



**Identification et caractérisation d'éléments génétiques
chez *Aeromonas salmonicida* permettant le suivi
géographique des souches causant la furunculose**

Mémoire

Jean-Guillaume Emond Rheault

Maîtrise en microbiologie

Maître ès sciences (M. Sc.)

Québec, Canada

© Jean-Guillaume Emond Rheault, 2015

Résumé

Aeromonas salmonicida sous-espèce *salmonicida* est un agent pathogène opportuniste responsable chaque année d'importantes pertes économiques pour les aquaculteurs de Salmonidés. Dans cette étude, plusieurs analyses ont été effectuées dans l'objectif de trouver une méthode pour distinguer entre différents isolats de cette bactérie au niveau génomique. Suite à l'alignement chromosomique des souches A449 et 01-B526, trente-deux altérations génomiques ont été répertoriées et elles peuvent être classées en cinq groupes : séquences d'insertions (13), séquences répétées en tandem (12), séquences insérées dans des gènes (5), site de polymorphisme multi-évènementiel (1) et îlot génomique (*AsaGEI1a*) (1). Des criblages PCR et des séquençages génomiques ont révélé l'existence de trois autres îlots génomiques (*AsaGEI1b*; *AsaGEI2a*; *AsaGEI2b*). Chaque *AsaGEI* est hautement spécifique à une région géographique. Les *AsaGEI(1a;2a)* sont seulement observés en Amérique du Nord et les *AsaGEI(1b;2b)* en Europe. Dans ces travaux, plusieurs marqueurs pouvant permettre d'identifier l'origine géographique des souches pathogènes ont été découverts.

Abstract

Aeromonas salmonicida subsp. *salmonicida* is an opportunistic pathogen, which causes significant economic loss in salmonid aquaculture. In this work, analyses were conducted with the objective to find a way to distinguish between different isolates of the bacterium at the genomic level. Following the chromosomal alignment between A449 and 01-B526 strains, thirty-two genomic alterations were found and were classified in five groups: insertion sequences (13), tandem repeat sequences (12), CDS-modeling sequences (5), multi-event polymorphism site (1) and genomic island (*AsaGEI1a*) (1). PCR assays and genomic sequencing revealed the existence of four forms (*AsaGEI(1a;1b;2a;2b)*) of the genomic island. Each GEI appeared to be strongly associated with a specific geographic region. *AsaGEI(1a;2a)* were exclusively found in North American isolates and *AsaGEI(1b;2b)* in those from Europe. In this study, several indicators useful to identify the geographical origin of pathogenic strains of this bacterium were discovered.

Table des matières

Résumé.....	iii
Abstract.....	v
Table des matières.....	vii
Liste des tableaux.....	xi
Liste des figures.....	xiii
Liste des abréviations.....	xv
Remerciements.....	xvii
1. Introduction.....	1
1.1. L'aquaculture.....	1
1.1.1. Dans le monde.....	1
1.1.2. Au Québec.....	1
1.2. Maladies infectieuses affectant les Salmonidés.....	2
1.3. La furunculose chez l'omble de fontaine.....	3
1.4. La bactérie : <i>Aeromonas salmonicida</i> sous-espèce <i>salmonicida</i>	4
1.4.1. Caractéristiques génomiques d'A449.....	6
1.4.2. Analyses génétiques d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i>	7
1.5. Un nouvel élément génétique chez <i>A. salmonicida</i> sous-espèce <i>salmonicida</i>	10
1.6. Les éléments mobiles de l'ADN.....	12
1.6.1. Les séquences d'insertion.....	12
1.6.2. Les bactériophages.....	14
1.6.3. Les îlots génomiques (GEI).....	16
1.7. L'instabilité génomique liée aux séquences répétées en tandem.....	17
2. Hypothèses.....	21
3. Objectifs spécifiques.....	23
4. Matériels et méthodes.....	25
4.1. Isolats bactériens et conditions de croissance.....	25
4.2. Extractions d'ADN bactérien.....	25
4.3. Conception d'amorces RAPD.....	26
4.4. RAPD.....	26
4.5. PCR.....	27
4.6. Séquençage et assemblage des génomes.....	27

4.7.	Analyses bio-informatiques	28
5.	Résultats.....	31
5.1.	Analyses RAPD	31
5.2.	Étude des variations chromosomiques entre 01-B526 et A449	34
5.2.1.	Dissemblances impliquant des IS	34
5.2.2.	Variations du nombre de répétitions dans des séquences TR.....	35
5.2.3.	Dissemblances impliquant l'insertion/délétion d'une séquence dans un gène	36
5.2.4.	Site de polymorphisme multi-évènementiel	37
5.2.5.	Différences impliquant l'insertion/délétion d'un GEI.....	38
5.2.6.	Analyses des sites de polymorphisme dans dix autres isolats d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i>	41
5.3.	Caractérisation des <i>AsaGEI</i> chez <i>A. salmonicida</i>	43
5.3.1.	Caractéristiques générales d'un nouveau GEI dans la souche 01-B526.....	43
5.3.2.	Structure d' <i>AsaGEI1b</i> et <i>AsaGEI2a</i> et leur relation avec <i>AsaGEI1a</i>	49
5.3.3.	Sites d'insertion des <i>AsaGEI</i>	52
5.3.4.	Prévalence des différents variants d' <i>AsaGEI</i> par criblage PCR	54
5.3.5.	Corrélation entre le génotype d' <i>AsaGEI</i> et l'origine géographique des isolats	55
5.3.6.	Analyse du GEI de la souche JF3224.....	56
5.3.7.	Annotation et comparaison génomique du prophage 3	58
6.	Discussion	61
6.1.	RAPD	61
6.2.	Différences chromosomiques entre A449 et 01-B526.....	63
6.2.1.	Modifications impliquant des IS	63
6.2.2.	Variations du nombre de répétitions dans des séquences TR.....	66
6.2.3.	Différences impliquant l'insertion/délétion d'une séquence dans un gène.....	68
6.3.	Caractérisation des <i>AsaGEI</i> chez <i>Aeromonas salmonicida</i>	69
6.3.1.	Relation évolutive entre les différents éléments <i>AsaGEI</i> ou reliée aux <i>AsaGEI</i>	70
6.3.2.	Relation entre l'origine géographique des isolats, le génotype d' <i>AsaGEI</i> et la présence/absence du prophage 3.....	72
6.4	Éléments génétiques responsables de la variabilité chromosomique chez <i>A. salmonicida</i> sous-espèce <i>salmonicida</i>	76
7.	Conclusion.....	77
8.	Perspectives	79
	Bibliographie.....	81

Annexe A : Souches d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i> utilisées dans cette étude	91
Annexe B : Amorces utilisées dans cette étude	95
Annexe C : Fonction et position des ORF dans <i>AsaGEI1a</i>	97
Annexe D : Matrice carrée des comparaisons génomique des trois <i>AsaGEI</i> chez les huit souches d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i>	99
Annexe E : Fonction et position des ORF dans <i>AsaGEI1b</i>	101
Annexe F : Fonction et position des ORF dans <i>AsaGEI2a</i>	103
Annexe G : Fonction et position des ORF dans <i>AsaGEI2b</i>	105
Annexe H : Fonction et position des ORF dans le prophage 3	107

Liste des tableaux

Tableau 1 : Les séquences d'insertion (IS) du génome de la souche A449 d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i>	7
Tableau 2 : Liste des amorces RAPD testées avec une quantité élevée de site d'hybridation dans les chromosomes d'A449 et de 01-B526 et leurs résultats respective.	32
Tableau 3 : Listes des dissemblances observées entre le chromosome d'A449, qui est utilisée comme référence, par rapport au chromosome de la souche 01-B526.....	39
Tableau 4 : Analyses bio-informatiques par l'outil BLASTn des dissemblances chromosomiques observées entre les souches A449 et 01-B526 dans les séquences génomiques assemblées en contig de dix autres isolats d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i>	42
Tableau 5 : Origine géographique des 139 isolats d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i> par rapport à leur génotype d' <i>AsaGEI</i>	55
Tableau 6 : Origine géographique des 139 isolats d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i> par rapport à leur génotype d' <i>AsaGEI</i> (Tableau 5 corrigé).	73

