenced in the framework of the *Thermococcus barophilus* genome project [3], is not related to other *Thermococcales* plasmids. Interestingly, pTBMP1 encodes a homologue of archaeal replication initiation protein Cdc6/Orc1 and thus replicates via a theta mechanism [60].

The pTN2 family comprising 7 plasmids from *Thermococcales* and one from *Methanococcales*, with sizes ranging from 8.5 kb to 13.3 kb, is divided into two subfamilies based on the gene content [60]. The pTN2-like subfamily comprises pTN2 isolated from *T. nautili*, pP12-1 from *Pyrococcus* sp., pCIR10 and pIRI48 both from *Thermococcus* sp. [40,60], whereas the pEXT9a-like subfamily is represented by pAMT7 and pIRI33, all isolated from *Thermococcus* sp. [60]. Bacteria with double-membranes, membrane vesicles (OMVs) are formed by budding of the outer membrane and contain not only periplasmic and outer membrane proteins, but also proteins characteristic of the inner primase superfamily PriS and PriL, strongly suggesting that membrane and cytoplasmic compartments in Archaea, they replicate via a theta mechanism [40]. pTN2-like (except for precisely in *Thermococcales*, MVs are produced by pIRI48 which encodes only the PriS domain) and pEXT9a-like proteins have the PriL domain in common, whereas the PriS domain is missing in pEXT9a-like proteins [60]. The pTN2-like S-layer [85].

4. Membrane vesicles (MVs) in deep-sea hydrothermal

4.1. General features of membrane vesicles in the three domains of life

The production of MVs is a universal physiological phenomenon shared by cells from the three domains of life [74,78]. MVs were recently reported in deep-sea hydrothermal vents by studying *Thermococcales* strains [74,79] (see below section 4.2). In the open ocean, MVs may reach up to 6.10 to 3.10¹⁰ MVs per mL of seawater, and notably play a key role in carbon cycling [80].

MVs are spherical closed compartments between ~50 and 200 nm in diameter with differing production depending on the domain of life [81,82] (Fig. 1). In *Eukarya*, different types of MVs have been described, some of which originated from the extrusion of intracellular multivesicular bodies and ectosomes, also called microparticles, directly or indirectly by budding from the cytoplasmic membrane [81,82]. Bacteria with double-membranes, membrane vesicles (OMVs) are formed by budding of the outer membrane and contain not only periplasmic and outer membrane proteins, but also proteins characteristic of the inner primase superfamily PriS and PriL, strongly suggesting that membrane and cytoplasmic compartments in Archaea, they replicate via a theta mechanism [40]. pTN2-like (except for precisely in *Thermococcales*, MVs are produced by pIRI48 which encodes only the PriS domain) and pEXT9a-like proteins have the PriL domain in common, whereas the PriS domain is missing in pEXT9a-like proteins [60]. The pTN2-like S-layer [85].

4.2. Role of MVs

MVs are involved in many exchanges between cells in the three domains of life, e.g., transporting toxin molecules, quorum-sensing agents, pathogenicity factors and metabolites [74]. Protection of MVs against viral infection has also been reported [86]. Lastly, MVs can carry different types of acids (Fig2).

Notably, some archaea produce proteolytic toxins against competing species [87,88]. Archaeal MVs were first described in the terrestrial and thermophilic order *Sulfolobales* as a conveyor of proteinaceous toxins [88e90]. Similarly, *Thermococcus* MVs were recently found to display toxicity towards competing *Thermococcales* strains suggesting the existence of “Thermococcins” (G. A. et al., unpublished work).

A recent study reports that marine bacterial extracellular vesicles are enriched in outer membrane components, suggesting that they harbor viral receptors that can act to bind viraparticles [84]. As reported for mesophilic bacteria [86], protection of MVs against viral infection could also be possible in hot deep-sea environments. *Thermococcales* virions were sometimes observed attached to MVs, which may result in less exposure to viral infection [23,79,85] and increase survival of the host population.

Interestingly, MVs have been universally reported to be able to carry and transfer different type of nucleic acids by fusing with recipient cells. Transfers between cells of chromosomal plasmid or viral DNAs packaged into MVs have indeed been reported in bacteria [91e94]. All these results lead us to consider that MVs could be involved in HGT and thus represent another type of MGE. The ability of archaeal MVs to carry and potentially transfer different types of DNA was also recently reported in *Thermococcales* strains isolated from deep-sea hydrothermal vents.

4.3. A focus on MVs isolated from *Thermococcales* and implicated in HGT

MV production actually appears to be widespread in *Thermococcales* isolated from deep-sea hydrothermal vents; 26 out of 34 *Thermococcales* strains, observed by transmission electron microscopy, revealed the production of MVs [79]. Some of these MVs were strongly associated with cellular DNA and consequently appeared as fluorescent by epifluorescence microscopy. Experiments showed that these MVs can protect exogenous DNA against DNase treatments and denaturation, suggesting intravesicular localization of the DNA [79]. Transformation with naked DNA might actually not be an option in deep-sea hydrothermal vents due to denaturation; it was therefore hypothesized that MVs, by protecting DNA against thermodegradation, could play an important role in HGT in these deep-sea environments [79]. The production of nanospheres and nanotubes was recently observed at the surface of MVs from *T. kodakaraensis* and *T. gammatolerans*. It has been speculated that this mechanism might be used to facilitate DNA transfer

MVs to cells [95]. The idea that MVs produced by *Thermococcales* can be considered as MGEs was recently strengthened when plasmid incorporation and transfer through MVs was demonstrated using the genetically tractable species *T. kodakaraensis* [85]. MVs containing plasmids were also observed in *Thermococcus* strains [36,74]. MVs produced by *Nautili* sp. 30-1 harbored the pTN3 plasmid, which corresponds to a viral genome, carries genes from a defective virus, has ineffective virion production [74]. The name “viral membrane vesicles” (vMV) was proposed to specify these particular vesicles containing viral/plasmid genomes, considering that virus and plasmid spaces often overlap [96]. These results reinforce the hypothesis of an evolutionary connection between MVs and viruses. This presumption was previously made by considering that the viruses produce virions that strikingly resemble MVs. For example, some archaeoviruses and bacteriophages produce pleomorphic virions consisting of lipidic vesicles into which viral proteins and corresponding nucleic acids are embedded [97]. It has been postulated that MVs produced by ancient cells or proto cells, anterior to LUCA, could have played a key role in the onset of some of the first viral lineages [98].

Despite copious evidence that MVs are produced by (hyper-)thermophilic prokaryotes in deep-sea biospheres, their abundance and functions remain largely undetermined to date. The study of MVs thus emerges as a new fundamental field of research in cell and evolutionary biology. Further work is needed to better understand the global impact of MVs on deep-sea hydrothermal vents.

5. Conclusion

MGEs may have been key players in the evolution of the biosphere since the very beginning. The microbial genome remains a driving force in the diversity and evolution of prokaryotes and eukaryotes. Although the real impact of MGEs is not clearly known in deep-sea hydrothermal vents, as reviewed in Fig. 1, they have the potential of becoming powerful agents that drive adaptation of their cellular hosts. For example, temperate viruses and their lysogenic hosts seek to coexist so as to ensure their intertwined fitness when confronted with the extreme conditions of deep-sea vents. Evidence of this via comparative analysis of cellular and metagenomes [26] lies in the relative enrichment of genes related to energy metabolism in the proviral fraction, by providing their hosts with new or supplementary features to enable adaptation to a challenging environment. These proviruses boost host fitness and in turn enhance their own survival.

Comparative studies of MGEs from Archaea and Bacteria will help us understand the dynamic genetic network of the microbiota communities in the deep biosphere. The study of how microbial populations are modulated by these genetic interactions will become a keystone area.

Although the ecological and evolutionary effects of MGEs are not clearly known in deep-sea vents, their production probably plays a major role in the lifestyle of extremophiles.

microorganisms making analysis of these processes a prerequisite for fully understanding their physiology and rationally exploiting their biotechnological potential (transfer of nucleic acids etc.).

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

This work was financially supported by the French Ministry of Higher Education to JL, J. Fremont SD and the Agence Nationale de la Recherche, project Thermovesicles (ANR-12-BSV3-0023-01) to CG and PF.

References

- [1] Prieur D, Erauso G, Geslin C, Lucas S, Gaillard M, Bidault A, et al. Genetic elements of *Thermococcales*. *Biochem Soc Trans* 2004;32:184e7.
- [2] Anderson RE, Brazelton WJ, Baross JA. Is the genetic landscape of the deep subsurface biosphere affected by viruses? *Front Microbiol* 2011;2:219.
- [3] Flores GE, Reysenbach A-Hydrothermal environment. In: Reiter J, Thiel V, editors *Encyclopedia of earth science series: encyclopedia of geobiology*. Netherlands: Springer.; 2015. p. 67.
- [4] Forterre P. The virocell concept and environmental microbiology. *ISME J* 2013;7:233e6.
- [5] Koonin EV, Dolja VV. Virus world as an evolutionary network of viruses and capsidless selfish elements. *Microbiol Mol Biol Rev* 2014;78:278e303.
- [6] Suttle CA Viruses in the sea. *Nature* 2005;437:356e61.
- [7] Suttle CA Marine viruses - major players in the global ecosystem. *Rev Microbiol* 2007;5:801e12.
- [8] Rohwer F, Thurber RV. Viruses manipulate the marine environment. *Nature* 2009;459:207e12.
- [9] Kristensen DM, Mushegian AR, Dolja VV, Koonin EV. New dimensions of the virus world discovered through metagenomics. *Microbiol* 2010;18:11e9.
- [10] McDaniel LD, Young E, Delaney J, Ruhnau F, Ritchie KB, Paul JH. High frequency of horizontal gene transfer in the oceans. *Science* 2010;330:50e50.
- [11] Cortez D, Forterre P, Giraldo SA hidden reservoir of integrative elements is the major source of recently acquired foreign genes and ORFs in archaeal and bacterial genomes. *Genome Biol* 2009;10:R65.
- [12] Forterre P, Prangishvili D. The major role of viruses in cellular evolution: facts and hypotheses. *Surv Opin Virol* 2013;3:558e65.
- [13] Ortmann A, Suttle CA High abundances of viruses in a deep-sea hydrothermal vent system indicate viral mediated microbial mortality. *Deep Sea Res Part* 2005;52:1515e27.
- [14] Williamson SJ, Cary SC, Williamson KE, Helton RR, Bench SR, Winget D, et al. Lysogenic virus-host interactions predominate at deep-sea diffuse-flow hydrothermal vents. *ISME J* 2008;2:1112e21.
- [15] Yoshida-Takashima M, Nunoura T, Kazama H, Noguchi T, Inoue K, Akashi H, et al. Spatial distribution of viruses associated with planktonic and attached microbial communities in hydrothermal environments. *Environ Microbiol* 2012;78:1311e20.
- [16] Wommack KE, Williamson SJ, Sundbergh A, Helton RR, Glazer BT, Portune K, et al. An instrument for collecting discrete large-volume water samples suitable for ecological studies of microorganisms. *Deep Sea Res Part* 2004;51:1781e92.
- [17] Juniper SK, Bird DF, Summit M, Vong MP, Baker ET. Bacterial and viral abundances in hydrothermal plumes over northern Gorda Ridge. *Deep Sea Res Part* 1998;45:2739e49.
- [18] Horvath P, Barrangou R. CRISPR/Cas: the immune system of bacteria and archaea. *Science* 2010;327:167e70.
- [19] Richter HR, Andau L, Plagens A. Exploiting CRISPR/Cas: interference mechanisms and applications. *Mol Sci* 2013;14:14518e31.
- [20] Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* 2010;11:181e9.
- [21] Makarova KS, Wolf YI, Koonin EV. Potential genomic determinants of hyperthermotolerance. *Trends Genet* 2003;19:172e6.
- [22] Anderson RE, Brazelton WJ, Baross JA. Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *ISME J* 2011;77:120e33.
- [23] Geslin C, Le Romancer M, Erauso G, Gaillard M, Perrot G, Prieur D. PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, *Pyrococcus abyssi*. *J Bacteriol* 2003;185:3888e94.
- [24] Geslin C, Gaillard M, Flament D, Rouault K, Le Romancer M, Prieur D, et al. Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from *Pyrococcus abyssi*. *J Bacteriol* 2007;189:4510e9.
- [25] Llanas A, Koonin EV, Biennu N, Prieur D, Geslin C. TPV1, the first virus isolated from the hyperthermophilic genus *Thermococcus*. *Environ Microbiol* 2012;14:503e16.
- [26] Anderson RE, Sogin ML, Baross JA. Evolutionary strategies of viruses, bacteria and archaea in hydrothermal vent ecosystems revealed through metagenomic analysis. *Bios One* 2014;9:e109696. <http://dx.doi.org/10.1371/journal.pone.0109696>.
- [27] Uldahl K, Peng X. Biology, biodiversity and application of thermophilic viruses. In: Satyanarayana L, Littlechild J, Kawarabayashi, editors. *Thermophilic microbes in environment and industrial biotechnology: biotechnology of thermophiles*. Netherlands: Springer; 2013. p. 271e304.
- [28] Ackermann HW, Prangishvili D. Prokaryote viruses studied by electron microscopy. *Arch Virol* 2012;157(10):1843e9.
- [29] Liu B, Wu S, Song Q, Zhang X, Xie L. Two novel bacteriophages of thermophilic bacteria isolated from deep-sea hydrothermal vent. *Microbiol* 2006;53:163e6.
- [30] Liu B, Zhang X. Deep-sea thermophilic *Geobacillus* bacteriophage GVE2: transcription profile and proteomic characterization of virions. *Appl Microbiol Biotechnol* 2008;80:697e707.
- [31] Wang Y, Zhang X. Genome analysis of deep-sea thermophilic phage D6E. *Appl Environ Microbiol* 2010;76:678e61.
- [32] Yoshida-Takashima T, Takai Y, Shimamura S, Nunoura T, Takai K. Genome sequence of a novel deep-sea vent Epsilonproteobacterial phage provides new insight into the co-evolution of Epsilonproteobacteria and their phages. *Extremophiles* 2013;17:405e19.
- [33] Nakagawa S, Takai K, Inagaki F, Hirayama H, Nunoura T, Horikoshi K, et al. Distribution, phylogenetic diversity and physiological characteristics of epsilon-proteobacteria in a deep-sea hydrothermal field. *Microbiol* 2005;7:1619e32.
- [34] Lossouarn J, Nesbø CL, Mercier C, Zhaxybayeva O, Johnson MS, Charchuck P, et al. «Menage a trois» a selfish genetic element uses a virus to propagate within *Thermotoga*. *Environ Microbiol* 2015.
- [35] Alain K, Marteinsson VT, Miroshnichenko ML, Bonch-Osmolovskaya EA, Prieur D, Birrien J-L. *Marinitoga piezophila* sp. nov., a rod-shaped thermo-piezophilic bacterium isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* 2002;52:1331e9.
- [36] Soler N, Gaudin M, Marguet E, Forterre P. Bacteriophages, viruses and virus-like membranes from *Thermococcales*. *Biochem Soc Trans* 2011;39:36e44.
- [37] Leijn JA, Albers S-V, Atomi H, Allers T. Model organisms for genetics in the domain Archaea: methanohalophiles, *Thermococcales* and *Sulfolobales*. *FEMS Microbiol Rev* 2011;35:577e608.
- [38] Fukui T, Atomi H, Kanai T, Matsumi R, Fujiwara S, Tanaka T. Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res* 2005;15:352e63.
- [39] Zivanovic Y, Armengaud J, Lagorce A, Leplat C, Dutertre M, et al. Genome analysis and genome-wide proteomics of *Thermococcus gammatolerans*, the most radioresistant organism known among the Archaea. *Genome Biol* 2009;10:R70.

- [40] Soler N, Marguet E, Cortez D, Desnoves N, Keller J, van Tilbeurgh H, et al. Two new families of plasmids from hyperthermophilic archaea encoding new families of replication proteins. *Nucleic Acids Res* 2010;38:5088–104.
- [41] Geslin C, Le Romancer M, Gaillard M, Erauso G, Prieur D. Observation of virus-like particles in high temperature enrichment cultures from deep-sea hydrothermal vents. *Res Microbiol* 2003;154:303–7.
- [42] Gorlas A, Geslin C. A simple procedure to determine the infectivity and host range of viruses infecting anaerobic and hyperthermophilic microorganisms. *Extremophiles* 2013;17:349–55.
- [43] Gorlas A, Alain K, Bienvenu N, Geslin C. *Thermococcus priuspi* sp. nov., a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *J Syst Evol Microbiol* 2013;63:2920–6.
- [44] Gorlas A, Krupovic M, Forterre P, Geslin C. Living side by side with a virus: characterization of two novel plasmids from *Thermococcus priuspi*, a host for the spindle-shaped virus TPV1. *Appl Environ Microbiol* 2013;79:3822–8.
- [45] Frost LS, Leplae R, Summers AO, Toussain A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 2005;3:722–32.
- [46] Lipps G. In: Lipps G, editor. *Plasmids: current research and future trends*. Norfolk: Caister Academic press; 2008.
- [47] Sobecky PA, Hazen TH. Horizontal gene transfer and mobile genetic elements in marine systems. *Methods Mol Biol* 2009;532:435–53.
- [48] Ji B, Gimenez G, Barbe V, Vacherie B, Rouy Z, Amrani A, et al. Complete genome sequence of piezophilic mesophilic sulfate-reducing bacterium *Desulfovibrio hydrothermalis* AM13^T. *Genome Announc* 2013;1(4):00226–12.
- [49] Reysenbach AN, Hamamura N, Podar M, Griffiths E, Ferreira S, Hochstein S, et al. Complete and draft genome sequences of six members of the *Aquificales*. *Bacteriol* 2009;191(6):1992–3.
- [50] Pati A, Zhang X, Lapidus A, Nolan M, Lucas S, Del Rio TG, et al. Complete genome sequence of *Ignifitimus profundus* type strain (506). *Stand Genomic Sci* 2011;4(2):210–20.
- [51] Giovannelli D, Ricci J, Pérez-Rodríguez Hügler M, O'Brien C, Keddis R, et al. Complete genome sequence of *Thermovibrio ammonifigans* HB-1^T, a thermophilic chemolithoautotrophic bacterium isolated from a deep-sea hydrothermal vent. *Stand Genomic Sci* 2012;7(1):82–90.
- [52] Takaki Y, Shimamura S, Nakagawa S, Fukuhara Y, Horikawa H, Ankai A, et al. Bacterial lifestyle in a deep-sea hydrothermal vent chimney revealed by the genome sequence of the thermophilic bacterium *Deferribacter desulfuricans* SSM1. *DNA Res* 2010;17:123–37.
- [53] Lucas S, Han J, Lapidus A, Cheng J-F, Goodwin LA, Pitluck S, et al. Complete genome sequence of the thermophilic, heterotrophic bacterium *Marinitoga piezophila* KA3. *J Bacteriol* 2012;194:5974–5.
- [54] Takai K, Kobayashi H, Nealson KH, Horikoshi K. *Deferribacter desulfuricans* sp. nov., a novel sulfur-, nitrate- and arsenate-reducing thermophile isolated from a deep-sea hydrothermal vent. *J Syst Evol Microbiol* 2003;53:839–46.
- [55] Akimkina T, Ivanov P, Kostrov S, Sokolova T, Bonch-Osmolovskaya E, Le Cam E, et al. A highly conserved plasmid from the extreme thermophile *Thermotoga maritima* MC24 is a member of a family of plasmids distributed worldwide. *Plasmid* 1999;42:236–40.
- [56] Harriott O, Huber R, Stetter KO, Betts P, Woll KM. A cryptic miniplasmid from the hyperthermophilic bacterium *Thermotoga* sp. RQ7. *J Bacteriol* 1994;176:2759–62.
- [57] Zhaxybayeva S, Withers KS, Foght J, Green AG, Bruce D, Detter C, et al. Genome sequence of the mesophilic *Thermotoga* bacterium *Mesotoga prima* MesG1.Ag.4.2 reveals the largest *Thermotogales* genome to date. *Genome Biol Evol* 2012;4:700–8.
- [58] Benbouzid-Rollat, López-García P, Watrin L, Erauso G, Prieur D, Forterre P. Isolation of new plasmids from hyperthermophilic Archaea. *Res Microbiol* 1997;148:767–75.
- [59] Lepage E, Marguet E, Geslin C, Matte-Tailliez C, Jilg V, Forterre P, et al. Molecular diversity of new *Thermococcales* isolates from a single extracellular DNA produced by hyperthermophilic archaea of the area of hydrothermal deep-sea vents as revealed by randomly amplified
- polymorphic DNA fingerprinting and 16S rRNA gene sequence analysis. *Appl Environ Microbiol* 2004;70:1277–86.
- [60] Krupovic M, Gonnet M, Hania WB, Forterre P, Erauso G. Insights into dynamics of mobile genetic elements in hyperthermophilic environments from five new *Thermococcus* plasmids. *PLoS One* 2013;8:e49044.
- [61] Soler N, Justome A, Quevillon-Cheruel S, Lorieux F, Le Cam E, Margue E, et al. The rolling-circle plasmid pTN1 from the hyperthermophilic archaeon *Thermococcus nautilus*. *Mol Microbiol* 2007;66:357–70.
- [62] Erauso G, Marsin S, Benbouzid-Rollat, Bauche MF, Barbeyron T, Zivanovic Y, et al. Sequence of plasmid pGT5 from the archaeon *Thermococcus abyssi*: evidence for rolling-circle replication in a hyperthermophilic bacterium. *Bacteriol* 1996;178:3232–7.
- [63] Ilyina TV, Koonin EV. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeobacteria. *Nucleic Acids Res* 1992;20:3279–85.
- [64] Lucas S, Toffin L, Zivanovic Y, Charlier D, Moussard H, Forterre P, et al. Construction of a shuttle vector for, and spheroplast transformation of hyperthermophilic archaeon *Pyrococcus abyssi*. *Appl Environ Microbiol* 2002;68:5528–36.
- [65] Waage I, Schmid G, Thumann S, Thomm M, Hausner W. Shuttle vector-based transformation system for *Pyrococcus furiosus*. *Appl Environ Microbiol* 2010;76:3308–13.
- [66] Santangelo TJ, Cubonov L, Reeve JN. Shuttle vector expression in *Thermococcus kodakaraensis*: contributions of cis elements to protein synthesis in a hyperthermophilic archaeon. *Appl Environ Microbiol* 2008;74:3099–104.
- [67] Burggraf J, Jannasch HW, Nicolaus B, Stetter KO. *Archaeoglobus profundus* sp. nov. Represents a New Species within the Sulfate-reducing Archaeobacteria. *Syst Appl Microbiol* 1990;13:24–8.
- [68] López-García P, Forterre P, van der Oost J, Erauso G. Plasmid pGS5 of the hyperthermophilic archaeon *Archaeoglobus profundus* is negatively supercoiled. *Bacteriol* 2000;182:4998–5000.
- [69] López-García P, Forterre P. DNA topology in hyperthermophilic archaea: reference states and their variation with growth phase, temperature and pressure. *DNA Res* 1997;23:1267–79.
- [70] Ward DE, Revett M, Nandakumar R, Little JH, de Vos WM, van der Oost J, et al. Characterization of plasmid pRT1 from *Pyrococcus* sp. strain JT1. *Bacteriol* 2002;184:2561–6.
- [71] Gonnet M, Erauso G, Prieur D, Le Romance M. pAMT11, a novel plasmid isolated from a *Thermococcus* strain closely related to the virus-like integrated element TKV1 of the *Thermococcus kodakaraensis* genome. *Res Microbiol* 2011;162:132–43.
- [72] Forterre P, Krupovic M, Raymann K, Soler N. Plasmids from *Euryarchaeota*. *Microbiol Spectr* 2014. 2:PLAS-0027–2014.
- [73] Vannier P, Marteinsson VT, Fridjonsson OH, Oger P, Jebbar M. Complete genome sequence of the hyperthermophilic, piezophilic, heterotrophic, and Carboxydophilic archaeon *Thermococcus barophilus* JMAc-teriol 2011;193:1481–2.
- [74] Gaudin M, Krupovic M, Marguet E, Gauliard E, Virkaite-Krupovic V, Le Cam E, et al. Extracellular membrane vesicles harbouring viral
- [75] Krupovic M, Bamford DH. Archaeal proviruses TKV4 and MVV extend the PRD1-adenovirus lineage to the phylum Euryarchaeota. *Virology* 2008;375:292–300.
- [76] Oberto J, Gaudin M, Cossu M, Gorlas A, Slesarev A, Marguet E, et al. Genome sequence of a hyperthermophilic archaeon, *thermococcus* 30e1, that produces viral vesicles. *Genome Announc* 2014. 2: e002
- [77] Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, et al. Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii*. *Science* 1996;273:1058–73.
- [78] György B, Szabolc TG, Pászti M, Pál Z, Misják P, Aradi B, et al. Membrane vesicles: current state-of-the-art and emerging role of extracellular vesicles. *Cell Mol Life Sci* 2011;68:2667–88.
- [79] Soler N, Margue E, Verbavatz J-M, Forterre P. Virus-like vesicles and extracellular DNA produced by hyperthermophilic archaea of the order *Thermococcales*. *Res Microbiol* 2008;159:390–9.

- [80] Scanlan DJ. Bacterial vesicles in the ocean. *Science* 2014;343:143e4.
- [81] Deatherage BL, Cookson BT. Membrane vesicle release in bacteria, eukaryotes and archaea: a conserved yet underappreciated aspect of microbial life. *Infect Immun* 2012;80:1948e57.
- [82] Raposo G, Stoorvogel W. Extracellular vesicles: exosomes, microvesicles and friends. *J Cell Biol* 2013;200:373e83.
- [83] Kulp A, Kuehn MJ. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annu Rev Microbiol* 2010;64:163e84.
- [84] Biller SJ, Schubotz R, Roggensack SE, Thompson AW, Summons RE, Chisholm SW. Bacterial vesicles in marine ecosystems. *Science* 2014;343:183e6.
- [85] Gaudin M, Gauliard E, Schouten S, Houel-Renault L, Lenormand P, Marguet E, et al. Hyperthermophilic archaea produce membrane vesicles that can transfer DNA. *Environ Microbiol Rep* 2013;5:109e16.
- [86] Manning AJ, Kuehn MJ. Contribution of bacterial outer membrane vesicles to innate bacterial defense. *PLoS Microbiol* 2011;11:258.
- [87] Haseltine CH, Hill T, Montalvo-Rodriguez R, Kemper SK, Shand RF, Blum P. Secreted Euryarchaeal microhalocins kill hyperthermophilic *Crenarchaea*. *J Bacteriol* 2001;183:287e91.
- [88] Prangishvili D, Holz I, Stieger E, Nickell S, Kristjansson JK, Zillig W. Sulfolobins, specific proteinaceous toxins produced by strains of the extremely thermophilic archaeal genus *Sulfolobus*. *J Bacteriol* 2000;182:2985e8.
- [89] Ellen AF, Albers S-V, Huibers W, Pitcher A, Hobel CFV, Schwarz H, et al. Proteomic analysis of secreted membrane vesicles from archaeal *Sulfolobus* species reveals the presence of a vesicle sorting complex component. *Extremophiles* 2009;13:67e79.
- [90] Ellen AF, Rohulya OV, Fusetti F, Wagner M, Albers S-V, Driessen AJM. The Sulfolobin genes of *Sulfolobus acidocaldarius* encode novel antimicrobial proteins. *J Bacteriol* 2011;193:4380e7.
- [91] Dorward DW, Garon CF, Judd RC. Export and intercellular transfer of DNA via membrane blebs of *Neisseria gonorrhoeae*. *J Bacteriol* 1989;171:2499e505.
- [92] Velimirov B, Hagemann S. Mobilizable bacterial DNA packaged into membrane vesicles induces serial transduction. *Mob Genet Elem* 2011;1:80e1.
- [93] Kolling GL, Matthews KR. Export of virulence genes and Shiga toxin by membrane vesicles of *Escherichia coli* O157:H7. *Appl Environ Microbiol* 1999;65:1843e8.
- [94] Yaron S, Kolling GL, Simon L, Matthews KR. Vesicle-mediated transfer of virulence genes from *Escherichia coli* O157:H7 to other enteric bacteria. *Appl Environ Microbiol* 2000;66:4414e20.
- [95] Margue E, Gaudin M, Gauliard E, Fourquaux I, Blond du Plouy S, Matsui I, et al. Membrane vesicles, nanopods and/or nanotubes produced by hyperthermophilic archaea of the genus *Thermococcus*. *Biochem Soc Trans* 2013;41:436e42.
- [96] Soler N, Krupovic M, Margue E, Forterre P. Membrane vesicles in natural environments: a major challenge in viral ecology. *J SMI* 2014;9:793e6.
- [97] Pietilä MK, Atanasova NS, Manole V, Liljeroos L, Butcher SJ, Oksanen HM, et al. Virion architecture unifies globally distributed pleolipoviruses infecting halophilic archaea. *Virob* 2012;86:5067e79.
- [98] Forterre P, Krupovic M. The origin of virions and virocells: the escape hypothesis revisited. Witzany G, editor. *Viruses essential agents life*. Netherlands: Springer; 2012. p. 123e60.

Chapitre II : Ménage à trois, caractérisation de MPV1

Présentation des travaux

Ces travaux ont consisté en la caractérisation du premier virus associé au phylum des Thermotogae, MPV1 (*Marinitoga piezophila* virus) et elle d'un élément génétique « pirate », le plasmide pMP1 (plasmide *Marinitoga piezophila* 1).

MPV1 est un siphovirus typique du titre à la température C. Les virus sont composés d'une capsid icosaédrique de 55 nm de diamètre à laquelle est attachée une longue queue flexible de 200 nm de long pour 10 nm de large. Ce virus est intégré sous forme de provirus dans le génome de *Marinitoga piezophila* KA3. Cette Thermotogae thermophile et anaérobie est piézophile mais peut être adaptée à la pression atmosphérique.

Ce « Ménage à trois » repose sur l'interaction d'une cellule hôte, d'un virus et d'un plasmide pMP1. Ce plasmide a la particularité d'être capable de se substituer aux capsides virales à la place du génome viral. En effet, lors du déclenchement du cycle lytique par induction à la mitomycine C, le nombre de copies du plasmide augmente drastiquement (nombre de copies des génomes viraux multiplié par 3, nombre de copies du plasmide multiplié par 3). Le mécanisme d'excision du plasmide n'est pas connu et si celui-ci participe à la survie et à la vie du phage viral potentiellement impliqué dans l'excision de l'ADN. D'autre part, le plasmide est aussi capable de se propager dans une cellule d'un autre genre : *Thermosipho* sp. AT1244-VC14. Ce ménage à trois semble être un nouvel exemple de piratage moléculaire où le pirate (ici le plasmide) prend la place du génome viral dans la capsid afin d'assurer sa propre propagation. De plus, profitant ainsi de la protection assurée par cette coque protéique.

Ces travaux ont été menés dans le cadre de la thèse de Julien Lossouarn (soutenue en mars

2014). Lors de ce stage de thèse, le rôle de l'élément génétique de souche de

Thermotogae été réalisé menant notamment à la découverte et à la caractérisation du

age à t ois. J'ai appo ta o ti utio à ette tude du atmon stage de master 2 et le début de ma thèse. Mes travaux ont porté sur différentes étapes de cette caractérisation. Dans un premier temps, la mise au point de la méthode de purification des particules virales a été améliorée via l'utilisatio d'u g adie t aie d'iodiol. Pa la suite,tude du contenu des particules virales par l'e t atio des aides u l i ues AN p se ts das les capsides purifiées a été réalisée, permettant le séquençage de ceux ci. Le séquençage et les PCR sur ces ANp se ts à l't ieu de laapside aat v l la p se e de l'ADN a i ais aussi de l'ADN plasidique, un suivi en qPCR de l'ADN bactérien (gène codant pour ARNr16s), viral (gène codant pour la capside) et plasmidique (gène hypothétique), avec et sans induction, a t alis , afi d' tudie l'volutio deeu -ci. E fi , des tests de t io o t t alis s, permettant de mettre en évidence la propagation du plasmide dans une souche appartenant au genre Thermosipho.

'Ménage à trois': a selfish genetic element uses a virus to propagate within *Thermotogales*

Julien Lossouarn,^{1,2,3} Camilla L. Nesbø,^{4,5}
Coraline Mercier,^{1,2,3} Olga Zhaxybayeva,⁶
Milo S. Johnson,⁶ Rhianna Charchuck,⁴
Julien Farasin,^{1,2,3} Nadège Bienvenu,^{1,2,3}
Anne-Claire Baudoux,^{7,8,9} Grégoire Michoud,^{1,2,3}
Mohamed Jebbar^{1,2,3} and Claire Geslin^{1,2,3*}

¹Université de Bretagne Occidentale (UBO, UEB),
Institut Universitaire Européen de la Mer (IUEM) – UMR
6197, Laboratoire de Microbiologie des Environnements
Extrêmes (LMEE), rue Dumont d'Urville, F-29280
Plouzané, France.

²CNRS, IUEM – UMR 6197, Laboratoire de
Microbiologie des Environnements Extrêmes (LMEE),
rue Dumont d'Urville, F-29280 Plouzané, France.

³Ifremer, UMR 6197, Laboratoire de Microbiologie des
Environnements Extrêmes (LMEE), Technopôle Pointe
du diablea, F-29280 Plouzané, France.

⁴CEES, Department of Biology, University of Oslo, Oslo
0316, Norway.

⁵Department of Biological Sciences, University of
Alberta, Edmonton, AB T6G2R3, Canada.

⁶Department of Biological Sciences, Dartmouth College,
Hanover, NH 03755, USA.

⁷Sorbonne Universités, UPMC Univ Paris 06, Paris
75005, France.

⁸UMR 7144, Equipe DIPO, Station Biologique de
Roscoff, Roscoff 29680, France.

⁹CNRS, UMR 7144, Adaptation et Diversité en Milieu
Marin, Station Biologique de Roscoff, Roscoff 29680,
France.

Summary

Prokaryotic viruses play a major role in the microbial ecology and evolution. However, the virosphere associated with deep-sea hydrothermal ecosystems remains largely unexplored. Numerous instances of lateral gene transfer have contributed to the complex and incongruent evolutionary history of *Thermotogales*, an order well represented in deep-sea hydrothermal vents. The presence of clustered regularly interspaced short palindromic repeats (CRISPR)

loci has been reported in all *Thermotogales* genomes, suggesting that these bacteria have been exposed to viral infections that could have mediated gene exchange. In this study, we isolated and characterized the first virus infecting bacteria from the order *Thermotogales*, *Marinitoga piezophila* virus 1 (MPV1). The host, *Marinitoga piezophila* is a thermophilic, anaerobic and piezophilic bacterium isolated from a deep-sea hydrothermal chimney. MPV1 is a temperate *Siphoviridae*-like virus with a 43.7 kb genome. Surprisingly, we found that MPV1 virions carry not only the viral DNA but preferentially package a plasmid of 13.3 kb (pMP1) also carried by *M. piezophila*. This 'ménage à trois' highlights potential relevance of selfish genetic elements in facilitating lateral gene transfer in the deep-sea biosphere.

Introduction

Deep-sea hydrothermal vents are home to many thermophilic and hyperthermophilic *Bacteria* and *Archaea*, whose genetic diversity, metabolic capacity and physiological adaptations are being actively explored (reviewed in Flores and Reysenbach, 2011). However, the virosphere associated with these deep-sea hydrothermal microorganisms remains largely unstudied. The observation of abundant populations of lysogens and temperate viruses indicates that virus-mediated gene exchange and recombination may be critical to the survival and stability of prokaryotes within extreme environments (Williamson *et al.*, 2008). Additionally, metagenomic studies suggest that a wide diversity of deep-sea hydrothermal microorganisms is exposed to viral infections (Williamson *et al.*, 2008; Anderson *et al.*, 2011a). Yet, published viromes mostly contain reads with no homologues in public databases and do not provide any information about the infected hosts (Anderson *et al.*, 2011b). Therefore, to understand the impact of viruses on the microbial diversity in the deep-sea ecosystems, it remains crucial to isolate new viruses and to characterize host–virus model systems.

To date, only seven viruses, isolated from deep sea hydrothermal vents, have been described. Two of them are lemon-shaped archaeoviruses, PAV1 and TPV1, which infect *Thermococcales*, obligate anaerobic sulfur-metabolizing hyperthermophiles (Geslin *et al.*, 2003a; 2007; Gorlas *et al.*, 2012). The other five isolates are tailed

Received 29 August, 2014; revised 13 January, 2015; accepted 13 January, 2015. *For correspondence. E-mail claire.geslin@univ-brest.fr; Tel. +33 (0)2 98 49 88 58; Fax +33 (0)2 98 49 87 05.

bacteriophages. Among these, four lytic bacteriophages infect aerobic, thermophilic and heterotrophic strains belonging to *Bacillus* and *Geobacillus* genera with optimal growth temperatures around 65°C. The bacteriophage BVW 1 is unclassified whereas GVE1 and GVE2 are affiliated to the *Siphoviridae* family (Liu *et al.*, 2006; Liu and Zhang, 2008), and D6E belongs to the *Myoviridae* family (Wang and Zhang, 2010). The remaining bacteriophage, NrS-1, is a temperate siphovirus that infects the chemolitho-autotrophic, anaerobic and microaerobic, moderately thermophilic *Epsilonproteobacterium Nitratiruptor* sp. cultivated at 55°C (Yoshida-Takashima *et al.*, 2013).

Here we report and characterize the first bacteriophage that infects cells from the bacterial order *Thermotogales*, a lineage broadly represented in deep biosphere ecosystems and deep-sea hydrothermal vents in particular. This order consists of mesophilic (Nesbø *et al.*, 2012; Hania *et al.*, 2013) and hyper/thermophilic, anaerobic, chemo-organotrophic microorganisms (Bonch-Osmolovskaya, 2008). Phylogenetic analyses of 16S rRNA gene and 29 concatenated ribosomal proteins place *Thermotogales* as a sister group to *Aquificales*, and consequently as the second deepest branching order within domain *Bacteria* (Zhaxybayeva *et al.*, 2009). However, numerous instances of lateral gene transfer have contributed to the evolutionary history of these bacteria: many genes were exchanged with *Firmicutes* and, to a lesser extent, with *Archaea* and especially *Thermococcales* (Nelson *et al.*, 1999; Zhaxybayeva *et al.*, 2009). Similarly to many thermophilic prokaryotes, clustered regularly interspaced short palindromic repeats (CRISPRs) are present in all sequenced *Thermotogales* genomes (Zhaxybayeva *et al.*, 2009), suggesting that *Thermotogales* have been exposed to viral infections. Yet, prior to this study, no viruses infecting *Thermotogales* had been isolated. The only known extrachromosomal genetic elements in these lineages are two cryptic miniplasmids in the genus *Thermotoga* (Harriott *et al.*, 1994; Akimkina *et al.*, 1999) and a 1724 bp miniplasmid in *Mesotoga prima* (Zhaxybayeva *et al.*, 2012).

The newly discovered virus infects *Marinitoga piezophila* KA3, a member of the *Thermotogales* isolated from a deep-sea hydrothermal vent chimney located at a depth of 2630 m on the East-Pacific Rise (Alain *et al.*, 2002). *Marinitoga piezophila* virus 1 (MPV1) is a temperate *Siphoviridae*-like virus and is mitomycin C inducible. Its genome of 43.7 kb is identical to the provirus integrated in *M. piezophila* KA3 chromosome (Lucas *et al.*, 2012). Additionally, we discovered that MPV1 virions mainly contain a plasmid of 13.3 kb (referred as pMP1), which corresponds not to the viral genome but to a second mobile genetic element also carried by *M. piezophila*. We demonstrate that pMP1 uses the MPV1 viral capsid to propagate, highlighting a complex

evolutionary relationship between a bacterial host, an extrachromosomal element and a virus.

Results

Virion morphology

Particles of MPV1 have a head with a hexagonal outline of ~55 nm in diameter and a flexible non-contractile tail of ~200 nm in length and ~10 nm in width (Fig. 1). Based on these morphological characteristics, we classify MPV1 as a siphovirus from the order *Caudovirales*. Despite the piezophilic character of *M. piezophila*, which grows optimally at high hydrostatic pressure (40 MPa), this isolate can adapt to the growth at atmospheric pressure (0.1 MPa) after several subcultures. The virion morphology was identical under these two growth conditions. Because assays indicated that the viral production is functional and similar at both atmospheric and high hydrostatic pressures of 40 MPa (Table S1), all subsequent experiments were carried out at atmospheric pressure.