Liste des figures

Figure 1 : Représentation de la taxinomie d'isolats d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i> sous forme schématique adaptée du <i>Bergey's Manual of Systematic Bacteriology - Volume 2</i>	5
Figure 2 : Représentations schématiques des trois profils d'amplification observés lors du criblage d'une collection d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i> avec 15 paires d'amorces pour la présence de l'élément génétique de 51 kb.	11
Figure 3 : Résultats de trois RAPD migrés en électrophorèse sur gel d'agarose	333
Figure 4 : Représentation des alignements nucléotidique et protéique de l' <i>ABC-type transporter, ATPase subunit</i> d'A449 par rapport à celle chez 01-B526.....	37
Figure 5 : Représentation d'une section de l'alignements nucléotidique et de l'alignement protéique de la <i>pyruvate formate lyase</i> de la souche A449 par rapport à la <i>pyruvate formate lyase</i> chez 01-B526.....	388
Figure 6 : Caractéristiques génétiques d' <i>AsaGEI1a</i>	455
Figure 7 : Comparaison colinéaire d' <i>AsaGEI1a</i> avec les éléments génétiques des souches d' <i>A. hydrophila</i> RB-AH et d' <i>A. jandaei</i> Riv2.	477
Figure 8 : Comparaison colinéaire d' <i>AsaGEI1a</i> avec l'élément génétique de la souche d' <i>A. salmonicida</i> sous-espèce <i>achromogenes</i> AS03.	488
Figure 9 : Comparaison colinéaire des trois <i>AsaGEI</i>	5151
Figure 10 : Représentation schématique des sites d'insertion des trois <i>AsaGEI</i> par rapport au génome de la souche d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i> A449.....	533
Figure 11 : Génotypage de 139 isolats d' <i>A. salmonicida</i> sous-espèce <i>salmonicida</i> pour la présence de neuf séquences trouvées chez <i>AsaGEI1a</i>	544
Figure 12 : Comparaison colinéaire des quatre <i>AsaGEI</i>	577
Figure 13 : Comparaison colinéaire du prophage 3 avec les éléments génétiques découvert chez les souches d' <i>A. veronii</i> B565 et d' <i>A. hydrophila</i> 4AK4.....	599
Figure 14 : Relation évolutive entre les quatre <i>AsaGEI</i> et les trois éléments apparentés aux <i>AsaGEI</i> d'après leurs sites d'insertion	7272
Figure 15 : Représentation de la provenance des 139 isolats d' <i>Aeromonas salmonicida</i> sous-espèce <i>salmonicida</i> de la collection Charette listé à l'Annexe A	755

Liste des abréviations

ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messenger
ARNt	Acide ribonucléique de transfert
ATPase	Adénosine-triphosphatase
<i>AsaGEI</i>	<i>Aeromonas salmonicida Genomic Island</i>
BIME	Éléments bactériens en mosaïque intercalée (<i>Bacterial Interspersed Mosaic Elements</i>)
BLAST	<i>Basic Local Alignment Search Tool</i>
CDS	Séquence codante (<i>Coding DNA Sequence</i>)
DR	Répétition directe (<i>Direct Repeat</i>)
EtBr	Bromure d'éthidium (<i>Ethidium Bromide</i>)
FAO	Organisation des Nations Unies pour l'alimentation et l'agriculture (<i>Food and Agriculture Organization</i>) (www.fao.org/home/en/)
GEI	Îlot génomique (<i>Genomic Island</i>)
HGT	Transfert horizontal de gènes (<i>Horizontal Gene Transfer</i>)
IS	Séquence d'insertion (<i>Insertion Sequence</i>)
kb	Kilobase
MAPAQ	Ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec
Mb	Million de base
MGE	Élément génétique mobile (<i>Mobile Genetic Element</i>)
MITE	Élément transposable miniature avec région répétée inversée (<i>Miniature Inverted-repeat Transposable-Element</i>)
MLVA	Analyse multi-locus VNTR (<i>Multilocus VNTR analysis</i>)
nt	Nucléotide
NCBI	<i>National Center for Biotechnology Information</i>
ORF	Cadre ouvert de lecture (<i>Open Reading Frame</i>)
pb	Paire de base
PCR	Réaction de polymérisation en chaîne (<i>Polymerase Chain Reaction</i>)
RAPD	Amplification aléatoire d'ADN polymorphe (<i>Random Amplified Polymorphic DNA</i>)
REP	séquence palindrome extragénique répétitive (<i>Repetitive Extragenic Palindromic sequence</i>)

RFLP	Polymorphisme de longueur des fragments de restriction (<i>Restriction Fragment Length Polymorphism</i>)
PFGE	Électrophorèse en champs pulsés (<i>Pulsed-Field Gel Electrophoresis</i>)
RAQ	Ressource aquatique Québec
TR	Répété en tandem (<i>Tandem Repeat</i>)
VNTR	<i>Variable-Number Tandem Repeat</i>

Remerciements

Plusieurs personnes ou organismes ont contribué de façons directes ou indirectes à l'étude présentée dans ce document.

Tout d'abord, j'aimerais débiter en remerciant tous les membres du laboratoire Charette, et spécialement Antony T. Vincent, qui m'a été d'une grande aide en me conseillant durant l'entièreté de mes analyses.

Ensuite, je voudrais remercier les membres de mon comité aviseur qui sont Michel Frenette, Brian Boyle, Nicolas Derome et Steve Charette. Donc, encore une fois, un gros merci à Michel Frenette pour ses nombreux conseils.

Troisièmement, merci à toutes et tous ceux qui m'ont supporté durant mes moments de stress. Donc, un merci particulier à Myriam Ouellet, ma copine, qui a été présente pour m'écouter et me rassurer lors de ces moments stressants. Aussi, merci à ma famille et mes amies pour m'avoir supporté durant ces deux années d'étude.

Quatrièmement, j'aimerais remercier tous les collaborateurs qui ont gracieusement envoyé des isolats d'*A. salmonicida* sous-espèce *salmonicida*, donc merci à Andrée Lafaille, Joachim Frey, Roselyne Stevenson et Mohamed Faisal.

Et dernièrement, j'aimerais remercier les organismes subventionnaires pour m'avoir permis de compléter mes études 2^e cycle. Merci au RAQ pour l'obtention d'une bourse FONCER (2013-2014) et d'un complément de bourse FONCER (2014-2015), et aussi merci au FQRNT pour l'obtention d'une bourse de maîtrise (2014-2015).

«La science consiste à passer d'un étonnement à un autre¹»

Bonne lecture,

JGER

¹ Aristote

1. Introduction

1.1. L'aquaculture

1.1.1. Dans le monde

Le terme aquaculture est utilisé pour toutes les activités de production végétale ou animale en milieu aquatique. Elle se pratique aussi bien en eau salée, au bord des mers et des côtes océaniques, qu'en eau douce, dans des lacs, des étangs, des rivières et des bassins. D'après l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO), l'aquaculture est probablement le secteur de production alimentaire qui a connu la croissance la plus rapide en un demi-siècle (Food and Agriculture Organization of the United Nations, 2014). En 2012, la production mondiale de l'aquaculture représentait 50% des poissons destinés à la consommation humaine. L'élevage de poissons semble une bonne alternative à la surpêche qui menace l'équilibre des écosystèmes marins et dulcicoles ainsi que la sécurité alimentaire par le déclin des populations aquatiques. Les données statistiques de 2012 analysées par la FAO indiquent que 41,4% de la production mondiale de l'aquaculture est dédiée à l'élevage d'espèces de poissons d'eau douce. Les poissons principalement élevés sont la carpe du roseau (*Ctenopharyngodon idellus*), la carpe argentée (*Hypophthalmichthys molitrix*) et la carpe commune (*Cyprinus carpio*) pour des parts de 7,6%, 6,3% et 5,7% de la production mondiale, respectivement. Les quatre pays élevant la plus grande quantité de produits aquicoles sont en ordre, la Chine, l'Inde, le Vietnam et le Bangladesh (Food and Agriculture Organization of the United Nations, 2014).

1.1.2. Au Québec

Dans la province de Québec, les piscicultures produisent principalement des poissons de la famille des Salmonidés, l'unique famille de l'ordre des Salmoniformes. Elle comprend les ombles, les saumons, les truites, les ombres (*Thyllamus*) et les corégones. La production piscicole de la province est principalement dédiée à l'omble de fontaine, aussi couramment appelé la truite mouchetée (*Salvelinus fontinalis*). Il est un poisson indigène qui est le choix le plus populaire lors de l'élevage pour le marché de l'ensemencement. La production aquacole est destinée à 78% à ce marché.

Une petite proportion de la production du marché de l'ensemencement est destinée aussi à l'élevage du doré jaune (*Sander vitreus*), de la perchaude (*Perca flavescens*) et de l'achigan à grande bouche (*Micropterus salmoides*) (MAPAQ, 2013). Ces trois espèces de l'ordre des Perciformes font l'objet d'une faible production en aquaculture, car les technologies connues et utilisées par l'industrie aquacole sont optimisées pour la production des Salmonidés. Mais ces espèces peuvent être une éventualité dans des conditions où la température de l'eau est trop chaude pour l'élevage de Salmonidés.

D'après les statistiques du ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec (MAPAQ), le marché de l'ensemencement de l'omble de fontaine totalisait des ventes de 8,5 M\$ en 2011 (MAPAQ, 2012). La deuxième espèce majoritairement élevée dans la province est la truite arc-en-ciel (*Oncorhynchus mykiss*), mais celle-ci pour le marché de table (MAPAQ, 2013). La production aquacole est étendue à toute la province, mais avec certaines concentrations dans les régions de l'Estrie, des Laurentides, de l'Outaouais et du Centre-du-Québec.

1.2. Maladies infectieuses affectant les Salmonidés

L'un des grands défis de l'aquaculture consiste à combattre les maladies infectieuses causant chaque année de grandes pertes économiques (Lafferty *et al.*, 2015). Cette réalité est aussi vraie pour tous les secteurs de production alimentaire, mais se complique en aquaculture, car les premiers symptômes sont souvent décelés au moment où la pathologie cause les premiers cas de mortalité.

Les maladies infectieuses sont causées par des agents biologiques microscopiques qui sont couramment nommés agents pathogènes. Ces maladies sont classées en différents groupes selon l'agent qui les causent. Ces microorganismes peuvent être présents dans l'eau ou directement sur les poissons et ils peuvent être soit présents naturellement dans l'environnement immédiat ou avoir été introduits suite à une contamination extérieure lors d'échanges commerciaux (Hedrick, 1996). Les agents étiologiques peuvent être soit des virus, des mycètes, des parasites ou des bactéries. Les agents pathogènes dits opportunistes infectent généralement les poissons lors d'une diminution de la réponse de leur système immunitaire qui peut être causée par des facteurs de stress.

Les différents stress subis par les Salmonidés endommagent la mince couche de mucus à la surface de leur épiderme qui est nécessaire à leur protection, donc les stress rendent plus vulnérables les poissons aux agents pathogènes (Uhland *et al.*, 1997). Comme évoqué à la section sur l'aquaculture au Québec (1.1.2), l'omble de fontaine est l'espèce de poisson élevée en plus grande proportion. L'AQUICOLE de septembre 2010 compile les maladies diagnostiquées entre 1999 et 2010 chez l'omble de fontaine. Durant cette période, les trois pathologies les plus fréquemment observées en ordre d'importance étaient la furonculose, les maladies branchiales et les parasitoses (Morin, 2010). Durant ces dix ans, chaque année la furonculose représentait entre 30% et 60% des maladies diagnostiquées chez l'omble de fontaine au Québec (Morin, 2010). De plus, les rapports de 2011 et de 2012 du service d'ichtyopathologie de la Faculté de médecine vétérinaire (FMV) de l'Université de Montréal révèlent que la furonculose est encore l'une des trois pathologies les plus régulièrement diagnostiquées entre les périodes d'avril 2010 et de décembre 2011 chez l'omble de fontaine (Lafaille, 2011, 2012). Donc, la furonculose est une pathologie ayant d'importantes répercussions économiques (Bernoth *et al.*, 1997). Il est donc nécessaire de mieux comprendre son agent étiologique pour ainsi apprendre à combattre cette infection.

1.3. La furonculose chez l'omble de fontaine

La furonculose a été pour la première fois documentée lors d'une éclosion de lésions ulcéreuses entre les années 1888 et 1889 dans une pisciculture en Allemagne qui élevait des truites brunes (Dworkin *et al.*, 2006). Suite à cette première description, plusieurs écrits ont fait part de l'émergence de la pathologie dans d'autres pays. Au Québec, les premières observations de furonculose remontent à la fin des années cinquante (MAPAQ, 2012). Plusieurs facteurs de stress sont connus pour favoriser les éclosions de la maladie, comme une augmentation de la température de l'eau durant l'été, un manque d'oxygène dissous, une suralimentation entraînant un surplus de matière organique dans l'eau ou une importante manipulation des poissons (déplacement, vaccination, etc.) (Uhland *et al.*, 1997).

La furonculose est causée par une bactérie pathogène nommée *Aeromonas salmonicida* (Janda et Abbott, 2010). Elle peut se présenter sous plusieurs formes : aiguë, subaiguë ou même chronique. Lors de la forme aiguë, qui affecte plus souvent les jeunes poissons, aucun furoncle n'est visible, l'intestin est généralement atteint et l'infection peut

se terminer par une septicémie. Elle cause entre 30% et 50% de mortalité pendant l'infection. Lors de la forme subaiguë, il y a deux stades à la maladie : 1) formation de protubérances dans la musculature qui se nomme furoncle et 2) une infection des intestins. Contrairement à la forme aiguë, la maladie s'étend sur une longue période. La forme chronique de la furunculose est caractérisée par la présence d'hémorragies intramusculaires et de furoncles sous la peau. Les poissons qui guérissent de cette pathologie peuvent demeurer porteurs du pathogène et devenir ainsi des vecteurs de transmission aux autres poissons (Uhland *et al.*, 1997).

Les sources d'où provient cette bactérie peuvent être variées. Celle-ci peut être introduite par le biais de poissons sains ou malades, leurs excréments ou liquide contenu dans les furoncles, du matériel de pêche, d'eau contaminée ou de vecteurs de transmission, tels les oiseaux et les parasites (Uhland *et al.*, 1997). Cependant, l'origine et le mouvement des souches pathogènes d'*Aeromonas salmonicida* sont encore mal connus (Bernoth *et al.*, 1997).

À la suite d'une éclosion de furunculose, l'infection peut être traitée avec succès au moyen d'antibiotiques incorporés dans la moulée. Pour éclairer sur le choix du médicament à utiliser, un antibiogramme est très fortement recommandé. Au Québec, les antibiotiques doivent être prescrits par le médecin vétérinaire pour que leur utilisation soit légale. Pour éviter l'apparition de la maladie, plusieurs moyens prophylactiques peuvent être pris : garder une bonne hygiène de l'équipement piscicole, éviter les sources de contaminations extérieures et contrôler les sources de stress infligés aux poissons (Uhland, 2004; Lafaille, 2010; Morin, 2010).

1.4. La bactérie : *Aeromonas salmonicida* sous-espèce *salmonicida*

Se situant dans le domaine des *Bacteria* et dans la classe des *Gammaproteobacteria*, l'ordre des *Aeromonadales* contient deux familles, soit celle des *Aeromonadaceae* et des *Succinivibrionaceae*, dont la dernière famille est de classification incertaine (*incertae sedis*) (Figure 1). Le genre type de la famille des *Aeromonadaceae* est *Aeromonas*, dont les caractéristiques sont des bacilles à Gram négatif, chimio-organotrophe, anaérobie facultatif, généralement positif pour la catalase et l'oxydase (Martin-Carnahan et Joseph, 2005). Le genre *Aeromonas* regroupe des espèces dont plusieurs sont des agents pathogènes causant de nombreuses infections chez les

humains et les animaux. Par exemple chez l'homme, *A. hydrophila*, *A. dakhensis*, *A. veronii*, *A. caviae*, *A. trota* et *A. jandaei* ont déjà été caractérisées comme des agents pathogènes causant des gastroentérites, des infections cutanées et des septicémies (Figueras, 2005; Chen *et al.*, 2014).