Virus–host relationship

MPV1 is a mitomycin C inducible temperate bacteriophage causing host cells lysis upon induction (Fig. 2). A concentration of 5 µg ml⁻¹ of mitomycin C was sufficient for optimal virus induction, as estimated by viral counts (Fig. S1). The viral production increased, reaching ~3 × 10⁹ viral particles ml⁻¹, 2 h after the induction by 5 µg ml⁻¹ of mitomycin C (Fig. 2). A concomitant drastic drop to 2.19 × 10⁶ cells ml⁻¹ in host cell counts was observed, suggesting the lysis of the cells (Fig. 2). Regrowth of bacteria was observed 5 h after induction, suggesting that some cells withstood the mitomycin C effects. Untreated cells produced 30-fold less viral



Fig. 1. Electron micrograph of two MPV1 particles negatively stained with 2% uranyl acetate. The viruses were cultivated at atmospheric pressure.

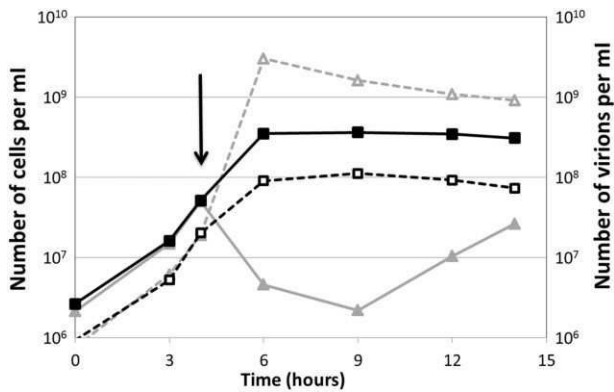


Fig. 2. Virus–host kinetics at atmospheric pressure monitored by flow cytometry. Typical growth curve of the host *M. piezophila* is shown as filled squares on continuous black line, and spontaneous production of MPV1 particles, reaching 1.11×10^8 viral particles per ml, is shown as open squares on dotted black line. Curves in grey correspond to the cultures grown with the addition of $5 \mu\text{g ml}^{-1}$ of mitomycin C (depicted with the arrow): filled triangles on continuous line show *M. piezophila* growth, whereas open triangles on dotted line refer to parallel MPV1 viral production, reaching $\sim 3 \times 10^9$ viral particles per ml.

particles than mitomycin-treated culture, with a maximal virus abundance of 1.11×10^8 viral particles ml^{-1} after 9 h of incubation (Fig. 2). This basal viral production could be explained by a spontaneous induction of MPV1 within a fraction of the growing host cells. Similar spontaneous induction was observed during the growth of *Nitratiruptor* sp., the deep-sea hydrothermal isolate lysogenized by the mitomycin C inducible temperate siphovirus NrS-1 (Yoshida-Takashima *et al.*, 2013).

DNA content of viral particles

The genome sequence of *M. piezophila* KA3 contains a provirus sequence of ~ 40 kb integrated into the strain's chromosome and a cccDNA of 13 386 bp denoted as pMARPI01 (Lucas *et al.*, 2012). pMARPI01 has an *in silico* RFLP pattern identical to that of pMP1, the plasmid that we isolated from *M. piezophila* KA3 by an alkaline lysis extraction (Fig. S2A). For brevity, we will refer to both pMARPI01 and the discovered plasmid as pMP1 throughout the rest of the manuscript. Analysis of DNA extracted from the purified MPV1 viral capsids confirmed that the provirus is the integrated form of the MPV1 genome. Surprisingly, RFLP analyses indicated that the viral DNA is much less frequently present in MPV1 virions (Fig. S2B). Instead, the viral capsids mainly contain pMP1 and hence preferentially package the plasmid. Sequencing of the DNA from the viral particles further confirmed this hypothesis, as not only both a 43 715 bp sequence corresponding to the provirus inserted in *M. piezophila* genome (Fig. 3) and a 13 386 bp plasmid DNA

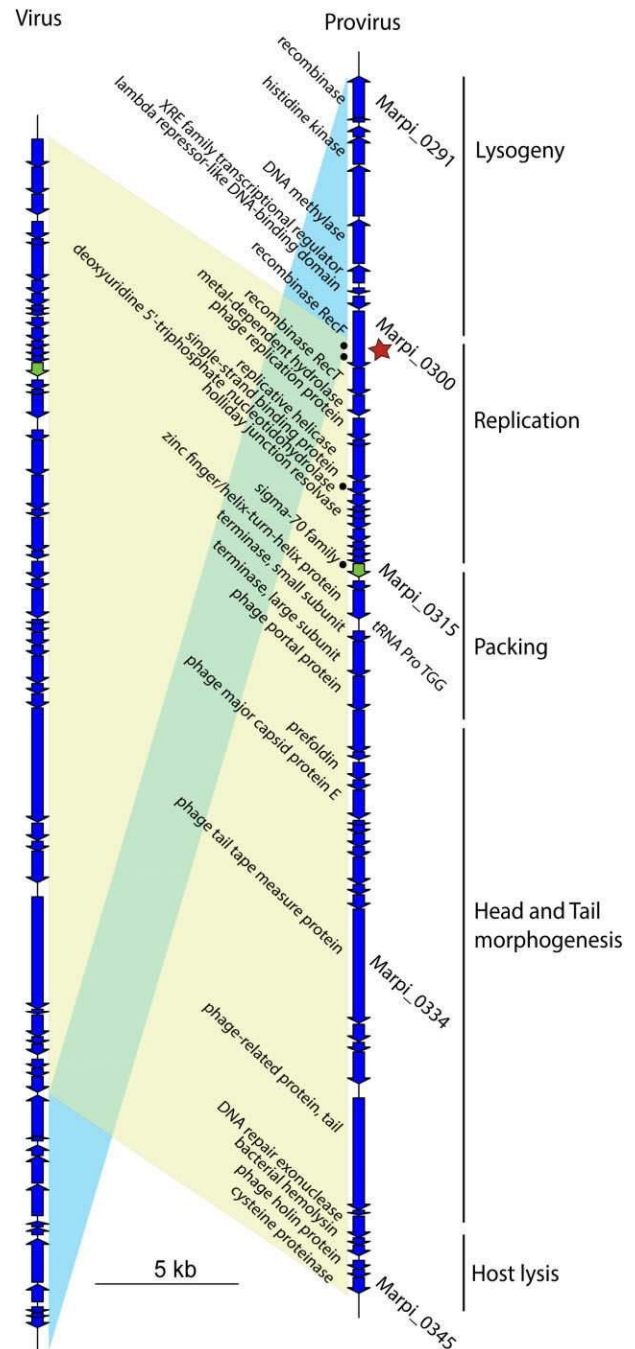


Fig. 3. Comparison of the MPV1 provirus in *M. piezophila* and viral DNA. The regions with significant pairwise BLASTN similarity scores are connected. Informative ORF annotations are shown. The ORF highlighted in green has a homologue in the pMP1 plasmid (see Fig. 4). The ORF marked with a star is disrupted in the linearized genome packaged in the virions. The black dots indicate the short sequences with significant sequence similarity to the regions of the pMP1 genome. The division of the virus into functional modules is approximate, as it is based only on gene annotations. The figure was produced using genoPlotR (Guy *et al.*, 2010).

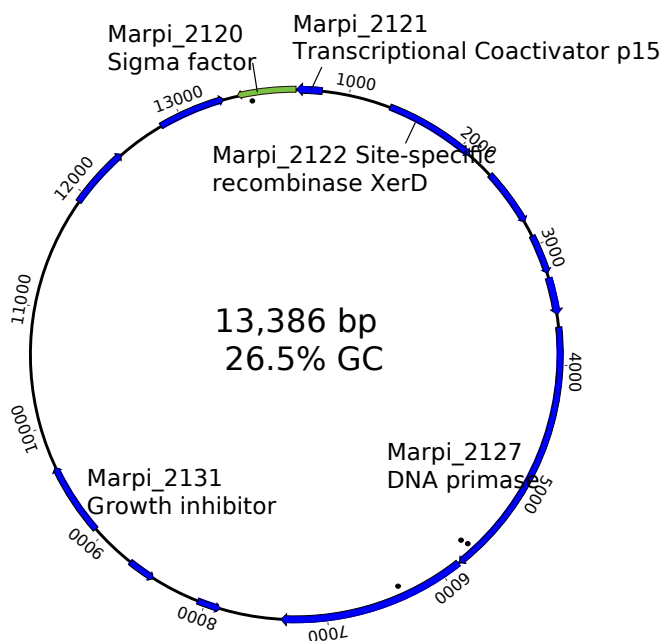


Fig. 4. Map of the pMP1 plasmid. Only informative ORF annotations are shown. The ORF highlighted in green has a homologue in the MPV1 genome (see Fig. 3). The black dots indicate the short sequences with significant similarity to the MPV1 genome. The figure was created in Geneious 7 (Biomatters Limited).

corresponding to pMP1 (Fig. 4) were fully assembled, but also a 20-fold higher coverage of the plasmid reads compared with the 'true viral' reads was observed. We rule out possible extracapsid DNA contamination, as the purified viral suspensions from induced cultures were treated with DNase before DNA extraction. Moreover, specific polymerase chain reaction (PCR) amplifications targeting the provirus, the plasmid and the 16S rRNA gene in the DNA extracted from the viral capsids, resulted in no PCR product corresponding to the 16S rRNA gene (Fig. S3). The same results were obtained with DNA extracted from a purified viral suspension of an uninduced culture (not shown). We concluded that even without induction both viral and plasmid DNA were incorporated into the viral capsids.

Attempts to infect a putative sensitive host

Despite several attempts, we never succeeded in curing *M. piezophila* KA3 strain of MPV1 or pMP1, thus we could not test MPV1's ability to re-infect its host. A purified viral suspension was added to growing cultures of two other potential host cells (*Thermosipho* sp. strain AT1244-VC14 and *M. camini* strain DV1142). For both cultures, no lysis was observed under a phase-contrast microscope. *Marinitoga camini* strain DV1142 did not get infected by

MPV1 (data not shown). After three subcultures of the *Thermosipho* sp. strain AT1244-VC14, no viral DNA was detected, but a 198 bp fragment from the pMP1 plasmid was amplified by PCR from the total DNA (Fig. 5). Moreover, the complete plasmid pMP1 was extracted by alkaline lysis from the *Thermosipho* sp. strain infected by MPV1 (Fig. S4). These findings highlight an inter-genus plasmid (pMP1) propagation by means of a viral capsid.

Effect of the mitomycin C induction on the evolution of the copy numbers of chromosomal, proviral and plasmid target genes by qPCR analyses

The relative quantity of the chromosomal DNA in *M. piezophila* KA3 induced with mitomycin C was two times lower than in the culture not exposed to the antibiotic (Fig. 6 and Fig. S5). This might be a result of a chromosomal DNA damaging effect of mitomycin C, which induces the lytic phase of MPV1. On the other hand, the relative concentration of viral DNA was almost threefold higher in mitomycin C treated cells than in untreated ones (Fig. 6 and Fig. S5). This post-induction increase of the viral DNA replication was consistent with the lysogenic relationship between MPV1 and its host. Finally, the relative concentration of plasmid DNA was 10-fold higher in mitomycin C treated cells than in the uninduced cells (Fig. 6 and Fig. S5). The induction of the plasmid was unexpected and could be a response, as a survival strategy, to escape to the host lysis.

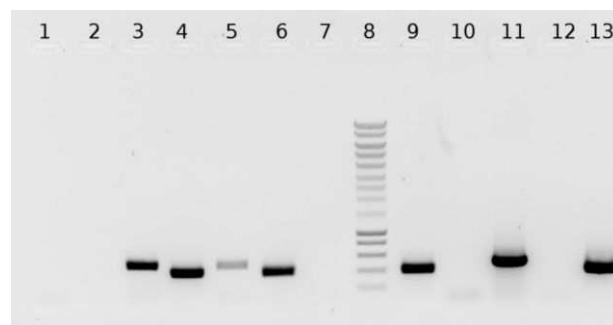


Fig. 5. MPV1 infectivity assays performed on *Thermosipho* sp. AT1244-VC14⁺ and checked by PCR experiments. Lanes 1, 2: PCR on total DNA of the uninfected *Thermosipho* strain targeting (1) MPV1 viral DNA or (2) pMP1 plasmid DNA. Lanes 3, 4: PCR on total DNA of the *Thermosipho* strain just after addition of MPV1 purified capsids targeting (3) MPV1 viral DNA or (4) pMP1 plasmid DNA. Lanes 5, 6: PCR on total DNA of *Thermosipho* strain 22 h after addition of MPV1 purified capsids targeting (5) MPV1 viral DNA or (6) pMP1 plasmid DNA. Lanes 7, 9: PCR on total DNA of the *Thermosipho* strain after 3 subcultures targeting (7) MPV1 viral DNA or (9) pMP1 plasmid DNA. Lanes 10, 12: Negative controls targeting (10) MPV1 viral DNA or (12) pMP1 plasmid DNA. Lanes 11, 13: Positive controls targeting (11) MPV1 viral DNA or (13) pMP1 plasmid DNA. Lane 8: Smart Ladder 1 kb (Eurogentec).

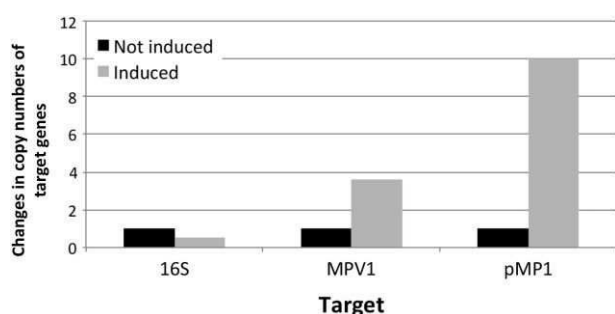


Fig. 6. qPCR assays to assess changes in copy number of *M. piezophila* chromosomal, proviral and plasmid DNA target genes in the presence or absence of mitomycin C induction (with 5 $\mu\text{g ml}^{-1}$ of mitomycin C). Assays without mitomycin C (the black bars) have been set to a value of 1. The light grey bars represent the relative changes in copy number after the induction.

Sequence analysis of the bacteriophage DNA and the plasmid DNA

The viral DNA sequences obtained from the purified MPV1 capsids was identical to the provirus sequence on *M. piezophila* KA3's chromosome (Fig. 3). *De novo* assembly of the viral reads revealed that while some of the packaged DNA molecules are circular, others are linearized, disrupting the Open Reading Frame (ORF) corresponding to Marpi_0300 in the provirus (marked by a star in Fig. 3). The viral genome has 54 predicted ORFs, which include all genes necessary for viral particle production (Fig. 3). The gene order in the viral genome is similar to that seen in *Enterococcus* viruses (Yasmin *et al.*, 2010). Based on BLAST searches, 23 of the 54 predicted viral ORFs have no detectable homologues in GenBank (Table S2). Among the 31 ORFs with homologues in GenBank, eight ORFs have homologues in proviruses identified by us in two *Thermosipho* sp. genomes (C. Nesbø, unpublished), and three ORFs have homologues in the genomes of other *Thermotogales* (Table S2). Therefore, it is likely that MPV1 and closely related viruses have exchanged genes with other *Thermotogales* in the past. However, in the available *Thermotogales* genomes, CRISPR arrays are > 85% identical to MPV1 DNA across the entire spacer in only one *Marinitoga* genome (*Marinitoga* sp. 1137, C. Nesbø, unpublished) (data not shown).

When BLAST 'self matches' to the *M. piezophila* genome and matches to the above-mentioned *Thermosipho* proviruses were excluded, 22 out of the 31 ORFs had *Firmicutes* genes as their top-scoring BLAST hit. Moreover, among these 31 ORFs, 12 were similar to genes in other bacteriophages. The majority of these bacteriophage homologues belong to siphoviruses, but some are from the *Myoviridae* family of bacteriophages with contractile tails (Krupovic *et al.*, 2011), exemplified by the thermophilic *Geobacillus bacteriophage* GBSV1 (Liu

et al., 2009) (Table S2). On phylogenetic trees of 11 out of 19 MPV1 ORFs with more than three GenBank homologues, the sister taxon of an MPV1 gene was either a *Firmicute* or a *Firmicute* bacteriophage (Table S2). These analyses suggest a close evolutionary relationship of MPV1 to *Firmicutes* bacteriophages.

Lysis in Gram-negative bacteria involves three types of proteins: (i) endolysins (to degrade murein); (ii) holins (to increase inner membrane permeability) and (iii) spanners (to disrupt the outer membrane) (Wang *et al.*, 2000). Based on sequence similarity and composition, we have identified candidate lysis genes in the MPV1 genome. We assigned Marpi_340 a putative endolysin function, as similarity searches place the gene with phosphoesterases and metallophosphatase superfamily. We designated Marpi_343 and Marpi_344 as holins: Marpi_343 has significant similarity to holin homologues in GenBank, whereas physicochemical features of Marpi_344 (102 amino acids in length, predicted transmembrane helix, dual start motif and a charged C-terminal sequence) are compatible with those of known holins (Young, 2013). Similarity searches placed Marpi_345 within the NlpC/P60 family (secreted cysteine proteinases that degrade cell walls), making this gene a candidate for encoding a protein that disrupts the outer membrane. Marpi_341 possibly encodes another membrane-degrading enzyme, as it exhibits a weak similarity to type I phosphodiesterase/nucleotide pyrophosphatase superfamily in *Helicobacterium modestum* lce1 (a *Firmicute*).

pMP1 plasmid contains 13 ORFs: six have Genbank homologues with a predicted function (Fig. 4 and Table S3), whereas remaining seven could be assigned to the gene families in the Superfamily database (Gough *et al.*, 2001). No known structural genes involved in virion formation were identified. Nine of the putative 13 proteins have a predicted pI above 9.0 (Table S3). Moreover, all seven protein-coding genes with homologues in the Superfamily database were assigned to gene families involved in DNA interactions (i.e. transcription, regulation and DNA metabolic processes; Table S3). Although pMP1 was not found integrated into the sequenced *M. piezophila* chromosome, one of its genes (Marpi_2122) encodes a site-specific recombinase XerD (an integrase family protein). Genes for DNA polymerases and primases are not commonly found on plasmids; however, Marpi_2127 encodes a protein with both a DnaB (helicase) and a DnaG (primase) domain. Recently, a DNA primase/polymerase protein (PolpTN2) encoded by the pTN2 plasmid from *Thermococcus nautilus* has been characterized (Gill *et al.*, 2014). To summarize, a large fraction of pMP1-encoded proteins might be involved in DNA binding and possibly replication and propagation of the plasmid.

Pairwise BLASTN comparison of the pMP1 and MPV1 genomes revealed a 17 bp region with 100% identity [E-value = 0.14; located in Marpi_300 (MPV1) and Marpi_2128 (pMP1)] and three short regions with > 80% identity (E-value < 0.05): a 28 bp region with 89% identity [located in Marpi_300 (MPV1) and Marpi_2127 (pMP1)], a 33 bp region with 85% identity [located in Marpi_315 (MPV1) and Marpi_2120 (pMP1)] and a 23 bp region with 91% identity [located in Marpi_306 (MPV1) and Marpi_2127 (pMP1)] (Figs 3 and 4, Tables S2 and S3). These matching sequence motifs could serve as recognition sequences during packing of pMP1 genome into the viral particles. Interestingly, one of the pMP1 ORFs (Marpi_2120) has significant similarity to the proviral gene (Marpi_0315; green in Fig. 3) located two ORFs upstream of a putative small terminase subunit, which is involved in packing of viral DNA (Feiss and Rao, 2012). Both Marpi_2120 and Marpi_0315 have a DNA binding helix–turn–helix domain and show significant similarity to Sigma-70 family RNA polymerase genes. Taken together, these observations strongly suggest that both Marpi_0315 and Marpi_2120 could be involved in packing DNA into the viral particles.

Discussion

Deep-sea hydrothermal vents are harsh environments, characterized by steep physicochemical gradients, high hydrostatic pressures, high temperatures, lack of solar energy and prevalence of microbial chemosynthesis. The dynamic, gradient-dominated nature of the vent environment provides niches for diverse communities of microorganisms and makes it a particularly attractive site for studies of viral diversity and evolution (Anderson et al., 2011a). In this ecosystem, *Thermotogales* share the ecological niche with diverse prokaryotes, such as frequently isolated bacteria from the *Firmicutes* phylum and archaea from the *Thermococcales* order (Priour et al., 2004). The bacterial order *Thermotogales* is known for its complex evolutionary history shaped by large amounts of gene transfer from the above prokaryotic lineages (Nelson et al., 1999; Zhaxybayeva et al., 2009). Viruses represent one possible vector for this gene exchange, but prior to this study no viruses infecting *Thermotogales* cells have been identified. In support of this hypothesis, the reported genome of the first *Thermotogales* bacteriophage, MPV1, reveals a connection to genomes of *Firmicutes* and bacteriophages known to infect them (Table S2). Because *Thermotogales* and *Firmicutes* appear to share large number of genes (Zhaxybayeva et al., 2009), our findings imply that some of the genes exchanges may have been mediated by viruses. Further studies of *Thermotogales* and *Firmicutes* viruses will be needed to explore this hypothesis.

MPV1 is a temperate virus and lysogeny is indeed a common lifestyle in extreme environments such as hydrothermal vents (Williamson et al., 2008; Anderson et al., 2011a). From the viral point of view, lysogeny could be favored due to reduced viability of viral particles outside of the cell. For the cell, on another hand, carrying a virus could increase its fitness by enhancing the survivability in fluctuating environmental conditions via introduced novel genes and/or mutations (Anderson et al., 2011a). Further analyses will be necessary to better understand the MPV1 viral cycle. In particular, the mechanisms of MPV1 adsorption and extrusion through the 'toga', an unusual cell envelope forming a loose sheet envelope around cells, remain unknown. The toga is mainly composed of two proteins, the anchor protein OmpA1 and the porin OmpB (Petrus et al., 2012). Porins are known to serve as viral receptors: In *E. coli*, for example, T4 viral particles use the OmpC porin as a receptor (Hashemolhosseini et al., 1994; Marti et al., 2013). Therefore, we hypothesize that OmpB may serve as MPV1 receptors in *Thermotogales*.

Peculiarly, MPV1 viral capsids mainly package a plasmid of 13.3 kb (pMP1), facilitating the inter-genus plasmid propagation. Thus, this system is the first virus-mediated plasmid exchange reported in deep-sea hydrothermal vents. Although the interaction between MPV1 and pMP1 is not fully understood, the ratio of packed plasmid/viral DNA is high, which suggests that pMP1/MPV1 represent a new example of molecular piracy. This term has been coined to describe a biological phenomenon in which one replicon (the pirate) uses the structural proteins encoded by another replicon (the viral helper) to pack its own genome and thus allow the pirate's propagation (Christie and Dokland, 2012). The type of relationship between a virus and another mobile genetic element within a same host, the so-called 'ménage à trois', has been described in organisms from all three domains of life. A distinct plasmid (a transpoviron) discovered in giant eukaryoviruses infecting amoebae depends on giant viruses for its replication and propagation (Desnues et al., 2012). Analogies can be made between the eukaryotic transpovirons and virus-associated plasmids described in archaea and bacteria. For example, non-conjugative plasmids pSSVi and pSSVx exploit fuselloviruses in order to propagate within strains of the hyperthermophilic archaeon *Sulfolobus* (Arnold et al., 1999; Wang et al., 2007). Similar strategies were described for the well-studied *E. coli* P4/P2 system in which the plasmid P4 uses the capsids of the bacteriophage helper P2 to propagate in *E. coli* cells (Six, 1975; Lindqvist et al., 1993; Briani et al., 2001) and the *Staphylococcus aureus* SaPIs pathogenicity islands/ bacteriophages helper systems (Ram et al., 2012). In the two bacterial systems described above, the pirate

element is mobilized and packed into viral-like transducing particles assembled from proteins supplied by a helper caudovirus. These particular relationships involve a complex web of interactions between the helper and the pirate that occur at several levels, from transcriptional control to macromolecular assembly (Christie and Dokland, 2012).

The MPV1/pMP1 system shares many features with the *E. coli* and *S. aureus* pirates. Both pMP1 and P4 are plasmids of similar sizes and are able to persist in their respective host cells in a self-replicating multicopy cccDNA state. Similar to P4 and SaPIs, pMP1 harbours an integrase gene, suggesting that it might recombine within host chromosome. Also, like the SaPIs pirates, pMP1 is induced by mitomycin C along with the provirus and responds by replicating and escaping from the bacterial hosts by using helper viral capsids. The induction of the plasmid was unexpected and, to our knowledge, has not yet been reported in prokaryotes.

It has been demonstrated that pirates have evolved different strategies of interference with their helpers to favor the packing of their own pirate genomes into the helper capsids (Christie and Dokland, 2012). Specific DNA packing strategy of the pMP1 element remains to be discovered. For example, we have not observed small MPV1 capsids that are likely to contain a single copy of the plasmid pMP1. We hypothesize that multiple copies of pMP1 (x3) are packaged in MPV1 capsids. Indeed, the discontinuous headful packaging of multiple plasmid DNAs in viral capsid has been demonstrated (Bravo and Alonso, 1990; Coren *et al.*, 1995; Leffers and Rao, 1996). It has been shown that the frequency of these plasmid-containing particles is enhanced when there is DNA similarity between a plasmid and a virus (Bravo and Alonso, 1990). We detected three regions of significant sequence similarity between pMP1 and MPV1 genomes. Among these, the strongest candidate for a common packing signal is a 28 bp region in Marpi_0300, which is disrupted in the linearized version of the MPV1 genome. The pMP1 protein encoded by Marpi_2120 might be involved in the binding to the putative packing signal.

Would there be any benefit for the host from this 'ménage à trois'? Carrying a provirus often confers resistance against related bacteriophages (Canchaya *et al.*, 2003), whereas bearing the pirate that hijacks the viral capsid could increase the survival of the bacterial host population, due to lesser exposure to the bacteriophage DNA (Christie and Dokland, 2012). Resistance of the helper bacteriophages against the pirate elements, which alter the propagation capacity of the bacteriophage for their own survival, is expected to evolve but have not yet been observed. Thus, the MPV1/pMP1 system likely represents an ongoing evolutionary 'arms race'.

Mobile genetic elements (MGEs), such as viruses, plasmids, membrane vesicles, gene transfer agents, transposons and transpovirons (all collectively referred as a mobilome) interact with cellular organisms from all three domains of life, including those thriving in most extreme environments. The MPV1/pMP1 system demonstrates the existence of an intricate network of interactions in the deep-sea vent mobilome, and further studies are likely to unravel unknown mechanisms for the host–MGEs interactions. The MGEs have a potential to be powerful drivers of cellular host adaptations to the extreme marine environments, and comparative studies of these elements from hyper/thermophilic *Archaea* and *Bacteria* will help us understand the dynamics of the microbial communities in the deep biosphere.

Experimental procedures

Screening of *Thermotogales* strains for mobile genetic elements

A total of 60 strains of *Thermotogales* were retrieved from the UBO culture collection 'UBOCC' (<http://www.univ-brest.fr/UBOCC>) and from our laboratory culture collections. As described in *Supporting information*, these strains were isolated from different deep-sea vents in the Lau Basin, the East Pacific Rise and the Middle Atlantic Ridge and were screened for mobile genetic elements as previously described (Geslin *et al.*, 2003a,b; Gorlas *et al.*, 2012). One of them, the reference strain *M. piezophila* KA3 (Alain *et al.*, 2002), was discovered to be the host of the virus MPV1 and the plasmid pMP1.

Culture conditions

Thermotogales strains were cultured in a modified Ravot medium as previously described (Ravot *et al.*, 1995; Geslin *et al.*, 2003a,b) with minor modifications as described in *Supporting information*. Most experiments were performed with the *M. piezophila* KA3 cultured at 65°C and 0.1 MPa. For the high hydrostatic pressure growth experiments, *M. piezophila* KA3 cells were grown at 65°C and 40 MPa pressure as described in *Supporting information*.

Viral induction assays by addition of mitomycin C

Attempts were made to increase the viral production by induction using mitomycin C (Yoshida-Takashima *et al.*, 2013). Several final concentrations of mitomycin C (0.1 µg ml⁻¹, 1 µg ml⁻¹, 5 µg ml⁻¹ and 10 µg ml⁻¹) were added to cultures of *M. piezophila* at mid-log growth phase. After 3 h of incubation with mitomycin C, aliquots were collected in duplicate to determine viral production by epifluorescence microscopy. A 'mitomycin C-free' culture was used as a control. Samples were fixed with 25% glutaraldehyde [0.5% final concentration, Electron microscopy (EM) grade] for 15 min at 4°C, flash frozen in liquid nitrogen and stored at -80°C until analysis. The viral counts were performed

according to Noble and Fuhrman (1998). Viral particles collected on anodisc 0.02 µm filters were counted in each of 20 random fields using an Olympus microscope BX60 (magnification $\times 100$, WIB filter).

Flow cytometry counts of bacteria and viruses

Bacteria and virus abundances were monitored by flow cytometry using a FACSCanto II flow cytometer (BD Bioscience) equipped with a laser with an excitation wavelength of 488 nm (15 mW). Samples were fixed as previously described. The thawed samples were diluted from 100- to 5000-fold in autoclaved 0.2 µm filtered TE buffer (10:1 Tris-EDTA, pH 8.0) and stained with the nucleic acid-specific dye SYBR Green I (Invitrogen-Molecular Probes) for 15 min at room temperature and 10 min at 80°C for bacteria and viruses respectively. The trigger was set on green fluorescence. The sample was delivered at a rate of 50 µl min⁻¹ and analysed during at least 1 min. Tris-EDTA (TE)-buffer with autoclaved 0.2 µm filtered seawater was used as a control.

Concentration and purification of MPV1

Three hundred millilitres of *M. piezophila* culture were induced by addition of mitomycin C at early log-phase (5 µg ml⁻¹). Two hours later, debris and cells were pelleted by centrifugation at 7500 g, 4°C for 15 min and discarded. The supernatant was ultracentrifuged at 80 000 g, 10°C for 2 h (Beckman Optima LE-80 K 70.1 Ti rotor). The viral pellet was resuspended in 1 ml of buffer (10 mM Tris-HCL, 100 mM NaCl, 5 mM CaCl₂, 20 mM MgCl₂). The viral suspension was purified by ultracentrifugation in a linear iodixanol gradient [OptiPrep, 30–45% diluted in a buffer (see above)] at 40 000 g and 10°C for 5 h (Beckman Optima LE-80 K SW 55 rotor). Following ultracentrifugation, the opaque virus band was recovered and stored at 4°C until use. To exclude the possibility of cellular nucleic acids contamination, the concentrated viral particles suspensions were treated with DNase and RNase (50 µg ml⁻¹ and 100 µg ml⁻¹ respectively). The viral particles were examined using a JEOL JEM 100 CX II transmission electron microscope as previously described (Geslin et al., 2003a,b; Gorlas et al., 2012).

Extraction of cellular and viral DNA

Total DNA from *M. piezophila* cells and viral DNA from purified virions were prepared as previously described (Geslin et al., 2003a,b; Gorlas et al., 2012). Plasmid DNA was extracted from cells in log growth phase by alkaline lysis method (Bimboim and Doly, 1979). To confirm the nature of the DNA packaged in the viral capsids and to check potential chromosomal contamination, PCR amplifications were performed using primer sets for 16S rRNA gene (*M. piezophila* chromosomal DNA marker), major capsid protein encoding gene (Marpi_326, provirus and virus marker) and hypothetical protein encoding gene (Marpi_2124, pMP1 plasmid marker). The primers and PCR protocol are described in Supporting information.

RFLP analyses

Marinitoga piezophila KA3 total DNA and plasmid DNA were restricted by the enzymes HhaI and HindIII (Promega®);

packaged viral DNA was restricted by HhaI, in accordance with the manufacturer's instructions. Restriction fragments obtained were separated on a 0.8% agarose gel electrophoresis.

Attempts to infect a putative sensitive host

Strains belonging to *Marinitoga* and *Thermosipho* genera, obtained from our laboratory cultures collections, were screened for the presence of genetic elements and strains scored negative were selected as potential hosts. As a result, *Thermosipho* sp. AT1244-VC14 and *M. camini* DV1142 strains were selected. A purified MPV1 suspension was added in exponentially growing host (virus/host = 30) cultured in 5 ml modified Ravot medium at 65°C. Non-infected cultures were used as negative control. After three subcultures, total and plasmid DNAs were extracted from each culture. PCR amplifications were performed on total DNA as previously described in Supporting information by using the specific primers targeting the bacteriophage MPV1 and the plasmid DNA pMP1.

QPCR assays for quantifications of the copy numbers of chromosomal, proviral and plasmid target genes

A culture of *M. piezophila* KA3 was split into two aliquots. One aliquot was induced with mitomycin C (5 µg ml⁻¹) at early-log growth phase, whereas the second was not induced. The cultures were incubated for 80 min at 65°C, and total DNA was extracted. The DNA concentration was estimated by NanoDrop ND-1000 and diluted to a concentration of 0.2 ng µL⁻¹. Real-time PCR amplifications were performed in triplicate using Syber Green Supermix (Taq Core, Qiagen) according to the manufacturer's instructions with modifications. Primers sets are the same than those previously used; real-time PCR protocol is described in Supporting information. The copy numbers of target genes were calculated from mitomycin C treated and non-treated cultures respectively.

Sequencing and assembly of bacteriophage and plasmid DNA

Purified viral DNA and plasmid DNA were sequenced on Ion Torrent Personal Genome Machine (PGM) using the Ion PGM sequencing 200 kit v.2 (LifeTechnologies) and the Ion Xpress Barcode adapters 1–16 kit (LifeTechnologies). The sequences were assembled using the CLC Genomics Workbench 6.5.1 (<http://www.clcbio.com>) as well as MIRA 4 (Chevreux et al., 1999). Sequence analyses of protein-coding genes of bacteriophage and pMP1 are described in Supporting information.

Acknowledgements

This work was financially supported by the French Ministry of Higher Education to JL and the Agence Nationale de la Recherche, project Thermovésicules (ANR-12-BSV3-0023-01) to CG, a Norwegian Research Council award (project no. 180444/V40) to CLN, and Dartmouth's Walter and

Constance Burke Research Initiation Award and the Dean of Faculty start-up funds to OZ. We thank S. Dupont for his technical help and G. Siquin for assistance in electron microscopy.

References

- Akimkina, T., Ivanov, P., Kostrov, S., Sokolova, T., Bonch-Osmolovskaya, E.A., Firman, K., *et al.* (1999) A highly conserved plasmid from the extreme thermophile *Thermotoga maritima* MC24 is a member of a family of plasmids distributed worldwide. *Plasmid* **42**: 236–240.
- Alain, K., Marteinsson, V.T., Miroshnichenko, M.L., Bonch-Osmolovskaya, E.A., Prieur, D., and Birrien, J.-L. (2002) *Marinitoga piezophila* sp. nov., a rod-shaped, thermo-piezophilic bacterium isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* **52**: 1331–1339.
- Anderson, R.E., Brazelton, W.J., and Baross, J.A. (2011a) Is the genetic landscape of the deep subsurface biosphere affected by viruses? *Front Microbiol* **2**: 219. doi: 10.3389/fmicb.2011.00219.
- Anderson, R.E., Brazelton, W.J., and Baross, J.A. (2011b) Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol Ecol* **77**: 120–133.
- Arnold, H.P., She, Q., Phan, H., Stedman, K., Prangishvili, D., Holz, I., *et al.* (1999) The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. *Mol Microbiol* **34**: 217–226.
- Bimboim, H., and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**: 1513–1523.
- Bonch-Osmolovskaya, E.A. (2008) Thermotogales. In *Encyclopedia of Life Sciences (ELS)*. Chichester: John Wiley & Sons, Ltd.
- Bravo, A., and Alonso, J.C. (1990) The generation of concatemeric plasmid DNA in *Bacillus subtilis* as a consequence of bacteriophage SPP1 infection. *Nucleic Acids Res* **18**: 4651–4657.
- Briani, F., Deho, G., Forti, F., and Ghisotti, D. (2001) The plasmid status of satellite bacteriophage. *Plasmid* **45**: P4.
- Canchaya, C., Proux, C., Fournous, G., Bruttin, A., and Brüssow, H. (2003) Prophage genomics. *Microbiol Mol Biol Rev* **67**: 238–276.
- Chevreaux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Müller, W.E., Wetter, T., and Suhai, S. (1999) Genome sequence assembly using trace signals and additional sequence information. In *Computer Science and Biology: German Conference on Bioinformatics (GCB)* 45, pp. 45–56.
- Christie, G.E., and Dokland, T. (2012) Pirates of the *Caudovirales*. *Virology* **434**: 210–221.
- Coren, J.S., Pierce, J.C., and Sternberg, N. (1995) Headful packaging revisited: the packaging of more than one DNA molecule into a bacteriophage P1 head. *J Mol Biol* **249**: 176–184.
- Desnues, C., La Scola, B., Yutin, N., Fournous, G., Robert, C., Azza, S., *et al.* (2012) Provirophages and transpovirons as the diverse mobilome of giant viruses. *Proc Natl Acad Sci U S A* **109**: 18078–18083.
- Feiss, M., and Rao, V.B. (2012) The bacteriophage DNA packaging machine. In *Viral Molecular Machines*. Rossmann, M.G., and Rao, V.B. (eds). New York City, USA: Springer, pp. 489–509.
- Flores, G.E., and Reysenbach, A.-L. (2011) Hydrothermal Environments, Marine. In *Encyclopedia of Earth Science Series: Encyclopedia of Geobiology*. Reitner, J., and Thiel, V. (eds). Dordrecht, The Netherlands: Springer, pp. 456–467.
- Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003a) PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, *Pyrococcus abyssi*. *J Bacteriol* **185**: 3888–3894.
- Geslin, C., Le Romancer, M., Gaillard, M., Erauso, G., and Prieur, D. (2003b) Observation of virus-like particles in high temperature enrichment cultures from deep-sea hydrothermal vents. *Res Microbiol* **154**: 303–307.
- Geslin, C., Gaillard, M., Flament, D., Rouault, K., Le Romancer, M., Prieur, D., *et al.* (2007) Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from *Pyrococcus abyssi*. *J Bacteriol* **189**: 4510–4519.
- Gill, S., Krupovic, M., Desnues, N., Béguin, P., Sezonov, G., and Forterre, P. (2014) A highly divergent archaeo-eukaryotic primase from the *Thermococcus nautilus* plasmid, pTN2. *Nucleic Acids Res* **42**: 3707–3719.
- Gorlas, A., Koonin, E.V., Biennu, N., Prieur, D., and Geslin, C. (2012) TPV1, the first virus isolated from the hyperthermophilic genus *Thermococcus*. *Environ Microbiol* **14**: 503–516.
- Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J Mol Biol* **313**: 903–919.
- Guy, L., Kultima, J.R., and Andersson, S.G. (2010) genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **26**: 2334–2335.
- Hania, W.B., Postec, A., Aüllo, T., Ranchou-Peyruse, A., Erauso, G., Brochier-Armanet, C., *et al.* (2013) *Mesotoga infera* sp. nov., a mesophilic member of the order *Thermotogales*, isolated from an underground gas storage aquifer. *Int J Syst Evol Microbiol* **63**: 3003–3008.
- Harriott, O.T., Huber, R., Stetter, K.O., Betts, P.W., and Noll, K.M. (1994) A cryptic miniplasmid from the hyperthermophilic bacterium *Thermotoga* sp. strain RQ7. *J Bacteriol* **176**: 2759–2762.
- Hashemolhosseini, S., Montag, D., Krämer, L., and Henning, U. (1994) Determinants of receptor specificity of coliphages of the T4 family: a chaperone alters the host range. *J Mol Biol* **241**: 524–533.
- Krupovic, M., Prangishvili, D., Hendrix, R.W., and Bamford, D.H. (2011) Genomics of bacterial and archaeal viruses: dynamics within the prokaryotic virosphere. *Microbiol Mol Biol Rev* **75**: 610–635.
- Leffers, G., and Rao, V.B. (1996) A discontinuous headful packaging model for packaging less than headful length DNA molecules by bacteriophage T4. *J Mol Biol* **258**: 839–850.
- Lindqvist, B.H., Dehò, G., and Calendar, R. (1993) Mechanisms of genome propagation and helper exploitation by satellite phage P4. *Microbiol Rev* **57**: 683–702.