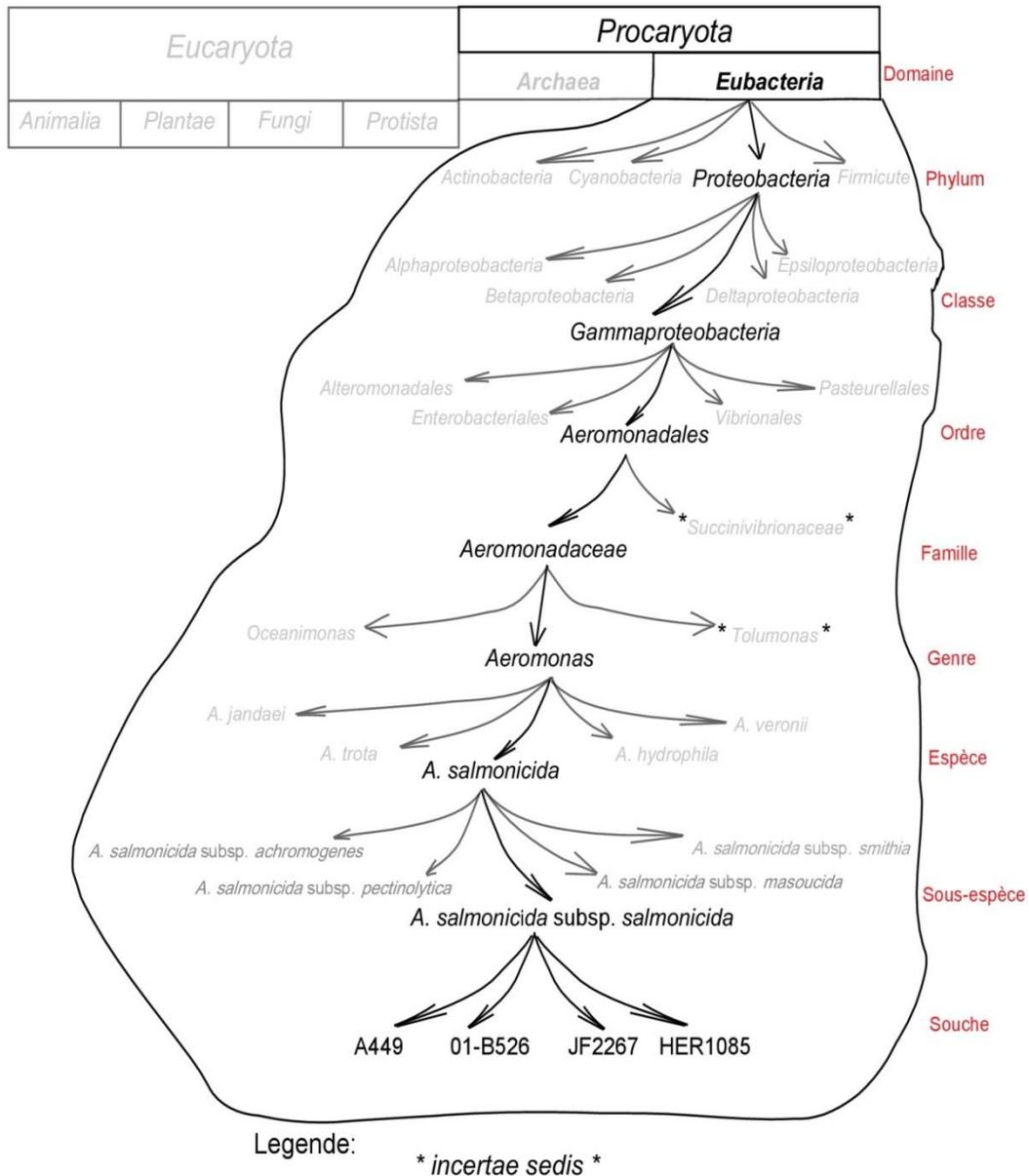


Figure 1 : Représentation de la taxinomie d'isolats d'*A. salmonicida* sous-espèce *salmonicida* sous forme schématique adaptée du *Bergey's Manual of Systematic Bacteriology - Volume 2* (Brenner *et al.*, 2005)

A. salmonicida est une espèce bactérienne non-motile causant la furonculose chez les poissons. Elle comporte cinq sous-espèces : *salmonicida*, *achromogenes*, *masoucida*, *smithia* et *pectinolytica* (Brenner *et al.*, 2005). La nomenclature divise aussi les sous-espèces d'*A. salmonicida* par leur capacité à causer une furonculose dite typique ou atypique. Les isolats d'*A. salmonicida* dits typiques regroupent les souches de la sous-espèce *salmonicida* et elles causent généralement des infections systémiques chez les Salmonidés. Elles produisent de larges colonies sur gélose sang, sont hémolytiques, ne fermentent pas le saccharose et elles produisent un pigment brun soluble dans la gélose sur un milieu contenant du tryptophane (Martin-Carnahan et Joseph, 2005).

Les caractérisations biochimiques des souches d'*A. salmonicida* sous-espèce *salmonicida* suggèrent que les isolats de cette bactérie sont phénotypiquement homogènes (Dalsgaard *et al.*, 1994; Diamanka *et al.*, 2013), et le même constat est vrai pour les caractéristiques génétiques lors d'études épidémiologiques par techniques moléculaires (Belland et Trust, 1988; García *et al.*, 2000; Hänninen *et al.*, 1995; McCormick *et al.*, 1990; Miyata *et al.*, 1995; O'hici *et al.*, 2000). Les isolats d'*A. salmonicida* classés comme atypiques causent des pathologies, dites furonculoses atypiques, autant chez les Salmonidés que les non-Salmonidés et elles sont caractérisées par la présence d'ulcères. Les colonies bactériennes dites atypiques ont des caractéristiques morphologiques et biochimiques différentes des typiques et sont produites par les quatre autres sous-espèces : *achromogenes*, *masoucida*, *smithia* et *pectinolytica*. Elles produisent très peu ou pas de pigments bruns solubles sur un milieu contenant du tryptophane (Burr et Frey, 2007; Lund *et al.*, 2002; Wiklund et Dalsgaard, 1998).

1.4.1. Caractéristiques génomiques d'A449

Le premier génome d'*A. salmonicida* sous-espèce *salmonicida* a été publié en 2008 et est celui de l'isolat A449, une souche pathogène provenant de France. Le génome d'A449 est constitué d'un seul chromosome circulaire, de deux plasmides de grandes tailles (pAsa4 et pAsa5) et de trois petits plasmides (pAsa1, pAsa2 et pAsa3) (Boyd *et al.*, 2003; Reith *et al.*, 2008). Le chromosome a une taille de 4,7 Mb et possède un contenu en G+C de 58,5%. Il est constitué de 4388 gènes qui encodent pour un total de 4086 protéines. La déviation en G+C du chromosome indique la présence d'une inversion de 195 kb (position 3,963,279 à 4,158,772) qui est encadrée entre deux séquences s'insertions (IS) ISAS4. Le chromosome et les deux plasmides de grandes tailles sont

caractérisés par la présence de nombreux éléments génétiques mobiles (MGE) (Reith *et al.*, 2008).

Les MGE identifiés chez A449 comprennent trois grandes familles : les IS, les prophages et les transposons. Ce génome comprend un total de 88 IS qui sont distribuées entre le chromosome, pAsa4 et pAsa5, à raison de 71, 7 et 10 IS, respectivement. Dix différents types d'IS sont présents dans le génome (Tableau 1, Katherine H. Tanaka, communication personnelle). L'annotation du chromosome a révélé la présence de deux prophages, nommé prophage 1 et prophage 2 (Reith *et al.*, 2008). En plus des 88 IS complètes, le génome d'A449 contient aussi 20 IS partielles qui semblent avoir été auparavant complètes, mais elles sont interrompues par l'intégration d'une autre IS dans leurs cadres de lecture ouverts (ORF), brisant ainsi la séquence codante (CDS) (Studer *et al.*, 2013).

Tableau 1 : Les séquences d'insertion (IS) du génome de la souche A449 d'*A. salmonicida* sous-espèce *salmonicida*.