- Liu, B., and Zhang, X. (2008) Deep-sea thermophilic *Geobacillus* bacteriophage GVE2 transcriptional profile and proteomic characterization of virions. *Appl Microbiol Biotechnol* **80**: 697–707.
- Liu, B., Wu, S., Song, Q., Zhang, X., and Xie, L. (2006) Two novel bacteriophages of thermophilic bacteria isolated from deep-sea hydrothermal fields. *Curr Microbiol* **53**: 163–166.
- Liu, B., Zhou, F., Wu, S., Xu, Y., and Zhang, X. (2009) Genomic and proteomic characterization of a thermophilic *Geobacillus* bacteriophage GBSV1. *Res Microbiol* **160**: 166–171.
- Lucas, S., Han, J., Lapidus, A., Cheng, J.-F., Goodwin, L.A., Pitluck, S., et al. (2012) Complete genome sequence of the thermophilic, piezophilic, heterotrophic bacterium *Marinitoga piezophila* KA3. *J Bacteriol* **194**: 5974–5975.
- Marti, R., Zurfluh, K., Hagens, S., Pianezzi, J., Klumpp, J., and Loessner, M.J. (2013) Long tail fibres of the novel broad-host-range T-even bacteriophage S16 specifically recognize *Salmonella* OmpC. *Mol Microbiol* **87**: 818–834.
- Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J., Haft, D.H., et al. (1999) Evidence for lateral gene transfer between *Archaea* and *Bacteria* from genome sequence of *Thermotoga maritima*. *Nature* **399**: 323–329.
- Nesbø, C.L., Bradnan, D.M., Adebisuyi, A., Dlutek, M., Petrus, A.K., Foght, J., et al. (2012) *Mesotoga prima* gen. nov., sp. nov., the first described mesophilic species of the *Thermotogales*. *Extremophiles* **16**: 387–393.
- Noble, R.T., and Fuhrman, J.A. (1998) Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquat Microb Ecol* **14**: 113–118.
- Petrus, A.K., Swithers, K.S., Ranjit, C., Wu, S., Brewer, H.M., Gogarten, J.P., et al. (2012) Genes for the major structural components of *Thermotogales* species' togas revealed by proteomic and evolutionary analyses of OmpA and OmpB homologs. *PLoS ONE* **7**: e40236.
- Prieur, D., Erauso, G., Geslin, C., Lucas, S., Gaillard, M., Bidault, A., et al. (2004) Genetic elements of *Thermococcales*. *Biochem Soc Trans* **32**: 184–187.
- Ram, G., Chen, J., Kumar, K., Ross, H.F., Ubeda, C., Damle, P.K., et al. (2012) *Staphylococcal* pathogenicity island interference with helper phage reproduction is a paradigm of molecular parasitism. *Proc Natl Acad Sci U S A* **109**: 16300–16305.
- Ravot, G., Magot, M., Fardeau, M.-L., Patel, B., Prensier, G., Egan, A., et al. (1995) *Thermotoga elfii* sp. nov., a novel thermophilic bacterium from an African oil-producing well. *Int J Syst Bacteriol* **45**: 308–314.
- Six, E.W. (1975) The helper dependence of satellite bacteriophage P4: which gene functions of bacteriophage P2 are needed by P4? *Virology* **67**: 249–263.
- Wang, I.-N., Smith, D.L., and Young, R. (2000) Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54**: 799–825.
- Wang, Y., and Zhang, X. (2010) Genome analysis of deep-sea thermophilic phage D6E. *Appl Environ Microbiol* **76**: 7861–7866.
- Wang, Y., Duan, Z., Zhu, H., Guo, X., Wang, Z., Zhou, J., et al. (2007) A novel *Sulfolobus* non-conjugative extrachromosomal genetic element capable of integration into the host genome and spreading in the presence of a fusellovirus. *Virology* **363**: 124–133.
- Williamson, S.J., Cary, S.C., Williamson, K.E., Helton, R.R., Bench, S.R., Winget, D., and Wommack, K.E. (2008) Lysogenic virus-host interactions predominate at deep-sea diffuse-flow hydrothermal vents. *Int J Syst Evol Microbiol* **2**: 1112–1121.
- Yasmin, A., Kenny, J.G., Shankar, J., Darby, A.C., Hall, N., Edwards, C., and Horsburgh, M.J. (2010) Comparative genomics and transduction potential of *Enterococcus faecalis* temperate bacteriophages. *J Bacteriol* **192**: 1122–1130.
- Yoshida-Takashima, Y., Takaki, Y., Shimamura, S., Nunoura, T., and Takai, K. (2013) Genome sequence of a novel deep-sea vent epsilonproteobacterial phage provides new insight into the co-evolution of *Epsilonproteobacteria* and their phages. *Extremophiles* **17**: 405–419.
- Young, R. (2013) Phage lysis: do we have the hole story yet? *Curr Opin Microbiol* **16**: 790–797.
- Zhaxybayeva, O., Swithers, K.S., Lapierre, P., Fournier, G.P., Bickhart, D.M., DeBoy, R.T., et al. (2009) On the chimeric nature, thermophilic origin, and phylogenetic placement of the *Thermotogales*. *Proc Natl Acad Sci U S A* **106**: 5865–5870.
- Zhaxybayeva, O., Swithers, K.S., Foght, J., Green, A.G., Bruce, D., Detter, C., et al. (2012) Genome sequence of the mesophilic *Thermotogales* bacterium *Mesotoga prima* MesG1. Ag. 4.2 reveals the largest *Thermotogales* genome to date. *Genome Biol Evol* **4**: 812–820.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Induction assays of MPV1 production by addition of several mitomycin C concentrations. Several final concentrations of mitomycin C (0.1 $\mu\text{g ml}^{-1}$, 1 $\mu\text{g ml}^{-1}$, 5 $\mu\text{g ml}^{-1}$ and 10 $\mu\text{g ml}^{-1}$) were added to cultures of *M. piezophila* at mid-log growth phase. Aliquots were collected in duplicate to determine viral production by epifluorescence microscopy. The high of bars correspond to the averages of the replicates. Error bars show one standard deviation.

Fig. S2. RFLP analyses of *M. piezophila* total DNA, plasmid DNA (pMP1) and viral DNA.

A. RFLP analyses of both *M. piezophila* total DNA and plasmid DNA (pMP1). 1: DNA Smart Ladder; 2: plasmid DNA restricted by HhaI; 3: total DNA restricted by HhaI; 4: plasmid DNA restricted by HindIII; 5: total DNA restricted by HindIII. B. RFLP analyses of the total DNA extracted from the purified viral capsids. 1: DNA Smart Ladder; 2: Packaged DNA uncut; 3: Packaged DNA restricted by HhaI. The two thickest bands (of ~10 000 bp and 2600 bp) correspond to the plasmid DNA restriction pattern (see lane 2 on the panel A). The weaker bands correspond to the viral sequence, since several of them are identical to those obtained after an *in silico* restriction of the *M. piezophila* provirus (not shown).

Fig. S3. PCR analysis of the total DNA extracted from purified MPV1 viral capsids from an induced culture. 1: Packaged DNA amplified by 16S rRNA primers; 2: Packaged DNA amplified by pMP1 plasmid primers; 3: Packaged DNA

amplified by MPV1 proviral primers; 4: 16S rRNA negative control; 5: 16S rRNA positive control; 6: DNA Smart Ladder.

Fig. S4. Agarose gel electrophoresis of plasmid DNA extracted by alkaline lysis from *Thermosipho* sp. strain AT1244-VC14, three subcultures after being infected by MPV1 virions. 1: DNA Smart Ladder; 2: plasmid DNA restricted by HhaI. The two bands indicated by arrows correspond to pMP1 restriction pattern obtained by HhaI (Fig. S2 panel A, lane 2).

Fig. S5. QPCR assays. Copy number of target genes in the presence or absence of mitomycin C induction. The black bars represent the number of copies of the 16S rRNA chro-

mosomal gene. The light gray bars represent the number of copies of the MPV1 viral gene. The dark gray bars represent the number of copies of the pMP1 plasmid gene.

Table S1. Comparison of viral production at 0.1 and 40 MPa monitored by flow cytometry.

Table S2. Summary of the similarity searches, protein properties' predictions and phylogenetic analyses of the MPV1 provirus.

Table S3. Summary of the similarity searches, protein properties' predictions and phylogenetic analyses of the pMP1 plasmid.

Appendix S1. Experimental procedures.

Supplementary experimental procedures

Screening of *Thermotogales* strains for mobile genetic elements. A total of 60 strains of *Thermotogales* were retrieved from the UBO culture collection "UBOCC" (www.univ-brest.fr/UBOCC) and from our laboratory cultures collections. These strains were isolated from deep-sea vents in the Lau Basin (latitude, 22°32'S; longitude, 176°43'W; depth, 1832 to 1887 m), the East Pacific Rise (latitude, 12°48'21"N; longitude, 103°56'35"W; depth, 2630 m) and the Middle Atlantic Ridge (latitude, 36°13'N; longitude, 33°54'W; depth 2775 m).

Culturing methods. *Thermotogales* strains were cultured in a modified Ravot medium which contained (per liter of distilled water) 0.2 g of NH_4Cl , 0.15 g of $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 0.1 g of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.5 g of KCl, 0.83 g of $\text{Na}_2\text{CO}_3 \cdot 3\text{H}_2\text{O}$, 2 g of yeast extract, 2 g of tryptone, 30 g of sea salt, 3.3 g of piperazine-N-bis(2-ethanesulfonic acid) (PIPES), 2 g of maltose and 0.001 g of resazurin. The pH was adjusted to 6 and the medium was sterilized by autoclaving. After the medium cooled down, the following solutions, separately sterilized by autoclaving, were added: 5 ml of a 7 % (wt/vol) FeSO_4 solution and 5 ml of 7 % (wt/vol) K_2HPO_4 solution. The medium was dispensed (50 ml) into 100-ml sterile vials and completed by adding 1% (wt/vol) elemental sulfur previously sterilized by steaming at 100°C for 1 h on three successive days. Anaerobiosis was obtained by applying vacuum to the medium and saturating it with N_2 . Finally, the medium was reduced by adding a sterile solution of $\text{Na}_2\text{S} \cdot 9\text{H}_2\text{O}$ (final concentration of 0.05% [wt/vol]). The medium inoculated, to the final concentration of 1%, was incubated at 60°C, 65°C and 70°C depending to the strains. *Marinitoga piezophila* KA3 is piezophile as its optimum growth occurring at 40 MPa. However, after several subcultures at 0.1 MPa cells are adapted to atmospheric pressure. Most of the experiments were consequently performed at atmospheric pressure. *Marinitoga piezophila* KA3 cultured at atmospheric pressure and 65°C with L-cystine added as electron acceptor. The high hydrostatic pressure experiments were performed in 5 ml syringes under 40 MPa pressure at 65 °C. The syringe was loaded anaerobically with 5 ml of reduced and modified Ravot medium and inoculated to a final concentration of 1% with *Marinitoga piezophila* KA3 culture in late exponential phase that was grown under high pressure. The syringes were incubated at high pressure (40 MPa) and temperature (65°C) in HP/HT systems, custom-built by Top Industrie.

Primers design and PCR protocol. To confirm the nature of the DNA packaged in the viral capsids and to check potential chromosomal contamination PCR amplifications were performed. Three primer sets were designed by using the Primer 3 software (Untergasser *et al.*, 2007). The first set is specific to *Marinitoga piezophila* KA3: 16S rRNA gene (forward: 5'-ACACATGCAAGTCGAACGAG-3' and reverse: 5'-CACTGGAAACGGTGGCTAAT-3'), the second set is specific to the major capsid protein gene of *Marinitoga piezophila* KA3 MPV1 provirus: (forward: 5'-AGGAGGACCTCAACCAACAA-3' and reverse: 5'-CAAGCACGAGATTTGAGTGG-3'). The last set is specific to "hypothetical protein" gene encoded on *Marinitoga piezophila* KA3 pMP1 plasmid: (forward: 5'-TATCTCGGGAACAGCTCCAA-3' and reverse: 5'-TGACAAAATGGAACATGTTTTTG-3'). The PCR reactions were carried out in a volume of 50 μL containing 100 ng template, 0.4 μM of each primer, 0.8 μM dNTPs, 1.5 mM MgCl_2 buffer and 0.024 U polymerase (Taq Core, Qiagen). The products sizes were 118 bp for 16S rRNA gene, 198 bp for MPV1 major capsid protein gene and 198 bp for pMP1 hypothetical protein gene.

Construction of standard curves for *Marinitoga piezophila* KA3 chromosome, plasmid pMP1 and provirus MPV1 quantifications by QPCR. QPCR analyses were performed to assess the changes in the replication rates, after mitomycin C induction, of both plasmid (by

targeting a gene encoding a hypothetical protein specific to pMP1, Marpi_2124) and viral DNA (by targeting the gene encoding the MPV1 major capsid protein, Marpi_326). In the different experiments with or without induction, the replication rates of the chromosome (by targeting the 16S rRNA gene) were also monitored as a control.

Primers sets are the same as those described above. The PCR products were obtained by conventional PCR and the products sizes were 118 bp for 16S rRNA gene, 198 bp for MPV1 major capsid protein gene and 198 bp for pMP1 hypothetical protein gene. The reactions were performed at 60°C in a volume of 50 µL containing 100 ng of *Marinitoga piezophila* KA3 total DNA, 0.4 µM of each primer, 0.8 µM dNTPs, 1.5 mM MgCl₂ buffer and 2 U Pfu polymerase. The PCR products were deposited on agarose gel, purified by "GeneJet Gel Extraction Kit" (ThermoScientific) and cloned in pUD plasmid used to transform recombinant *E. coli* DH5α strains that were cultivated in LB medium at 37°C with addition of 100 µg/mL of ampicillin. The cloned plasmids were extracted using "GeneJET Plasmid Midiprep" (ThermoScientific) kit and purified as previously described. The concentration of the three different plasmids was measured using a NanoDrop ND-1000 and the corresponding copies numbers were calculated using the following equation: DNA copy = $6.02 \times 10^{23} \times \text{DNA concentrations} / 660 \times \text{DNA length (pUD + insert)}$. 10-fold serial dilutions were performed for the 3 types of cloned plasmids pUD, ranging from 2×10^0 to 2×10^7 copies/µL to construct the standard curves of genes targeting the chromosome and the provirus and ranging from 2×10^0 to 2×10^7 copies/µL, for the gene targeting the cccDNA. The standard curves obtained were up to 0.997 and efficiency of the reactions up to 97%.

Preparation of template DNA for QPCR analysis. Two cultures of *Marinitoga piezophila* KA3 were established using the same inoculum. The first one was induced by addition of mitomycin C (5 µg/mL) in the early exponential growth phase whereas the second culture was not induced. The two cultures were incubated for 80 minutes after induction and then total DNA extractions were performed. The concentration of the extracted DNA was measured by using a Nanodrop ND-1000 and diluted in order to reach 0.2 ng/µL.

QPCR assays by using Sybr Green I dye. The QPCR assays were performed using a StepOnePlus Instrument (Applied Biosystem). The QPCR reactions were performed in a volume of 25 µL containing 1 ng template, 12.5 µL of SybrGreen Supermix (Taq Core, Qiagen) and 0.9 µM of each primers. The thermal cycling protocol was as follows : 10 min at 95°C then 40 cycles of 15 s at 95°C, 1 min at 60°C and 10 s at 72°C.

Sequencing and assembly of bacteriophage and plasmid DNA. The provirus was detected using Prophinder (Lima-Mendez 2008). Homologs of the predicted proteins in both MPV1 bacteriophage and pMP1 plasmid genomes were extracted from the RefSeq and UniProt RefSeq databases using BLASTP (Altschul, 1997) with an E-value cutoff of 10^{-4} . For ORFs with no or few matches in RefSeq database, a local database was also searched. Matches spanning at least 70% of a query were retained. For datasets with > 100 hits, representative per genus was kept. Homologs were aligned using ClustalW 2 (Larkin et al. 2007), and phylogenetic trees were reconstructed in FastTree2 (Price et al. 2009) under JTT+CAT model. Trans-membrane helices, helix-turn-helix motifs, coiled coils and isoelectric point for each ORF were predicted using programs from the EMBOSS (Rice et al., 2000). ORFs were assigned to gene families using Superfamily (Gou et al. 2001) and Conserved Domains (Marchler-Bauer et al. 2009) databases. Predicted ORFs of the MPV1 genome were also screened for the features of the genes involved in known pathways of double-stranded bacteriophage DNA of Gram-negative bacteria (Wang 2000; Young, 2013). Analysis of amino acid composition and predictions of protein charge

and transmembrane helices were performed using the PEPSTATS, CHARGE and T programs of the EMBOSS package (Rice 2000). Protein secondary structure was predicted using JPred 3 (Cole 2008).

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.

Cole C, Barber JD, Barton GJ. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* 36: 197–201.

Gough J, Karplus K, Hughey R, Chothia C. (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J Mol Biol* 313: 903–919.

Larkin M, Blackshields G, Brown N, Chenna R, McGettigan PA, McWilliam M, et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

Lima-Mendez G, Van Helden J, Toussaint A, Leplae R. (2008). Prophinder: a computational tool for prophage prediction in prokaryotic genomes. *Bioinformatics* 24: 863–865.

Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, et al. (2009). CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 37: 205–210.

Price MN, Dehal PS, Arkin AP. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26: 1641–1650.

Rice P, Longden I, Bleasby A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet* 16: 276–277.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* 35: 71–74.

Wang I-N, Smith DL, Young R. (2000). Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* 54: 799–825.

Young R. (2013). Phage lysis: do we have the whole story yet? *Opin Microbiol* 16: 790–797.

Supporting information

Figures and tables

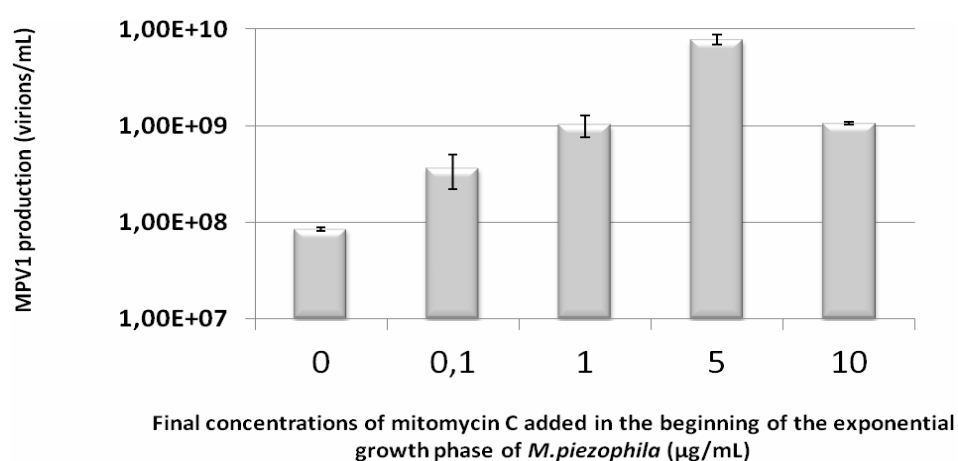


Fig. S1. Induction assays of MPV1 production by addition of several mitomycin C concentrations. Several final concentrations of mitomycin C (0.1 µg/mL, 1 µg/mL, 5 µg/mL and 10 µg/mL) were added to cultures of *M. piezophila* at mid-log growth phase. Aliquots were collected in duplicate to determine viral production by epifluorescence microscopy. The high of bars correspond to the averages of the replicates. Error bars show one standard deviation.

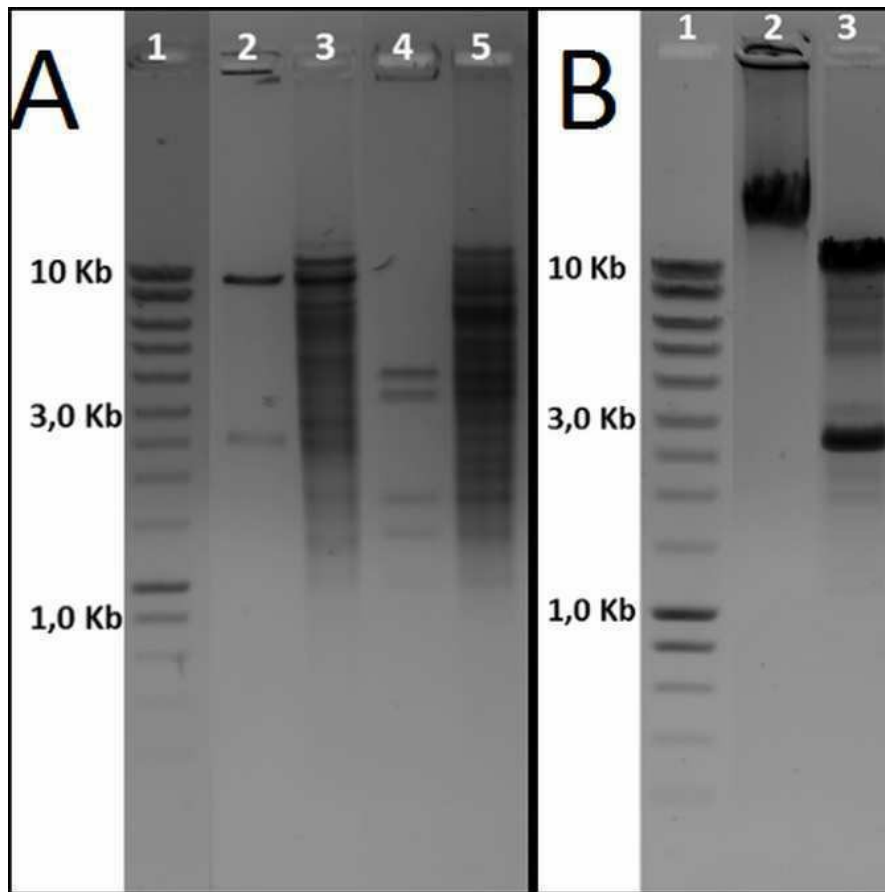


Fig. S2. RFLP analyses of *M. piezophila* total DNA, plasmid DNA (pMP1) and viral DNA.

A. RFLP analyses of both *M. piezophila* total DNA and plasmid DNA (pMP1). 1: DNA Smart Ladder; 2: plasmid DNA restricted by HhaI; 3: total DNA restricted by HhaI; 4: plasmid DNA restricted by HindIII; 5: total DNA restricted by HindIII.

B. RFLP analyses of the total DNA extracted from the purified viral capsids. 1: DNA Smart Ladder; 2: Packaged DNA uncut; 3: Packaged DNA restricted by HhaI. The two thickest bands (of ~10,000 bp and 2,600 bp) correspond to the plasmid DNA restriction pattern (see lane 2 on the panel A). The weaker bands correspond to the viral sequence, since several of them are identical to those obtained after an *in silico* restriction of the *M. piezophila* provirus (not shown).

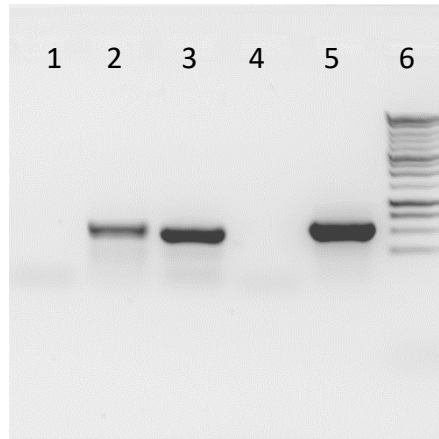


Fig. S3. PCR analysis of the total DNA extracted from purified MPV1 viral capsids from an induced culture. 1: Packaged DNA amplified by 16S rRNA primers; 2: Packaged DNA amplified by pMP1 plasmid primers; 3: Packaged DNA amplified by MPV1 proviral primers; 4: 16S rRNA negative control; 5: 16S rRNA positive control; 6: DNA Smart Ladder.

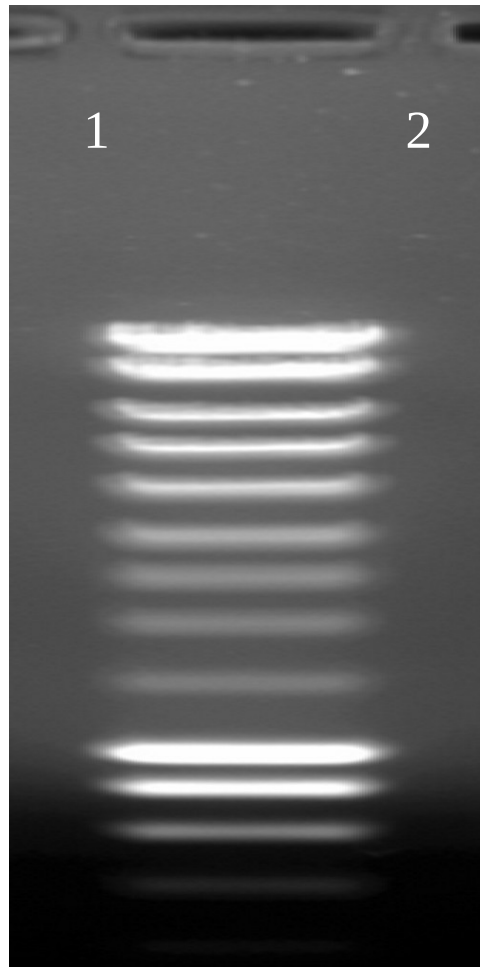


Fig. S4. Agarose gel electrophoresis of plasmid DNA extracted by alkaline lysis from *Thermosipho* sp. strain AT1244-VC14, three subcultures after being infected by MPV1 virions. 1: DNA Smart Ladder; 2: plasmid DNA restricted by HhaI. The two bands indicated by arrows correspond to pMP1 restriction pattern obtained by HhaI (Figure S2 panel A, lane 2).

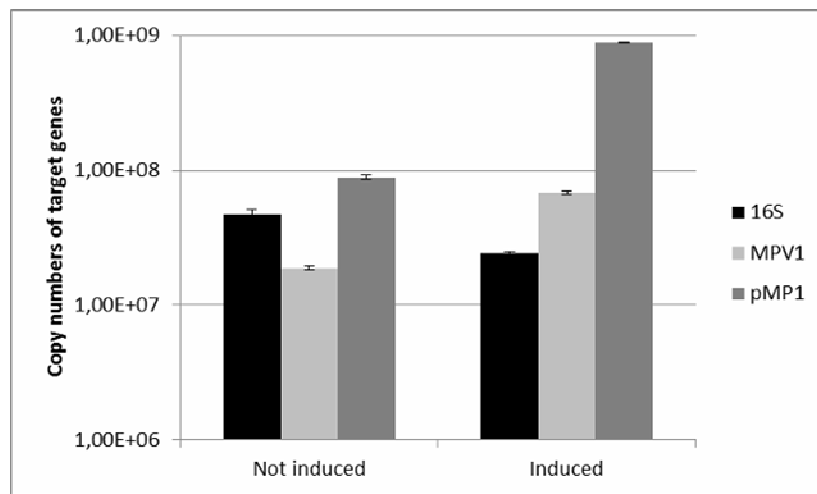


Fig. S5. QPCR assays. Copy number of target genes in the presence or absence of mitomycin C induction. The black bars represent the number of copies of the 16S rRNA chromosomal gene. The light grey bars represent the number of copies of the MPV1 viral gene. The dark grey bars represent the number of copies of the pMP1 plasmid gene.

Table S1 Comparison of viral production at 0.1 and 40 MPa monitored by flow cytometry.

Time	Bacterial counts (cells/mL)		Viral counts (virions/mL)	
	0.1 MPa	40 MPa	0.1 MPa	40 MPa
t = 0h	2.63×10^6	1.41×10^6	9.22×10^5	5.79×10^5
t = 4h	5.13×10^7	4.31×10^7	2.01×10^7	9.14×10^6
t = 9h	3.63×10^8	3.74×10^8	1.10×10^8	3.39×10^7

Table S2. Annotation of the MPV1 provirus.

ORF ^a	length ^b	annotation ^c	BLAST analyses			Protein predictions						
			BlastP match in nr ^d	phage ^e	<i>Thermotogae</i> ^f	Superfamily	CD Acc. ^h	CD name	Tmap	HTH	CC ⁱ	pI
291	1,470	Site-specific recombinases, DNA invertase Pin homologs	WP 006569700 site-specific recombinase, DNA invertase Pin <i>Thermoanaerobacter siderophilus</i>	x	Tmel 0397 (Tmel 1486)	Resolvase-like	cd00338	Ser_Recombinase			1	9.7682
292	162	hypothetical protein	No match									4.3022
293	372	hypothetical protein	YP002728383 hypothetical protein SULAZ 0390 <i>Sulfurihydrogenibium azorense</i> Az-Fu1				pfam14213	DUF4325				10.3206
294	966	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase./STAS domain.	YP004027214 ATP-binding region ATPase domain-containing protein <i>Caldicellulo siruptor kristjanssonii</i>	x		SpolIaa-like	pfam13466	STAS_2	1			4.8478
295	1,800	T5orf172 domain.	YP065302 hypothetical protein DP1566 <i>Desulfotalea psychrophila</i>	x			smart00974	T5orf172	2		5	9.3993
296	1,557	Adenine specific DNA methylase Mod	WP008474462 DNA methylase N-4/N-6 domain protein <i>Nitrolancetus hollandicus</i>	x		S-adenosyl-L-methionine-dependent methyltransferases	pfam01555	N6_N4_Mtase				9.5555
297	621	SOS-response transcriptional repressors (RecA-mediated autopeptidases)	WP006874092 hypothetical protein <i>Anaerotruncus colihominis</i>	x	(Tmel 1479, H17ap60334 04802)	LexA/Signal peptidase	cd00093	HTH_XRE		1		5.8055
298	210	Predicted transcriptional regulators	WP008783175 DNA-binding protein <i>Bifidobacterium</i> sp. 12 1 47BFAA	x	(Tmel 1478, H17ap60334 04807)	lambda repressor-like DNA-binding domains	cd00093	HTH_XRE		1		10.8154
299	450	hypothetical protein, contains possible packing site	No match		(H17ap60334 04812)						1	9.873
300	2,046	hypothetical protein	WP003364581 recombinase RecF <i>Clostridium botulinum</i>	x		P-loop containing nucleoside triphosphate hydrolases	pfam13476	AAA_23			4	5.4559
301	369	Recombinational DNA repair protein (RecE pathway)	WP003407752 recombinase RecT <i>Clostridium butyricum</i>	x	(H17ap60334 04827, Tmel 1472)		pfam03837	RecT		1		5.1063
302	729	Metal-dependent hydrolases of the beta-lactamase superfamily I	YP003781312 phage-like hydrolase <i>Clostridium ljungdahlii</i>	x		Metallo-hydrolase/oxidoreductase "Winged helix" DNA-binding domain^g	cl00446	Lactamase_B superfamily				7.7311
303	768	phage conserved hypothetical protein, C-terminal domain	ADA79920 putative phage replication protein <i>Lactobacillus</i> phage phiPYB5	x			pfam09524	Phg_2220_C				9.4949
304	255	hypothetical protein	No match									8.8039
305	1,275	Replicative DNA helicase	YP001568017 replicative DNA helicase <i>Petrogla mobilis</i>	x	Theba 0328, Pmob 0972, Marpi 1890, Kole 1515, Theth 0376, Kole 1515, THA 1353, THA 826, Tlet 0531, Fnod 0667, Ferpe 0238, Marpi 0443, Theba 0951,	P-loop containing nucleoside triphosphate hydrolases	cd00984	DnaB_C		1		6.0647
306	468	single stranded DNA-binding protein (ssb)	WP016288767 single-stranded DNA-binding protein <i>Lachnospiraceae</i> bacterium 3- 1	x	Theth 1460, Kole 0314, CTN 0058, Pmob 0559, Tlet 0829, Tmel 0637	Nucleic acid-binding proteins	cd04496	SSB_OBF				4.4113
307	420	deoxyuridine 5'-triphosphate nucleotidohydrolase (dut)	XP005142568: deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial <i>Melospitacus undulatus</i>	x		dUTPase-like	cd07557	trimeric_dUTPase	1			9.8363

308	201	hypothetical protein	No match										9.8649
309	249	hypothetical protein	No match					cl06547	Shugoshin_N superfamily	2	1		10.4978
310	366	Holliday junction resolvase	YP003183942 endodeoxyribonuclease RusA <i>Alicyclobacillus acidocaldarius</i>	x		Holliday junction resolvase RusA	pfam05866	RusA					10.1709
311	465	hypothetical protein	No match										4.55
312	273	hypothetical protein	No match										4.5088
313	276	hypothetical protein	No match								1		5.6601
314	231	hypothetical protein	No match			Fzo-like conserved region⁸ Sigma3 and sigma4 domains of RNA polymerase sigma factors					1		8.979
315	483	hypothetical protein	YP005093729 sigma-70 family RNA polymerase sigma factor <i>Marinitoga piezophila</i> KA3 Marpi 2120 (Plasmid)		Marpi 2120, (Tmel 1467, H17ap60334 04862)		TIGR02937	sigma70-ECF		1			9.4095
316	327	hypothetical protein	No match		Tnap 1136								4.303
317	1,029	putative zinc finger/helix-turn-helix protein, YgiT family	YP003841324 helix-turn-helix domain-containing protein <i>Caldicellulosiruptor obsidiansis</i>	x	Pmob 1215 (best match), Tnap 1135	lambda repressor-like DNA-binding domains	cd00093	HTH_XRE	2	1	1		6.6932
318	75	tRNA Pro TGG											
319	381	hypothetical protein	WP004096102 hypothetical protein <i>Acetonebma longum</i> , Among Phi blast matches: WP016077038.1 phage terminase small subunit <i>Lactobacillus</i> phage AQ113			Homeodomain-like	cl17451	HTH_23 superfamily		1			9.9187
320	1,248	phage terminase, large subunit, PBSX family	YP002804814 phage terminase large subunit, pbsx family <i>Clostridium botulinum</i>	x			cl12054	Terminase_3 superfamily					7.066
321	1,251	hypothetical protein	WP021769687 hypothetical protein <i>Leptotrichia</i> sp. oral taxon 22 second hit: WP006773947 phage portal protein <i>Clostridium hathewayi</i>	x			pfam05133	Phage_prot_Gp6	1		1		5.0359
322	1,506	hypothetical protein	WP003432484 hypothetical protein <i>Clostridium butyricum</i>	x			cl01681	DUF2115 superfamily	1		2		9.4814
323	282	hypothetical protein	No match										9.4531
324	603	hypothetical protein	No match	(x)		Prefoldin⁸					2		4.3817
325	354	hypothetical protein	WP017687985 hypothetical protein <i>Paenibacillus</i> sp. PAMC 26794	x			pfam02924	HDPD					4.7569
326	1,020	Phage major capsid protein E.	WP004445976 capsid protein <i>Clostridium botulinum</i>	x			pfam03864	Phage_cap_E					8.7531
327	183	hypothetical protein	No match										10.0744
328	321	hypothetical protein	YP001320641 hypothetical protein Am et 2832 <i>Alkaliphilus metalliredigens</i>										4.38
329	474	hypothetical protein	WP006306523 hypothetical protein <i>Centipeda periodontii</i>	x									10.7346
330	387	hypothetical protein	No match			TRAP-like⁸							4.4597
331	966	hypothetical protein	WP002670872 hypothetical protein <i>Treponema denticola</i>		(Tmel 1452, H17ap60334 04952)				1				4.3066
332	351	hypothetical protein	No match		(Tmel 1451, H17ap60334 04957)								4.7969
333	498	hypothetical protein	No match		(Tmel 1450, H17ap60334 04962)								9.0927
334	4,080	phage tail tape measure protein, TP901 family, core region	YP001285821 putative tail tape measure protein <i>Geobacillus</i> virus E2	x	(Tmel 1449, H17ap60334 04967)	ARM repeat	pfam10145	PhageMin_Tail	9		2		7.0184
335	615	hypothetical protein	No match		Theba 1821								6.9955

336	339	hypothetical protein	No match		(Tmel 1448, H17ap60334 04972)					4.5118	
337	1,131	hypothetical protein	No match		(Tmel 1447, H17ap60334 04977)	Concanavalin A-like lectins/glucanases⁵				4.6324	
338	4,011	hypothetical protein	WP016610917 phage minor structural protein <i>Enterococcus casseliflavus</i>	x	(Tmel 1446, H17ap60334 04982)	Fibronectin type III	cd00063	FN3	2	4.4423	
339	162	hypothetical protein	No match			tRNA-binding arm⁶			1	9.0385	
340	768	DNA repair exonuclease	WP018671617 hypothetical protein <i>Brevibacillus laterosporus</i>		(Tmel 1444) THA 1537, THA 1538	Metallo-dependent phosphatases	cd00838	MPP_superfamily		9.3745	
341	261	hypothetical protein	No match			Bacterial hemolysins	pfam11853	DUF3373	1	6.5311	
342	378	hypothetical protein	No match		(Tmel 1439)	Duffy binding domain-like⁶				9.1701	
343	330	Phage holin protein (Holin LLH).	WP016226830 LL-H family phage holin <i>Lachnospiraceae</i> bacterium 10-1	x			cl09890	Holin_LLH superfamily	1	1	6.2626
344	300	hypothetical protein	No match						1	4.861	
345	564	Orthopoxvirus protein of unknown function (DUF830).	YP002454578 hypothetical protein BC G9842 A0016 <i>Bacillus cereus</i> G9842	x		Cysteine proteinases				5.6907	

a) ORF is number in the genome sequence, i.e. 291 corresponds to Maripi_0291.

b) Length in bp.

c) Annotation in genome.

d) Best BlastP match outside the Thermotogales except for the ORF with match to the plasmid ORF which had no significant matches outside the Thermotogales. We also list additional matches where these added some information.

e) Presence in other phage.

f) Presence in other Thermotogae genomes. If this was in another predicted prophage the ORF id is listed in parenthesis.

g) Weak HMM library classifications, with an E-value greater than 0.0001.

h) CD accession.

i) Number of Coiled Coils.

j) * indicates that no phylogenetic analysis could be done as there were no matches in Genbank. ** indicates that no phylogenetic analysis could be done as there were too few hits or the sequences were too divergent to perform a reliable analysis.

Table S3. Summary of the similarity searches, protein properties' predictions and phylogenetic analyses of pMP1.

ORF ^a	length ^b	annotation ^c	BLAST analyses		Protein predictions						
			BlastP match in nr ^d		Superfamily	CD Acc. ^f	CD name	Tmap	HTH	CC ^g	pI
2120	480	Sigma-70_region_4.	YP_005096064 hypothetical protein Marpi_0315 <i>Marinitoga piezophila</i> KA3 (Prophage)		Sigma3 and sigma4 domains of RNA polymerase sigma factors	cl17315	LuxR_C_like superfamily			0.886	9.80
2121	204	Transcriptional_Coactivator_p15_(PC4)	YP_008403878 hypothetical protein JCM7686_1010 <i>Paracoccus aminophilus</i>		ssDNA-binding transcriptional regulator domain	cl03492	PC4 superfamily				9.63
2122	750	Site-specific_recombinase_XerD	YP_001739834 integrase family protein <i>Thermotoga</i> sp. R		DNA breaking-rejoining enzymes	cl00213	DNA_BRE_C superfamily				10.30
2124	480	hypothetical_protein	No match								4.6
2125	330	hypothetical_protein	No match								5.29
2126	294	hypothetical_protein	No match							0.623	5.71
2127	2148	DNA_primase_(bacterial_type)	YP_004026647 DnaB domain containing protein <i>Caldicellulosiruptor kristjanssonii</i>		Zinc beta-ribbon Sigma2 domain of RNA polymerase sigma factors^e						7.02
2128	1527	hypothetical_protein	No match			cd01029	TOPRIM_primases			0.884	9.93
2129	198	hypothetical_protein	No match						1		10.35
2130	231	hypothetical_protein	No match						1		10.12
2131	603	Growth_inhibitor	YP_002335318 hypothetical protein THA_1536 <i>Thermosipho africanus</i> TCF52B		Cell growth inhibitor/plasmid maintenance toxic component						9.93
2132	522	hypothetical_protein	No match		Ribbon-helix-helix	pfam02452	PemK			2	10.26
2133	552	hypothetical_protein	YP_001410905 hypothetical protein Fnod_1402 <i>Fervidobacterium nodosum</i> Rt17-B1						2		9.4

a) ORF is number in the genome sequence, i.e. 2120 corresponds to Maripi_2120.

b) Length in bp.

c) Annotation in genome.

d) Best BlastP match.

e) Weak HMM library classifications, with an E-value greater than 0.0001.

f) CD accession.

g) Coiled Coils probability.

h) * indicates that no phylogenetic analysis could be done as there were no matches in Genbank. ** indicates that no phylogenetic analysis could be done as there were too few hits or the sequences were too divergent to

Chapitre III : Deux nouveaux virus,
MCV1 et MCV2, isolés bactéries du
phylum des Thermotogae

Présentation des travaux

Lors du criblage des souches de *Thermotogae* ayant pour objectif la recherche de potentiels éléments génétiques mobiles leur étant associés, deux nouveaux virus ont été identifiés. Thèse de J. Lossouarnas. Dans le cadre de la thèse, j'ai effectué la caractérisation fonctionnelle et génomique de ces virus isolés de deux souches de *Marinitoga camini* (*M. camini*), la souche DV1197 porteuse du bactériovirus MCV1 (*M. camini* virus 1) et la souche DV1155 porteuse de MCV2 (*M. camini* virus 2). Ces virus ont été obtenus par le processus de lysogénie et sont intégrés au sein des chromosomes de ces deux hôtes. Les génomes des souches de *Marinitoga camini* DV1197 et DV1155 ont été séquencés, faisant ainsi l'objet d'un Genome Announcement en 2016 (voir p. 130).

MCV1 et MCV2 sont deux siphovirus tempérés inductibles à la mitomycine C. Ils ont des capsides icosaédriques de 70 nm de diamètre environ et de longues queues flexibles de 260 nm pour MCV1 et 270 nm pour MCV2 pour une largeur de 10 nm.

Les génomes des deux virus sont à ADN double brin, d'une longueur de 48,5 kb pour MCV1 et 50,7 kb pour MCV2. Ils contiennent respectivement 77 et 70 ORFs prédits. Ces génomes viraux ont été comparés au génome de MPV1, qui infecte aussi une bactérie du genre *Marinitoga*. Ces virus sont retrouvés dans le même type d'environnement (sources hydrothermales profondes) mais ont été isolés de souches provenant de lieux géographiquement éloignés (Océans Pacifique et Atlantique). Néanmoins, les comparaisons réalisées entre les génomes ont permis de constater qu'ils partageaient de nombreuses similitudes tant au niveau de l'organisation génomique qu'au niveau des ORFs eux-mêmes, et ce malgré la distance géographique les séparant. Ils partagent aussi des similarités avec les Firmicutes ainsi que les virus les infectant.

Des essais d'induction au cadmium ont aussi été réalisés sur la souche potat MCV, M. camini DV 97. Le but était d'étudier les possibilités d'induction virales via des stress physico-chimiques et l'impact de la souche. Néanmoins, l'induction n'a pas entraîné la production de particules virales mais celle de vésicules membranaires (MVs). En effet, l'ajout de cadmium ou de nickel, à des concentrations de 100 µg/ml, a provoqué la production de MVs de différentes tailles apparaissant souvent avec un contenu opaque en MET. Ces essais restent préliminaires, mais il est possible que cette production de MVs soit un mécanisme de détoxification des métaux lourds pour la cellule.

L'ensemble de ces travaux fait l'objet d'une publication scientifique en préparation qui sera prochainement soumise au journal « Environmental Microbiology ».

Draft Genome Sequences of Two *Marinitoga camini* Isolates Producing Bacteriophages

Coraline Mercier^a, Julien Lossouarn^{a*}, Thomas Haverkamp^a, Nadège Bienvenu^a, Anne Godfroy^b,
Valérie Cueff-Gauchard^c, Claire Geslin^d, Camilla Nesbo^e

^aUBO, Laboratoire de Microbiologie des Environnements Extrêmes, UMR 6197, IUEM, Rue Dumont d'Urville, 17048 Plouzané, France; ^bUMR 6197, IUEM, Rue Dumont d'Urville, 17048 Plouzané, France; ^cUMR 6197, IUEM, Rue Dumont d'Urville, 17048 Plouzané, France; ^dCentre for Ecological and Evolutionary Synthesis, University of Oxford, Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

* Present address: Julien Lossouarn, Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France.

Here, we present the draft genome sequences of two thermophilic *Marinitoga* strain members of the *Thermotoga* genus. *Marinitoga camini* DV1155 and *Marinitoga camini* DV1197. These strains were isolated from deep-sea hydrothermal vents of the Mid-Atlantic Ridge.

Received 15 September 2016; Accepted 21 September 2016; Published 10 November 2016

Citation: Mercier C, Lossouarn J, Haverkamp T, Bienvenu N, Godfroy A, Cueff-Gauchard V, Geslin C, Nesbo C. 2016. Draft genome sequences of two *Marinitoga* species producing bacteriophages. *Genome Announc* 4(6):e01261-16. doi:10.1128/genomeA.01261-16.

Copyright: © 2016 Mercier et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Address correspondence to Claire Geslin, claire.geslin@univ-brest.fr.

Bacteria from the *Marinitoga* genus are thermophilic, anaerobic, and organotrophic microorganisms isolated from deep-sea hydrothermal vents. The genome of *Marinitoga* DV1155 was obtained using Prophinder (11). For *M. camini* DV1197, we obtained 51 contigs totaling 2,274,557 bp with a GC content of 27.4%. No thermal springs. Among the five *Marinitoga* species described to date, only the *M. piezophila* KA3 genome has been sequenced. Here, we present the draft genome sequences of two novel *Marinitoga* isolates, DV1155 and DV1197, both containing both draft genomes were annotated in the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) (12).

The two strains were sampled during the DIVA 2 cruise in 1992. *M. camini* DV1155 was isolated from a black smoker chimney at the Menez Gwen site. This hydrothermal field is located on the Mid-Atlantic Ridge at a depth of 840 to 870 m. *M. camini* DV1197 was isolated from a colonization module deployed at the Lucky strike hydrothermal vent field. This site is located at a depth of 1700 m in the Mid-Atlantic Ridge. Strains DV1197 and DV1155 and *M. camini* DV1197, respectively. The versions described in this paper are the first versions of *M. camini* DV1155 and DV1197. Both whole-genome shotgun projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers AZAY000000000 and AZAY000000000. The versions described in this paper are the first versions of *M. camini* DV1155 and DV1197.

Genomic DNA was extracted following the protocol of Geslin et al. (10). The purity and quantity of the DNA were measured using Nanodrop and Qubit instruments (Thermo, Fisher Scientific). Shotgun libraries were constructed using the Nextera XT kit and sequenced as one of ten pooled, barcoded libraries on a MiSeq (all from Illumina) using 500 cycles paired-end reads. The genomes assembled *de novo* by CLC Genomics Workbench 7.0.4 using trimming default settings, automatic word size, a bubble size corresponding to the average input reads, a minimum contig length of 1000 bp and reads mapped back to the contigs.

For *M. camini* DV1155 this resulted in 56 contigs totaling 2,435,399 bp with an N_0 of 90,885 bp, longest contig size of 254.99 bp and GC content of 27.3%. No extrachromosomal DNA was observed but a proviral sequence of 50,700 bp was detected using the same method as for *M. camini* DV1155. Here, we present the draft genome sequences of two novel *Marinitoga* isolates, DV1155 and DV1197, both containing both draft genomes were annotated in the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) (12).

ACKNOWLEDGMENTS We thank Daniel Desbruyères, chief scientist of the French oceanographic vessel "DIVA 2" (1994), the captain and crew of the *Nautile* and support crew. FUNDING INFORMATION This work, including the efforts of Coraline Mercier, was funded by the French Ministry of Higher Education. This work, including the efforts of Coraline Mercier, was funded by Brittany Region. This work, including the efforts of Camilla Lothe Nesbo, was funded by the Norwegian Research Council (award 180444/V40). This work, including the efforts of Claire Geslin, was funded by Agence Nationale de la Recherche (ANR) (ANR-12-BSV3-0023-01).

REFERENCES Wery N, Lesongeur F, Fignat B, Derennes C, Cambon-Bonavita MA, Godfroy A, Barbier G. 2001. *Marinitoga camini* gen. nov., sp. nov., a

- rod-shaped bacterium belonging to the order *Thermotogales* isolated from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* 51:495–504. <http://dx.doi.org/10.1099/00207713-51-2-495>.
2. **Alain M, Marteinsson W, Eroshnichenko M, Much-Osmolovskaya EA, Prieur D, Birrien J-L** 2002 *Marinitoga piezophila* sp. nov., a rod-shaped thermo-piezophilic bacterium isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* 52:1331–1339. <http://dx.doi.org/10.1099/00207713-52-4-1331>.
 3. **Nunoura T, Oida H, Miyazaki M, Suzuki Y, Taka K, Horikoshi K** 2007. *Marinitoga okinawensis* sp. nov., a novel thermophilic and anaerobic heterotroph isolated from a deep-sea hydrothermal field, Southern nawa trough. *Int J Syst Evol Microbiol* 57:467–471. <http://dx.doi.org/10.1099/ijs.0.64640-0>.
 4. **Postec A, Le Breton C, Fardeau M-L, Lesongeur F, Pignet P, Jollivet J, Ollivier B, Godfroy A** 2005 *Marinitoga hydrogenitolerans* sp. nov., a novel member of the order *Thermotogales* isolated from a black smoker chimney on the Mid-Atlantic Ridge. *Int J Syst Evol Microbiol* 55:1217–1221. <http://dx.doi.org/10.1099/ijs.0.63550-0>.
 5. **Postec A, Ciobanu M, Birrien JL, Bienvenu N, Prieur D, Le Romancer M** 2010. *Marinitoga litoralis* sp. nov., a thermophilic, heterotrophic bacterium isolated from a coastal thermal spring on Ile Saint-Paul, southern Indian Ocean. *Int J Syst Evol Microbiol* 60:1778–1782. <http://dx.doi.org/10.1099/ijs.0.017574-0>.
 6. **Lucas S, Han J, Lapidus A, Cheng J-F, Goodwin LA, Pitluck S, Peters L, Mikhailova N, Teshima H, Detter JC, Han C, Tapia R, Land M, Huuhtanen L, Kyrpides NC, Ivanova N, Pagani I, Vannier P, Oger P, Barton AC, Nothmann N, Woyke T, Jebbar M** 2011. Complete genome sequence of *Marinitoga piezophila* KA3. *J Bacteriol* 193:5974–5975. <http://dx.doi.org/10.1128/JB.01430-11>.
 7. **Desbruyeres D, Alayse A-M, Antoine E, Barbier G, Barriga F, Biscoll M, Briand P, Brulport J-P, Comtet T, Cornec L, Crassous P, Dando Marie-claire F, Felbeck H, Lallier FH, Fiala-Medioni A, Gonçalves M, Ménard F, Kerdoncuff JP, Saldanha L, Pierre-Marie S** 1994. Information on the ecology of deep-sea vent communities in the Azores triple junction area: preliminary results of the Diva 2 cruise. *InterRidge News* 3:18–19.
 8. **Colaço A, Desbruyeres D, Comtet T, Alayse A** 1998. Ecology of the Menez Gwen hydrothermal vent field (Mid-Atlantic Ridge/Azores triple junction). *Cah Biol Mar* 39:237–240.
 9. **Lossouarn N, Nesbø CL, Mercier C, Zhaxybayeva D, Johnson MS, Charchuck R, Farasin J, Bienvenu N, Baudoux A-C, Michoud G, Jollivet J, Geslin C** 2015. “Ménage à trois”: a selfish genetic element uses a virus to propagate within *Thermotogales*. *Genes* 6:173278–3288. <http://dx.doi.org/10.1111/1462-2920.12783>.
 10. **Geslin C, Le Romancer M, Mauso G, Gaillard M, Perrot G, Grier D** 2003. PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote *Pyrococcus abyssi*. *J Bacteriol* 185:3888–3894. <http://dx.doi.org/10.1128/JB.185.13.3888-3894.2003>.
 11. **Lima-Mendes C, Van Helden JT, Toussain A, Lepoint R** 2008. Prophinder: A computational tool for prophage prediction in prokaryotic genomes. *Bioinformatics* 24:863–865. <http://dx.doi.org/10.1093/bioinformatics/btn043>.
 12. **Amend DA, Gussman A, Klimke W, Cochrane G, Field D, Garrity GM, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thornton N, White O** 2008. Toward an online repository of standard operating procedures (SOPs) for (meta) genomic annotation. *Genome Biol* 9:R1237–141. <http://dx.doi.org/10.1089/omi.2008.0017>.

Functional and genomic overview of two new viruses, MCV1 and MCV2, which infect *Marinitoga* bacteria isolated from deep-sea hydrothermal vents.

C. Mercier^{a,b,c}, J. Lossouarn^{a,b,c*}, C.L. Nesbø^{d,e}, T.H.A Haverkamp^d, A.C. Baudoux^f, M. Jebbar^{a,b,c}, N. Bienvenu^{a,b,c}, S. Thiroux^{a,b,c}, S. Dupont^{a,b,c}, C. Geslin^{a,b,c}

a UBO, Laboratoire de Microbiologie des Environnements Extrêmes UMR6197, IUEM, Rue Dumont d'Urville, 29280 Plouzané, France

b CNRS, Laboratoire de Microbiologie des Environnements Extrêmes UMR 6197, IUEM, Rue Dumont d'Urville, 29280 Plouzané, France

c Ifremer, Laboratoire de Microbiologie des Environnements Extrêmes UMR 6197, CS10070, 29280 Plouzané, France

d Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway

e University of Oslo and University of Alberta, Edmonton AB Canada

f Sorbonne Universités, UPMC Univ Pierre et Marie Curie (Paris 06), CNRS, Adaptation et Diversité en Milieu Marin UMR7144, Station Biologique de Roscoff, 29680 Roscoff, France

Correspondence : Dr Claire Geslin

Université de Bretagne Occidentale, CNRS, Ifremer, UMR 6197-Laboratoire de Microbiologie des Environnements Extrêmes (LM2E), Institut Universitaire Européen de la Mer (IUEM), Technopôle Brest-Iroise, BP70, 29 280 Plouzané, France

e-mail : claire.geslin@univ-brest.fr

Tél : +33 (0)2 98 49 88 58

Fax : +33 (0)2 98 49 87 05

Running title: Viruses in *Thermotogae*

*Current address: J. Lossouarn: Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

Abstract

Viruses represent a driving force in the evolution of microorganisms including those thriving in extreme environments. However, our knowledge of the viral diversity associated to microorganisms inhabiting the deep-sea hydrothermal vents remains limited. The phylum of *Thermotogae*, including thermophilic bacteria, is well represented in this environment. Only one virus was described in this phylum, MPV1 carried by *Marinitoga piezophila*. In this study, we report on the functional and genomic characterization of two new bacteriophages that infect bacteria from the *Marinitoga* genus. *Marinitoga camini* virus 1 and 2 (MCV1 and MCV2) are temperate siphoviruses with a linear dsDNA genome of 53.4 kb and 50.5 kb respectively. Here we present a comparative genomic analysis of the MCV1 and MCV2 viral genomes with that of MPV1. The results indicate that even if the host strains come from geographically distant sites, their genomes share numerous similarities. Interestingly, heavy metals did not induce viral production instead the host of MCV1 produce membrane vesicles. This study highlights interaction of mobile genetic elements (MGE) with their hosts and the importance of including hosts-MGEs' relationships in microbiological studies.

Key words : Hydrothermal deep-sea vent/ *Marinitoga*/ thermophile/ siphovirus/ membrane vesicles

Introduction

Deep sea hydrothermal vents are one of the most extreme environments on Earth. They are minerals edifices with massive sulphide deposits and high concentration of various metals (e.g. Cu, Cd, Ni, Zn, Mn, Co, Ag, Au) (Auguste *et al.*, 2016). These ecosystems are characterized by steep physicochemical gradients, high hydrostatic pressures, very high to low temperatures, obscurity and prevalence of microbial chemosynthesis (Flores and Reysenbach, 2011). They are home to a vast diversity of psychrophiles, mesophiles, hyper/thermophiles, acidophiles, piezophiles and even moderate halophiles microorganisms belonging to the Bacteria and Archaea domains (Jebbar *et al.*, 2015). The virosphere associated with these deep-sea microorganisms remains largely unexplored; to date only 8 viruses have been characterized (Lossouarn *et al.*, 2015a). Two of these viruses infect *Thermococcales*, one of the largest archaeal orders that comprises hyperthermophilic and chemoorganoheterotrophic microorganisms. These are non-virulent lemon-shaped archaeoviruses, and were isolated from *Pyrococcus abyssi* (PAV1) and *Thermococcus prieurii* (TPV1) (Geslin *et al.*, 2003a,b; Geslin *et al.*, 2007; Gorlas *et al.*, 2012). The remaining six are head and tailed bacteriophages. Four of these, BVW1, GVE1, GVE2 and D6E, are virulent (i.e. they replicate through a lytic cycle) and infect thermophilic and heterotrophic bacteria belonging to the *Bacillus* and *Geobacillus* genera (Liu *et al.*, 2006; Liu and Zhang, 2008; Wang and Zhang, 2010). The other two are temperate siphoviruses (i.e. they replicate through a lysogenic cycle), NrS1 infects a thermophilic ϵ proteobacterium belonging to the *Nitratiruptor* genus (Yoshida-Takashima *et al.*, 2013) and MPV1 infects *Marinitoga piezophila*, a bacterium from the *Petrotogales* order and *Thermotogae* phylum (Lossouarn *et al.*, 2015a,b).

The phylum *Thermotogae* is composed of anaerobic, chemo-organotrophic, sulfur reducing thermophilic and hyperthermophilic bacteria, as well as a newly characterized mesophilic genus (Gupta and Bhandari, 2011; Nesbø *et al.*, 2012; Zhaxybayeva *et al.*, 2012; Pollo *et al.*, 2015).

The *Thermotogae* are rod-shaped and characterized by an atypical outer sheath-like envelope, the toga (Kuwabara and Igarashi, 2012). They are found in hot ecosystems all over the world, including hydrothermal vents, thermal springs, oil reservoirs and more recently they have also been isolated from lower temperature hydrocarbon-impacted sites, such as polluted sediments (Pollo *et al.*, 2015).

In universal phylogenetic trees based on 16S rRNA sequences, the *Thermotogae* branch deep close to *Aquificae* (Hug *et al.*, 2016). However, this localization is only supported by a small subset of their genes (e.g. ribosomal protein encoding genes), while the majority of their genes place them with *Firmicutes* and with *Archaea*, especially *Thermococcales* (Zhaxybayeva *et al.*, 2009; Gupta and Bhandari, 2011; Bhandari and Gupta, 2015). Comparative genomic analyses of *Thermotogae* have revealed complex evolutionary histories with extensive horizontal gene transfer (HGT) (Zhaxybayeva *et al.*, 2009; Nesbo *et al.*, 2015; Pollo *et al.*, 2015). Similarly to many thermophilic microorganisms, clustered regularly interspaced short palindromic repeats (CRISPRs) are present in numerous *Thermotogae* genome sequences (Zhaxybayeva *et al.*, 2009; Nesbo *et al.*, 2015; Haverkamp *et al.*, unpublished), suggesting that they were exposed to viral infections during the course of their evolutionary history. However, only one virus has been described in this phylum, MPV1 carried by *Marinitoga piezophila* KA3 (Alain *et al.*, 2002; Lossouarn *et al.*, 2015b). MPV1 is a temperate siphovirus with a genome of 43.7 kb matching the provirus integrated in *M. piezophila* KA3 chromosome (Lucas *et al.*, 2012). This host carries a second mobile genetic element, a plasmid of 13.3 kb (referred as pMP1). Remarkably, pMP1 uses the MPV1 viral capsid to propagate, highlighting a complex evolutionary relationship between a bacterial host, an extrachromosomal element and a virus (Lossouarn *et al.*, 2015b).

Here, we report on the functional and genomic characterization of two new bacteriophages that infect bacteria from the *Marinitoga* genus. *Marinitoga camini* virus 1 (MCV1) and *Marinitoga*

camini virus 2 (MCV2) are siphoviruses that infect *Marinitoga camini* DV1197 and *Marinitoga camini* DV1155, respectively (Mercier *et al.*, 2016). These bacteria are thermophiles and were isolated from deep-sea hydrothermal vents in the Mid-Atlantic Ridge. We further present the results of comparative analysis of these two complete viral genomes (MCV1 and MCV2) with the viral genome of MPV1 infecting *M. piezophila* (Lossouarn *et al.*, 2015b).

Results

Virions morphology.

MCV1 virions produced by *M. camini* 97 have a head with a hexagonal outline of ~50 nm in diameter, a flexible non-contractile tail of ~200 nm in length and ~10 nm in width (Fig. 1 A). MCV2 virions produced by *M. camini* 55 have a head with a hexagonal outline of ~60 nm in diameter, a flexible non-contractile tail of ~230 nm in length and ~10 nm in width (Fig. 1 B). Based on these morphological characteristics, we classified both virions in the *Siphoviridae* family from the order *Caudovirales*.

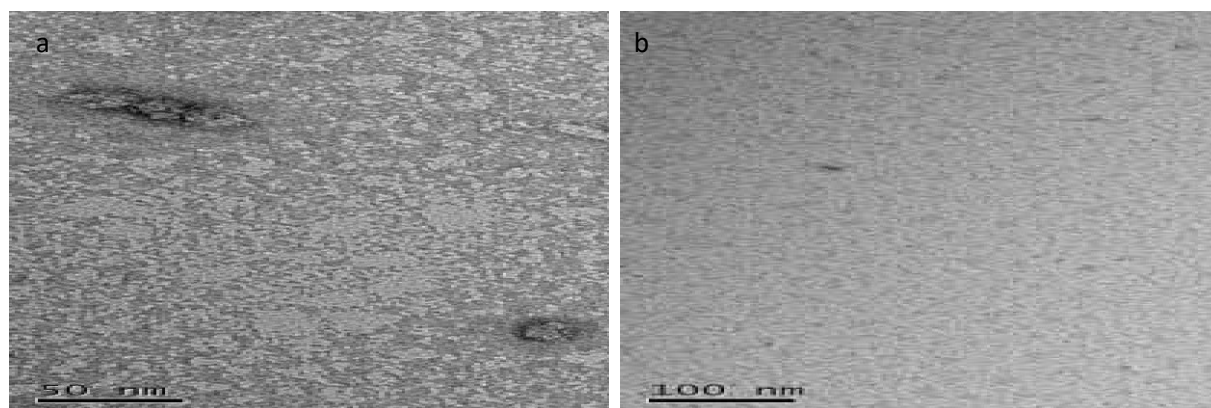


Figure 1: Electron micrographs of (a) a MCV1 virion and (b) a MCV2 virion, stained with 2% uranyl acetate.

Virus-host relationships and induction assays.

The basal production of MCV1, i.e. the virion production without noticeable lysis of host cells, peaked at the same time as the host. After 16 hours of cultures, maximal concentrations of MCV1 and *M. camini* 97 were observed, with 5.10^7 virions mL^{-1} and 10^8 cells mL^{-1} ,

respectively. After induction, the viral production drastically increased, reaching a peak of $\sim 10^8$ virions mL^{-1} , 4h after mitomycin addition. A concomitant drop in host abundance to $\sim 5 \cdot 10^6$ cells mL^{-1} was observed, indicating that host cell lysis occurred upon viral production (Fig. 2. A). Thus, induced cells produce 20-fold more virions than untreated culture. Regrowth of bacteria was detected 8h after induction ($2 \cdot 10^6$ to 10^7 cells mL^{-1}) suggesting that some cells could withstand the mitomycin C treatment.

As observed for MCV1, the spontaneous MCV2 production in untreated cells was around $5 \cdot 10^7$ viral particles mL^{-1} after 8 hours of growth when the cellular production reached a peak to $\sim 7 \cdot 10^7$ cells mL^{-1} . Four hours after induction, the viral production increased, reaching $\sim 5 \cdot 10^9$ viral particles mL^{-1} , which is 100-fold more virions than for untreated cells. A concomitant drastic drop to $\sim 10^7$ cells mL^{-1} was observed in host cell counts, suggesting host cell lysis. Weak regrowth of cells was observed 6 hours after induction, probably due to some mitomycin C resistant cells (Fig. 2 B).

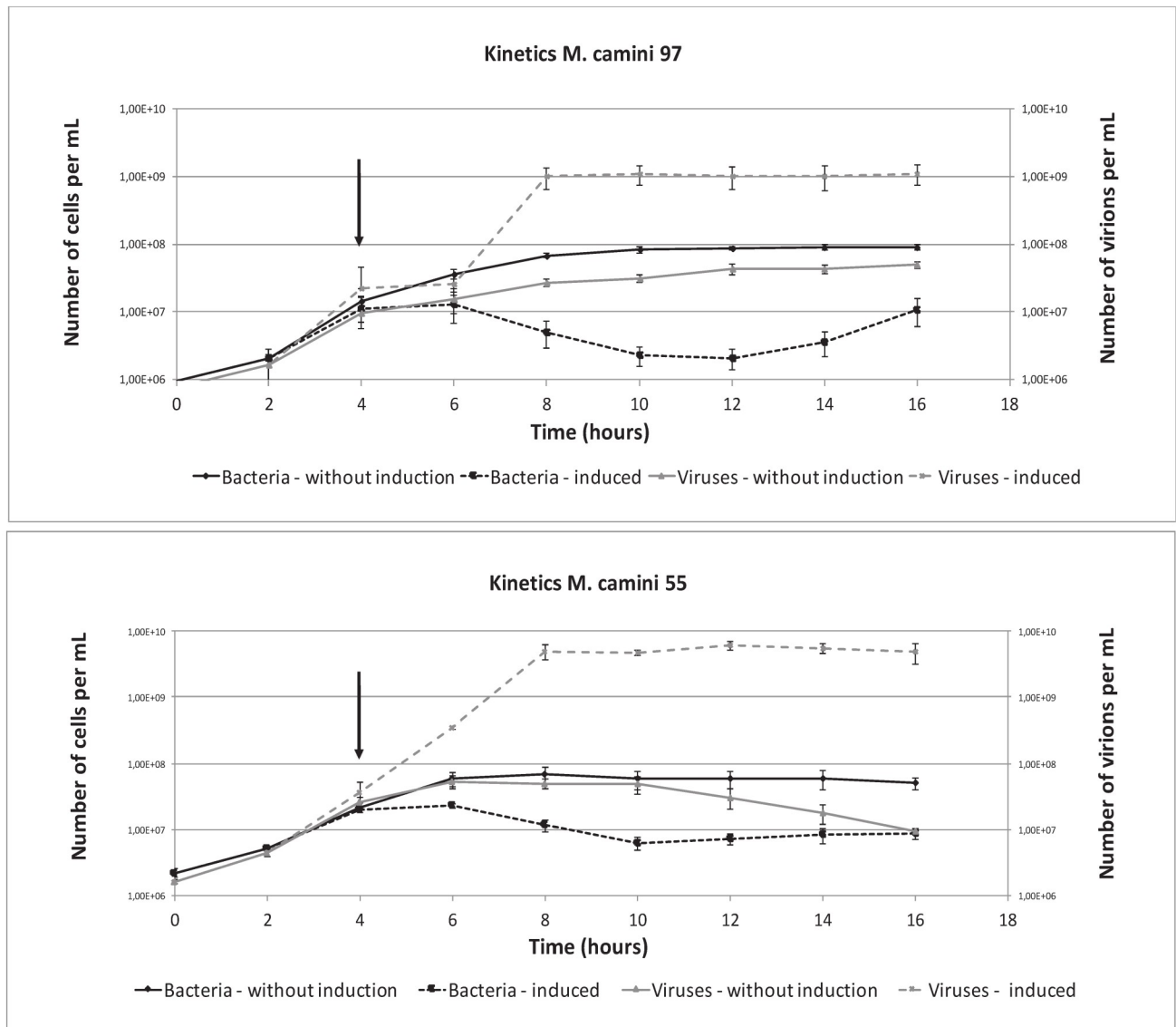


Figure 2: Virus-host kinetics of *M. camini* 97 and *M. camini* 55 monitored by flow cytometry.

Typical bacterial growth is shown as continuous black line whereas bacterial growth upon induction with mitomycin C is shown as discontinuous black line. Basal viral production is shown as continuous grey line (reaching $\sim 5.10^7$ virions/ml for both viruses) whereas viral production upon induction is shown as discontinuous grey line (reaching, 4 hours after induction, $\sim 10^9$ virions/ml for MCV1 and $\sim 6.10^9$ virions/ml for MCV2). Induction is indicated with an arrow.

In order to check the potential viral induction with environmental parameters, heavy metals (copper, cobalt, zinc and cadmium) were added to culture. Surprisingly, viral production was not induced but membrane vesicles (MVs) were observed. After addition of copper or cadmium (0.5 mM) in an early log culture of *M. camini* 97, TEM examination revealed a production of abundant pleomorphic and heterogenous MVs with sizes ranging from 50 to 250 nm (Fig. 3). Double membranes were observed on the majority of MVs pointing their cellular provenance. The high MV's production observed (with a very low concomitant viral production) suggested that this may be a way for cells to survive metals exposure. At high concentration of metals (up to 1mM of copper or cadmium), cell lysis was noticed and less MVs were observed.

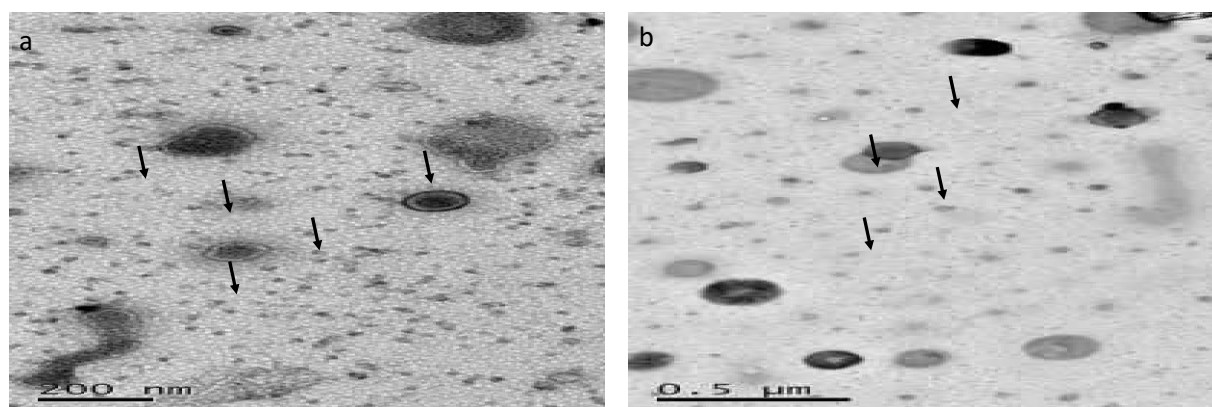


Figure 3: Electron micrographs of concentrated membrane vesicles, indicated with an arrow, produced by *M. camini* 97 (a) after adding 0,1mM of copper (b) or 0,5mM of cadmium, stained with 2% uranyl acetate.

Attempts to infect a putative sensitive host.

Purified MCV1 and MCV2 virions were added to strains belonging to *Marinitoga* and *Thermosipho* genera (*Marinitoga camini* strain DV1075, *Marinitoga okinawensis*, *Marinitoga hydrogenotolerens* and *Thermosipho* sp. strain AT1244-VC14). After the addition of virions no lysis was observed under a phase-contrast microscope for these potential hosts. Potential infection was checked by PCR using primers targeting the major capsid protein encoding gene.

Thus, infectivity has not been shown, neither for MCV1 nor for MCV2, although phylogenetically related strains were tested.

High temperature stability of virions.

Viral particles stability was checked by adding virions in a culture medium, without bacterial host, and incubated at 60°C. The results show that the stability of the virions is low; in less than 5 hours of incubation, the flow cytometry signal is lost (i.e loss of 22%.h⁻¹). With their hosts, the number of virions is stable and forms a stationary phase during at least 8 hours (see Fig. 2 A and B). This suggests that with hosts present there is equilibrium between the continuously viral production and viral decay.

DNA content of MCV1 and MCV2 and sequences analysis.

The genome sequence of *M. camini* 97 contains a predicted provirus sequence of ~53,400 bp integrated adjacent to the tRNA-glu gene. Analysis of DNA extracted from the purified MCV1 viral capsids confirmed that the provirus is the integrated form of the MCV1 genome (Fig. S1a). The MCV1 viral genome, assembled as a linear dsDNA sequence of 53,412 bp, has 77 predicted ORFs, including 21 annotated genes. These predicted genes covered 90% of the genome sequence. The majority of the putative genes have the same orientation while 22 are encoded on the complementary strand. The overall G + C content of the MCV1 genome (28.1%) is similar to the overall G + C content of the host genome (27.4%).

For *M. camini* 55, the bacterial chromosome contains a predicted proviral sequence of ~50,300bp also integrated adjacent to the tRNA-glu gene. Sequencing of viral DNA from purified capsids confirmed that the packaged viral DNA has the same sequence as the predicted provirus with a similar size (Fig. S1b). MCV2 has a genome length of 50,307 bp, with 70 predicted ORFs, including 17 annotated genes. These predicted genes covered 88.6 % of the genome sequence. The majority of the putative genes have the same orientation with 22 present

on the complementary strand. The overall G + C content of the MCV2 genome (28.7%) is comparable with the overall G + C content of the host genome (27.3%).

Results from BLASTN and BLASTP comparisons of the predicted ORFs and proteins from the two viruses to each other and sequences in Genbank are listed in [Supplemental Tables 2 and 3](#).

MCV1 and MCV2 appear to contain all genes necessary for viral particles synthesis. Several genes involved in DNA metabolism were identified, including, a replicative DNA helicase, single stranded DNA binding protein, recombinational DNA repair protein RecT, holliday junction resolvase and another resolvase. For the head and tail assembly, classical proteins like terminase large subunit, major capsid protein and tail tape measure protein (a chaperon protein allowing the tail assembly) were identified. Finally, for the lytic and lysogenic cycles, holin protein, anti-repressor proteins, transcriptional regulators (six total, including one from XRE family), SOS response repressor and protease *lexA* were identified.

Fifty-one ORFs in each virus had a BLASTN match in pairwise comparisons (expected value $< e-0.5$). Twenty-four ORFs in each of the virus genomes have no match in Genbank, outside the *Marinitoga* genus, when using an expected value cut-off at 0.05 (Supplemental Table 2 and 3). The MCV1 and MCV2 genomes also contain, respectively, 21 and 17 unique ORFs not found in the other viral genome nor in the MPV1 genome (Fig. 4, Supplemental Tables 2 and 3). The majority of the genes specific to one of the viruses encode hypothetical proteins, and these viral specific ORFs tend to be found in clusters of three or more genes in the same relative location in the two genomes (Fig. 4). Interestingly, one of the MCV1 specific genes, UF08_51 encoding a hypothetical protein, has no match in the MCV2 virus, but a match in the *Marinitoga* sp. 55 genome (MCV2's host) outside the *proV204_0211*). The viral specific proteins with predicted functions mainly encode proteins with different DNA binding domains (e.g. HTH- and zinc ribbons domains) and or typically mobile genes (e.g. toxin-antitoxin genes).

For both MCV1 and MCV2, most *Thermotogae* matches outside the genus *Marinitoga*, were to genes from *Thermosipho* genomes (N=13 and N=14, respectively). When comparing the two genomes to sequences outside the *Thermotogae*, 37% were affiliated with *Firmicutes*, 18%-20% with *Proteobacteria* and 6%-7% with *Archaea* (only *Euryarchaea*).

Comparing the predicted proteins from the two viruses to viral proteins in Genbank, revealed that 28 MCV1 ORFs and 24 MCV2 had one or more significant matches (expected value < 0.05). The majority of these bacterioviral homologues belong to siphoviruses and myoviruses isolated from *Firmicutes* phylum (*Bacilli* and *Clostridia* classes) but also from *Thermus-Deinococcus* and *Proteobacteria* phyla. Among these viruses, four were isolated from extreme environments, *Geobacillus* bacteriovirus D6E and *Epsilonproteobacteria* bacteriovirus Nsr1 from hydrothermal deep-sea vents in the Pacific, *Thermus* bacteriovirus P23-77 from alkaline hot spring in New Zealand and *Thermus* bacteriovirus phi OH2 isolated from a hot spring in Japan. The other bacterioviral homologues were isolated from diverse types of samples as soil and even from fermented foods.

Comparative genomics of 3 viruses infecting 3 strains and 2 species of Marinitoga.

A genome comparison of MCV1 and MCV2 to the previously described MPV1 bacteriovirus that infects *M. piezophila* KA3 is shown in Figure 4, which shows that three genomes have similar gene synteny, with module organization similar to that seen in various siphoviruses (Yasmin *et al.*, 2010). The modules are, 5' to 3', replication and transcription, packing, head and tail assembly, host lysis and finally lysogeny.

The genome sequence of MPV1 is considerably shorter (10 kb) than the two others, and has a total of 54 predicted ORFs, compared to 77 ORFs in MCV1 and 70 ORFs in MCV2. A set of conserved “core genes” shared by all the three viruses was identified and includes 35 genes among them are proteins for DNA metabolism like replicative DNA helicase and holliday junction resolvase. Proteins involved in the head and tail assembly are also present in the three

genomes, e.g. terminase large subunit or major capsid protein. Several hypothetical proteins are found in all three genomes, suggesting that these proteins probably provide important functions for these viruses (Lossouarn *et al.*, 2015b, Supplemental Tables 2 and 3). Both MCV1 and MCV2 have large clusters of unique genes (9 genes each) in the 3' genome, in their putative lysogeny modules. Similarly, in the lysogeny module, MPV1 has six genes not found in MCV1 and MCV2, among them a DNA methylase and a histidine kinase. In these three viral genomes the lysogeny modules appeared relatively more variable.

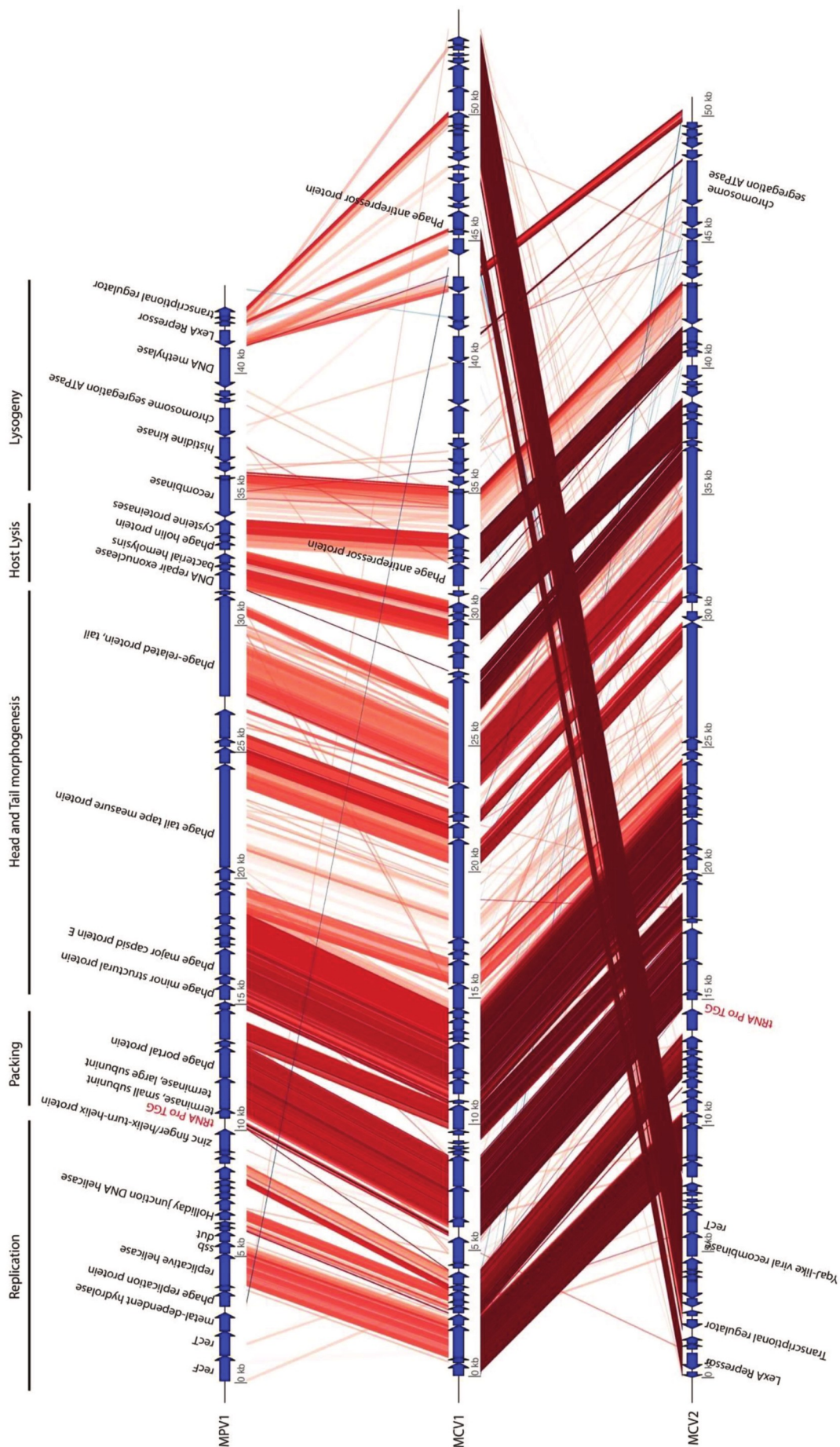


Figure 4: Genome comparison of MPV1, MCV1 and MCV2. The genomes were compared using tblastx the resulting percent identity values were plotted in GenoplottR (Guy *et al.*, 2010). Direct matches are colored in red hues, and the reversed ones in blue hues. The division of the virus into functional modules is approximate, as it is based only on gene annotations.