Nom	Longueur (pb)	Famille	Nombre de copies complètes	Activité démontrée	Implication	Autres noms	Source
ISAS1	1223	ISAs1	2	Oui	Inactivation, <i>A-layer</i> (Gustafson <i>et al.</i> , 1994)	ISAs1 (Reith <i>et al.</i> , 2008)	(Gustafson <i>et al.</i> , 1994)
ISAS2	1084	IS30	5	Oui	Inactivation, <i>A-layer</i> (Gustafson <i>et al.</i> , 1994)	ISAs2 (Reith <i>et al.</i> , 2008)	(Gustafson <i>et al.</i> , 1994)
ISAS3	1233	IS3	12	-	-	ISAs5 (Reith <i>et al.</i> , 2008), ISAS7 (NC_009348)	(Merino <i>et al.</i> , 2003)
ISAS4	1165	IS630	37	Suspectée	Transfert horizontal (Studer <i>et al.</i> , 2013)	ISAs7 (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
ISAS5	2614	IS21	12	Oui	Plasmide pAsa1B (Trudel <i>et al.</i> , 2013)	ISAs11 (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
ISAS6	1240	IS3	7	-	-	ISAs6 (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
ISAS8	754	IS1	3	-	-	ISAs8 (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
ISAS9	1624	IS3	4	-	-	ISAs9 (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
ISAS10	1229	IS30	2	-	-	ISAs10 (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
ISAS11	1326	IS256	4	Oui	Délétion, locus SSTT (Tanaka <i>et al.</i> , 2012)	ISAs3 (Fehr <i>et al.</i> , 2006) et NC_004338	(Fehr <i>et al.</i> , 2006)

1.4.2. Analyses génétiques d'*A. salmonicida* sous-espèce *salmonicida*

Pour mieux comprendre et étudier l'origine et le déplacement des souches pathogènes causant les éclosions de furunculose, plusieurs études avec des techniques

moléculaires variées ont déjà été réalisées afin de classer des isolats d'*A. salmonicida* sous-espèce *salmonicida* d'origines géographiques diversifiées (Europe, Japon, États-Unis, Canada, Écosse, Irlande, Grande-Bretagne, etc.). Certaines de ces études ne décèlent aucun polymorphisme entre les isolats analysés (Belland et Trust, 1988; Miyata et al., 1995), alors que les autres démontrent la présence de quelques différences génétiques (McCormick et al., 1990; Hänninen et al., 1995; García et al., 2000; O'hici et al., 2000), mais les divergences génétiques ne sont pas suffisantes pour permettre de mener des études épidémiologiques.

En 1995, Miyata et ses collaborateurs ont utilisé la technique de génotypage par amplification aléatoire d'ADN polymorphe, ou *random amplification polymorphism DNA* (RAPD) (Miyata et al., 1995), mais n'ont observé aucun polymorphisme. La même technique a été utilisée par Hänninen et ses collaborateurs en 1995, mais quant à eux, ils ont identifié des différences génétiques (Hänninen et al., 1995). La RAPD a été développée par Williams et ses collaborateurs afin de construire des cartes génétiques et a été publiée pour la première fois en 1990 (Williams et al., 1990). Elle consiste en l'utilisation de courts oligonucléotides d'environ 10 nucléotides (nt) et d'une température d'hybridation basse, entre 33°C (Shin et al., 2007) et 45°C (Domde et al., 2007), lors d'une réaction par polymérisation en chaîne, ou en anglais *polymerase chain reaction* (PCR). Ces conditions entraînent un appariement sur l'ADN à analyser et peuvent ainsi mener à l'amplification de fragments d'ADN. Puisque certains gènes et leur ordre peuvent être différents dans les génomes d'espèce ou sous-espèce différentes, les patrons de bandes devraient être théoriquement distinctifs d'une espèce à l'autre si l'amorce utilisée cible des régions génomiques divergentes. La RAPD a aussi été utilisée pour distinguer entre des isolats de la même espèce, comme pour *Pseudomonas aeruginosa* (Mahenthalingam et al., 1996), *Mycoplasma hyopneumoniae* (Artiushin et Minion, 1996) et *Salmonella enteritidis* (Lin et al., 1996). Étant une technique demandant très peu de matériel tout en étant peu coûteuse, elle est régulièrement utilisée pour observer des variations entre populations microbiennes (Atienzar et Jha, 2006). Mais la RAPD est régulièrement critiquée, car elle comporte peu de reproductibilité d'un laboratoire à l'autre. Malgré tout, dans l'étude menée par Hänninen et al. en 1995, les conditions choisies ont permis d'identifier des polymorphismes entre des isolats d'*A. salmonicida* sous-espèce *salmonicida*. Cependant, il y avait trop peu de divergence génétique pour étudier l'épidémiologie de la furonculose en Suède (Hänninen et al., 1995).

Dans deux études publiées en 2000, la technique de polymorphisme de longueur des fragments de restriction, ou en anglais *restriction fragment length polymorphism* (RFLP), a été utilisée pour mener des études épidémiologiques entre des isolats d'*A. salmonicida* sous-espèce *salmonicida*. Cette technique a été utilisée par Gracia et O'hici et leurs collaborateurs (García *et al.*, 2000; O'hici *et al.*, 2000). Concrètement, la RFLP consiste à digérer l'ADN par une ou des enzymes de restriction, puis à séparer l'ADN digéré par électrophorèse. Dans les deux cas, les profils de digestion suite à la RFLP ont été visualisés par électrophorèse en champs pulsés, ou en anglais *pulsed-field gel electrophoresis* (PFGE). Cette technique consiste à faire migrer l'ADN digéré sur gel d'agarose en champs électriques alternés. La PFGE, contrairement à l'électrophorèse traditionnelle, possède une résolution pouvant aller jusqu'à 10 Mb, alors que la résolution de l'électrophorèse traditionnelle ne peut dépasser 50 kb. Les techniques RFLP-PFGE permettent d'observer les polymorphismes affectant les fragments d'ADN de plus grandes tailles (Herschleb *et al.*, 2007).

Dans la première étude, Gracia et ses collaborateurs ont analysé les profils de restriction par RFLP-PFGE de 133 souches d'*A. salmonicida* sous-espèce *salmonicida* isolées dans 8 pays à la suite d'éclosions de furonculose. Les résultats indiquent l'existence de 16 profils de restriction, dont 15 profils possèdent entre 1 et 3 bandes de différentes, ce qui est insuffisant pour affirmer que les isolats étaient génétiquement hétérogènes (Tenover *et al.*, 1995). Dans l'étude de Gracia en 2000, 132 des 133 isolats d'*A. salmonicida* sous-espèce *salmonicida* étaient considérés génétiquement homogènes (García *et al.*, 2000). O'hici et ses collaborateurs, quant à eux, ont comparé la RAPD à la RFLP-PFGE en analysant 17 isolats d'*A. salmonicida* sous-espèce *salmonicida* avec les deux méthodes. Les résultats de l'étude de O'hici ont été analysés par la formule statistique d'indice et distance de Jaccard (Morgan *et al.*, 1993). Ces analyses statistiques ont été utilisées pour la construction des dendrogrammes de similarité par l'algorithme UPGMA disponible dans l'interface de phylogénie PHYLIP. Toujours dans l'étude de O'hici, les 17 isolats d'*A. salmonicida* sous-espèce *salmonicida* testés démontraient plus de 70% de similarité suite aux analyses statistiques pour chacune des deux techniques. En se référant à des valeurs de pourcentage de P supérieures au seuil de 70%, les isolats forment deux grands groupes (10 et 7 isolats pour chaque groupe). 8 des 10 isolats qui proviennent de l'Amérique du Nord se regroupent ensemble (les 2 autres proviennent de France), alors que 6 sur 7 isolats de l'Europe du Nord se regroupaient aussi ensemble (le dernier provenait des États-Unis). Outre les deux groupes qui sont observés prédominants

sur chaque continent, les isolats semblaient génétiquement homogènes, car ils présentaient tous des valeurs de pourcentage de P supérieures au seuil de similarité de 70%.

1.5. Un nouvel élément génétique chez *A. salmonicida* sous-espèce *salmonicida*

Le premier génome assemblé et annoté fut celui de l'isolat A449 (Reith *et al.*, 2008). En 2012, les séquences contiguës (*contigs*) d'un second génome d'*A. salmonicida* sous-espèce *salmonicida* 01-B526 ont été publiés (Charette *et al.*, 2012). Une analyse des séquences chromosomiques des deux isolats a permis la découverte d'un nouvel élément génétique de 51 kb présent dans la souche 01-B526, isolée au Québec, mais absent dans la souche A449, isolée en France. La prévalence de l'élément génétique a été étudiée par PCR dans une collection d'isolats d'*A. salmonicida* sous-espèce *salmonicida* en utilisant 15 paires d'amorces ciblant 13 ORF (*orf11*, *orf19*, *orf33*, *orf45*, *orf54*, *orf55*, *orf57*, *orf58*, *orf64*, *orf66*, *orf67*, *orf72* et *orf78*) et les deux sites d'insertion de l'élément (*Ins1* et *Ins2*). Les cibles PCR ont été choisies en se basant sur la possibilité que certains des ORF pussent coder pour de potentiels facteurs de virulence (*orf11*, *orf19*, *orf33*, *orf45*, *orf55* et l'*orf67*) ou pour des protéines avec des fonctions de connues (*orf54*, *orf57*, *orf58*, *orf64*, *orf66*, *orf72* et *orf78*) (travaux préliminaires de Francis Brochu, communication personnelle). Les résultats du criblage indiquaient que l'élément génétique de 51 kb était aussi présent chez d'autres isolats d'*A. salmonicida* sous-espèce *salmonicida* (Figure 2A), mais qu'il existait aussi deux autres variants (Mélanie Trudel, communication personnelle) (Figure 2B et C). L'un des deux variants semblait avoir le même site d'insertion que l'élément de 51 kb (PCR positives pour *Ins1* et *Ins2*) (Figure 2B), alors que le second semblait avoir un site d'insertion différent (PCR négatives pour *Ins1* et *Ins2*) (Figure 2C).

Puisque l'élément génétique de 51 kb découvert dans 01-B526 est absent dans la souche A449 et que deux variants de cet élément ont été observés dans d'autres isolats d'*A. salmonicida* sous-espèce *salmonicida*, il a été rapidement envisagé que tous les trois puissent appartenir à la même famille d'éléments mobiles de l'ADN.

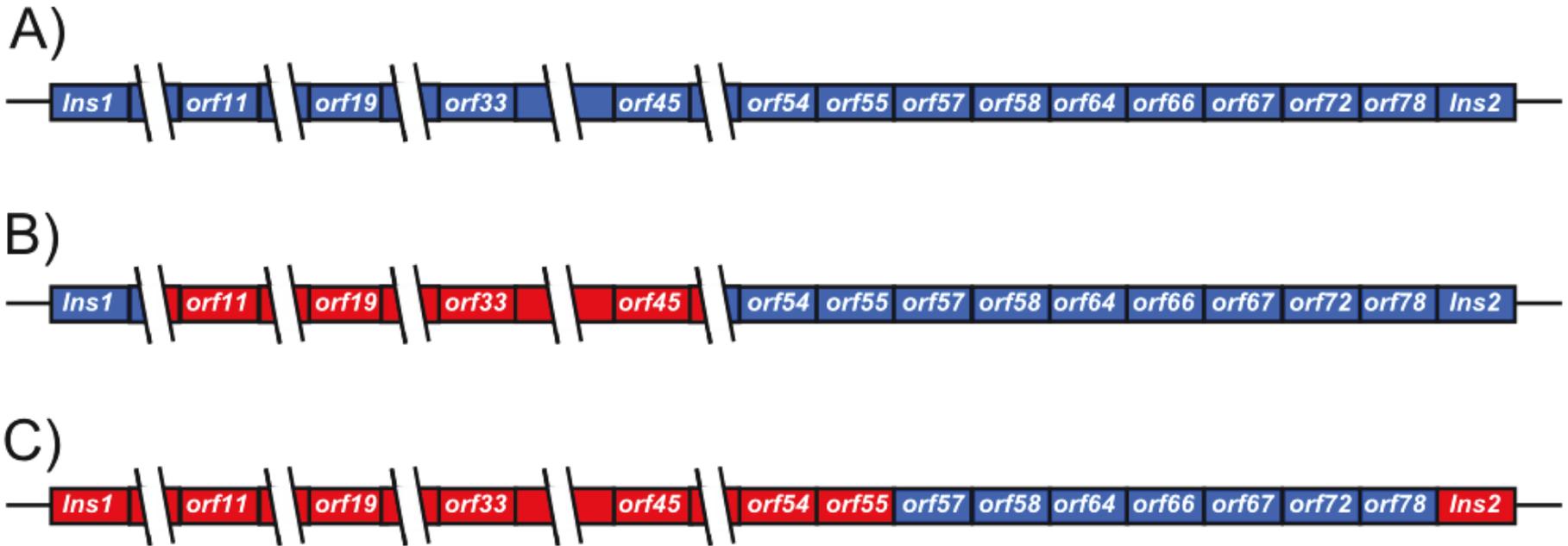


Figure 2 : Représentations schématiques des trois profils d'amplification observés lors du criblage d'une collection d'*A. salmonicida* sous-espèce *salmonicida* avec 15 paires d'amorces pour la présence de l'élément génétique de 51 kb. Les rectangles en bleu ou en rouge représentent les PCR pour lesquelles il y avait amplification ou non, respectivement. Les acronymes «orf» et «Ins» signifient cadre de lecture ouvert et site d'insertion, respectivement. A) Représentation des résultats des 15 cibles PCR de souches semblant posséder l'élément génétique de 51 kb. B) Représentation d'un variant de l'élément de 51 kb découvert chez certaines souches d'*A. salmonicida* sous-espèce *salmonicida*. C) Représentation des résultats PCR d'un deuxième variant de l'élément génétique de 51 kb.

1.6. Les éléments mobiles de l'ADN

L'ensemble des gènes d'une espèce bactérienne est appelé le pan-génome. L'assortiment de ces gènes comprend deux sous-ensembles, soit le génome de base (*core genome*) et le génome accessoire. Le génome de base englobe tous les gènes communs à toutes les souches constituant une espèce, alors que le génome accessoire comprend l'ensemble de tous les gènes non-essentiels généralement acquis par transfert horizontal de gènes (HGT). Le génome accessoire peut coder pour des fonctions apportant un bénéfice à la bactérie et ainsi augmenter la survie bactérienne dans certaines conditions. Le HGT contribue à la diversification ainsi qu'à l'adaptation des bactéries. Ce transfert s'effectue la plupart du temps via trois mécanismes d'échanges appelés la transformation, la conjugaison et la transduction. Dans certaines situations, les MGE peuvent être acquis par le HGT. Les MGE sont généralement responsables de l'instabilité génomique chez les bactéries. Cette instabilité peut entraîner des mutations ponctuelles ou des réarrangements génomiques, tels que des délétions, des insertions, des duplications, des amplifications, des inversions et des translocations (Darmon et Leach, 2014). La suite de cette section discutera seulement des instabilités chromosomiques résultant des réarrangements génomiques régis par les MGE.

Les MGE sont présents dans tous les domaines de la vie, aussi bien chez les organismes eucaryotes que chez les procaryotes, et il en résulte une impressionnante diversité. Le point commun entre les MGE est qu'ils sont tous des éléments transposables. Ce groupe comprend les séquences d'insertions (IS), les éléments transposables à répétitions inversées miniatures (MITE), les séquences palindromes extragénomiques répétitives (REP), les éléments bactériens en mosaïque intercalée (BIME), transposons, bactériophages et les îlots génomiques (GEI) (Darmon et Leach, 2014). Cette revue de littérature expliquera que les grandes lignes de la théorie sur les IS, les prophages et les GEI, des informations qui seront nécessaires pour comprendre les travaux de recherche effectués durant cette maîtrise.

1.6.1. Les séquences d'insertion

Les IS sont de courtes séquences d'ADN, variant normalement entre 0,7 et 2,5 kb. Elles contiennent un ou deux ORF qui codent pour les protéines responsables de leur propre transposition (transposase). Les IS sont flanquées par des séquences répétées

inversées (IR). L'insertion d'une IS dans le chromosome bactérien modifie le génome de l'organisme en dupliquant généralement le site d'insertion créant une séquence répétée directe (DR). La séquence DR du génome hôte reconnue par l'IS varie entre 2 et 14 nucléotides, dépendant de la nature de l'IS (Mahillon et Chandler, 1998). De plus, l'insertion d'une IS soit dans un ORF ou dans une région inter-génique peut inactiver un gène et/ou modifier l'expression de ce gène ou d'autres gènes situés en aval du site d'insertion. Tout dépendant de la nature du gène inactivé ou de l'expression génique modifiée, les conséquences cellulaires directes ou indirectes peuvent être un avantage ou un inconvénient pour la bactérie (Darmon et Leach, 2014; Mahillon et Chandler, 1998).