Discussion

The *Thermotogae* comprise a widespread phylum of mostly thermophilic or hyperthermophilic bacteria. Placed deep in the SSU rRNA based tree of life, their phylogeny is complex owing to numerous HTG-events in their evolutionary history. An important HGT vector is viruses, nevertheless, only one virus, MPV1 isolated from *Marinitoga piezophila*, has been characterized from this phylum (Lossouarn *et al.*, 2015b). The characterization of MCV1 and MCV2, also infecting *Marinitoga* strains, adds to our knowledge on the viral diversity in deep-sea hydrothermal vents.

MCV1 and MCV2 were isolated from *M. camini* 97 and *M. camini* 55, respectively, which were obtained from deep-sea hydrothermal vents in the Mid Atlantic Ridge. *M. camini* 97 was isolated from a colonization module deployed at the Lucky strike hydrothermal vent, located at 1700m of depth. *M. camini* 55 was extracted from a black smoker chimney at the Menez Gwen site, located at a depth of 840-870m (Colaço *et al.*, 1998). MPV1 was isolated from *Marinitoga piezophila* KA3, a piezophilic strain isolated from a deep-sea hydrothermal chimney sample collected at a depth of 2630 m on the East-Pacific Rise (13S N). Though these host strains come from geographically distant hydrothermal environments, they colonize the same type of niches with equivalent physical/chemical qualities, which allow them a similar metabolism based on anaerobiosis, chemoorganotrophy and thermophily.

In a recent study it was shown that, for the hyperthermophilic *Thermotogae*, similarity of environment is more important for gene exchanges than the geographical proximity (Nesbo *et al.*, 2015). The authors compared 11 *Thermotoga maritima*-like genomes, isolated from two environment types (oil reservoirs and marine sites). They reported that the 'same type of environment tended to be more similar, and exchanged more genes with each other than with geographically close isolates from different types of environments' (Nesbo *et al.*, 2015). Another recent study on *Thermosipho* spp. provides additional support for this

hypothesis underlining the importance of gene flow for the maintenance of distinct lineage and species (Haverkamp *et al.*, unpublished). Furthermore, one likely vector for such exchanges could be viruses.

MCV1 and MCV2 are temperate siphoviruses containing double stranded DNA similar to the previously described MPV1 (Lossouarn *et al.*, 2015b). A set of “core genes” shared by all three viruses was identified and includes proteins involved in DNA metabolism, head and tail assembly and lytic and lysogenic cycles. An ORF (UFO08_9 in MCV1 and UFO09_17 in MCV2) is a part of the core genes and is also found in some *Thermosipho* spp. genomes (putative proviruses) (Haverkamp *et al.*, unpublished). It also shares similarities with a *Marinitoga piezophila* ORF (outside proviral sequence) encoding for a sigma factor 70. This ORF could be an example of HGT mediated by viruses.

Among the core genes, shared hypothetical proteins were also identified, suggesting that these unknown proteins probably provide important functions for these viruses and represent a reservoir of new genes. Some of these hypothetical genes could represent virus-encoded auxiliary metabolic genes (VAMGs). These types of specialized viral genes, with a cellular origin, are however, likely to move back to their hosts where they could be involved in the microbial adaptation or evolution in these extreme environments as observed in aquatic ecosystems. Aquatic viral populations serve as gene reservoirs that contribute to niche partitioning of microbial species (Sime-Ngando, 2014). Moreover, some of VAMGs could be found in viromes (for example in the Pacific Ocean Virome, POV), improving host fitness, as genes involved in DNA repair and replication or chemotaxis and motility (Hurwitz *et al.*, 2015). Thereby, hypothetical proteins present in the core genes of the *Marinitoga* viruses, could represent such auxiliary genes useful to help hosts fitness in extreme environmental conditions. Interestingly, for genes with BLASTP matches in Genbank, over 60% have their top match, outside the *Thermotogae*, to genes from *Firmicutes* and bacteriophages associated to

Firmicutes (Supplemental Table 2 and 3). Among those viruses, are phi OH2 and D6E infecting *Firmicutes* isolated from extreme environment such as hot springs and hydrothermal vents (Wang and Zhang, 2010; Doi *et al.*, 2013). *Thermotogae* and *Firmicutes* appear to have exchanged large number of genes (Zhaxybayeva *et al.*, 2009), and our findings suggest that mobile genetic elements may have mediated some of these genes exchanges.

It is notable that most of the genes specific to each of the three viruses are found in the lysogeny module which contains genes expressed during the part of their lifecycle when they integrate into, or are integrated in, the host genome. This suggests that acquiring new genes probably adds a selective advantage for the virus, perhaps making it better at controlling or interacting with the host genome. The impact of lysogeny on genes dispersal and cell physiology is probably underestimated. Temperate viruses are often considered as a pool of genetic elements implicated in microbial genomes evolution and adaptation, but their ecological implication is so far not well understood (Paul, 2008; Knowles *et al.*, 2016; Thingstad and Bratbak, 2016).

For viruses, being incorporated into the host genome appears to be a strategy to survive harsh environmental conditions. Under laboratory conditions, at high temperature, the viral decay for MCV1 virions is near 22%.h⁻¹. Dell Anno *et al.* showed that in deep-sea ecosystems an important fraction of viruses (about 25%), once they are released by cell lysis, undergo fast decomposition (Dell'Anno *et al.*, 2015). They observed that viral decay is faster in deep-sea ecosystem and that the stability of virions is lower than in other marine ecosystems. In comparison, in marine environments in the absence of new production, virions decay presents an average half-life of 48h (Wommack and Colwell, 2000). It is therefore likely that it is helpful for the extremophile virus to stay incorporated in the host genome rather than to shift to a lytic cycle where virions will have a very short time to infect another host.

An interesting feature of the host-viruses kinetics for all three *Marinitoga* viruses are high spontaneous viral productions (more than 10⁷ virions ml⁻¹) which may be due to the fluctuant

and extreme physicochemical conditions they encounter. This was also described for another bacteriophage Nrs-1, infecting a thermophilic *εProteobacterium* from the same type of environment (Yoshida-Takashima *et al.*, 2013), where virions appeared to be released continuously and spontaneously, probably resulting from stochasticity in gene expression or a low induction of the hosts SOS response. Spontaneous provirus induction, i.e. virions' production, could thus be advantageous for both lysogenic hosts and their proviruses. On the one hand, a portion of non-lysogenized bacteria, potentially competitive, could be killed and on the other hand survivors, undergoing lysogenic conversion, could be able to spread the viral DNA. This spontaneous virions release, could also have an important ecological impact by releasing organic matter, which may contribute to the nutrient cycling in hydrothermal deep-sea vents.

Environmentally relevant stressors (i.e. high concentrations of heavy metals) did not induce the production of virions but we detected an increase in membrane vesicles (MVs) production. The production of MVs is a universal physiological phenomenon shared by cells from all three domains of life (Kulp and Kuehn, 2010; Deatherage and Cookson, 2012). MVs are involved in numerous physiological processes such as cell-to-cell communications, nucleic acid or toxins transport, stress response and interactions with viruses (Soler *et al.*, 2008; Manning and Kuehn, 2011; Gaudin *et al.*, 2013; Biller *et al.*, 2014). For *Thermococcales* it was demonstrated that MVs could be a defense system for cells by preventing toxic accumulation of sulfur (Gorlas *et al.*, 2015). Similarly, several studies on eukaryotes and bacteria have shown that membrane vesicles can play a role in detoxification (Sterling *et al.*, 2007; Jagannadham and Chattopadhyay, 2015; Slavic *et al.*, 2016). The phenomenon was first observed in eukaryotic marine organisms such as mollusks and crustaceans, which accumulate cadmium and where excretion of cadmium is performed by MVs as a detoxifying mechanism (Sterling *et al.*, 2007). The observation of increased MVs production with addition of copper or cadmium suggests

that the MVs produced by *Marinitoga* could be involved in detoxification. The fact that environments where these organisms thrive, hydrothermal vents, are characterized by the presence of high levels of heavy metals (cadmium, copper, zinc...) supports this notion. Further studies will be needed in order to test if MVs can protect the cells against heavy metals toxicity. Moreover, since these strains produce both MVs and viruses they could serve as models to investigate mechanisms involved in virus/MVs/hosts' interaction. Globally, viruses and other mobile genetic elements (MGEs) have a strong impact on microbial ecology, especially in oceans. The analyses presented here extend our knowledge about these newly discovered mobile genetic elements in the deeply branching bacterial phylum *Thermotogae*. This bacterial order and associated MGEs are significant for addressing long-term evolutionary adaptation to fluctuant and extreme physicochemical conditions. Our work confirms a strong evolutionary connection of *Thermotogae* viruses to *Firmicutes* and their bacteriophages. The impact of MGEs on their hosts is very important and it remains crucial to study new host-MGEs' relationships in an ecological context to well understand the reasons of this old and deep co-evolution.

Experimental procedures

Screening of Thermotogae strains for mobile genetic elements

A total of 55 strains of *Thermotogae* were retrieved from the UBO culture collection 'UBOCC' (<http://www.univ-brest.fr/UBOCC>) and from our laboratory culture collections. These strains were isolated from different deep-sea vents in the Lau Basin, the East Pacific Rise and the Middle Atlantic Ridge and were screened for mobile genetic elements as previously described (Lossouarn *et al.*, 2015b). Two of them, *Marinitoga camini* DV1197 and *Marinitoga camini* DV1155, were discovered to be the hosts of viruses MCV1 and MCV2 respectively.

Culture conditions.

Marinitoga camini DV1197 (hereafter referred to as *M. camini* 97) was cultured at 60°C in a modified Ravot medium as previously described (Geslin *et al.*, 2003a,b; Lossouarn *et al.*, 2015b). The medium was completed by adding 1% (wt/vol) elemental sulfur. *Marinitoga camini* DV1155 (hereafter referred to as *M. camini* 55) was cultured in the same medium at 65°C.

Viral induction assays.

Attempts were made to increase the viral production by induction using an antibiotic treatment that damage DNA (mitomycin C) as well as metal stresses (Llanos *et al.*, 2000; Yoshida-Takashima *et al.*, 2013). A concentration of 4µg/mL of mitomycin C was used as described in Lossouarn *et al.*, 2015b. Induction assays by metals were done on 25 mL of *M. camini* 97 culture in early log-phase, by adding various metals (copper, cadmium, zinc, cobalt) with different concentrations (1 to 0.001 mM for copper and cadmium and 0.1 to 0.001 mM for zinc and cobalt). After 24h incubation period, debris and cells were pelleted by centrifugation at 7500g, 4°C for 20 min and discarded. Supernatants were ultracentrifuged at 80,000g, 10°C for 1h (Beckman Optima LE-80 K 70.1 Ti rotor). Pellets were resuspended in 50µL of buffer (10 mM Tris-HCL, 100 mM NaCl, 5 mM CaCl₂, 20 mM MgCl₂) and observed using a JEOL JEM 100 CX II® transmission electron microscope as described previously (Geslin *et al.*, 2003a,b; Gorlas *et al.*, 2012).

Flow cytometry counts of bacteria and viruses.

Samples were fixed with 25% glutaraldehyde (0.5% final concentration, Electron microscopy (EM) grade, Sigma-Aldrich) for 15 min at 4°C, flash frozen in liquid nitrogen and stored at -80°C until analysis. The thawed samples were diluted from 100 to 10,000-fold in autoclaved 0.2 µm filtered TE buffer (10:1 Tris-EDTA, pH 8.0) and stained with the nucleic acid-specific dye SYBR Green I (Life technologies®) for 15 min at room temperature for bacteria and 10 min at 80°C for viruses as described in Brussaard (2004). Samples were analyzed using a

FACSCanto II flow cytometer (BD Bioscience®) equipped with a laser with an excitation wavelength of 488 nm (15 mW). The trigger was set on green fluorescence. The samples were delivered at a flow rate of 30 or 50 $\mu\text{L}\cdot\text{min}^{-1}$ and analyzed during at least 1 min. TE-buffer with autoclaved 0.2 μm filtered seawater was used as a control.

Viral concentration and purification.

Five hundred milliliters of *M. camini* 97 and 55 were induced by addition of mitomycin C (4 $\mu\text{g}/\text{mL}$) in early log-phase. Three hours after induction, debris and cells were pelleted by centrifugation at 7,500 g, 4°C for 15 min and discarded. The supernatant was ultracentrifuged at 80,000 g, 10°C for 1 h (Beckman Optima LE-80 K 70.1 Ti rotor®). The viral pellet was resuspended in 1 mL of buffer (10 mM Tris-HCL, 100 mM NaCl, 5 mM CaCl_2 , 20 mM MgCl_2). The viral suspension was purified by ultracentrifugation in a linear iodixanol gradient [OptiPrep®, 25–40% diluted in a buffer (see above)] at 200,000 g and 10°C for 10 h (Beckman Optima LE-80 K SW 41 rotor®). After ultracentrifugation, an opaque virus band was recovered and stored at 4°C until use. To exclude the possibility of cellular nucleic acids contamination, the concentrated viral particles suspensions were treated with exo-DNase RQ1 and RNase (Promega®, 2 U/ml each) for 45 min at 37°C, followed by addition of DNase-stop and a heat treatment (10 min at 70°C). The viral particles were examined using a JEOL JEM 100 CX II® transmission electron microscope as previously described (Geslin *et al.*, 2003a,b; Gorlas *et al.*, 2012).

Attempts to infect a putative sensitive host.

Marinitoga and *Thermosipho* spp. isolates from our culture collection were screened for the presence of virions and strains scored negatively were selected as potential hosts. As a result, *Thermosipho* sp. AT1244-VC14, *Marinitoga camini* DV1075, *Marinitoga hydrogenotolerans* and *Marinitoga okinawensis* were tested. Different levels of multiplicity of infection (MOI), of purified MCV1 and MCV2, were tested (50, 30, 10, 1, 0.1, 0.001 and 0.0001 virions per cell).

Non-infected cultures were used as negative control. After 8 subcultures, total DNA was extracted from each subculture. PCR amplifications were performed on total DNA as described in supporting information by using primers that target the major capsid protein encoding gene of MCV1 and 16S rRNA gene (Table S1).

High temperature stability of virions.

Purified virions of MCV1 (3.108 virions/mL) were added in modified Ravot medium and incubated during 22 hours at 60°C. 500µl were sampled, in triplicate, every 2 hours (0h to 8h) and two supplementary samples were obtained at 20 and 22 hours. Each sample was flash frozen in liquid nitrogen and fixed with glutaraldehyde and stored at -80°C, until flow cytometry analysis as described before.

Extraction of cellular and viral DNA.

Total DNA from *Thermosipho sp.* AT1244-VC14, *M. camini* 97 and *M. camini* 55 cells and viral DNA from purified virions (MCV1 and MCV2) were prepared as described previously (Geslin *et al.*, 2003a,b; Gorlas *et al.*, 2012). In order to obtain viral genomes, we eliminated extracapsid DNA contamination, by treating the purified viral suspensions from induced cultures with DNases and RNase (DNase RQ1 2U.ml⁻¹, ExoIII 2U.ml⁻¹, RNase A 5mg.ml⁻¹) before DNA extraction.

Sequencing and assembly of bacteriophages DNA.

Purified viral DNA of MCV1 and MCV2 were sequenced as barcoded libraries on a MiSeq using 500 cycles generating 2x250 bp paired-end reads, using the Nextera XT kit for library construction. The sequences were assembled using the CLC Genomics Workbench 6.5.1 (<http://www.clcbio.com>). Sequence analyses of protein-coding genes of bacteriophages (MCV1 and MCV2) are described in Supporting information.

Acknowledgments

This work was financially supported by the French Ministry of Higher Education and the Brittany region to CM and the Agence Nationale de la Recherche, project Thermovesicles (ANR-12-BSV3-0023-01) to CG and JL, a Norwegian Research Council award (project no. 180444/V40) to CLN. We thank G. Siquin and P. Elies for assistance in electron microscopy.

Conflict of interest

The authors declare no conflict of interest.

References

- Alain, K., Marteinsson, V.T., Miroshnichenko, M., Bonch-osmolovskaya, E., Prieur, D., and Birrien, J. (2002) *Marinitoga piezophila* sp. nov., a rod-shaped, thermo-piezophilic bacterium isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int. J. Syst. Evol. Microbiol.* 52: 1331–1339.
- Auguste, M., Mestre, N.C., Rocha, T.L., Cardoso, C., Cueff-Gauchard, V., Le Bloa, S., *et al.* (2016) Development of an ecotoxicological protocol for the deep-sea fauna using the hydrothermal vent shrimp *Rimicaris exoculata*. *Aquat. Toxicol.* 175: 277–285.
- Bhandari, V. and Gupta, R.S. (2015) Erratum to: Molecular signatures for the phylum (class) *Thermotogae* and a proposal for its division into three orders (*Thermotogales*, *Kosmotogales* ord. nov. and *Petrotogales* ord. nov.) containing four families (*Thermotogaceae*, *Fervidobacteriaceae* fam. nov., *Kosmotogaceae* fam. nov. and *Petrotogaceae* fam. nov.) and a new genus *Pseudothermotoga* gen. nov. with five new combinations. *Antonie Van Leeuwenhoek* 108: 1281.
- Biller, S.J., Schubotz, F., Roggensack, S.E., Thompson, A.W., Summons, R.E., and Chisholm, S.W. (2014) Bacterial vesicles in marine ecosystems. *Science*. 343: 183–186.
- Brussaard, C. P. (2004). Optimization of procedures for counting viruses by flow cytometry. *Applied and Environmental Microbiology*, 70: 1506-1513.

- Colaco, A., Desbruyeres, D., Comtet, T., & Alayse, A. M. (1998). Ecology of the Menez Gwen hydrothermal vent field (Mid-Atlantic Ridge Azores Triple Junction). *Cahiers de Biologie Marine*, 39: 237-240.
- Deatherage, B.L. and Cookson, B.T. (2012) Membrane vesicle release in Bacteria, Eukaryotes, and Archaea: a conserved yet underappreciated aspect of microbial life. *Infect. Immun.* 80: 1948–1957.
- Dell'Anno, A., Corinaldesi, C., and Danovaro, R. (2015) Virus decomposition provides an important contribution to benthic deep-sea ecosystem functioning. *Proc. Natl. Acad. Sci.* 112: E2014–E2019.
- Doi, K., Mori, K., Martono, H., Nagayoshi, Y., Fujino, Y., Tashiro, K., *et al.* (2013) Draft genome sequence of *Geobacillus kaustophilus* GBlys, a lysogenic strain with bacteriophage ϕ OH2. *Genome Announc.* 1: e00634-13.
- Flores, G.E. and Reysenbach, A.-L. (2011) Hydrothermal environments, marine. In, *Encyclopedia of Geobiology*. Springer, pp. 456–467.
- Gaudin, M., Gaudiard, E., Schouten, S., Houel-Renault, L., Lenormand, P., Marguet, E., and Forterre, P. (2013) Hyperthermophilic archaea produce membrane vesicles that can transfer DNA. *Environ. Microbiol. Rep.* 5: 109–16.
- Geslin, C., Gaillard, M., Flament, D., Rouault, K., Le Romancer, M., Prieur, D., and Erauso, G. (2007) Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from *Pyrococcus abyssi*. *J. Bacteriol.* 189: 4510–4519.
- Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003a) PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, "*Pyrococcus abyssi*." *J Bacteriol* 185: 3888–3894.
- Geslin, C., Le Romancer, M., Gaillard, M., Erauso, G., and Prieur, D. (2003b) Observation of virus-like particles in high temperature enrichment cultures from deep-sea hydrothermal

- vents. *Res. Microbiol.* 154: 303–307.
- Gorlas, A., Koonin, E. V, Bienvenu, N., Prieur, D., and Geslin, C. (2012) TPV1, the first virus isolated from the hyperthermophilic genus *Thermococcus*. *Environ. Microbiol.* 14: 503 – 516.
- Gorlas, A., Marguet, E., Gill, S., Geslin, C., Guigner, J.M., Guyot, F., and Forterre, P. (2015) Sulfur vesicles from *Thermococcales*: A possible role in sulfur detoxifying mechanisms. *Biochimie* 118: 356–364.
- Gupta, R.S. and Bhandari, V. (2011) Phylogeny and molecular signatures for the phylum *Thermotogae* and its subgroups. *Antonie Van Leeuwenhoek* 100: 1–34.
- Guy, L., Kultima, J.R., Andersson, S.G.E. (2010) genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* 26 : 2334–2335.
- Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., *et al.* (2016) A new view of the tree of life. *Nat. Microbiol.* 1: 16048.
- Hurwitz, B.L., Brum, J.R., and Sullivan, M.B. (2015) Depth-stratified functional and taxonomic niche specialization in the “core” and “flexible” Pacific Ocean microbiome. *ISME J.* 9: 472 – 484.
- Jagannadham, M. V and Chattopadhyay, M.K. (2015) Role of outer membrane vesicles of bacteria. *Resonance* 20: 711–725.
- Jebbar, M., Franzetti, B., Girard, E., and Oger, P. (2015) Microbial diversity and adaptation to high hydrostatic pressure in deep-sea hydrothermal vents prokaryotes. *Extremophiles* 19: 721–740.
- Knowles, B., Silveira, C.B., Bailey, B.A., Barott, K., Cantu, V.A., Cobián-Güemes, A.G., *et al.* (2016) Lytic to temperate switching of viral communities. *Nature* 531: 466–470.
- Kulp, A. and Kuehn, M.J. (2010) Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annu. Rev. Microbiol.* 163–184.

- Kuwabara, T. and Igarashi, K. (2012) Microscopic studies on *Thermosipho globiformans* implicate a role of the large periplasm of *Thermotogales*. *Extremophiles* 16: 863–870.
- Liu, B., Wu, S., Song, Q., Zhang, X., and Xie, L. (2006) Two novel bacteriophages of thermophilic bacteria isolated from deep-sea hydrothermal fields. *Curr. Microbiol.* 53: 163–166.
- Liu, B. and Zhang, X. (2008) Deep-sea thermophilic *Geobacillus* bacteriophage GVE2 transcriptional profile and proteomic characterization of virions. *Appl. Microbiol. Biotechnol.* 80: 697–707.
- Llanos, J., Capasso, C., Parisi, E., Prieur, D., and Jeanthon, C. (2000) Susceptibility to heavy metals and cadmium accumulation in aerobic and anaerobic thermophilic microorganisms isolated from deep-sea hydrothermal vents. *Curr. Microbiol.* 41: 201–205.
- Lossouarn, J., Dupont, S., Gorlas, A., Mercier, C., Bienvenu, N., Marguet, E., *et al.* (2015a) An abyssal mobilome: Viruses, plasmids and vesicles from deep-sea hydrothermal vents. *Res. Microbiol.* 166: 742–752.
- Lossouarn, J., Nesbø, C.L., Mercier, C., Zhaxybayeva, O., Johnson, M.S., Charchuck, R., *et al.* (2015b) “Ménage à trois”: a selfish genetic element uses a virus to propagate in *Thermotogales*. *Environ. Microbiol.* 17: 3278–3288.
- Lucas, S., Han, J., Lapidus, A., Cheng, J.-F., Goodwin, L.A., Pitluck, S., *et al.* (2012) Complete genome sequence of the thermophilic, piezophilic, heterotrophic bacterium *Marinitoga piezophila* KA3. *J. Bacteriol.* 194: 5974–5975.
- Manning, A.J. and Kuehn, M.J. (2011) Contribution of bacterial outer membrane vesicles to innate bacterial defense. *BMC Microbiol.* 11: 258.
- Mercier, C., Lossouarn, J., Haverkamp, T., Bienvenu, N., Godfroy, A., Cueff-Gauchard, V., *et al.* (2016) Draft genome sequences of two *Marinitoga camini* isolates producing bacteriophages. *Genome announc.* 4: 6.

- Nesbø, C.L., Bradnan, D.M., Adebisuyi, A., Dlutek, M., Petrus, A.K., Foght, J., *et al.* (2012) *Mesotoga prima* gen. nov., sp. nov., the first described mesophilic species of the *Thermotogales*. *Extremophiles* 16: 387–393.
- Nesbo, C.L., S Swithers, K., Dahle, H., Haverkamp, T.H.A., Birkeland, N.-K., Sokolova, T., *et al.* (2015) Evidence for extensive gene flow and *Thermotoga* subpopulations in subsurface and marine environments. *ISME J.* 9: 1532–1542.
- Paul, J.H. (2008) Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J.* 2: 579–589.
- Pollo, S.M.J., Zhaxybayeva, O., and Nesbø, C.L. (2015) Insights into thermoadaptation and the evolution of mesophily from the bacterial phylum *Thermotogae*. *Can. J. Microbiol.* 61: 655–670.
- Sime-Ngando, T. (2014) Environmental bacteriophages: viruses of microbes in aquatic ecosystems. *Front. Microbiol.* 5: 355.
- Slavic, K., Krishna, S., Lahree, A., Bouyer, G., Hanson, K.K., Vera, I., *et al.* (2016) A vacuolar iron-transporter homologue acts as a detoxifier in *Plasmodium*. *Nat. Commun.* 7.
- Soler, N., Marguet, E., Verbavatz, J.M., and Forterre, P. (2008) Virus-like vesicles and extracellular DNA produced by hyperthermophilic archaea of the order *Thermococcales*. *Res. Microbiol.* 159: 390–399.
- Sterling, K.M., Mandal, P.K., Roggenbeck, B.A., Ahearn, S.E., Gerencser, G.A., and Ahearn, G.A. (2007) Heavy metal detoxification in crustacean epithelial lysosomes: role of anions in the compartmentalization process. *J. Exp. Biol.* 210: 3484–3493.
- Thingstad, T.F. and Bratbak, G. (2016) Microbial oceanography: Viral strategies at sea. *Nature* 531: 454–455.
- Wang, Y. and Zhang, X. (2010) Genome analysis of deep-sea thermophilic phage D6E. *Appl. Environ. Microbiol.* 76: 7861–7866.

- Wommack, K.E. and Colwell, R.R. (2000) Virioplankton: Viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64: 69–114.
- Yasmin, A., Kenny, J.G., Shankar, J., Darby, A.C., Hall, N., Edwards, C., and Horsburgh, M.J. (2010) Comparative genomics and transduction potential of *Enterococcus faecalis* temperate bacteriophages. *J. Bacteriol.* 192: 1122–1130.
- Yoshida-Takashima, Y., Takaki, Y., Shimamura, S., Nunoura, T., and Takai, K. (2013) Genome sequence of a novel deep-sea vent epsilonproteobacterial phage provides new insight into the co-evolution of *Epsilonproteobacteria* and their phages. *Extremophiles* 17: 405–419.
- Zhaxybayeva, O., Swithers, K.S., Foght, J., Green, A.G., Bruce, D., Detter, C., *et al.* (2012) Genome sequence of the mesophilic *Thermotogales* bacterium *Mesotoga prima* MesG1.Ag.4.2 reveals the largest *Thermotogales* genome to date. *Genome Biol. Evol.* 4: 700–708.
- Zhaxybayeva, O., Swithers, K.S., Lapierre, P., Fournier, G.P., Bickhart, D.M., DeBoy, R.T., *et al.* (2009) On the chimeric nature, thermophilic origin, and phylogenetic placement of the *Thermotogales*. *Proc. Natl. Acad. Sci. U. S. A.* 106: 5865–5870.

Supporting Information

Figure S1: Comparison of viral genomes isolated from capsids and provirus integrated in the host genomes. The sequences were aligned using MAUVE 2.3.1 and the comparison was plotted using Genoplots (Guy *et al.*, 2010) a: MCV1 and *M. camini* 97 and b: MCV2 and *M. camini* 55.

Table S1: Primers targeting ARNr 16S, MCV1 and MCV2 major capsid proteins.

Table S2: Annotations and comparison tables of MCV1.

Table S3: Annotations and comparison tables of MCV2.

Supporting informations

Supplementary Material and Methods

Sequence assembly and analysis

The MCV1 and MCV2 genomes were assembled de novo by CLC Genomics Workbench 7.0.4, using trimming default settings, automatic word size, a bubble size corresponding to the average length of the input reads, a minimum contig length of 1000 bp, and reads mapped back to the contigs. Coding sequences (CDS) were identified and annotated in the Geneious 9 (www.geneious.com). BLASTP searches against Genbanks none redundant database (nr) and pairwise BLASTN searches to compare the three viruses (MPV1, MCV1 and MCV2) were also performed in Geneious 9 (www.geneious.com). Blast searches against virus sequences in Genbank were done at NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) limiting the search to Viruses (taxid:10239). Conserved domain (CD) searches were also done at NCBI. Both Genome sequences have been deposited at DDBJ/EMBL/GenBank under the accession no. KP836355 and KP836356 for MCV1 and MCV2, respectively.

Table S1: Primers targeting ARNr 16S and MCV1 & MCV2 major capsid proteins

Target			Forward/reverse	Sequence
Bacterial 16S rRNA gene			Forward	AGAGTTTGATCATGGCTCAG
Bacterial 16S rRNA gene			Reverse	GGTTACCTTGTTACGACT T
Major	Capsid	Protein	Forward	AGCTCAGCCAACAAAAGCAG
MCV1&MCV2				
Major	Capsid	Protein	Reverse	AGAAGGAATTTGGGCTGCTG
MCV1&MCV2				

locus tag	annotation	Length	Best match in MCV22	% Identity	Match in MPV1	% identity2	expected	best match outside the Marinitoga in BLASP searches of Genbank
UF08_1	hypothetical protein	169	UF09_6_hypothetical protein	95%				WP_042729055 phage replication protein
UF08_2	hypothetical protein	70	UF09_7_hypothetical protein	99%				no match
UF08_3	replicative DNA helicase	427	UF09_8_replicative DNA helicase	98%	Marpi_0305 replicative DNA helicase	66%	0	WP_031484309 replicative DNA helicase
UF08_4	single-stranded DNA-binding protein	142	UF09_9_single-stranded DNA-binding protein	97%	Marpi_0443 single stranded DNA-binding protein	57%	2,34E-52	WP_052670309 hypothetical protein
UF08_5	hypothetical protein	92	UF09_13_hypothetical protein	99%	Marpi_0308 hypothetical protein	47%	5,06E-08	WP_015616411 hypothetical protein
UF08_6	hypothetical protein	82	UF09_14_hypothetical protein	100%	Marpi_0309 hypothetical protein	53%	6,58E-15	XP_008855251 hypothetical protein EF022120
UF08_7	Holliday junction resolvase RusA	122	UF09_15_hypothetical protein	94%	Marpi_0310 Holliday junction DNA helicase	72%	3,10E-58	WP_054198714 hypothetical protein
UF08_8	hypothetical protein	77	UF09_16_hypothetical protein	99%	Marpi_0314 hypothetical protein	86%	6,39E-38	no match
UF08_9	hypothetical protein	161	UF09_17_hypothetical protein	95%	Marpi_2120 DNA-directed RNA polymerase sigma-70 factor	38%	1,11E-25	WP_012057673 RNA polymerase subunit ma-24
UF08_10	hypothetical protein	71						WP_002672686 hypothetical protein
UF08_11	hypothetical protein	337						WP_025278489 hypothetical protein
UF08_12	hypothetical protein	127	UF09_19_hypothetical protein	100%	Marpi_0319 hypothetical protein	83%	2,09E-60	WP_047809280 hypothetical protein
UF08_13	terminase large subunit	416	UF09_20_terminase large subunit	99%	Marpi_0320 phage terminase, large subunit, PBSX famil	97%	0	WP_039259229 terminase
UF08_14	hypothetical protein	405	UF09_21_hypothetical protein	99%	Marpi_0321 hypothetical protein	93%	0	WP_027339946 hypothetical protein
UF08_15	hypothetical protein	63						no match
UF08_16	hypothetical protein	51						XP_006580293 PREDICTED: regulation clear pre-mRNA domain-containing protein in 1B-like isoform X3
UF08_17	hypothetical protein	42						WP_052353345 hypothetical protein
UF08_18	hypothetical protein	62	UF09_22_hypothetical protein	98%				XP_008914408 hypothetical protein PI4293
UF08_19	hypothetical protein	319	UF09_23_hypothetical protein	99%	Marpi_0322 hypothetical protein	88%	0	WP_038226045 hypothetical protein
UF08_20	hypothetical protein	45	UF09_24_hypothetical protein	100%	Marpi_0323 hypothetical protein	91%	1,59E-19	no match
UF08_21	Chaperone protein DnaK	200	UF09_25_hypothetical protein	99%	Marpi_0324 hypothetical protein	81%	6,53E-98	WP_063491264 DUF4355 domain-containing protein
UF08_22	hypothetical protein	118	UF09_26_hypothetical protein	99%	Marpi_0325 hypothetical protein	93%	3,54E-72	WP_013866486 hypothetical protein
UF08_23	major capsid protein	340	UF09_27_major capsid protein	99%	Marpi_0326 phage capsid protein	91%	0	EGO87149 capsid protein
UF08_24	hypothetical protein	61	UF09_28_hypothetical protein	100%	Marpi_0327 hypothetical protein	80%	1,38E-24	KLD98821 hypothetical protein AF74_
UF08_25	hypothetical protein	107	UF09_29_hypothetical protein	100%	Marpi_0328 hypothetical protein	86%	2,41E-59	WP_012063953 hypothetical protein
UF08_26	hypothetical protein	158	UF09_30_hypothetical protein	100%	Marpi_0329 hypothetical protein	97%	3,14E-106	WP_006306523 hypothetical protein
UF08_27	hypothetical protein	127	UF09_31_hypothetical protein	98%	Marpi_0330 hypothetical protein	88%	8,57E-75	WP_010381216 MULTISPECIES: cell division topological specificity factor MinE

locus tag	annotation	Length	Best match in MCV22	% Identity	Match in MPV1	% identity2	expected	blast match outside the Marinitoga in BLASP searches of Genbank
UF08_28	hypothetical protein	327	UF09_32_hypothetical protein	70%	Marpi_0331 hypothetical protein	34%	1,76E-46	WP_004102183 hypothetical protein
UF08_29	hypothetical protein	313	UF09_32_hypothetical protein	67%	Marpi_0331 hypothetical protein	65%	2,87E-150	WP_004102183 hypothetical protein
UF08_30	hypothetical protein	125	UF09_33 hypothetical protein	39%	Marpi_0332 hypothetical protein	38%	4,79E-11	WP_004102185 hypothetical protein
UF08_31	hypothetical protein	163	UF09_34_hypothetical protein	29%	Marpi_0333 hypothetical protein	25%	2,47	WP_004102187 hypothetical protein
UF08_32	tail length tape-measure protein	1 304	UF09_35_tail length tape measure protein	89%	Marpi_0334 phage tail tape measure protein, TP901 family	28%	7,57E-41	WP_012057655 phage tail tape measure otein
UF08_33	hypothetical protein	201			Marpi_0335 hypothetical protein	76%	2,06E-92	WP_021570654 hypothetical protein
UF08_34	hypothetical protein	115	UF09_37_hypothetical protein	93%	Marpi_0336 hypothetical protein	66%	6,87E-47	WP_004102189 hypothetical protein
UF08_35	hypothetical protein	393	UF09_38_hypothetical protein	91%	Marpi_0337 hypothetical protein	32%	2,68E-37	WP_004102191 hypothetical protein
UF08_36	Hypothetical protein	1 374	UF09_39_phage realted protein tail	92%	Marpi_0338 hypothetical protein	53%	2,30E-40	WP_004102192 hypothetical protein
UF08_37	hypothetical protein	79	UF09_40_hypothetical protein	100%				KYC55724 hypothetical protein APG09_01
UF08_38	hypothetical protein	213						KKT87314 hypothetical protein UW86_04G0005
UF08_39	hypothetical protein	156						KKT21313 hypothetical protein UW04_06G0016
UF08_40	hypothetical protein	254	UF09_41_DNA repair exonuclease	98%	Marpi_0340 DNA repair exonuclease	78%	5,56E-152	WP_004102197 DNA repair exonuclease
UF08_41	hypothetical protein	97	UF09_42_hypothetical protein	100%	Marpi_0341 hypothetical protein	41%	5,37E-07	KYC55578 hypothetical protein APG08_07
UF08_42	hypothetical protein	127	UF09_43_hypothetical protein	99%	Marpi_0342 hypothetical protein	75%	2,64E-62	WP_012057645 hypothetical protein
UF08_43	transcriptional regulator	71						KXS38990 hypothetical protein AWU54_27
UF08_44	antirepressor protein	294						WP_020411111 antirepressor
UF08_45	Phage holin protein	100	UF09_47_phage holin protein	98%	Marpi_0343 phage holin protein (Holin_LLH)	54%	4,68E-30	WP_016226830 LL-H family phage holin
UF08_46	hypothetical protein	97	UF09_48_hypothetical protein	97%	Marpi_0344 hypothetical protein	57%	2,69E-31	WP_032557479 succinate dehydrogenase
UF08_47	hypothetical protein	188	UF09_49_hypothetical protein	99%	Marpi_0345 Orthopoxvirus hypothetical protein	80%	1,06E-109	CUB50865 hypothetical protein BN2127_10_00382
UF08_48	resolvase	485	UF09_50_Resolvase N-terminal domain	69%	Marpi_0291 site-specific recombinase DNA invertase Pin	44%	2,26E-121	WP_036226465 hypothetical protein
UF08_49	hypothetical protein	56			Marpi_0292 hypothetical protein	77%	8,09E-22	WP_059109184 ribosome biogenesis GTPase Der
UF08_50	hypothetical protein	121						WP_057574369 hypothetical protein
UF08_51	hypothetical protein	153						KUK80138 Uncharacterized protein XD96_0662
UF08_52	hypothetical protein	201						KUK80137 Uncharacterized protein XD96_0661
UF08_53	hypothetical protein	134						KUK80136 Uncharacterized protein XD96_0660
UF08_54	hypothetical protein	361						KPV61867 hypothetical protein AOA66_19