L'interaction d'une IS avec une autre molécule d'ADN possédant une séquence identique, par exemple la même IS, peut entraîner des réarrangements génomiques. La recombinaison de deux éléments qui sont placés dans la même orientation sur le chromosome entraîne une délétion de séquence, alors que la recombinaison entre deux IS d'orientation opposée entraîne une inversion de la séquence encadrée par ces éléments. Un fait important est que les IS peuvent soit s'insérer dans le chromosome bactérien, dans un bactériophage ou même dans un plasmide. Bien que les IS soient de courtes séquences d'ADN, leurs insertions ou excisions peuvent entraîner d'importantes instabilités génomiques suite à leur implication dans la recombinaison avec une autre molécule d'ADN possédant les mêmes segments. Dépendant de la finalité des modifications entraînées par les IS, ces MGE peuvent être considérés soit néfastes ou avantageux pour la cellule hôte. Sommes toutes, elles sont profitables à la cellule, car elles aident le génome de la cellule à évoluer (Darmon et Leach, 2014; Mahillon et Chandler, 1998).

Plusieurs études ont démontré des exemples de plasticité génomique chez *A. salmonicida* sous-espèce *salmonicida* due aux mouvements d'IS pouvant ainsi causer soit l'inactivation d'un gène ou le réarrangement de son génome (Gustafson *et al.*, 1994; Reith *et al.*, 2008; Tanaka *et al.*, 2012; Trudel *et al.*, 2013). Dans deux des cas, les IS ont été activées lors d'une incubation des souches à une température égale ou plus élevée que 25°C (Gustafson *et al.*, 1994; Tanaka *et al.*, 2012). Gustafson et ses collaborateurs ont observé l'inactivation du gène *vapA*, qui code pour une protéine qui forme la *A-layer*, suite à des incubations prolongées à une température de 30°C chez A449. Les IS impliquées dans cette inactivation du gène *vapA* sont l'ISAS1 et ISAS2 (Gustafson *et al.*, 1994). Il a aussi été montré qu'une incubation prolongée d'A449 à 25°C peut conduire à des réarrangements entre trois des ISAS11 du plasmide de grande taille pAsa5 entraînant

la délétion du système de sécrétion de type trois (profil de délétion de type 1) et dans certains cas, la délétion du locus avec une partie du plasmide pAsa5 (profil de délétion de type 2) (Tanaka *et al.*, 2012). Dans cette même étude, une incubation à 25°C des souches 01-B526 et 01-B516 a entraîné presque seulement des profils de délétion de type 1, mais les ISAS11 ne sont pas responsables de ces profils de délétion contrairement à ce qui est observé chez A449. Le mécanisme de délétion chez les souches 01-B526 et 01-B516 reste encore indéterminé (Tanaka *et al.*, 2012). La mobilité d'une des ISAS5 a été constatée chez la souche HER1104 suite à l'observation de la présence de cette IS dans le petit plasmide pAsal1B, un variant de pAsal1 qui ne possède pas d'ISAS5 (Trudel *et al.*, 2013). Finalement, il a été observé suite à l'analyse de la déviation du contenu en G+C du génome d'A449 de la présence d'une inversion chromosomique de 195 kb encadrée par deux ISAS4 (Reith *et al.*, 2008).

1.6.2. Les bactériophages

Les bactériophages, ou simplement nommés phages, sont des virus infectant spécifiquement les bactéries. Leur matériel génétique peut être soit de l'ARN ou de l'ADN et peuvent être dans les deux cas soit en simple ou en double brin (quatre possibilités). Cependant, la majorité des génomes phagiques répertoriés à ce jour possède un génome d'ADN double brin (Sherwood *et al.*, 2010). La quantité de particules phagiques est estimée dans la biosphère à 10^{31} , ce qui est supérieur d'un facteur dix à la quantité de bactéries peuplant ces mêmes environnements, ce qui rend le contact phage-bactérie incontournable (Bru et Hendrix, 2002).

La majorité des phages peuvent être classés en deux grandes catégories d'après leur cycle de vie : les phages lytiques et les phages tempérés. L'infection d'une bactérie par un phage lytique résultera en la lyse bactérienne. Il y aura ainsi la libération d'une grande quantité de particules phagiques dans l'environnement qui seront prêtes à perpétuer un nouveau cycle infectieux. Le phage tempéré a la capacité quant à lui de s'intégrer dans le génome de la bactérie qu'il infecte et d'entrer en dormance (phase lysogénique), il sera appelé à ce moment prophage. Lors d'une situation de stress extracytoplasmique, par exemple suite à une augmentation de la température ou à la présence d'un antibiotique dans le milieu, les prophages actifs sortiront de leur léthargie et ils débiteront la formation de particules phagiques (phase lytique) (Garcia-Russell *et al.*, 2009).

Les génomes bactériens peuvent contenir plusieurs prophages qui augmenteront de ce fait la quantité de gènes codés par les cellules. Par exemple, certaines bactéries sont composées jusqu'à 20% de gènes phagiques (Casjens *et al.*, 2000). Les phages peuvent coder pour des protéines qui augmenteront le *fitness* de son hôte (Canchaya *et al.*, 2004; Fortier et Sekulovic, 2013; Jackson *et al.*, 2011). De plus, il est connu que les prophages peuvent augmenter la virulence de plusieurs bactéries telles que *Corynebacterium diphtheriae* (Cerdano-Tarraga, 2003), *Vibrio cholerae* (Waldor et Mekalanos, 1996) et *Escherichia coli* O157:H7 (Plunkett III *et al.*, 1999). Les bactériophages peuvent aussi causer de l'instabilité génomique suite à la recombinaison entre deux prophages possédant des séquences semblables. La finalité de la recombinaison dépend de l'orientation des éléments l'un par rapport à l'autre dans le génome. Pour faire un parallèle, le résultat de la recombinaison entre deux prophages suit la même logique qu'une recombinaison entre deux IS (inversion/délétion).

Actuellement, neuf génomes assemblés et fermés d'*Aeromonas* sont disponibles sur la base de données publique GenBank (*A. salmonicida* sous-espèce *salmonicida* A449 (Reith *et al.*, 2008), *A. hydrophila* sous-espèce *hydrophila* ATCC 7966 (Seshadri *et al.*, 2006), *A. hydrophila* 4AK4 (Gao *et al.*, 2013), *A. hydrophila* AL09-71 (Pridgeon *et al.*, 2014b), *A. hydrophila* pc104A (Pridgeon *et al.*, 2014a), *A. hydrophila* YL17 (GenBank: CP007518.1), *A. hydrophila* ML09-119 (Tkedar *et al.*, 2013), *A. veronii* B565 (Li *et al.*, 2011) et *A. media* WS (Chai *et al.*, 2012) (GenBank: CP007567.1)). Dans ces génomes, trois prophages ont clairement été identifiés par prédictions bio-informatiques, deux dans le génome d'A449 (Reith *et al.*, 2008) et un dans le génome d'ATCC 7966 (Seshadri *et al.*, 2006), mais aucune étude fonctionnelle n'a été effectuée dans les trois cas. La présence de prophages par analyses bio-informatiques dans les sept autres génomes n'ont pas été investiguées (Li *et al.*, 2011; Chai *et al.*, 2012; Gao *et al.*, 2013; Tkedar *et al.*, 2013; Pridgeon *et al.*, 2014a, 2014b).

Pour compléter les prédictions bio-informatiques, des techniques expérimentales peuvent être utilisées pour induire les prophages. Certains auteurs ont induit le cycle lytique de phages lysogènes en exposant *A. hydrophila* ML09-119 (Hossain *et al.*, 2013) et *A. media* O18 (Beilstein et Dreiseikelmann, 2008), soit à la mitomycine C (antibiotique) ou aux UV, respectivement. L'induction démontrait la présence de cinq phages lysogènes chez *A. hydrophila* (AH1, AH2, AH3, AH4 et AH5) et d'un phage lysogène chez *A. media* (ϕ O18P). D'un autre côté, plusieurs phages lytiques infectant des

bactéries du genre *Aeromonas* ont été identifiés et seize ont été séquencés² (Petrov *et al.*, 2010; Comeau *et al.*, 2012; J H Kim *et al.*, 2012; Ji Hyung Kim, Son, Choi, *et al.*, 2012; Ji Hyung Kim, Son, Choresca, *et al.*, 2012; Shen *et al.*, 2012; Jun *et al.*, 2014). Les analyses révèlent qu'ils font majoritairement partie de la famille des *Myoviridae*, famille de phage reconnue pour posséder une queue contractile, dont le représentant type est le bactériophage T4 infectant *E. coli*. En bref, les phages lytiques sont davantage décrits dans la littérature que les prophages, car ces derniers ont été peu caractérisés chez les bactéries du genre *Aeromonas*.