locus tag	annotation	Length	Best match in MCV22	% Identity	Match in MPV1	% identity2	expected	blast match outside the Marinita BLASP searches of Genbank
UF08_55	bipolar DNA helicase HerA	553						CUU36910 hypothetical protein 973
UF08_56	hypothetical protein	345						WP_027719510 hypothetical p
UF08_57	hypothetical protein	119						XP_010260093 PREDICTED: unc protein LOC104599305
UF08_58	hypothetical protein	369						CDB45357 putative uncharacte
UF08_59	transcriptional regulator	202	UF09_61_SOS-response repressor and protease LexA	78%	Marpi_0297 SOS response transcriptional repressor, RecA-mediated autopeptidase	45%	1,76E-52	WP_053001073 hypothetical p
UF08_60	SOS-response repressor and protease LexA	216	UF09_61_SOS-response repressor and protease LexA	99%	Marpi_0297 SOS response transcriptional repressor, RecA-mediated autopeptidase	43%	1,10E-37	WP_053001073 hypothetical p
UF08_61	transcriptional regulator	71	UF09_62_hypothetical protein	99%	Marpi_0298 transcriptional regulator	54%	5,46E-11	KKQ85063 Helix-turn-helix dom
UF08_62	antirepressor protein	241						WP_015397656 anti-repressor
UF08_63	hypothetical protein	64						KUO58776 cobalt ABC transpor binding protein
UF08_64	hypothetical protein	241						WP_045166120 haloacid dehal
UF08_65	transcriptional regulator, XRE family	113	UF09_64_Transcriptional regulator XRE family	98%				WP_062805019 hypothetical p
UF08_66	transcriptional regulator	70	UF09_65_hypothetical protein	86%				WP_057001098 hypothetical p
UF08_67	hypothetical protein	111	UF09_66_hypothetical protein	99%				KKW41634 Foldase protein Prs
UF08_68	hypothetical protein	232	UF09_67_hypothetical protein	100%				XP_005178606 PREDICTED: lam alpha
UF08_69	hypothetical protein	78	UF09_68_hypothetical protein	99%				KKW21360 Alanine-tRNA ligase
UF08_70	hypothetical protein	54	UF09_69_hypothetical protein	98%				XP_008625909 hypothetical pr 044
UF08_71	hypothetical protein	154			Marpi_0299 hypothetical protein	58%	3,30E-47	no match
UF08_72	Viral nuclease	313	UF09_70_Yqaj-like viral recombinase	100%				WP_004102128 hypothetical p
UF08_73	recombinational DNA repair protein RecT	294	UF09_1_recombinational DNA repair protein RecT	100%				WP_004102129 DNA recombin RecT
UF08_74	hypothetical protein	78	UF09_2_hypothetical protein	100%				WP_047597684 hypothetical p
UF08_75	hypothetical protein	43	UF09_3_hypothetical protein	100%				CZZ10673 Uncharacterised prot
UF08_76	transcriptional regulator	63	UF09_4_hypothetical protein	100%				WP_049499605 MULTISPECIES nal regulator
UF08_77	hypothetical protein	103	UF09_5_hypothetical protein	100%				WP_059175353 hypothetical p

locus tag	annotation	best match outside Thermotogae	taxonomic name of best match outside the Thermotogae	expected (outside Thermotogae)	% identity (outside Thermotogae)	taxonomy of best match outside Thermotogae	Match in other V
UF08_1	hypothetical protein	WP_042729055 phage replication protein O	Pseudomonas fluorescens	7,07E-27	47%	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas	AMO43596.1hyp [Pseudomonas ph
UF08_2	hypothetical protein						
UF08_3	replicative DNA helicase	WP_031484309 replicative DNA helicase KIJ95750 Single-stranded DNA-binding protein	Desulfovibrio frigidus	5,03E-79	37%	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales; Desulfovibrionaceae; Desulfovibrio	NP_795382.1 put subunit [Streptoc
UF08_4	binding protein		Desulfonauticus sp. 38 4375	3,03E-28	39%	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales; Desulfohalobiaceae; Desulfonauticus	YP_009193875.1 [Paenibacillus ph
UF08_5	hypothetical protein	WP_015616411 hypothetical protein XP_008855251 hypothetical protein ENU1_022120	Clostridium pasteurianum	2,01	36%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium	
UF08_6	hypothetical protein		Entamoeba nuttalli P19	1,78	34%	Eukaryota; Amoebozoa; Archamoebae; Entamoebidae; Entamoeba	
UF08_7	Holliday junction resolvase RusA	WP_054198714 hypothetical protein	Helicobacter pullorum	3,11E-09	35%	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter	ANS05815.1 holl Meditteranean ph
UF08_8	hypothetical protein						
UF08_9	hypothetical protein	WP_027121895 RNA polymerase sigma factor	Mycoplasma imitans	0,656	25%	Bacteria; Tenericutes; Mollicutes; Mycoplasmataceae; Mycoplasma	AEW47696.1 put
UF08_10	hypothetical protein	WP_002672686 hypothetical protein	Treponema denticola	0,378	38%	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema	ABE69212.1 nucl
UF08_11	hypothetical protein	WP_025278489 hypothetical protein	Barnesiella viscericola	1,01E-09	25%	Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Porphyromonadaceae; Barnesiella	AET73077.1 hypo globosa virus 12T
UF08_12	hypothetical protein	WP_047809280 hypothetical protein	Desulfosporosinus acididurans	1,95E-08	32%	Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae; Desulfosporosinus	AKH46794.1 hypo
UF08_13	terminase large subunit	WP_039259229 terminase	Clostridium botulinum	1,83E-154	58%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium	YP_009216669.1 virus] [Clostridium
UF08_14	hypothetical protein	WP_027339946 hypothetical protein	Halonatronum saccharophilum	2,78E-14	25%	Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halobacteroidaceae; Halonatronum	
UF08_15	hypothetical protein						
UF08_16	hypothetical protein	XP_006580293 PREDICTED: regulation of nuclear pre-mRNA domain-containing protein in 1B-like isoform X3	Glycine max	7,93	37%	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Gunneridae; Pentapetales; rosids; fabids; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine; Soja	BAQ85920.1 hypo Meditteranean ph
UF08_17	hypothetical protein	WP_052353345 hypothetical protein	Methermicoccus shengliensis	1,05	41%	Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methermicoccaceae; Methermicoccus	
UF08_18	hypothetical protein	XP_008914408 hypothetical protein PPTG_24293	Phytophthora parasitica INRA-310	1,32	32%	Eukaryota; Stramenopiles; Oomycetes; Peronosporales; Phytophthora	
UF08_19	hypothetical protein	WP_038226045 hypothetical protein	Vibrio sp. ER1A	5,27E-19	22%	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio	YP_006990482.1 phage phiMMP02
UF08_20	hypothetical protein						
UF08_21	Chaperone protein DnaK	WP_063491264 DUF4355 domain-containing protein	Lactobacillus plantarum	9,11E-04	32%	Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus	YP_004301280.1
UF08_22	hypothetical protein	WP_013866486 hypothetical protein	Methanothermococcus okinawensis	5,08E-16	51%	Archaea; Euryarchaeota; Methanococci; Methanococcales; Methanococcaceae; Methanothermococcus	AMM44193.1 hyp [Arthrobacter ph
UF08_23	major capsid protein	EGO87149 capsid protein	Clostridium botulinum C str. Stockholm	5,55E-72	40%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium	YP_008858642.1 Davies]
UF08_24	hypothetical protein	KLD98821 hypothetical protein AF74_02140	Arcobacter butzleri L349	3,87	38%	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Arcobacter	
UF08_25	hypothetical protein	WP_012063953 hypothetical protein	Alkaliphilus metalliredigens	3,37E-03	37%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Alkaliphilus	
UF08_26	hypothetical protein	WP_006306523 hypothetical protein	Centipeda periodontii	5,22E-06	30%	Bacteria; Firmicutes; Negativicutes; Selenomonadales; Veillonellaceae; Centipeda	BAN05345.1 tail
UF08_27	hypothetical protein	WP_010381216 MULTISPECIES: cell division topological specificity factor MinE	Pseudoalteromonas	5,85	32%	Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Pseudoalteromonadaceae	AJP09134.1 hypo ascovirus 3f]

locus tag	annotation	best match outside Thermotogae	taxonomic name of best match outside the Thermotogae	expected (outside Thermotogae)	% identity (outside Thermotogae)	taxonomy of best match outside Thermotogae	Match in other databases
UF08_28	hypothetical protein	WP_035430862 hypothetical protein	Bacillus sp. UNC322MFChir4.1	8,16E-08	23%	Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus	YP_008770616 phage Glitter
UF08_29	hypothetical protein	WP_002670872 hypothetical protein	Treponema denticola	9,37E-12	23%	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema	BAQ84218.1 Mediterranean
UF08_30	hypothetical protein	XP_002837480 hypothetical protein	Tuber melanosporum Mel28	0,25	32%	Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina; Pezizomycetes; Pezizales; Tuberaceae; Tuber	CBZ06023.1 congregata b
UF08_31	hypothetical protein	WP_010856156 T9SS C-terminal target domain-containing protein	Lunatimonas lon arensis	1,51	38%	Bacteria; Bacteroidetes; Cytophagia; Cytophagales; Cyclobacteriaceae	YP_00805975 phage pSF-1]
UF08_32	tail length tape-measure protein	WP_027339931 phage tail tape measure protein	Halonatronum saccharophilum	4,24E-30	25%	Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halobacteroidaceae; Halonatronum	
UF08_33	hypothetical protein	WP_021570654 hypothetical protein	Escherichia coli	3,02E-06	33%	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia	
UF08_34	hypothetical protein	no match					
UF08_35	hypothetical protein	OAD21159 Pentaxin	Candidatus Thiomargarita nelsonii	0,305	35%	Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales; Thiotrichaceae; Thiomargarita	
UF08_36	Hypothetical protein	WP_026486803 hypothetical protein	Caldanaerobius polysaccharolyticus	5,48E-18	28%	Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales; Thermoanaerobacteraceae; Caldanaerobius	YP_00900315 vB_RleS_L33
UF08_37	hypothetical protein	KYC55724 hypothetical protein APG09_01571	Arc I group archaeon ADurb1213 BinO 2801	5,57E-16	46%	Archaea; Euryarchaeota; Methanomicrobia; unclassified Methanomicrobia; Arc I group	
UF08_38	hypothetical protein	KKT87314 hypothetical protein UW86_C0004G0005	Microgenomates group bacterium GW2011 GWA1	2,70E-43	42%	Bacteria	YP_00920222
UF08_39	hypothetical protein	KKT21313 hypothetical protein UW04_C0006G0016	Parcubacteria group bacterium GW2011 GWB1 43 8	1,50E-07	34%	Bacteria; unclassified Parcubacteria group	
UF08_40	hypothetical protein	KXX67042 Ser/Thr phosphatase family protein	Christensenella minuta	2,37E-20	31%	Bacteria; Firmicutes; Clostridia; Clostridiales; Christensenellaceae; Christensenella	YP_00877070 phage Staley
UF08_41	hypothetical protein	KYC55578 hypothetical protein APG08_01587	Arc I group archaeon ADurb1113 BinO 1801	5,60E-05	31%	Archaea; Euryarchaeota; Methanomicrobia; unclassified Methanomicrobia; Arc I group	BAQ94054.1 Mediterranean
UF08_42	hypothetical protein	ACT80355 maturase K, partial (chloroplast)	Calceolaria germainii	0,209	31%	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Calceolariaceae; Calceolaria	
UF08_43	transcriptional regulator	KXS38990 hypothetical protein AWU54_2187	Halobacteroidaceae bacterium T328-2	1,30E-20	62%	Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halobacteroidaceae	EES90331.1 [Lactococcus]
UF08_44	antirepressor protein	WP_020411111 antirepressor	Microbulbifer variabilis	8,16E-36	58%	Bacteria; Proteobacteria; Gammaproteobacteria; Cellvibrionales; Microbulbiferaceae; Microbulbifer	YP_00832014 [Lactococcus]
UF08_45	Phage holin protein	WP_016226830 LL-H family phage holin	Lachnospiraceae bacterium 10-1	1,17E-04	34%	Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae	
UF08_46	hypothetical protein	WP_032557479 succinate dehydrogenase CUB50865 hypothetical protein BN2127_JRS10_00382	Prochlorococcus sp. scB245a 521B10	1,59	29%	Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae; Prochlorococcus	
UF08_47	hypothetical protein		Bacillus subtilis	1,67E-06	27%	Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus	
UF08_48	resolvase	ABY92005 Recombinase	Thermoanaerobacter sp. X514	5,67E-57	32%	Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales; Thermoanaerobacter	
UF08_49	hypothetical protein	WP_059109184 ribosome biogenesis GTPase Der	Catabacter hongkongensis	1,09	36%	Bacteria; Firmicutes; Clostridia; Clostridiales; Catabacteriaceae; Catabacter	
UF08_50	hypothetical protein	WP_057574369 hypothetical protein	Clostridium sordellii	3,58E-21	41%	Bacteria; Firmicutes; Clostridia; Clostridiales; Peptostreptococcaceae; Peptoclostridium	
UF08_51	hypothetical protein	WP_013788522 hypothetical protein	Thermoanaerobacterium xylanolyticum	1,62E-08	28%	Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales; Thermoanaerobacterium	
UF08_52	hypothetical protein	ETR73156 2-hydroxyglutaryl-CoA dehydratase	Candidatus Magnetoglobus multicellularis str. Araruama	1,66	26%	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales; Desulfobac	
UF08_53	hypothetical protein	WP_008695396 plasmid stabilization protein	Fusobacterium ulcerans	8,07E-05	37%	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae; Fusobacterium	YP_00734851 [Escherichia]
UF08_54	hypothetical protein	KPV61867 hypothetical protein AOA66_1539	Candidatus Bathyarchaeota archaeon B A2	4,95E-111	48%	Archaea; Archaea candidate phyla; Candidatus Bathyarchaeota	YP_00921797 (endogenous)

locus tag	annotation	best match outside Thermotogae	taxonomic name of best match outside the Thermotogae	expected (outside Thermotogae)	% identity (outside Thermotogae)	taxonomy of best match outside Thermotogae	Match in
UF08_55	bipolar DNA helicase HerA	CUU36910 hypothetical protein DCOP10_1 1973	Armatimonadetes bacterium DC	2,57E-100	35%	Bacteria; Armatimonadetes	
UF08_56	hypothetical protein	WP_027719510 hypothetical protein	Desulfoviregula thermocuniculi	7,83E-23	27%	Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales; Thermoanaerobacteraceae; Desulfoviregula	
UF08_57	hypothetical protein	XP_010260093 PREDICTED: uncharacterized protein LOC104599305	Nelumbo nucifera	0,918	31%	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Nelumbonaceae; Nelumbo	
UF08_58	hypothetical protein	CDB45357 putative uncharacterized protein	Phascolarctobacterium sp. CAG:207	8,84	22%	Bacteria; Firmicutes; Negativicutes; Selenomonadales; Acidaminococcaceae; Phascolarctobacterium	
UF08_59	transcriptional regulator	WP_058964196 transcriptional regulator	Ruminococcaceae bacterium AM2	9,76E-41	40%	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; unclassified Ruminococcaceae	YP_24003
UF08_60	SOS-response repressor and protease LexA	WP_009006200 MULTISPECIES: SOS-response transcriptional repressor	Fusobacterium	1,90E-28	36%	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae	YP_00701 thermoph
UF08_61	transcriptional regulator	KKQ85063 Helix-turn-helix domain protein	Parcubacteria group bacterium GW2011 GWF2 38 8	3,05E-10	42%	Bacteria; unclassified Parcubacteria group	YP_00916 [Lactobac
UF08_62	antirepressor protein	WP_015397656 anti-repressor protein	Bartonella australis	1,53E-50	43%	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bartonellaceae; Bartonella	YP_00812 PG1]
UF08_63	hypothetical protein	KUO58776 cobalt ABC transporter ATP-binding protein	Gracilibacter sp. BRH c7a	5,48	43%	Bacteria; Firmicutes; Clostridia; Clostridiales; Gracilibacteraceae; Gracilibacter	
UF08_64	hypothetical protein	WP_045166120 haloacid dehalogenase	Thermoanaerobacter cellulyticus	0,502	33%	Bacteria; Firmicutes; Clostridia; Clostridiales; Anaerocellum group	
UF08_65	transcriptional regulator, XRE family	WP_062805019 hypothetical protein	Enterococcus pernyi	3,38E-16	43%	Bacteria; Firmicutes; Bacilli; Lactobacillales; Enterococcaceae; Enterococcus	YP_00919
UF08_66	transcriptional regulator	WP_057001098 hypothetical protein	Carnobacterium maltaromaticum	5,38E-09	38%	Bacteria; Firmicutes; Bacilli; Lactobacillales; Carnobacteriaceae; Carnobacterium	YP_00146
UF08_67	hypothetical protein	KKW41634 Foldase protein PrsA	Candidatus Magasanikbacteria bacterium GW2011 G WA2 56 11	2,63E-02	30%	Bacteria; Candidatus Magasanikbacteria	
UF08_68	hypothetical protein	XP_005178606 PREDICTED: laminin subunit alpha	Musca domestica	0,612	27%	Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Muscomorpha; Muscoidea; Muscidae; Musca	
UF08_69	hypothetical protein	KKW21360 Alanine-tRNA ligase	Candidatus Adlerbacteria bacterium GW2011 GWC1 5 0 9	6,25	36%	Bacteria; Candidatus Adlerbacteria	
UF08_70	hypothetical protein	XP_008625909 hypothetical protein YYE_04044	Plasmodium vinckei vinckei	8,07	33%	Eukaryota; Alveolata; Apicomplexa; Aconoidasida; Haemosporida; Plasmodium	YP_00767 [Staphyloc
UF08_71	hypothetical protein						YP_00824 phi OH2]
UF08_72	Viral nuclease	WP_015358110 phage-type endonuclease	Clostridium stercorarium	2,85E-117	54%	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminiclostridium	
UF08_73	recombinational DNA repair protein RecT	WP_060476225 recombinase RecT	Geobacillus sp. PA-3	9,59E-101	55%	Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Geobacillus	YP_0082403
UF08_74	hypothetical protein	WP_047597684 hypothetical protein	Pseudomonas sp. RIT288	0,888	38%	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas	
UF08_75	hypothetical protein	CZZ10673 Uncharacterised protein	Enterobacter cloacae	1,68E-02	60%	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Enterobacter; cloacae complex	
UF08_76	transcriptional regulator	WP_049499605 MULTISPECIES: transcriptional regulator	Streptococcus	1,94E-13	53%	Bacteria; Firmicutes; Bacilli; Lactobacillales; Streptococcaceae	YP_00877
UF08_77	hypothetical protein	WP_059175353 hypothetical protein	Thermodesulfovibrio aggregans	0,571	29%	Bacteria; Nitrospirae; Nitrospirales; Nitrospiraceae; Thermodesulfovibrio	

locus tag	annotation	Length (aa)	Match in MCV1	% identity MCV1 ORF Match in MPV1	% identity	expected best match outside the Marintoga in BLASP searches of Genbank	ta
UF09_1	recombinational DNA repair protein RecT	294	UF08_73 recombinational DNA repair protein RecT	100%		WP_004102129 DNA recombination protein RecT	TH
UF09_2	hypothetical protein	78	UF08_74 hypothetical protein	100%		WP_047597684 hypothetical protein	Pa
UF09_3	hypothetical protein	43	UF08_75 hypothetical protein	100%		CZ210673 Uncharacterised protein	Er
UF09_4	transcriptional regulator	63	UF08_76 hypothetical protein	100%		WP_008377672 transcriptional regulator	Er
UF09_5	hypothetical protein	103	UF08_77 hypothetical protein	100%		WP_059175353 hypothetical protein	TH
UF09_6	hypothetical protein	211	UF08_1 hypothetical protein	95%		WP_042729055 phage replication protein O	Pa
UF09_7	hypothetical protein	70	UF08_2 hypothetical protein	99%		EHY68330 hypothetical protein SEH00A_02810	Sa
UF09_8	replicative DNA helicase	427	UF08_3 replicative DNA helicase	98%	Marpi_0305 replicative DNA helicase	66% 0 WP_031484309 replicative DNA helicase	De
UF09_9	single-stranded DNA-binding protein	149	UF08_4 single-stranded DNA-binding protein	97%	Marpi_0443 single-stranded DNA-binding protein	59% 2,61E-51 WP_052670309 hypothetical protein	De
UF09_10	hypothetical protein	183					
UF09_11	hypothetical protein	114				WP_044141503 helicase	Br
UF09_12	deoxyuridine 5'-triphosphate nucleotidohydrolase	140			Marpi_0307 deoxyuridine 5'-triphosphate nucleotidohydrolase Dut	81% 1,57E-77 WP_040347745 deoxyuridine 5'-triphosphate nucleotidohydrolase	Br
UF09_13	hypothetical protein	92	UF08_5 hypothetical protein	99%	Marpi_0308 hypothetical protein	47% 4,64E-08 WP_015616411 hypothetical protein	CI
UF09_14	hypothetical protein	82	UF08_6 hypothetical protein	100%	Marpi_0309 hypothetical protein	53% 6,58E-15 XP_008855251 hypothetical protein ENU1_022120	Er
UF09_15	hypothetical protein	122	UF08_7 hypothetical protein	94%	Marpi_0310 Holliday junction DNA helicase	69% 7,97E-57 WP_054198714 hypothetical protein	Ha
UF09_16	hypothetical protein	77	UF08_8 hypothetical protein	99%	Marpi_0314 hypothetical protein	83% 9,55E-38 WP_047264626 lipoprotein	CI
UF09_17	hypothetical protein	161	UF08_9 hypothetical protein	95%	Marpi_0315 hypothetical protein	50% 3,15E-46 WP_012057673 RNA polymerase subunit sigma-24	TH
UF09_18	hypothetical protein	277				WP_022851452 hypothetical protein	Gi
UF09_19	Homeodomain-like	127	UF08_12 hypothetical protein	100%	Marpi_0319 hypothetical protein	82% 8,98E-60 WP_047809280 hypothetical protein	De
UF09_20	terminase large subunit	416	UF08_13 terminase large subunit	99%	Marpi_0320 phage terminase, large subunit, PBSX family	97% 0 WP_053284682 terminase	CI
UF09_21	hypothetical protein	405	UF08_14 hypothetical protein	99%	Marpi_0321 hypothetical protein	93% 0 WP_027339946 hypothetical protein	Ha
UF09_22	hypothetical protein	62	UF08_18 hypothetical protein	98%		XP_008914408 hypothetical protein PPTG_24293	Ph
UF09_23	hypothetical protein	502	UF08_19 hypothetical protein	99%	Marpi_0322 hypothetical protein	85% 0 WP_038226045 hypothetical protein	Vi
UF09_24	hypothetical protein	94	UF08_20 hypothetical protein	100%	Marpi_0323 hypothetical protein	88% 4,78E-79 WP_047441240 peptidoglycan L-alanyl-D-glutamate endopeptidase	CI
UF09_25	hypothetical protein	200	UF08_21 protein of unknown function DUF4355	99%	Marpi_0324 hypothetical protein	82% 1,62E-97 WP_009531408 hypothetical protein	Pa
UF09_26	structural protein	118	UF08_22 hypothetical protein	99%	Marpi_0325 hypothetical protein	94% 1,10E-72 WP_036595750 hypothetical protein	Pa
UF09_27	major capsid protein	340	UF08_23 major capsid protein	99%	Marpi_0326 Phage major capsid protein E	91% 0 WP_039218408 phage capsid protein	CI
UF09_28	hypothetical protein	61	UF08_24 hypothetical protein	100%	Marpi_0327 hypothetical protein	80% 2,01E-23 KLD98821 hypothetical protein AF74_02140	Ar
UF09_29	hypothetical protein	107	UF08_25 hypothetical protein	100%	Marpi_0328 hypothetical protein	86% 7,36E-58 WP_012063953 hypothetical protein	Al
UF09_30	hypothetical protein	158	UF08_26 hypothetical protein	100%	Marpi_0329 hypothetical protein	98% 1,24E-106 WP_006306523 hypothetical protein	Ce
UF09_31	hypothetical protein	127	UF08_27 hypothetical protein	98%	Marpi_0330 hypothetical protein	89% 2,51E-74 WP_053940024 5-methyltetrahydropteroylribiglutamate-homocysteine S-methyltrans	Br
UF09_32	hypothetical protein	322	UF08_29 hypothetical protein	67%	Marpi_0331 hypothetical protein	88% 0 WP_012057658 hypothetical protein	TH
UF09_33	hypothetical protein	117	UF08_30 hypothetical protein	39%	Marpi_0332 hypothetical protein	94% 9,10E-72 WP_004102185 hypothetical protein	TH
UF09_34	hypothetical protein	167	UF08_31 hypothetical protein	29%	Marpi_0333 hypothetical protein	70% 2,89E-78 WP_004102187 hypothetical protein	TH
UF09_35	tail length tape measure protein	1 516	UF08_32 tail length tape-measure protein	89%	Marpi_0334 phage tail tape measure protein, TP901 family	72% 0 WP_004102188 phage tail ta pe measure protein	TH

locus tag	annotation	Length (aa)	Match in MCV1	% identity MCV1 ORF Match in MPV1		% identity	expected blast match outside the Marinitoga in BLASP searches of Genbank
UF09_36	hypothetical protein	131					YP_003169739 hypothetical protein
UF09_37	hypothetical protein	115	UF08_34_hypothetical protein	93%	Marpi_0336_hypothetical protein	66%	8,29E-48 WP_004102189 hypothetical protein
UF09_38	hypothetical protein	390	UF08_35_hypothetical protein	91%	Marpi_0337_hypothetical protein	31%	3,20E-36 WP_041426083 hypothetical protein
UF09_39	phage reallted protein tail	1549	UF08_36_chitin binding protein	92%	Marpi_0338_hypothetical protein	65%	0 WP_004102192 hypothetical protein
UF09_40	hypothetical protein	79	UF08_37_hypothetical protein	100%			KYC55724 hypothetical protein APG09_01571
UF09_41	DNA repair exonuclease	254	UF08_40_hypothetical protein	98%	Marpi_0340_DNA repair exonuclease	80%	1,85E-154 WP_004102197 DNA repair exonuclease
UF09_42	hypothetical protein	97	UF08_41_hypothetical protein	100%	Marpi_0341_hypothetical protein	41%	1,63E-06 KYC5578 hypothetical protein APG08_01587
UF09_43	hypothetical protein	127	UF08_42_hypothetical protein	99%	Marpi_0342_hypothetical protein	75%	4,27E-62 WP_012057645 hypothetical protein
UF09_44	hypothetical protein	134					WP_012057691 hypothetical protein
UF09_45	hypothetical protein	64					WP_015463858 ATP synthase subunit alpha
UF09_46	hypothetical protein	195					no match
UF09_47	phage holin protein	107	UF08_45_hypothetical protein	98%	Marpi_0343_phage holin protein (Holin_LLH)	54%	1,07E-31 WP_016226830 LL-H family phage holin
UF09_48	hypothetical protein	98	UF08_46_hypothetical protein	97%	Marpi_0344_hypothetical protein	56%	3,63E-32 no match
UF09_49	hypothetical protein	194	UF08_47_hypothetical protein	99%	Marpi_0345_Orthopoxvirus protein of unknown function (DUF830)	80%	1,25E-109 CUB50865 hypothetical protein BN2127_JRS10_00382
UF09_50	Resolvase N-terminal domain	490	UF08_48_resolvase	69%	Marpi_0291_site-specific recombinase, DNA invertase Pi	88%	0 WP_047753688 hypothetical protein
UF09_51	hypothetical protein	162					WP_002975830 hypothetical protein
UF09_52	hypothetical protein	344					no match
UF09_53	hypothetical protein	141					WP_028559036 hypothetical protein
UF09_54	hypothetical protein	265					no match
UF09_55	Chromosome segregation ATPase	583			Marpi_0295_T5orf172 domain-containing protein	67%	0 WP_011968228 hypothetical protein
UF09_56	hypothetical protein	122					WP_015336104 hypothetical protein
UF09_57	hypothetical protein	133					WP_036093225 mRNA interferase MazF2
UF09_58	hypothetical protein	117					no match
UF09_59	hypothetical protein	88					WP_015898402 hypothetical protein
UF09_60	hypothetical protein	68					no match
UF09_61	SOS-response repressor and protease LexA	216	UF08_60_SOS-response repressor and protease LexA3	99%	Marpi_0297_SOS response transcriptional repressor, RecA-mediated autopeptidase	44%	6,09E-39 WP_053001073 hypothetical protein
UF09_62	hypothetical protein	71	UF08_61_hypothetical protein	99%	Marpi_0298_transcriptional regulator	54%	3,44E-10 KQ85063 Helix-turn-helix domain protein
UF09_63	hypothetical protein	87					WP_063517479 hypothetical protein
UF09_64	Transcriptional regulator XRE family	117	UF08_65_transcriptional regulator, XRE family	98%			WP_009618532 transcriptional regulator
UF09_65	hypothetical protein	69	UF08_66_hypothetical protein	86%			WP_019132685 transcriptional regulator
UF09_66	hypothetical protein	111	UF08_67_hypothetical protein	99%			KKW41634 Foldase protein PrsA
UF09_67	hypothetical protein	232	UF08_68_hypothetical protein	100%			XP_005178606 PREDICTED: laminin subunit alpha
UF09_68	hypothetical protein	78	UF08_69_hypothetical protein	99%			KKW21360 Alanine-tRNA ligase
UF09_69	hypothetical protein	54	UF08_70_hypothetical protein	98%			XP_008625909 hypothetical protein YYE_04044
UF09_70	YqaI-like viral recombinase	313	UF08_72_recombination-related protein	100%			WP_036226574 hypothetical protein

locus tag	annotation	best match outside Thermotogae	taxonomic name of best match outside the Thermotogae	expected (outside Thermotogae)	% identity (outside Thermotogae)	taxonomy of best match outside Thermotogae	Matches in other Viruses	% identity2
UF09_1	recombinational DNA repair protein RecT	WP_060476225 recombinase RecT	Geobacillus sp. PA-3	9,59E-101	55%	Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Geobacillus	YP_008240347.1 RecT protein [Thermus phage phi OH2]	55%
UF09_2	hypothetical protein	WP_047597684 hypothetical protein	Pseudomonas sp. RIT288	0.888	38%	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas		
UF09_3	hypothetical protein	CZZ10673 Uncharacterised protein	Enterobacter cloacae	1.68E-02	60%	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Enterobacter; Enterobacter cloacae complex		
UF09_4	transcriptional regulator	WP_008377672 transcriptional regulator	Enterococcus sp. C1	3,68E-13	52%	Bacteria; Firmicutes; Bacilli; Lactobacillales; Enterococcaceae; Enterococcus	YP_008770535.1 DNA binding protein [Bacillus enterococcus	45%
UF09_5	hypothetical protein	WP_059175353 hypothetical protein	Thermodesulfobivrio aggregans	1.21	29%	Bacteria; Nitrospirae; Nitrospirales; Nitrospiraceae; Thermodesulfobivrio		
UF09_6	hypothetical protein	WP_042729055 phage replication protein O	Pseudomonas fluorescens	3.02E-26	47%	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas	AM043596.1 hypothetical protein PAEP54_00450 [Pseudomonas phage YMC11/07/P54_PAE_BP]	38%
UF09_7	hypothetical protein	EHY68330 hypothetical protein SEH00A_02810	Salmonella enterica subsp. houtena e str. ATCC BAA-1581	2.68	28%	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella		
UF09_8	replicative DNA helicase	WP_031484309 replicative DNA helicase	Desulfobivrio frigidus	1.84E-79	37%	Bacteria; Proteobacteria; Desulfobivriales; Desulfobivriaceae; Desulfobivrio	NP_795382.1 putative DNA polymerase III delta prime subunit [Streptococcus pyogenes phage 315.1]	33%
UF09_9	single-stranded DNA-binding protein	WP_007489185 single-strand-binding family protein	Flavonifractor plautii	3.80E-28	40%	Bacteria; Firmicutes; Clostridia; Clostridiales; Flavonifractor	YP_009193875.1 single-stranded DNA-binding protein [Paenibacillus phage Harrison]	40%
UF09_10	hypothetical protein							
UF09_11	hypothetical protein	WP_044141503 helicase	Bacillus pumilus	4.04	26%	Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus		
UF09_12	deoxyuridine 5'-triphosphate nucleotidohydrolase	WP_040347745 deoxyuridine 5'-triphosphate nucleotidohydrolase	Brevibacterium mcbrellneri	5.34E-25	42%	Bacteria; Actinobacteria; Micrococcales; Brevibacteriaceae; Brevibacterium	AD230445.1 dUTPase [Cowpox virus]	41%
UF09_13	hypothetical protein	WP_015616411 hypothetical protein	Clostridium pasteurianum	2.01	36%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium		
UF09_14	hypothetical protein	XP_008855251 hypothetical protein ENU1_022120	Entamoeba nuttalli P19	1.78	34%	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Helicobacter	YP_007111562.1 holliday-junction resolvase RusA [Enterobacteria phage mep043 c-1]	28%
UF09_15	hypothetical protein	WP_054198714 hypothetical protein	Helicobacter pullorum	2.40E-09	35%	Bacteria; Proteobacteria; Helicobacteriales; Helicobacteraceae; Helicobacter		
UF09_16	hypothetical protein	WP_047264626 lipoprotein	Candidatus Liberibacter africanus	6.96	35%	Bacteria; Proteobacteria; Rhizobiales; Rhizobiaceae; Candidatus Liberibacter		
UF09_17	hypothetical protein							
UF09_18	hypothetical protein	WP_057677254 RNA polymerase sigma factor RpoD	Curvibacter sp. PAE-UM	0.19	29%	Betaproteobacteria; Burkholderiales; Comamonadaceae; Curvibacter	AEW47696.1 putative tail fiber 1 [Bacillus phage BCU4]	28%
UF09_19	Homeodomain-like	WP_047809280 hypothetical protein	Geovibrio sp. L21-Ace-BES	1.66E-10	25%	Bacteria; Deferribacteres; Deferribacteriales; Deferribacteraceae; Geovibrio		
UF09_20	terminase large subunit	WP_053284682 terminase	Desulfosporosinus acidurans	2.13E-08	52%	Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae; Desulfosporosinus	AKH46794.1 hypothetical protein [uncultured marine virus]	52%
UF09_21	hypothetical protein	WP_027339946 hypothetical protein	Desulfosporosinus				YP_009216669.1 terminase large subunit (endogenous virus) [Clostridium phage phiCT453A]	59%
UF09_22	hypothetical protein	WP_053284682 terminase	Clostridium botulinum	4.35E-152	58%	Clostridium		
UF09_23	hypothetical protein	WP_038226045 hypothetical protein	Halonatronum saccharophilum	2.27E-14	25%	Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halobacteroidaceae; Halonatronum		
UF09_24	hypothetical protein	WP_047441240 peptidoglycan L-alanyl-D-glutamate endopeptidase	Phytophthora parasitica IN RA-310	3.3	32%	Eukaryota; Stramenopiles; Oomycetes; Peronosporales; Phytophthora		
UF09_25	hypothetical protein	WP_009531408 hypothetical protein	Vibrio sp. ER1A	1.43E-18	22%	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio	YP_006990482.1 head morphogenesis protein [Clostridium phage phiMMP02]	27%
UF09_26	structural protein	WP_036595750 hypothetical protein	Chryseobacterium m sp. YR561	6.09	40%	Bacteria; Bacteroidetes; Flavobacteriia; Flavobacteriales; Flavobacteriaceae; Chryseobacterium		
UF09_27	major capsid protein	WP_039218408 phage capsid protein	Peptostreptococcaceae bacterium O8Kc.8	6.06E-02	27%	Bacteria; Firmicutes; Clostridia; Clostridiales; Peptostreptococcaceae	YP_004301280.1 gp5 [Brochothrix phage NFS]	30%
UF09_28	hypothetical protein	KLD98821 hypothetical protein AF74_02140	Paenibacillus sophorae	5.79E-14	38%	Bacteria; Firmicutes; Bacilli; Bacillales; Paenibacillaceae; Paenibacillus	AMM44193.1 hypothetical protein KELLEZO_23 [Arthrobacter phage Kellezio]	35%
UF09_29	hypothetical protein	WP_012063953 hypothetical protein	Clostridium novyi	8.00E-72	40%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium	YP_008858642.1 major capsid protein [Brevibacillus phage Davies]	32%
UF09_30	hypothetical protein	WP_006306523 hypothetical protein	Arcobacter butzleri L349	4.77	38%	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Arcobacter		
UF09_31	hypothetical protein	WP_053940024 S-methyltetrahydropteroylglutamate--homocysteine S-methyltransferase	Alkaliphilus metalliredigens	3.24E-03	37%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Alkaliphilus		
UF09_32	hypothetical protein	WP_002670872 hypothetical protein	Centipeda periodontii	3.58E-04	29%	Bacteria; Firmicutes; Negativicutes; Selenomonadales; Veillonellaceae; Veillonella	BAN05345.1 tail protein [Nitratiruptor phage Nr5-1]	36%
UF09_33	hypothetical protein	WP_032077236 metal-dependent RNase	Buchnera aphidicola	3.45	24%	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Buchnera	AJP09134.1 hypothetical protein [Heliothis virescens ascovirus 3f]	23%
UF09_34	hypothetical protein	WP_053069208 hypothetical protein	Treponema denticola	5.23E-14	25%	Bacteria; Spirochaetes; Spirochaetiales; Spirochaetaceae; Treponema	YP_009188581.1 putative major tail protein [Gordonia phage GMA3]	26%
UF09_35	tail length tape measure protein	WP_027339931 phage tail tape measure protein	Clostridium drakei	3.25	31%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium		
			Bizionia psychrotolerans	3.65	24%	Bacteria; Bacteroidetes; Flavobacteriia; Flavobacteriales; Flavobacteriaceae; Bizionia		
			Halonatronum saccharophilum			Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halobacteroidaceae; Halonatronum	YP_008318494.1 tail protein [Shigella phage Sfil]	20%

locus tag	annotation	best match outside Thermotogae	taxonomic name of best match outside the Thermotogae	expected (outside Thermotogae)	% identity (outside Thermotogae)	taxonomy of best match outside Thermotogae	Matches in other Viruses	%
UF09_36	hypothetical protein	YP_003169739 hypothetical protein	Thermus thermophilus bacteriophage P23-77	1,56E-08	49%	Viruses; Virus families not assigned to an order; Sphaerolipoviridae; Gammasphaerolipovirus	YP_003169739.1 hypothetical protein [Thermus thermophilus bacteriophage P23-77]	
UF09_37	hypothetical protein	no match						
UF09_38	hypothetical protein	WP_027339930 hypothetical protein	Halonatronum saccharophilum	1,05	52%	Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halobacteroidaceae; Halonatronum		
UF09_39	phage reallted protein tail	WP_034858455 phage minor structural protein	Enterococcus casseliflavus	1,01E-49	32%	Bacteria; Firmicutes; Bacilli; Lactobacillales; Enterococcaceae; Enterococcus	BA029161.1 hypothetical protein [uncultured Mediterranean phage uvMED]	
UF09_40	hypothetical protein	KYC55724 hypothetical protein APG09_01571	Arc1 group archaeon ADur b1213 BinD 2801	5,57E-16	46%	Archaea; Euryarchaeota; Methanomicrobia; unclassified Methanomicrobia; Arc1 group		
UF09_41	DNA repair exonuclease	KXK67042 Ser/Thr phosphatase family protein	Christensenella minuta	6,93E-21	31%	Bacteria; Firmicutes; Clostridia; Clostridiales; Christensenellaceae; Christensenella	YP_008770743.1 metal-dependent hydrolase [Bacillus phage Staley]	
UF09_42	hypothetical protein	KYC55578 hypothetical protein APG08_01587	Arc1 group archaeon ADur b1113 BinD 1801	5,23E-04	31%	Archaea; Euryarchaeota; Methanomicrobia; unclassified Methanomicrobia; Arc1 group	BA094054.1 hypothetical protein [uncultured Mediterranean phage uvMED]	
UF09_43	hypothetical protein	ACT80355 maturase K, partial (chloroplast)	Calceolaria germainii	0,211	31%	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; C		
UF09_44	hypothetical protein	no match						
UF09_45	hypothetical protein	WP_015463858 ATP synthase subunit alpha	Psychromonas sp. CNPT3	1,44E-02	39%	Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Psychromonadaceae; Psychromonas		
UF09_46	hypothetical protein							
UF09_47	phage holin protein	WP_016226830 LL-H family phage holin	Lachnospiraceae bacterium 10-1	6,60E-06	34%	Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae		
UF09_48	hypothetical protein							
UF09_49	hypothetical protein	CU850865 hypothetical protein BN2127_JRS10_00382	Bacillus subtilis	1,48E-06	27%	Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus		
UF09_50	Resolvase N-terminal domain	ABY92005 Recombinase	Thermoanaerobacter sp. X514	1,22E-40	30%	Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales; Thermoanaerobacteraceae; Thermoanaerobacter	YP_009217897.1 serine recombinase (endogenous virus) [Clostridium phage phiCT4538]	
UF09_51	hypothetical protein	WP_002975830 hypothetical protein	Leptospira vanthielii	2,32E-02	34%	Bacteria; Spirochaetes; Leptospirales; Leptospiraceae; Leptospira		
UF09_52	hypothetical protein							
UF09_53	hypothetical protein	WP_028559036 hypothetical protein	Paenibacillus pinihumi	0,41	28%	Bacteria; Firmicutes; Bacilli; Bacillales; AEX62303.1 vacuolar sorting protein [Moumouvirus Paenibacillaceae; Paenibacillus Morve]		
UF09_54	hypothetical protein							
UF09_55	Chromosome segregation ATPase	WP_011968228 hypothetical protein	Clostridium beijerinckii	1,09E-148	56%	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium	EES90330.1 phage protein [Clostridium phage D-1873]	
UF09_56	hypothetical protein	WP_015336104 hypothetical protein	Desulfovibrio hydrothermalis	3,40E-02	51%	Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales; Desulfovibrionaceae; Desulfovibrio	AIQ08090.1 polyprotein [Hepatitis C virus]	
UF09_57	hypothetical protein	WP_036093225 mRNA interferase MazF2	Leptospira weilii	9,29E-04	32%	Bacteria; Spirochaetes; Leptospirales; Leptospiraceae; Leptospira		
UF09_58	hypothetical protein							
UF09_59	hypothetical protein	WP_015898402 hypothetical protein	Acidobacterium capsulatum	1,39E-09	37%	Bacteria; Acidobacteria; Acidobacteriales; Acidobacteriaceae; Acidobacterium	YP_004539068.1 unnamed protein product [Mycobacterium phage Faith1]	
UF09_60	hypothetical protein							
UF09_61	SOS-response repressor and protease LexA	WP_009006200 MULTISPECIES: SOS-response transcriptional repressor	Fusobacterium	4,77E-29	36%	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae	YP_007010947.1 transcriptional regulator [Deep-sea thermophilic phage D6E]	
UF09_62	hypothetical protein	KKQ85063 Helix-turn-helix domain protein	Parcubacteria group bacterium GW2011 GWF2 38 8	1,91E-09	42%	Bacteria; unclassified Parcubacteria group	YP_009168545.1 putative transcription regulator [Lactobacillus phage LBR48]	
UF09_63	hypothetical protein	WP_063517479 hypothetical protein	Vibrio parahaemolyticus	0,526	40%	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio		
UF09_64	Transcriptional regulator XRE family	WP_009618532 transcriptional regulator	Desulfosporosinus sp. OT	5,58E-13	44%	Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae; Desulfosporosinus	ALV8461.1 transcriptional regulator [Bacillus phage vB_BceS-MY192]	
UF09_65	hypothetical protein	WP_019132685 transcriptional regulator	Peptoniphilus obesi	9,79E-13	52%	Bacteria; Firmicutes; Tissierellales; Peptoniphilales; Peptoniphilus	WP_044984172.1 transcriptional regulator, partial [Thermoanaerobacterium phage THSA-485A]	
UF09_66	hypothetical protein	KKW41634 Foldase protein PrsA	Candidatus Magasanikbacteria bacterium GW2011 GWA2 56 11	2,09E-02	30%	Bacteria; Candidatus Magasanikbacteria		
UF09_67	hypothetical protein	XP_005178606 PREDICTED: laminin subunit alpha	Musca domestica	0,65	27%	Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachy		
UF09_68	hypothetical protein	KKW21360 Alanine-tRNA ligase	Candidatus Adlerbacteria bacterium GW2011 GWC1 50 9	2,41	38%	Bacteria; Candidatus Adlerbacteria		
UF09_69	hypothetical protein	XP_008625909 hypothetical protein YYE_04044	Plasmodium vinckei vinckei	8,07	33%	Eukaryota; Alveolata; Apicomplexa; Aconoidasida; Haemosporidia; Plasmodiidae; Plasmodium; Plasmodium (Vinckei)		
UF09_70	YqaI-like viral recombinase	WP_015358110 phage-type endonuclease	Clostridium stercorarium	2,85E-117	54%	Bacteria; Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminiclostridium	P_008240349.1 phage-type endonuclease [Thermus phage phi OH2]	

Chapitre IV : Etude du génome de Thermosipho sp. AT1244-VC14

Présentation des travaux

Les données relatives à l'isolation des têtes usées sont, en dessous, des tests effectués sur la souche *Thermosiphon* sp. AT1244-VC14. Cette souche est la seule qui permet la propagation du plasmide pMP1. Néanmoins, celle-ci est pas sensible à l'infection par les différents virus de *Marinitoga* caractérisés (MPV1, MCV1 et MCV2).

C'est dans le cadre de l'étude du génome de cette souche ont été réalisés. Cette étude a porté sur le système de défense bactérien : le système CRISPR-Cas. Il s'agit du système CRISPR-Cas de classe 1 Type III-A. Les résultats des essais au laboratoire et du système sont présentés, celui-ci paraît donc fonctionnel.

Les spacers du système ont aussi été étudiés et ceux-ci ne montrent aucune homologie de séquence avec pMP1 ou MPV1. Néanmoins, deux spacers ont des similarités avec les virus MCV1 et MCV2 (>85%). Ceci pourrait expliquer pourquoi ces deux virus sont incapables d'infecter cette souche. D'autre part, la présence de ces spacers pourrait signifier que *Thermosiphon* sp. AT1244-VC14 a déjà été en contact avec MCV1, MCV2, ou un virus proche de ceux-ci. Ces résultats font l'objet du *Genome* en cours de publication.

Draft Genome Sequences of a thermophilic and anaerobic bacterium, *Thermosipho* sp. AT1244-VC14, isolated from a deep sea hydrothermal vent.

Coraline Mercier^{a,b,c}, Julien Lossouarn^{a,b,c}, Camilla Nesbø^{d,e}, Thomas H.A Haverkamp^d, Nadège Bienvenu^{a,b,c}, Claire Geslin^{a,b,c}.

a UBL, Laboratoire de Microbiologie des Environnements Extrêmes UMR6197, IUEM, Rue Dumont d'Urville, 29280 Plouzané, France

b CNRS, Laboratoire de Microbiologie des Environnements Extrêmes UMR 6197, IUEM, Rue Dumont d'Urville, 29280 Plouzané, France

c Ifremer, Laboratoire de Microbiologie des Environnements Extrêmes UMR 6197, CS10070, 29280 Plouzané, France

d Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway

e Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

*Present address: Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

Running title: Draft genome of a *Thermosipho* isolate

Address correspondence to: Claire Geslin (claire.geslin@univ-brest.fr)

Abstract

Here, we present the draft genome sequences of a thermophilic *Thermosipho* strain member of the Thermotogales order. This strain was isolated from an active chimney in a deep sea hydrothermal vent, Rainbow site, located on the Mid Atlantic Ridge. This genome notably contains a complete and expected functional CRISPR-cas system, classified as Classe I, Type III-A. Spacers were also found in 4 loci, no match with *Marinitoga piezophila* mobile genetic elements were found whereas two spacers match with two new viruses of *Marinitoga camini*. It suggests that this strain has already been infected by those viruses or other closely related viruses.

Introduction

The bacterial order Thermotogales, consists of mesophilic (Nesbø et al., 2012; Ben Hania et al., 2013) and hyper/thermophilic, anaerobic, chemo-organotrophic microorganisms (Bonch-osmolovskaya, 2008). These bacteria are rod-shaped characterized by an outer sheath-like envelope, the toga. This lineage is broadly represented in deep biosphere ecosystems and deep-sea hydrothermal vents in particular. Phylogenetic analyses of 16S rRNA gene and 29 concatenated ribosomal proteins place Thermotogales as a sister group to Aquificales, and consequently as one of the deepest branching order within domain Bacteria (Zhaxybayeva et al., 2009; Hug et al., 2016). Genomes of this bacterial order contain numerous evidence of lateral gene transfer and notably, CRISPR-cas systems are present in all Thermotogales genome sequences (Nesbø et al., 2009; Zhaxybayeva et al., 2009). CRISPR-cas is a bacterial and archaeal defense system against invading mobile genetic elements, with a wide distribution in the prokaryotic domain. Indeed, approximatively 47% of analyzed prokaryotes

encode CRISPR cas loci distributed in 87% of archaeal genomes and 50% of bacterial genomes (not necessarily complete loci) (Makarova et al., 2015).

At the time of writing, over 80 completed and ongoing Thermotogae genome projects are counted, comprised 10 of the 12 described Thermotogae genera (Pollo et al., 2015). Among the eight Thermosipho species described (*T. africanus*, *T. activus*, *T. affectus*, *T. atlanticus*, *T. geolei*, *T. globiformans*, *T. japonicus* and *T. melanesiensis*), only two complete genome sequences are currently available (*T. africanus* and *T. melanesiensis*).

Thermosipho sp. AT1244-VC14 (hereafter referred to as *Thermosipho* sp. 1244) was isolated from the hydrothermal site Rainbow (Portugal, Azores) in the Mid-Atlantic Ridge at a depth of 2275m. It was sampled from an active chimney in 2001 during the ATOS cruise. This rod-shaped anaerobic thermophilic chemoorganotrophic bacterium grows optimally at 60°C under atmospheric pressure and anaerobic conditions in a rich medium. Here we describe the genome of *Thermosipho* sp. 1244 and study his CRISPR-cas system. The aim of this study is to extend our knowledge about this genus and its relationships with mobile genetic elements.

Results and discussion

The chromosome consists in 1,818,555 bp in 34 contigs, with the longest contig size of 301,822 bp, and a G+C content of 31.2 %. No extrachromosomal element has been detected. On the basis of the 16S rRNA gene sequence extracted from this genome, the closest strain of *Thermosipho* sp. 1244 is *Thermosipho affectus* strain ik275mar (Podosokorskaya et al., 2011)(99% identity, 99% query coverage, e-value 0).

RAST server counted 1,851 coding DNA sequences (CDSs), 42 tRNAs and 5 rRNAs. The genome comprises 70.1% of assigned putative functions, while the rest was annotated as encoding hypothetical proteins. Several genes were identified and associated with resistance to toxic compounds, as well as stress response to heat shock, or oxidation, which is coherent with the environmental origin of the strain.

Clustered regularly interspaced short palindromic repeats (CRISPRs) are present in all sequenced Thermotogales genomes (Zhaxybayeva et al., 2009), suggesting that Thermotogales have been exposed to numerous viral infections. In this context, a transposase and CRISPR loci associated with cas genes were found in *Thermosipho* sp. 1244 genome. Twenty-eight different CRISPRs associated genes were annotated, including cas 1, cas 2, cas 3, cas 4, cas 6 and cas 10 genes. This system seems to be functional indeed all essential genes were found in the genome (Makarova et al., 2015).

According to the Makarova classification (Makarova et al., 2015), this CRISPR cas system could be classified as a Class 1, characterized by the presence of a multi subunit CRISPR-RNA-effector complex (crRNA-effector complex). It also could be a good candidate for the Type III (like other sequenced and studied *Thermosipho* genomes, Haverkamp et al in prep.) because it contains Cas10 Large Subunit (LS) involved in the target cleavage of the system and also protein containing CARF domains which could be implicated in the regulation of the system. Supporting the hypothesis of Type III classification, several protein sequences involved in the multiprotein cr-RNA effector complex were founded. Notably numerous Csm and Cmr which are found in Type III A and B CRISPR cas systems. Concerning the subtypes, it could be classified in the Type III-A, supported by the presence of cas1, cas2, and cas6 genes (absent from III-B). Important genes with their putative functions are summarized in the table 1.

Table 1: Summary of important genes implicated in the CRISPR-cas system of *Thermosipho* sp.

1244

Protein	Function
Endonuclease Cas1	Spacer acquisition
Endoribonuclease Cas2	Spacer acquisition
Endonuclease/helicase Cas3	Nuclease – helicase, possible implication in the spacer acquisition and target degradation
Exonuclease Cas4, RecB family	Possible implication in the spacer acquisition and DNA reparation
Cas5, RAMP superfamily	Process and/or pre-crRNA stabilization
Endoribonuclease Cas6, RAMP superfamily	crRNA maturation and possible implication in target degradation
Cas10, large subunit of type III CRISPR-Cas systems, contains HD superfamily nuclease domain	Effector complex
Csm2, small subunit	Effector complex
Csm3, group 7 of RAMP superfamily	Effector complex
Csm4, group 5 of RAMP superfamily	Effector complex
Csm5, group 7 of RAMP superfamily	Effector complex
Cmr3, group 5 of RAMP superfamily	Effector complex
Cmr4, Cas7 group, RAMP superfamily	Effector complex
Cmr5, small subunit	Effector complex
Cmr6, Cas7 group, RAMP superfamily	Effector complex
Csh2, Cas7 group, RAMP superfamily	Effector complex
Csx1, contains CARF domain	Effector complex

Interestingly, a plasmid (pMP1, carried by *Marinitoga piezophila* KA3) could propagate itself in *Thermosipho* sp. 1244 using a viral capsid (molecular piracy of MPV1, the virus infecting *Marinitoga piezophila* KA3 (Lossouarn et al., 2015)). CRISPI server found 67 spacers in 4 loci. Those spacers were blasted against the nr virus database and also specifically against the three described *Thermotogales* viruses (MPV1, MCV1, MCV2). No perfect matches were shown but two spacers show matches with MCV1 and MCV2 (85% of identity between spacer and viral genes). These results could be an explanation for the impossibility for MCV1 or MCV2 to infect

this strain. It also shows that this *Thermosipho* strain has already been infected by MCV1, MCV2 or a similar virus. No matches were found corresponding to MPV1 or pMP1 which is consistent with the possibility for pMP1 to propagate itself in this *Thermosipho* strain.

Nucleotide sequence accession numbers.

Genome has been deposited at DDBJ/EMBL/ GenBank under the accession number X.

Materials and methods

Culture conditions

Thermosipho sp. 1244 was cultured at 60°C in a modified Ravot medium as previously described (Geslin et al., 2003; Lossouarn et al., 2015). The medium was completed by adding 1% (wt/vol) elemental sulfur.

Genome sequence and annotation

The genome was sequenced by using a combination of the Illumina and Ion Torrent platforms with a genome coverage of 30 X.

The annotation was done with RAST server (Overbeek et al., 2014) and CRISPR cas system was studied by CRISPI web server (Rousseau et al., 2009).

Acknowledgements

We thank the chief scientist of the French oceanographic ship "Atalante", Pierre-Marie Sarradin, the captain and crew of the NO L Atalante and the DSV Nautille pilots and support crew.

Funding informations

This work was financially supported by the French Ministry of Higher Education and the Brittany region to CM and the Agence Nationale de la Recherche, project Thermovesicles (ANR-12-BSV3-0023-01) to CG, a Norwegian Research Council award (project no. 180444/V40) to CLN.

References

- Bonch-osmolovskaya, E. (2008) Thermotogales: Phylogeny and Taxonomy of the Order. *Encycl. Life Sci.* 1–7.
- Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003) PAV1, the first virus-like particle isolated from a hyperthermophilic euryhalophilic cyanobacterium. *J. Bacteriol.* 185: 3888–3894.
- Ben Hania, W., Postec, A., Aüllo, T., Ranchou-Peyruse, A., Erauso, G., Brochier-Armanet, C., et al. (2013) *Mesotoga infera* sp. nov., a mesophilic member of the order Thermotogales, isolated from an underground gas storage aquifer. *Int. J. Syst. Evol. Microbiol.* 63: 3003–3008.
- Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., et al. (2016) A new view of the tree of life. *Nat. Microbiol.* 1: 16048.
- Lossouarn, J., Nesbø, C.L., Mercier, C., Zhaxybayeva, O., Johnson, M.S., Charchuck, R., et al. (2015) Mesotoga: a selfish genome that uses a virus to propagate. *Thermotogales. Environ. Microbiol.* 17: 3278–3288.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., et al. (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13: 722–736.
- Nesbø, C.L., Baptiste, E., Curtis, B., Dahle, H., Lopez, P., Macleod, D., et al. (2009) The genome of *Thermosiphon africanus* TCF52B: lateral genetic connections to the Firmicutes and Archaea. *J. Bacteriol.* 191: 1974–1978.
- Nesbø, C.L., Bradnan, D.M., Adebisoye, A., Dlutek, M., Petrus, A.K., Foght, J., et al. (2012) *Mesotoga prima* gen. nov., sp. nov., the first described mesophilic species of the Thermotogales. *Extremophiles* 16: 387–393.

- Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., et al. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42: D206–D214.
- Podosokorskaya, O.A., Kublanov, I. V., Reysenbach, A.L., Kolganova, T. V., and Bonch-Osmolovskaya, E.A. (2011) *Thermosipho affectus* sp. nov., a thermophilic, anaerobic, cellulolytic bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent. *Int. J. Syst. Evol. Microbiol.* 61: 1160–1164.
- Pollo, S.M.J., Zhaxybayeva, O., and Nesbø, C.L. (2015) Insights into thermoadaptation and the evolution of mesophily from the bacterial phylum Thermotogae. *Can. J. Microbiol.* 1–16.
- Rousseau, C., Gonnet, M., Le Romancer, M., and Nicolas, J. (2009) CRISPI: a CRISPR interactive database. *Bioinformatics* 25: 3317–3318.
- Zhaxybayeva, O., Swithers, K.S., Lapierre, P., Fournier, G.P., Bickhart, D.M., DeBoy, R.T., et al. (2009) On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. U. S. A.* 106: 5865–5870.

Chapitre V : Virome : validation et
normalisation d'un protocole (étude
préliminaire)

x Présentation des travaux

Les familles virales sont aussi nombreuses que diversifiées et ne disposent pas de gènes à eux seuls sous forme de l'ARN. Cependant, en utilisant des technologies de séquençage haut débit, il est maintenant possible de produire et séquencer des viromes et sélectionner des virus d'intérêt.

Peu de viromes sont disponibles dans les environnements marins profonds, où la diversité virale reste encore faiblement caractérisée (Williamson et al., 2008; Anderson et al., 2014; Brum and Sullivan, 2015). Ces premières études montrent que la diversité virale semble moins importante que dans d'autres systèmes comme la colonne d'eau pélagique. Par ailleurs, de nombreux gènes viraux découverts dans ces viromes ne montrent aucune similarité dans les bases de données, mettant en évidence le manque de connaissances sur ces virus mais aussi le fait qu'ils soient une source potentielle de nouveaux gènes à explorer (Williamson et al., 2008; Anderson et al., 2014).

C'est dans ce contexte qu'un projet transversal a été mis en place entre le LM E et le LUBEM (Laboratoire Universitaire de Biodiversité et Ecologie Microbienne) de Brest pour développer un protocole permettant d'obtenir des virus provenant de divers virus et de leur hôte. L'objectif au LM E a été de produire des virus provenant de souches hôtes au profond afin d'enrichir les connaissances sur ces environnements peu explorés. Au LUBEM, l'objectif est d'étudier les virus présents dans les épithéliums de patients atteints d'ulcère de l'œsophage pour évaluer l'effet des virus sur les communautés microbiennes.

Des cohortes de patients sains et de patients atteints de la mucoviscidose seront analysées afin d'étudier les variations de populations afin d'obtenir une vision globale de la présence de celles-ci dans l'écosystème pulmonaire humain.

Les thématiques de travail du LM2E et du LUBEM sont différentes tout comme les échantillons à traiter. Néanmoins, la volonté commune d'obtenir de nouvelles données sur des hôpitaux viau provient d'environnements complexes a mené à la réalisation de ce projet, permettant le partage des idées et des méthodologies. La base de ce projet était de tester deux protocoles différents (basés sur les méthodes de purification virale utilisées dans les deux laboratoires), l'un utilisant la filtration et l'autre utilisant la précipitation par l'acide trichloroacétique. Le but était d'identifier la méthode qui servira de base pour l'optimisation de virologie.

L'objectif de ce travail était de valider et d'optimiser un protocole qui servira de base pour l'optimisation de virologie.

x Origine des échantillons testés

Les sédiments hydrothermaux étudiés proviennent du bassin de Guaymas dans le Golfe du Mexique, au large de la Californie. Ils ont été échantillonnés dans le cadre de la campagne océanique IFREMER « BIG » (Biodiversité à Guaymas) en Juillet 2010 et conservés à -80°C. Afin d'avoir suffisamment de matériel à étudier, des échantillons ont été regroupés. Ils proviennent de 4 « horizons » différents, à partir de carottes échantillonnées à 4 profondeurs différentes dans la couche sédimentaire (0 à 2 cm : 1 échantillon, 2 à 4 cm : 2 échantillons, 6 à 8 cm : 1 échantillon et de 8 à 10 cm : 1 échantillon). Le pool contient donc 30 g de sédiments en suspension dans du tampon (10 mM Tris-HCl, 100 mM NaCl, 5 mM CaCl₂, pH 8), pour un volume total de 90 ml (soit une concentration de 0,33 g.l⁻¹).

Les expectorations choisies proviennent de patients des 2 sexes et de tous âges (de 18 mois à 80 ans), par conséquent le pool étudié représente davantage un intérêt technique du point de vue de la difficulté à travailler avec une substance visqueuse. Du fait de la viscosité des expectorations, celles-ci doivent être préalablement fluidifiées à volume avec du Sputagel® Mast Diagnostic (contenant du dithiothréitol) pour pouvoir prélever des volumes. Ce pool d'expectorations est conservé à -80°C.

x Micro-organismes et virus contrôles

Les échantillons ont ensuite été artificiellement chargés en particules virales et micro-organismes sous forme de suspensions et titrés afin de les suivre et d'évaluer l'efficacité des protocoles tests. Il s'agit notamment de levures (*Candida albicans*), de bactéries (*Pseudomonas aeruginosa*), de virus à ADN (le siphovirus MPV1 et un adénovirus) et à ARN (rhinovirus).

Organisme chargé	<i>P. aeruginosa</i>		<i>C. albicans</i>		<i>Adénovirus</i>		<i>MPV1</i>		<i>Rhinovirus</i>	
Concentration finale dans l'échantillon	Haute	Basse	Haute	Basse	Haute	Basse	Haute	Basse	Haute	Basse
Filtration	5,8.10 ⁶	5,8.10 ³	1,0.10 ⁵	1,0.10 ²	1,23.10 ⁶	1,23.10 ³	1,0.10 ⁷	1,0.10 ⁴	2,18.10 ³	2,18.10 ⁰
Gradient	5,24.10 ⁷	5,24.10 ⁴	8,0.10 ⁵	8,0.10 ²						

Tableau 1 : Concentrations en microorganismes cibles avec lesquelles les échantillons ont été chargés (en UFC/ml (unité formant colonie) ou TCID₅₀/mL (tissue culture infective dose 50))

- Protocoles

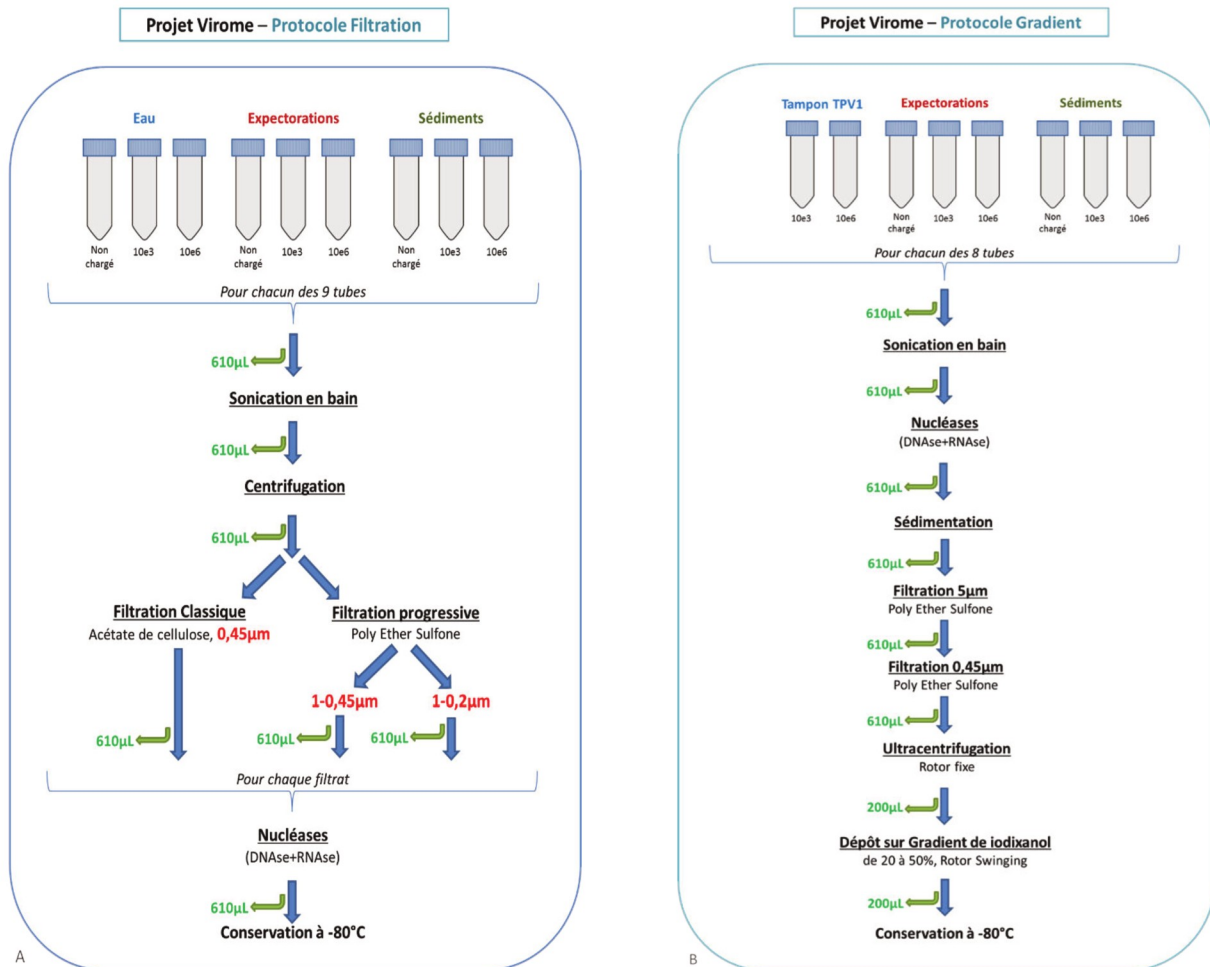


Figure 1 : Schéma général du protocole « Filtration » (A) et du protocole « Gradient » (B) Master Lisa Billard (juin 2015).

Afin de réaliser le suivi à la fois qualitatif et quantitatif, des contrôles ajoutés, des qPCR ont été réalisées aux différentes étapes du protocole. La technologie SYBR a été utilisée et la spécificité de l'amplification a donc été assurée par l'analyse de la courbe de fusion (ou « *Melt Curve* »). L'intérêt, ici, n'est pas d'obtenir une quantification précise du nombre de copies d'ADN, mais plutôt de comparer les Ct (cycle seuil) entre eux, afin d'étudier l'impact de chaque étape des deux protocoles étudiés sur l'enrichissement viral.

Les défis majeurs résident en la conservation et la détection des virus, faiblement abondants et fragiles. Les virus à ARN, les plus fragiles, sont ceux qui présentent le plus grand défi technique. Ces travaux auxquels j'ai pu participer du début à la fin, lors de l'admission et du stage de Master 1 de Julia Sanchez, sont maintenant réalisés dans le cadre de la thèse de Lisa Billard au LUBEM (résultats présentés ci-dessous, Projet Virome porté par le Pr C. Payan et le Dr S. Vallet, LUBEM).

x Premier résultats

Un premier essai de séquençage haut débit (HiSeq 2X125 bp, GATC, 5 millions de reads) a été réalisé sur des échantillons d'expectorations humaines et d'organismes contrôles (cités ci-dessus). Il s'agit d'une expérience réalisée sur des échantillons bruts, non purifiés, mais chargés artificiellement avec les contrôles. Ainsi, les extractions et rétro-transcriptions ont été menées sur ces échantillons bruts chargés afin d'observer les résultats du séquençage sur les échantillons bruts. Ce séquençage permettra par la suite de comparer ses résultats aux échantillons ayant subi les protocoles de purifications, pour observer leur efficacité. Ces premiers résultats de séquençages hauts débit ont montré que, pour des raisons techniques, les données de séquençage des virus à ADN sont de meilleure qualité que celles des virus à ARN. Ces résultats concordent avec ceux obtenus dans la littérature (Losada et al., 2015; Wang et al., 2016). Les quelques reads pour lesquels une assignation taxonomique a été possible ont permis d'identifier des virus à ADN, au contraire des virus à ARN qui n'ont pu être identifiés hormis des rétrovirus probablement intégrés au génome humain.

Il s'agit également de atiovirus appartenant à la famille Siphoviridae infectant des bactéries habituellement retrouvées dans le système respiratoire des patients atteints de mucoviscidose, mais aussi de quelques virus humains (Cystic Fibrosis Foundation, Patient Registry, Annual Data Report 2012).

x Avancement des travaux

Après ce premier essai, le protocole a évolué et la méthode retenue se déroule en 5 étapes. Tout d'abord, les échantillons sont soumis pour détecter les pathogènes. Ensuite ceux-ci sont centrifugés et filtrés, afin d'éliminer la majorité des micro-organismes et particules présents. La méthode choisie pour l'utilisation de gènes, a été retenue car sa mise en œuvre est facile et les premiers résultats de suivi PCR étaient pas concluant. Un traitement aux nucléases est aussi réalisé pour éliminer les acides nucléiques. Enfin, après l'étape de RT, une transcription et une amplification aléatoire (par SIA : sequence independant amplification) sont réalisées permettant une meilleure détection des différents virus présents (Données communiquées et issues du travail de master 2 et de thèse de Lisa Billard).

L'objectif est ainsi de valider le protocole de diagnostic et d'étudier les suites issues d'échantillons naturels (non chargés en micro-organismes et virus contrôles) pour, par la suite, appliquer ce protocole aux collections de sécrétions respiratoires de patients atteints de mucoviscidose. Ces études permettront d'avoir une meilleure vision des populations présentes dans le microbiote des patients et de prédire l'impact des interactions potentielles entre les micro-organismes et les virus présents au sein de ce microbiote.

Les avancées de ce travail seront mises à profit, par le LM2E, afin de réaliser de nouveaux travaux à partir d'échantillons humains.

x Bibliographie

- Anderson, R.E., Sogin, M.L., and Baross, J.A. (2014) Evolutionary strategies of viruses, bacteria and archaea in hydrothermal vent ecosystems revealed through metagenomics. *PLoS One* 9: e109696.
- Brum, J.R. and Sullivan, M.B. (2015) Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat. Rev. Microbiol.* 13: 147–159.
- Losada, P.M., Chouvarine, P., Schulz, A., Hedtfeld, S., Mielke, S., Dorda, M., et al. (2015) The microbial metagenome of cystic fibrosis lower airways. *J. Cyst. Fibros.* 14: S66.
- Wang, Y., Zhu, N., Li, Y., Lu, R., Wang, H., Liu, G., et al. (2016) Metagenomic analysis of viral genetic diversity in respiratory samples from children with severe acute respiratory infection in China. *Clin. Microbiol. Infect.* 22: 458-e1.
- Williamson, S.J., Rusch, D.B., Yooseph, S., Halpern, A.L., Heidelberg, K.B., Glass, J.I., et al. (2008) The Sorcerer II Global Ocean Sampling Expedition: Metagenomic Characterization of Viruses within Aquatic Microbial Samples. *PLoS One* 3: e1456.

Discussion et Perspectives

Les virus sont les entités biologiques les plus abondantes de la planète (Suttle, 2005). Cette abondance est particulièrement étudiée dans les océans où ils sont un important réservoir de diversité génétique. En effet, il est estimé que les océans contiennent plus de 10²³ particules virales et que 1 kilogramme de sédiment marin en comprendrait 1 million. Nous, aujourd'hui, ne connaissons qu'une petite partie de la diversité virale (Breitbart and Rohwer, 2005).

Par ailleurs, même si la diversité locale des génotypes viraux est très importante, il apparaît que celle-ci est aussi élevée à l'échelle globale (Breitbart and Rohwer, 2005; Chow and Suttle, 2015). Diverses études en métagénomique montrent en effet que beaucoup de génotypes viraux sont ubiquistes, et qu'ils sont distribués dans différents types d'écosystèmes et lieux géographiquement éloignés, et ceci via le transfert des particules virales ou des provirus. Cette ubiquité des virus est l'aboutissement de l'adaptation, celle-ci étant dépendante des conditions environnementales et des hôtes présents (Breitbart and Rohwer, 2005; Brum et al., 2015; Paez-Espino et al., 2016).

Chez les bactéries et notamment chez les Thermotogae, il a été montré que les génomes provenant d'un même type d'écosystème avaient tendance à partager plus de gènes et à être plus similaires que ceux issus d'écosystèmes géographiquement proches mais aux conditions physico-chimiques différentes (Nesbo et al., 2015). En effet, dans cette étude, des souches hyperthermophiles provenant de puits de pétrole et de sites marins éloignés de sources hydrothermales situés dans l'océan Pacifique ou l'Atlantique ont été comparées. Les résultats présentés montrent que les génomes provenant du même type d'écosystème (puits de pétrole ou sites hydrothermaux marins) ont tendance à montrer plus de similarité et à échanger plus de gènes que ceux qui sont géographiquement proches mais issus d'écosystèmes différents (Figure 1).

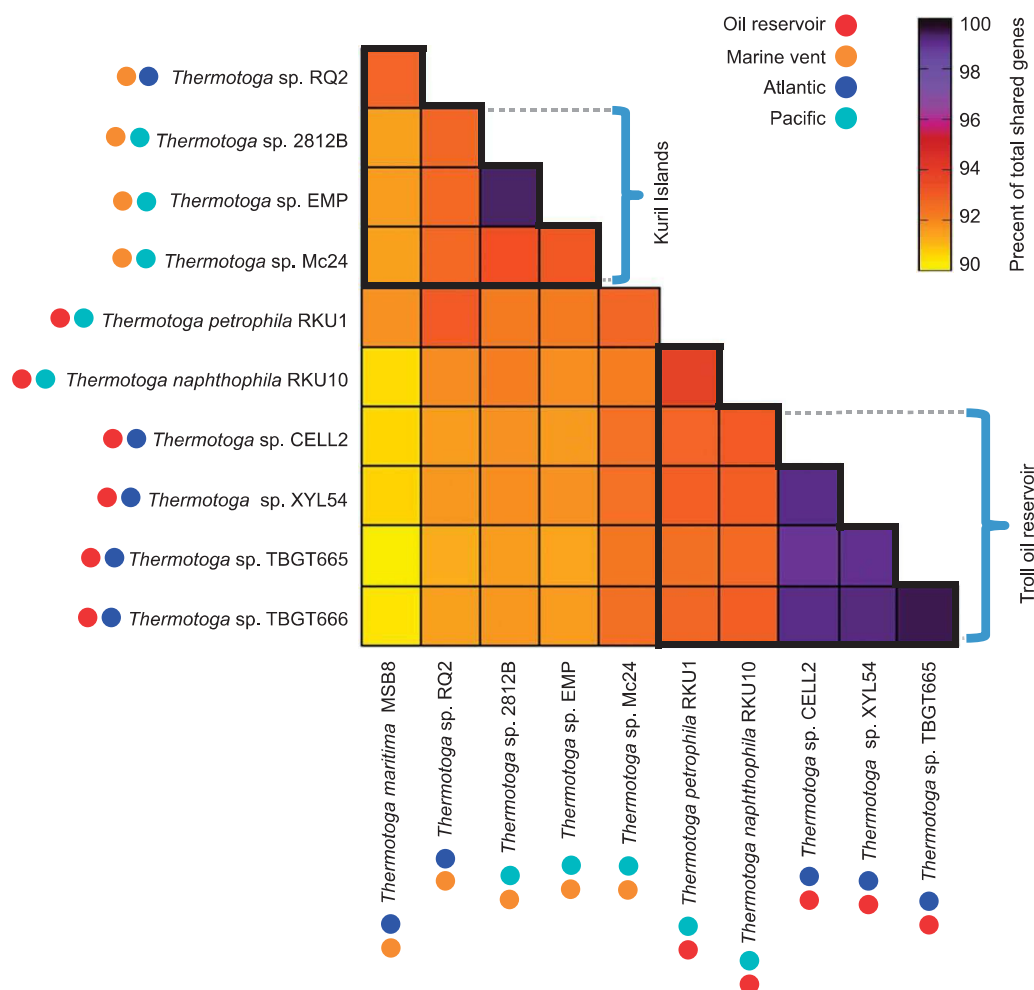


Figure 1: Comparaison de 11 génomes de *Thermotoga* maritima-like. Les génomes sont classés par rapport à la position géographique du lieu où ils ont été isolés (Océan Atlantique ou Pacifique) et par rapport à l'environnement où ils ont été prélevés (puits de pétrole ou site hydrothermale) (Nesbo et al., 2015).

Il semble donc que, pour les virus comme pour leurs hôtes, la distance géographique ne soit pas un facteur limitant les échanges. Le même constat a été réalisé lors de nos travaux concernant la comparaison des trois génomes viraux issus de différentes souches de *Marinitoga*.

I. Deux nouveaux virus, MCV1 et MCV2, isolés du phylum des Thermotogae et comparaison avec MPV1

Ces trois virus infectent des bactéries appartenant au phylum des Thermotogae, phylum bien représenté notamment dans les sources hydrothermales océaniques profondes. Néanmoins ce phylum comprend aussi des micro-organismes et ouvre des portes à d'autres sites chauds répartis sur l'ensemble de la planète comme les sources chaudes terrestres, les puits de pétroles, ou les sédiments pollués (Pollo et al., 2015). Malgré les nombreuses preuves de transferts latéraux de gènes dans les génomes de Thermotogae, peu d'éléments génétiques mobiles ont été décrits pour cet ordre (Lossouarn, Dupont, et al., 2015). Ainsi les virus MPV1, et maintenant MCV1 et MCV2, représentent les seuls virus décrits comme infectant une bactérie appartenant à ce phylum (Lossouarn, Nesbø, et al., 2015). Ces trois virus, MPV1, MCV1 et MCV2 proviennent de trois lieux géographiquement éloignés. MPV1 qui infecte *M. piezophila* provient d'une source hydrothermale océanique profonde située le long de la dorsale du Pacifique Est à 2630 m de profondeur (Alain et al., 2002). MCV1 et MCV2 infectent deux souches de *M. camini* (respectivement DV1197 et DV1155) isolées de sites hydrothermaux profonds situés dans le long de la dorsale médio-atlantique. *M. camini* DV1197 provient du site Lucky Strike situé à une profondeur de 2630 m et *M. camini* DV1155 provient du site Menez Gwen situé à une profondeur allant de 840 à 870 m (Colaço et al., 1998). Néanmoins, ces souches thermophiles proviennent du même type d'environnement, c'est-à-dire des sources hydrothermales océaniques profondes, notamment caractérisées par de forts gradients physico-chimiques, l'absence de lumière et une importante pression hydrostatique.

Ainsi, ces trois virus, malgré leur éloignement géographique, présentent de nombreuses similarités comme décrit précédemment dans la partie résultats (Chapitre 3) de ce manuscrit. En se basant sur des critères morphologiques, ils ont été classés dans la famille Siphoviridae appartenant à l'ordre des Caudovirales. Ce sont trois virus tempérés mettant en œuvre un cycle de vie lysogène. Ils partagent aussi de nombreuses similitudes génomiques. Ils possèdent tous un ADN double brin d'une taille allant de 15 à 25 kbp. L'organisation génomique est similaire, de 5' à 3' : réplication et transcription, encapsidation, assemblage des virions, lyse de l'hôte et éjection. Un « core genome », composé de 35 ORFs, a été identifié au sein des trois génomes. La majeure partie de ces ORFs codent pour des protéines hypothétiques mais certains ont pu être annotés. Parmi ces protéines est retrouvée une protéine de structure : la protéine majeure de capsid. Celle-ci se localise sur la surface de la particule virale car elle est constituée à plus de 50 % des trois protéines. Une autre protéine impliquée dans la fabrication des virions a aussi été identifiée avec une forte similarité : 80 %, il s'agit de la gène de la terminase. Les terminases virales sont des enzymes qui permettent l'assemblage du génome viral et permettent l'éjection des nouveaux virions (Feiss and Catalano, 2005). La petite sous-unité de la terminase n'a pu être identifiée dans aucun des génomes, il est possible qu'elle soit présente mais trop différente de celles retrouvées dans les bases de données pour être identifiée. Des protéines impliquées dans le métabolisme des acides nucléiques font aussi partie de ce « core genome », notamment une hélicase. Une résolvase de jonctions de Holliday de la famille RusA a été identifiée dans les trois génomes viraux. Ces enzymes sont notamment impliquées dans le processus de recombinaison homologue où elles clivent les jonctions de Holliday liant deux brins d'ADN (Macmaster et al., 2006). Une **protéine de liaison à l'ADN** a été mise en évidence sans pouvoir être identifiée.

Une protéine impliquée dans la lyse et l'œuvre du *le* *Iti* *ue* est aussi *it* *ouv* *e*, il s'agit de la protéine holine qui permet la lyse cellulaire en créant des pores dans la membrane. Enfin un régulateur de transcription semble être partagé par ces trois génomes.

L'identité de ce core genome comprenant entre 45 et 65% des ORFs de ces génomes viraux laisse supposer un lien évolutif entre ceux-ci, malgré la distance géographique qui les sépare. Par ailleurs, ces travaux laissent aussi supposer la possibilité d'une co-évolution entre ces virus et les Firmicutes ainsi que les virus les infectant. Des travaux portant sur les transferts latéraux de gènes via l'utilisation de *de* *o* *a* *teu* *s* *pouaie* *t* *t* *e* *effe* *tu* *s*. En plaçant des co-cultures de différents micro-organismes (comme par exemple des *Thermotogae* portant des provirus et des Firmicutes), il se ait possible d'étudier les relations potentielles ainsi que les échanges de gènes entre les différentes espèces via les éléments génétiques mobiles tels que les virus. Différentes conditions pourraient être testées (présence ou absence de virus, d'inoculum, etc...). Il est possible d'étudier systématiquement toutes les interactions possibles et leurs effets.

II. Thermotogae et CRISPR

Les Thermotogae ont été la cible de nombreux transferts latéraux de gènes au cours de leur évolution. Leur phylogénie particulière et la composition de leurs génomes en portent de nombreuses traces. En effet, ces bactéries thermophiles sont proches des Aquificales selon les études basées sur l'ARN ribosomal et les protéines ribosomales. Néanmoins, elles présentent de nombreuses similarités avec les Firmicutes mais aussi, dans une moindre mesure, avec les archées (Zhaxybayeva et al., 2009; Gupta and Bhandari, 2011). Il est connu que les infections virales sont des vecteurs majeurs des transferts latéraux de gènes. D'ailleurs, les études de des génomes viraux présentées dans le chapitre 4, d'importance en relation avec des gènes d'autres phylotaxons et notamment avec des Firmicutes et les virus qui leurs sont associés. Parmi les gènes présentant des similarités avec ceux retrouvés dans les bases de données et après exclusion des « self-matches » avec des Thermotogae, plus de 60% des ORFs de ces génomes ont montré une similarité importante avec les Firmicutes ou leurs virus. Ces résultats mettent en évidence une possible co-évolution entre les Firmicutes et les Thermotogae. D'autre part, cela implique aussi que les virus présents dans les échantillons pourraient être en grande partie responsables des nombreux transferts horizontaux de gènes observés au sein des échantillons de l'océan des Thermotogae. De manière générale, il semble que, dans les sources hydrothermales profondes les transferts latéraux de gènes soient plus nombreux que dans d'autres environnements marins. En effet, presque 60% des génomes séquencés sont issus de ces sources hydrothermales profondes. D'autre part, de nombreux gènes impliqués dans les transferts latéraux de gènes (comme les intégrases) sont retrouvés dans les métagénomes issus de ces environnements extrêmes (Anderson et al., 2014; Anantharaman et al., 2016).

Même si au uvis 'avaitt is e vide e avat MPV , des p euves d'ions passées avaient été détectées. En effet, des systèmes CRISPR-cas ont été retrouvés dans tous les génomes de Thermotogae séquencés à ce jour (Nesbø et al., 2009; Zhaxybayeva et al., 2009; Pollo et al., 2015).

Le système CRISPR- cas est u s st e de d fe se desat ies et a h es o te l'i vasio par des éléments génétiques mobiles et notamment les virus (Horvath and Barrangou, 2010). Ces systèmes sont très répandus au sein des génomes bactériens en général, puisque 50% d'e t e eu e o tie e t, et e o e plus da s les g es d'a h es ui so t % à po te un système CRISPR-cas (Makarova et al., 2015). Il a aussi été décrit que ce système de défense était particulièrement important pour les thermophiles qui contiennent proportionnellement plus souvent des systèmes CRISPR-cas que les mésophiles (Sorek et al., 2008; Chellapandi and Ranjani, 2015). Ils sont caractérisés par la présence de courtes séquences répétées entre les uelles so t s des spae s. Ces spae s so t des paties d'ADNes qe ettat la reconnaissance des éléments génétiques mobiles envahissant la cellule. Ces séquences sont associées à des gènes cas CRISPR asso iated se ue ui pe ette t l'iativatio des acides nucléiques exogènes. Ils sont aussi utilisés pour classer les différents types de systèmes en classes, types et sous types (Horvath and Barrangou, 2010; Chellapandi and Ranjani, 2015; Makarova et al., 2015). Les types I et III sont les plus répandus. Chez les thermophiles les sous-types les plus représentés sont les I-A, I-B, III-A et III-B (Lange et al., 2013; Makarova et al., 2015).

Lors de la caractérisation de MPV1, virus isolé de *M. piezophila*, il a été montré que le plasmide pirate pMP1, également porté par *M. piezophila*, était capable de se propager dans la souche AT1244-VC14 appartenant au genre *Thermosiphon* (Lossouarn, Nesbø, et al., 2015).

Ce constat est partiel et il est intéressant de noter que des tests fonctionnels par pMP et MPV1 ont été menés sur diverses souches de *Marinitoga*, plus proches phylogénétiquement de leur hôte naturel, mais ceci sans succès. Le plasmide pMP1 est donc capable de se propager dans une souche d'une génétique différente de son hôte initial *Thermosiphon* sp. AT1244-VC14. Cette souche a été isolée de la dorsale médio-atlantique à 2000 m de profondeur au niveau d'une source hydrothermale active. Les résultats montrant la capacité de propagation de pMP1 dans cette souche *Thermosiphon* sp. AT1244-VC14 (contrairement aux *Marinitoga* testées) ont conduit au séquençage de celle-ci. Cette étude a permis de mettre en évidence la présence d'un système CRISPR-Cas, comme pour toutes les autres souches séquencées de *Thermotogae*. Ce système paraît fonctionnel puisqu'il oriente tous les génomes candidats à sa mise en œuvre.

Le processus de fonctionnement du système CRISPR-Cas se déroule en deux grandes étapes (Figure 1). La première est l'acquisition du spacer dans laquelle sont intégrés les protospacers et Cas. Lorsqu'une aide moléculaire est requise et que dans la cellule, il a eu lieu l'acquisition d'un motif PAM protospacer adjoint (motif) qui est adjacent au protospacer (spacer encore dans l'aide moléculaire). Cette acquisition permet de livrer le protospacer pour la suite. Il s'agit de la séquence des protospacers du locus CRISPR. La seconde étape est l'intégration avec l'aide nucléaire exogène. Lorsque ce dernier est reconnu par le système de défense, il y a transcription de la séquence CRISPR correspondante. Le produit généré correspond à un long transcrit précurseur appelé pre-crRNA qui est ensuite modifié pour donner le crRNA CRISPR RNA. L'existence est attestée et impliquée dans le processus. Le RNA est associé ensuite à diverses protéines Cas (notamment les protéines issues des gènes cas10, csm et cmr pour le type III) pour donner un complexe stable.

C'est le principe qui va se faire à l'aide d'une enzyme pour permettre sa dégradation (Horvath and Barrangou, 2010; van der Oost et al., 2014; Rollie et al., 2015).

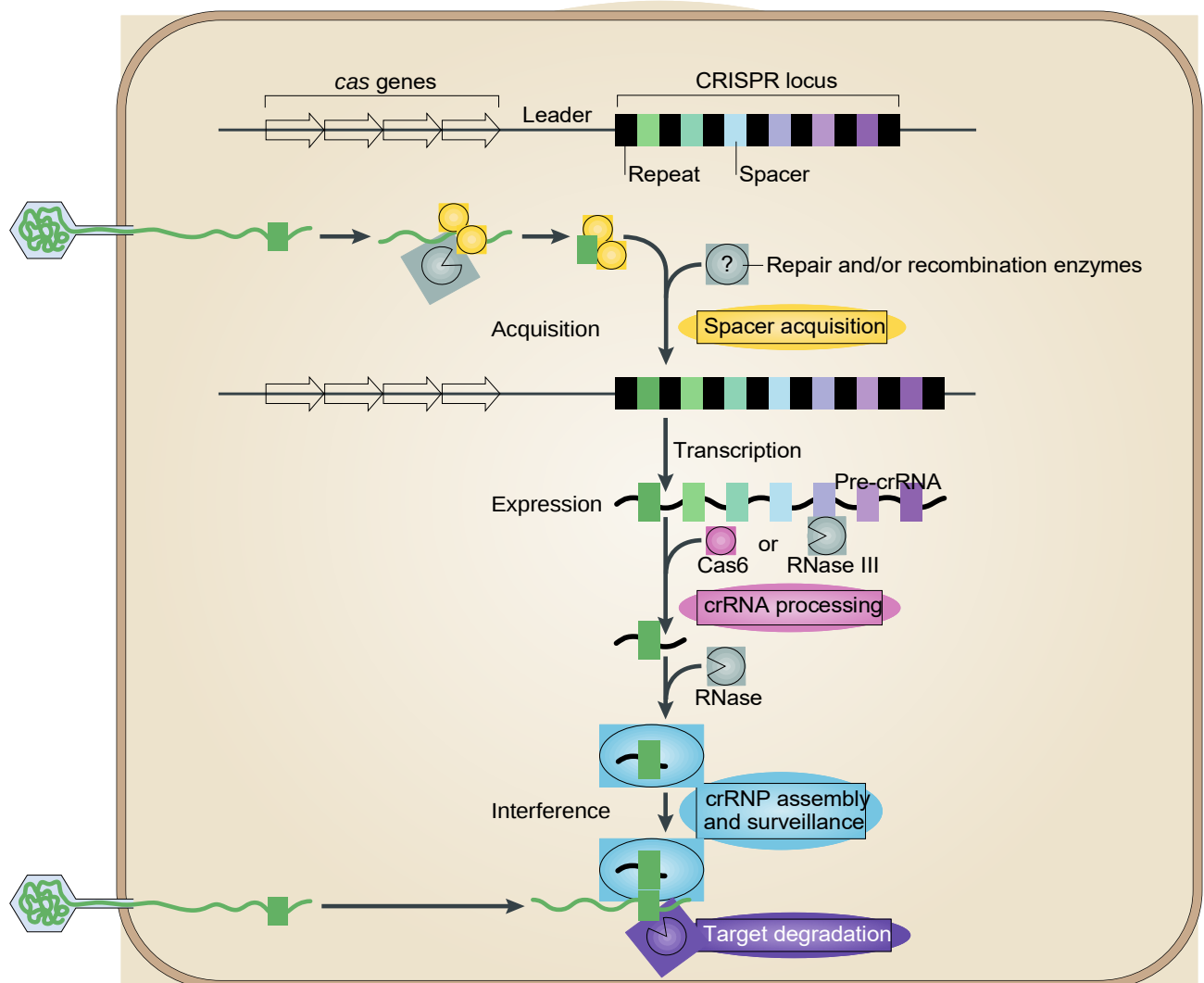


Figure 2 : Schéma du principe de fonctionnement générale d'un système CRISPR. La première partie du schéma représente l'acquisition d'un nouveau spacer et la seconde partie représente l'interférence avec un acide nucléique exogène. Les gènes cas et leurs produits protéiques sont colorés selon leurs fonctions (jaune) acquisition du spacer, (rose) mise en place (« processing ») du crRNA, (bleu) assemblage et stabilité du complexe effecteur, (violet) dégradation de la cible. Les éléments gris sont des produits qui ne sont pas issus des gènes cas. Le génome viral est vert et le protospacer est symbolisé par un rectangle vert sur celui-ci.

Le système présent dans la souche *Thermosipho* sp. AT1244-VC14 peut être classé comme système de Classe 1, type III-A (Makarova et al., 2015). Le génome contient en effet le gène signature de la classe III : *cas10* (grande sous unité). Le séquençage a aussi révélé la présence des deux gènes signatures des types III-A et III-B, respectivement *csm2* et *cmr5*. Les gènes codant pour les protéines *cas5* et plusieurs protéines du groupe *cas7* (comme *cmr4* ou *cmr6*) ont aussi été identifiés et font partie des éléments retrouvés dans tous les CRISPR-cas III-A et III-B. Néanmoins la présence des gènes *cas1*, *cas2* et *cas6*, classe ce système dans le type III-A car les autres systèmes de types III ne contiennent généralement pas ces gènes (Makarova et al., 2015).

Les spacers contenus dans ce système ont été étudiés notamment grâce au server CRISPR (Rousseau et al., 2009). Au uat h 'a t ide tifi pou MPV et pMP qui estoh ent avec la possibilité d'i fe tio de la sou he pas l e ts g ti ues o iles. Pa o t e, matchs au dessus de 85% de similarités ont été retrouvés pour MCV1 et MCV2. Ce constat est oh e t ave l'i possiilit pou es vius d'i fe te *Thermosipho* sp. AT1244-VC14. Toutefois, nous ne pouvons conclure que la non-infection soit le seul fait de la présence de ces séquences spacers. Ceci met aussi en évidence que cette souche de *Thermosipho* a sans doute déjà été en contact avec ces virus ou des virus proches de ceux-ci.

III. Importance de la lysogénie et virome

L'importance de la lysogénie dans les océans a longtemps été sous-estimée du fait des moyens de détection utilisés et qui consistaient globalement en l'utilisation de plaques. Par ailleurs, le faible nombre de bactéries et archées cultivables (moins de 1% de la diversité estimée) et l'absence de gènes aux séquences variables pour l'ARN Spa et le ple de l'exploitation de cette diversité. Néanmoins, la rapide évolution des méthodes de séquençage a permis d'obtenir de plus en plus de gènes de attaches et d'adhésions aussi de nouveaux tag variables appelés viroes. Malgré le reste des défis à jouer dans l'exploitation de ceux-ci, ils permettent une meilleure compréhension et une vision plus globale de la répartition des virus dans les environnements (Brum and Sullivan, 2015; Chow and Suttle, 2015; Paez-Espino et al., 2016).

Il y a encore peu de viromes disponibles au niveau des environnements marins profonds. Néanmoins, quelques études ont permis de mettre en évidence l'abondance et la diversité des génomes viraux dans ces environnements (Williamson, Rusch, et al., 2008; Anderson et al., 2014). Les études menées sur des métagénomes par Williamson, Anderson et leurs collaborateurs ont aussi permis de mettre en évidence l'importance de la lysogénie dans les sources hydrothermales océaniques profondes. En effet, il apparaît que ce cycle viral serait plus répandu que celui des ostéoblastes hydrothermalaux au sein de la colonne d'eau. Ainsi, dans ces conditions extrêmes les virus auraient tendance à rester intégrés de façon stable dans les génomes cellulaires (Williamson, Cary, et al., 2008; Williamson, Rusch, et al., 2008; Anderson et al., 2013, 2014). La prépondérance de virus tempérés au niveau des sources hydrothermales, notamment révélée par des études métagénomiques, peut être considérée comme une stratégie de survie face aux conditions environnementales extrêmes et variables.

Sachant que la dégradation des virions en environnement marin profond (viral decay) est importante (pour MCV1, survie de moins de 5 heures hors de son hôte), il paraît cohérent que les virus ont plus de chances de survie à l'intérieur de leur hôte plutôt qu'à l'extérieur. En restant intégrés dans les génomes cellulaires les virus se permettent de la possibilité d'être dégradés s'ils se trouvent pas un autre hôte facile à atteindre. Ils pourraient ainsi être aussi protégés du broutage et des conditions environnementales difficiles (Paul, 2008). Cette relation de type lysogénique peut être vue comme symbiotique puisqu'elle apporte aussi des

benefices à l'hôte qui porte le prophage. En effet, l'adaptation de la cellule à son environnement (fitness) peut en être améliorée et des échanges de gènes menant à une coévolution positive dans l'évolution peut être mis en oeuvre. Le fait de porter un prophage offre aussi à la cellule un type d'immunité contre les autres virus proches de celui qui est intégré dans son génome (Paul, 2008; Breitbart, 2012). Les virus découverts dans ces métagénomes portent aussi souvent des gènes appelés AMG (Auxiliary Metabolic Genes, soit gènes auxiliaires du métabolisme, qui offrent l'hostilité de l'immunité à l'interface entre les deux entités (Anderson et al., 2014; Rosenwasser et al., 2016). Ces gènes sont impliqués dans le métabolisme énergétique, notamment ceux du soufre et du méthane mais aussi dans la biosynthèse des acides aminés et la réplication/réparation/recombinaison de l'ADN (Anderson et al., 2014; Anantharaman et al., 2016).

C'est dans ce contexte qu'un projet transversal a été mis en place entre le LM E et le LUBEM (Laboratoire Universitaire de Biodiversité et Ecologie Microbienne) de Brest pour développer un protocole permettant d'obtenir des virus provenant de divers environnements. Au LM E, l'objectif est de produire des virus à partir de sédiments hydrothermaux profonds pour élargir les connaissances sur les communautés virales présentes dans ces environnements peu explorés.

Au LUBEM l'objectif est d'étudier les différents virus et les rôles de ces communautés sur les patients ainsi que sur leur microbiote.

J'ai participé, via l'encadrement du stage de master de Julia Sanchez, à la mise en place d'un premier protocole. Ces travaux portés, par le Pr Payan et le Dr S. Vallet se poursuivent maintenant dans le cadre de la thèse de Lisa Billard au LUBEM. Le protocole pour produire un virus est en passe d'être mis au point. Il permettra ainsi d'obtenir des échantillons où la majorité des cellules sont éliminées tout en conservant au maximum les particules virales présentes. Le séquençage des échantillons permettra de mettre en évidence les virus présents dans les divers échantillons traités et ainsi améliorer notre compréhension de leur impact dans les environnements étudiés.

C'est dans cette direction qu'au LM E, de nouveaux échantillons hydrothermaux marins profonds vont être étudiés via une approche de métagénomique. En effet, afin de prospecter la biodiversité des écosystèmes extrêmes, des membres du laboratoire embarquent régulièrement à bord de navires océanographiques pour des campagnes à la mer. Les échantillons issus de ces campagnes océanographiques permettront donc de réaliser de nouveaux viromes, en bénéficiant du travail mené en amont. Cette approche permettra d'élargir nos connaissances sur le mobilome abyssal, qui reste encore à ce jour assez peu connu. Ainsi, ce type d'étude permet d'avoir accès à une vision globale de ces séquences génétiques mobiles présentes dans l'environnement. Par ailleurs, il est possible d'associer des travaux de caractérisation de nouveaux modèles hôtes virus. Ces travaux permettront de faire le lien entre les hôtes et les EGMs qui leur sont associés et de fournir de nouvelles données pour étudier les rôles de ces derniers.

IV. Vésicules membranaires et détoxification

Les vésicules extracellulaires MVs font partie d'un plus grand sécrét appelé omobilome (comprenant les virus, les plasmides, les MVs, les GTAs, les transposons). Elles sont nombreuses notamment dans les océans où leur abondance a récemment été estimée à $10^5/10^6$ MVs/l d'eau de mer (Scanlan, 2014). La production de vésicules membranaires est un mécanisme universel de communication cellulaire, connu chez les eucaryotes et les bactéries. Elles jouent un rôle dans le transfert de matériel génétique, d'agents de défense, de facteurs de pathogénicité, de transcrits d'ARN et d'ADN, dans la défense contre les attaques virales, mais très peu documenté chez les bactéries thermophiles et les archées (Scanlan, 2014; Jagannadham and Chattopadhyay, 2015; Lossouarn, Dupont, et al., 2015).

C'est lors de divers essais in vitro viraux au laboratoire que nous avons observé la production de MVs chez la souche *Marinitoroga camini* DV1197, hôte du virus MCV1.

Initialement, l'objectif était d'étudier la production virale en présence de stress environnementaux. Les stress pouvaient être appliqués et présentés dans l'environnement où la souche a été isolée). En effet, les inductions étaient, jusqu'alors, réalisées à l'aide d'un antibiotique, la mitomycine C. Hors, cet antibiotique ne représente en aucun cas un stress environnemental. C'est à partir de ce constat que nous avons essayé d'étudier la production virale à l'aide de différents traitements appliqués et présentés dans les souches hydrothermales profondes. Le cadmium, le zinc, le cuivre et cobalt ont été testés à diverses concentrations de l'ordre du mM) sur les cultures en début de phase exponentielle. Les échantillons induits ont ensuite été concentrés et observés en MET (Figure 1). C'est alors qu'a été découvert non pas une induction de la production virale mais une forte production de MVs en réponse à l'ajout de cuivre et de cadmium.

Pour que cette production soit induite, il a fallu utiliser des concentrations en métaux beaucoup plus importantes (0.5 mM) que celles généralement mesurées dans l'environnement de l'océan (Charlou et al., 2000)). Néanmoins, ces concentrations élevées de métaux peuvent vaie d'un site à l'autre, et les facteurs de stress déclenchant ces réactions de défense peuvent donc être multiples.

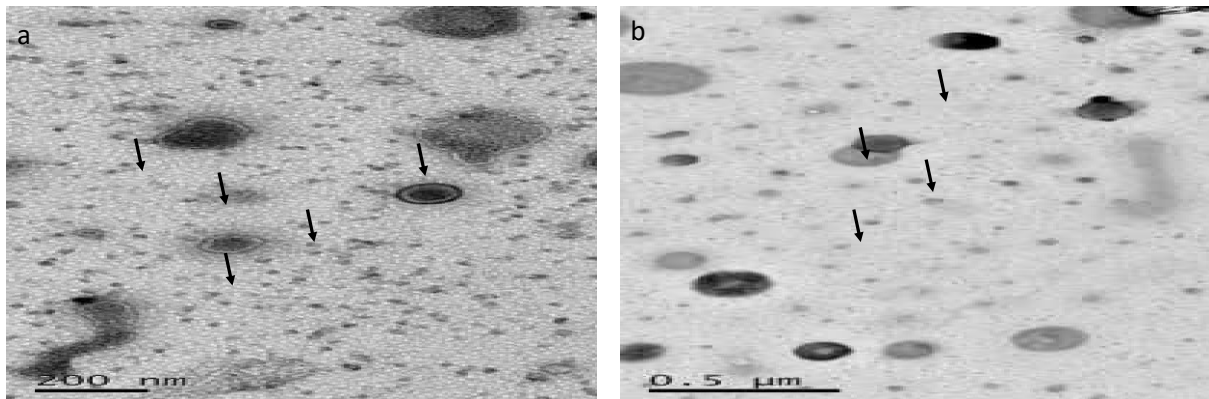


Figure 3 : Observation en microscopie électronique de MVs produites par *M. camini* DV1197 suite à l'ajout de métaux lourds. Concentration de MVs suite à l'ajout de cadmium (b) 0,1 mM de cuivre.

L'importance des MVs est démontrée par des études réalisées sur des organismes modèles. Elles sont, par exemple, impliquées dans un mécanisme de détoxification du fer chez divers eucaryotes, notamment des plantes ou des protozoaires en séquestrant celui-ci dans de petits compartiments intracellulaires (Slavic et al., 2016). Chez les bactéries, elles peuvent être impliquées dans la résistance aux antibiotiques ou aux désinfectants (Jagannadham and Chattopadhyay, 2015). Dans les océans, elles peuvent participer à la détoxification de divers métaux lourds tels que le cadmium chez des crustacés ou des mollusques (Sterling et al., 2007).

Au niveau des sources hydrothermales profondes, il a été montré que divers symbiontes pouvaient, grâce à la production de MVs, jouer le rôle d'hôtes de l'accumulation de composés soufrés tout en utilisant celles-ci pour le stockage de gènes (Pflugfelder et al., 2005; Eichinger et al., 2014). Les MVs contenant du soufre (aussi appelées globules de soufre) est un phénomène décrit de longue date et bien connu chez les bactéries (Dahl and Prange, 2006). Ce type de MVs a aussi été retrouvé chez des archées de l'ordre des Thermococcales provenant de sources hydrothermales profondes, environnement riche en composés soufrés et en divers composés toxiques (Gorlas et al., 2015). Dans ce cas aussi, les MVs seraient impliquées dans un mécanisme de détoxification.

C'est donc à noter que l'hypothèse selon laquelle ces vésicules produites en réaction à l'ajout de soufre se seraient constituées de détoxification a été mise en doute. D'autres tests ont pu être réalisés pour évaluer la production de MVs chez des souches d'archées thaïogènes (Post-doctorat de Samuel Dupont, Thèse de Sarah Thiroux). D'autre part, il se leait que lorsque cette production de MVs est déclenchée, tant chez les Thermotogae que chez les archées méthanogènes, celle-ci inhibe la production de particules virales. En effet, lorsque de nombreuses MVs sont observées en MET suite à l'ajout de soufre, très peu, voire aucune particule virale n'est détectée. Il est donc possible que les productions soient régularisées via un effet potentiellement antagoniste.

Afin de poursuivre l'effort de caractérisation de nouveaux éléments génétiques mobiles (EGM) issus des sources hydrothermales profondes au laboratoire de nouvelles études vont être réalisées notamment via la thèse de Sarah Thiroux débutée en octobre 2016. Celles-ci porteront sur la production de MVs ainsi que leur potentielle implication dans la détoxification des métaux. Il s'agit notamment d'exploiter les capacités de séparation (ICP-MS) pour mettre en évidence la présence de métaux lourds

Ces expériences permettront de vérifier la présence des métaux au sein des MVs et ainsi de valider leur implication dans un mécanisme de détoxification. La présence d'aides uli ues dans les vésicules sera aussi recherchée afin d'étudier leur potentielle implication dans les transferts latéraux de gènes et de l'évolution des communautés microbiennes qui les produisent. Enfin, des analyses en microscopies seront réalisées. Celles-ci auront pour objectif de connaître l'influence des MVs sur les relations symbiotiques et les interactions possibles entre ces différents éléments.

Par ce nouveau projet nos connaissances sur la diversité moléculaire des éléments génétiques mobiles (EGMs : virus, vésicules, plasmides) seront améliorées. Cela permettra également de comprendre leurs impacts sur les communautés microbiennes (adaptation, détoxification, évolution) mais aussi plus globalement leur influence sur les écosystèmes chauds où encore trop peu de données sont disponibles à l'heure actuelle. Ces études pourront aussi mettre en évidence le potentiel biotechnologique de ces EGMs (transfert des toxines, de nutriments, d'aides uli ues...).

V. Bibliographie

- Alain, K., Marteinsson, V.T., Miroshnichenko, M., Bonch-osmolovskaya, E., Prieur, D., and Birrien, J. (2002) *Marinitoga piezophila* sp. nov., a rod-shaped, thermo-piezophilic bacterium isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int. J. Syst. Evol. Microbiol.* 52: 1331-1339.
- Anantharaman, K., Breier, J.A., and Dick, G.J. (2016) Metagenomic resolution of microbial functions in deep-sea hydrothermal plumes across the Eastern Lau Spreading Center. *ISME J.* 10: 225-239.
- Anderson, R.E., Brazelton, W.J., and Baross, J.A. (2013) The deep virosphere: assessing the viral impact on microbial community dynamics in the deep subsurface. *Rev. Mineral. Geochemistry* 75: 649-675.
- Anderson, R.E., Sogin, M.L., and Baross, J.A. (2014) Evolutionary strategies of viruses, bacteria and archaea in hydrothermal vent ecosystems revealed through metagenomics. *PLoS One* 9: e109696.
- Breitbart, M. (2012) Marine Viruses: Truth or Dare. *Ann. Rev. Mar. Sci.* 4: 425-448.
- Breitbart, M. and Rohwer, F. (2005) Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13: 278-284.
- Brum, J.R., Ignacio-Espinoza, J.C., Roux, S., Doulier, G., Acinas, S.G., Alberti, A., et al. (2015) Patterns and ecological drivers of ocean viral communities. *Science* (80-.). 348:.
- Brum, J.R. and Sullivan, M.B. (2015) Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat. Rev. Microbiol.* 13: 147-159.
- Charlou, J.L., Donval, J.P., Douville, E., Jean-Baptiste, P., Radford-Knoery, J., Fouquet, Y., et al. (2000) Compared geochemical signatures and the evolution of Menez Gwen ° ' N and Lucky Strike ° ' N hydrothermal fluids, south of the Azores Triple Junction on the Mid-Atlantic Ridge. *Chem. Geol.* 171: 49-75.
- Chellapandi, P. and Ranjani, J. (2015) Knowledge-based discovery for designing CRISPR-CAS systems against invading mobilomes in thermophiles. *Syst. Synth. Biol.* 9: 97 - 106.
- Chow, C.-E.T. and Suttle, C.A. (2015) Biogeography of Viruses in the Sea. *Annu. Rev. Virol.* 2: 41-66.
- Colaço, A., Desbruyères, D., Comtet, T., and Alayse, A. (1998) Ecology of the Menez Gwen hydrothermal vent field (Mid-Atlantic Ridge / Azores Triple Junction). *Cah. Biol. Mar.* 39: 237-240.
- Dahl, C. and Prange, A. (2006) Bacterial sulfur globules: occurrence, structure and metabolism. In, *Inclusions in prokaryotes*. Springer, pp. 21-51.
- Eichinger, I., Schmitz-Esser, S., Schmid, M., Fisher, C.R., and Bright, M. (2014) Symbiont - driven sulfur crystal formation in a thiotrophic symbiosis from deep-sea hydrocarbon seeps. *Environ. Microbiol. Rep.* 6: 364-372.

- Feiss, M. and Catalano, C.E. (2005) Bacteriophage lambda terminase and the mechanism of viral DNA packaging. In, *Viral Genome Packaging Machines: Genetics, Structure, and Mechanism*. Springer, pp. 5–39.
- Gorlas, A., Marguet, E., Gill, S., Geslin, C., Guigner, J.M., Guyot, F., and Forterre, P. (2015) Sulfur vesicles from Thermococcales: A possible role in sulfur detoxifying mechanisms. *Biochimie* 118: 356–364.
- Gupta, R.S. and Bhandari, V. (2011) Phylogeny and molecular signatures for the phylum Thermotogae and its subgroups. *Antonie Van Leeuwenhoek* 100: 1–34.
- Horvath, P. and Barrangou, R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science* (80-.). 327: 167–170.
- Jagannadham, M. V and Chattopadhyay, M.K. (2015) Role of outer membrane vesicles of bacteria. *Resonance* 20: 711–725.
- Lange, S.J., Alkhnbashi, O.S., Rose, D., Will, S., and Backofen, R. (2013) CRISPRmap: an automated classification of repeat conservation in prokaryotic adaptive immune systems. *Nucleic Acids Res.* gkt606.
- Lossouarn, J., Dupont, S., Gorlas, A., Mercier, C., Bienvenu, N., Marguet, E., et al. (2015) An abyssal mobilome: Viruses, plasmids and vesicles from deep-sea hydrothermal vents. *Res. Microbiol.* 166: 742–752.
- Lossouarn, J., Nesbø, C.L., Mercier, C., Zhaxybayeva, O., Johnson, M.S., Charchuck, R., et al. Ménage à trois : a selfish genetic element uses a virus to propagate within Thermotogales. *Environ. Microbiol.* 17: 3278–3288.
- Macmaster, R., Sedelnikova, S., Baker, P.J., Bolt, E.L., Lloyd, R.G., and Rafferty, J.B. (2006) RusA Holliday junction resolvase: DNA complex structure--insights into selectivity and specificity. *Nucleic Acids Res.* 34: 5577–84.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., et al. (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13: 722–736.
- Nesbø, C.L., Baptiste, E., Curtis, B., Dahle, H., Lopez, P., Macleod, D., et al. (2009) The genome of *Thermosiphon africanus* TCF52B: lateral genetic connections to the Firmicutes and Archaea. *J. Bacteriol.* 191: 1974–1978.
- Nesbo, C.L., S Swithers, K., Dahle, H., Haverkamp, T.H.A., Birkeland, N.-K., Sokolova, T., et al. (2015) Evidence for extensive gene flow and Thermotoga subpopulations in subsurface and marine environments. *ISME J.* 9: 1532–1542.
- van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 12: 479–492.
- Paez-Espino, D., Eloie-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., et al. Uncovering Earth's virome. *Nature* advance on:

- Paul, J.H. (2008) Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J.* 2: 579–589.
- Pflugfelder, B., Fisher, C.R., and Bright, M. (2005) The color of the trophosome: elemental sulfur distribution in the endosymbionts of *Riftia pachyptila* (Vestimentifera; Siboglinidae). *Mar. Biol.* 146: 895–901.
- Pollo, S.M.J., Zhaxybayeva, O., and Nesbø, C.L. (2015) Insights into thermoadaptation and the evolution of mesophily from the bacterial phylum Thermotogae. *Can. J. Microbiol.* 1–16.
- Rollie, C., Schneider, S., Brinkmann, A.S., Bolt, E.L., and White, M.F. (2015) Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *Elife* 4: e08716.
- Rosenwasser, S., Ziv, C., Creveld, S.G. van, and Vardi, A. (2016) Virocell Metabolism: Metabolic Innovations During Host-Virus Interactions in the Ocean. *Trends Microbiol.*
- Rousseau, C., Gonnet, M., Le Romancer, M., and Nicolas, J. (2009) CRISPI: a CRISPR interactive database. *Bioinformatics* 25: 3317–3318.
- Scanlan, D. (2014) Bacterial vesicles in the ocean. *Science* (80-.). 343: 143–144.
- Slavic, K., Krishna, S., Lahree, A., Bouyer, G., Hanson, K.K., Vera, I., et al. (2016) A vacuolar iron-transporter homologue acts as a detoxifier in *Plasmodium*. *Nat. Commun.* 7:.
- Sorek, R., Kunin, V., and Hugenholtz, P. (2008) CRISPR —a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6: 181–186.
- Sterling, K.M., Mandal, P.K., Roggenbeck, B.A., Ahearn, S.E., Gerencser, G.A., and Ahearn, G.A. (2007) Heavy metal detoxification in crustacean epithelial lysosomes: role of anions in the compartmentalization process. *J. Exp. Biol.* 210: 3484–3493.
- Suttle, C.A. (2005) Viruses in the sea. *Nature* 437: 356–361.
- Williamson, S.J., Cary, S.C., Williamson, K.E., Helton, R.R., Bench, S.R., Winget, D., and Wommack, K.E. (2008) Lysogenic virus host interactions predominate at deep-sea diffuse-flow hydrothermal vents. *ISME J.* 2: 1112–1121.
- Williamson, S.J., Rusch, D.B., Yooseph, S., Halpern, A.L., Heidelberg, K.B., Glass, J.I., et al. (2008) The Sorcerer II Global Ocean Sampling Expedition: Metagenomic Characterization of Viruses within Aquatic Microbial Samples. *PLoS One* 3: e1456.
- Zhaxybayeva, O., Swithers, K.S., Lapierre, P., Fournier, G.P., Bickhart, D.M., DeBoy, R.T., et al. (2009) On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. U. S. A.* 106: 5865–5870.

Résumé : Nos connaissances sur la diversité virale associée aux micro-organismes présents dans les sources hydrothermales océaniques profondes restent encore limitées. Seules quelques études concernant l'impact de ceux-ci sur la mortalité microbienne dans ces écosystèmes sont disponibles. En effet, seuls 6 bactériovirus et 2 archéovirus provenant de ces écosystèmes ont été caractérisés à l'heure actuelle. Les deux archéovirus des archées anaérobies hyperthermophiles appartiennent à l'ordre *Thermococcales* et ont été décrits au laboratoire. Afin d'étendre nos connaissances sur la diversité virale associée aux micro-organismes colonisant ces environnements, il a été décidé d'élargir les recherches à l'ordre bactérien. L'ordre bactérien est composé de bactéries chimio-organotrophes anaérobies en majorité thermophiles ou hyperthermophiles. De nombreux transferts latéraux de gènes ont contribué à l'histoire évolutive des bactéries, supposant une forte implication des virus dans celle-ci. Ces travaux de thèse ont permis la caractérisation fonctionnelle et génomique de deux nouveaux siphovirus, MCV1 et MCV2, infectant deux souches de *Marinitoga camini*. Ces souches ont été isolées de deux sites hydrothermaux profonds (Menez Gwen et Lucky Strike) au niveau de la dorsale médio-atlantique. Ces virus mettent en œuvre un cycle lysogénique avec une production basale sans induction relativement haute ($>10^7$ virions/mL). Une comparaison de ces deux génomes viraux à celui de MPV1, virus précédemment isolé de *Marinitoga piezophila*, a été réalisée, révélant la présence de nombreuses similarités. Un core genome de 35 ORFs partagé par ces trois génomes a été identifié, incluant des protéines impliquées dans le métabolisme de l'ADN, l'assemblage des virions et le cycle lysogénique. Des protéines hypothétiques ont aussi été identifiées parmi ces gènes communs, elles portent donc probablement des fonctions importantes pour ces bactériovirus. Par ailleurs, 60% des gènes de ces virus ayant une correspondance dans les bases de données, après exclusion des *Thermotogales*, partagent des similarités avec les *Firmicutes* et les bactériovirus qui leur sont associés. Le génome d'une autre *Thermotogales*, *Thermosiphon* sp. AT1244-VC14 a été étudié ainsi que son système CRISPR-cas. Ces résultats indiquent que cette souche, qui porte un système CRISPR-cas qui semble complet et fonctionnel, a probablement déjà été infectée par MCV1, MCV2 ou un virus similaire. Ces travaux permettent d'étendre nos connaissances portées par les bactéries du phylum *Thermotogae*, encore peu décrits à ce jour. Les éléments génétiques mobiles associés à ce phylum sont particulièrement intéressants car ils ont probablement eu un impact important dans l'évolution de ces microbiennes ainsi que dans leur adaptation aux conditions physico-chimiques extrêmes et fluctuantes présentes dans les écosystèmes qu'elles colonisent.

Mots clés : bactériovirus, *Thermotogales*, sources hydrothermales profondes, éléments génétiques mobiles, océans profonds.

New host-virus systems from deep sea hydrothermal vents

Abstract : Our knowledge of the viral diversity associated to microorganisms inhabiting the deep-sea hydrothermal vents is still limited. Only a few studies have focused on viral abundance and impact on microbial mortality within these ecosystems. A limited number of viruses (6 bacteriophages and 2 archaeophages) were isolated from these environments and characterized. Two viruses associated to hyperthermophilic anaerobic *Archaea*, from the *Thermococcales* order, have been described in our laboratory. In order to deepen our knowledge on the viral diversity of these extreme environments, we have extended our investigation to the bacterial order of *Thermotogales*. This order is composed of anaerobic chemoorganotrophic bacteria that are, for the most part, hyper/thermophilic. Numerous lateral gene transfers have contributed to the evolutionary history of the *Thermotogales*, implying the potential involvement of viruses. Here, we will report the characterization of two new siphoviruses MCV1 and MCV2 that infect two strains of *Marinitoga camini*. Those bacterial strains were isolated from two deep-sea hydrothermal vents sites (Menez Gwen and Lucky strike) in the Mid Atlantic Ridge. These viruses are temperate with a high basal production of virions ($>10^7$ virions/mL). Comparative genomics with MPV1, a virus isolated from *M. piezophila*, was performed and show that those bacteriophages share numerous similarities. A set of "core genes" shared by all these three viruses was identified and includes proteins involved in DNA metabolism, head and tail assembly and lysogenic cycle. Shared hypothetical proteins were also identified, suggesting that these unknown proteins probably provide important functions for these viruses. Interestingly, for genes with blastp matches in Genbank, over 60% have their top matches, outside the *Thermotoga*, to genes from *Firmicutes* and bacteriophages associated to *Firmicutes*. We also analyzed the genome of *Thermosiphon* sp. 1244 and studied his CRISPR-cas system. Our results indicated that this *Thermosiphon* strain, with a complete and functional CRISPR-cas system, had already been infected by MCV1, MCV2 or a similar virus. The analyses presented here extend our knowledges about these newly discovered viruses in the deeply branching bacterial phylum *Thermotogae*. This bacterial order and associated mobile genetic elements are significant for addressing long-term evolutionary adaptation to fluctuant and extreme physicochemical conditions.

Key words: bacteriophages, *Thermotogales*, deep sea, hydrothermal vents, mobile genetic elements.