1.6.3. Les îlots génomiques (GEI)

Le troisième groupe de MGE qui sera discuté dans ces travaux de recherche est les GEI. Ils ont été identifiés dans de nombreux organismes. Les GEI ont été observés pour la première fois suite à la comparaison des génomes d'un isolat pathogène avec un non-pathogène qui appartenaient tous deux à la même espèce bactérienne (Dobrindt *et al.*, 2004). Concrètement, ces MGE sont de grandes séquences génomiques, variant entre 10 à 200 kb, qui sont présents chez certains isolats bactériens, mais absents chez d'autres isolats de la même espèce. Les GEI sont des éléments mobiles de l'ADN qui ont été acquis par HGT. Leurs gènes codent normalement pour plusieurs protéines de mobilité pouvant être des gènes d'origines aussi variées que phagiques, plasmidiques, d'IS et de transposons. Cependant, au fil des générations, certains GEI ont perdu leur capacité de mobilité (Darmon et Leach, 2014). Ils s'intègrent généralement dans le chromosome bactérien, mais on peut aussi les repérer dans des plasmides ou dans des prophages. Ils sont flanqués de séquences répétées directes variant entre quelques nucléotides jusqu'à une centaine. L'insertion des GEI dans les génomes bactériens peut être décelée suite aux analyses comparatives du contenu en G+C, de la déviation en G+C et de l'usage des codons. Ils s'intègrent généralement dans des ARNt (Ho Sui *et al.*, 2009).

En plus des gènes de mobilités, les GEI encodent habituellement des gènes accessoires offrant un avantage à son hôte, lesquels augmentent les chances de survie et de colonisation de nouveaux environnements. Dépendant des fonctions des protéines codées par le GEI, il peut être caractérisé comme étant un îlot de pathogénicité (comme l'îlot SPI1 chez *Salmonella*) (Que *et al.*, 2013), un îlot métabolique (comme l'îlot de

² JF974294, NC_008208, NC_007022, NC_015251, NC_005135, NC_005260, NC_020879, NC_019528, NC_014635, KJ858521, NC_019543, NC_019538, NC_014636, NC_009542, NC_019527, NC_023688.

production de la gomme de xanthane chez *Xanthomonas*) (Lima *et al.*, 2008), un îlot de résistance (comme l'îlot de résistance aux antibiotiques AbaR7 chez *Acinetobacter baumannii*) (Sung *et al.*, 2012), un îlot de symbiose (comme chez *Mesorhizobium loti*) (Sullivan *et al.*, 2002), un îlot saprophytique (comme l'îlot encodant des adhésines chez des isolats d'*E. coli*) (Hacker et Carniel, 2001), un îlot écologique (comme l'îlot permettant la dégradation du phénol chez *Pseudomonas putida*) (Ravatn *et al.*, 1998), et un îlot de défense (comme chez *Shewanella* sp. ANA-3) (Makarova *et al.*, 2011). Finalement, les GEI sont des éléments génétiques qui ont été repérés dans plusieurs génomes bactériens et dont les fonctions variées concèdent à son hôte des avantages de survie dans certaines conditions de croissance.

Aucun GEI n'a encore été identifié dans la littérature chez *A. salmonicida* sous-espèce *salmonicida*. En date de novembre 2014, deux souches de cette sous-espèce bactérienne étaient présentes sur la base de données publique NCBI (A449 et 01-B526). L'alignement chromosomique démontrait la présence d'un élément génétique de 51 kb chez 01-B526 (Charette *et al.*, 2012). Les caractéristiques générales de cet élément doivent être investiguées pour déterminer à quel groupe de MGE il appartient, car il pourrait s'agir du premier GEI observé chez *A. salmonicida* sous-espèce *salmonicida*.

1.7. L'instabilité génomique liée aux séquences répétées en tandem

Les séquences répétées en tandem, ou en anglais *tandem repeats sequences* (TR *sequences*), sont des séquences nucléotidiques répétées en orientation directe qui sont placées l'une à la suite de l'autre. Tout dépendant si le motif nucléotidique constituant la séquence TR est conservé ou non dans chacune de ses répétitions, elle peut être classifiée soit par les termes identique/parfaite ou dégénérative/imparfaite, respectivement. Le terme identique/parfaite signifie que le motif de la séquence TR est identique dans toutes ses répétitions. Alors que le terme dégénérative/imparfaite caractérise une séquence TR où s'est produit une ou plusieurs mutations, donc le motif n'est pas identique dans toutes ses répétitions (Zhou *et al.*, 2014).

Une caractéristique intéressante des séquences TR est leur grande diversité entre les différentes espèces bactériennes, même les plus étroitement liées. Cette observation indique que les séquences TR sont soumises à des changements évolutifs rapides (Yang *et al.*, 2003; Mrázek, 2006; Kassai-Jáger *et al.*, 2008). Une analyse bio-informatique de

plus de 300 génomes de procaryotes démontrait que la distribution des séquences TR variait d'après l'espèce bactérienne, la taille du génome et le pourcentage en G+C (Mrázek *et al.*, 2007). Plus précisément, selon certains auteurs (Moxon *et al.*, 2006; Treangen *et al.*, 2009), les séquences TR avec de courts motifs (de 2 à 4 pb) sont plus abondantes dans les petits génomes (< 2 Mb) de bactéries qui sont des agents pathogènes adaptés à leur hôte avec un faible contenu en G+C (< 40%). Au contraire, les séquences TR avec des motifs de 5 à 11 pb sont observées plus fréquemment dans les génomes de grandes tailles (> 4 Mb) de bactéries non-pathogènes et opportunistes avec un pourcentage en G+C élevé (> 60%). Des analyses bio-informatiques ont démontré que le nombre de répétitions du motif d'une même séquence TR pouvait varier entre différentes souches de la même espèce (Yang *et al.*, 2003). Elles sont appelées séquences VNTR pour l'acronyme du terme anglais *variable-number tandem repeats*. Ces régions hyper-variables entre plusieurs individus sont utilisées pour effectuer des études de *multilocus VNTR analysis* (MLVA) dans le but de retracer le lien de parenté entre diverses souches de la même espèce, par exemple chez *Salmonella enterica* (Ramisse *et al.*, 2004) et chez *Xanthomonas oryzae* (Zhao *et al.*, 2012).

Les séquences TR sont l'une des causes de l'instabilité génomique chez les bactéries. Les régions contenant une séquence TR peuvent potentiellement subir une mutation soit par contraction (délétion) ou par expansion (insertion) de leur motif répété. Ces séquences peuvent se situer dans des régions inter-géniques ou dans des régions intra-géniques. Encore peu d'études fonctionnelles ont été effectuées sur les effets de ces séquences dans des régions inter-géniques. Certaines évidences indiquent qu'elles peuvent soit modifier l'affinité des protéines régulatrices de la transcription (Metruccio *et al.*, 2009), affecter l'initiation de la transcription en altérant la distance entre les éléments promoteurs -35 et -10 (van der Ende *et al.*, 2000; Carson *et al.*, 2002) ou encore modifier la stabilité de l'ARNm (Attia et Hansen, 2006).

D'un autre côté, plusieurs études relatent les effets de la variation des séquences TR intra-géniques chez des bactéries à Gram positif et négatif, qui sont fréquemment observées dans des gènes codant soit pour des facteurs de virulence, des composantes de la surface cellulaire et pour des enzymes de restriction-modification (Moxon *et al.*, 2006; Mrázek *et al.*, 2007; Power *et al.*, 2009; Janulczyk *et al.*, 2010). Le mécanisme d'action des séquences TR dépend de la nature de celle-ci. Plus précisément, si le motif de la VNTR n'est pas un multiple de trois (1, 2, 4, 5, etc.), les réarrangements de son nombre de répétitions auront pour effet soit de rétablir (phénotype ON) ou de briser

(phénotype *OFF*) le cadre de lecture du gène. Ce phénomène appelé variation de phase a été répertorié dans plusieurs gènes chez différentes espèces bactériennes (Zhou *et al.*, 2014). Dans une autre situation, si le motif de la séquence TR intra-génique est un multiple de trois (3, 6, 9, 12, etc.), les altérations de son nombre de répétitions semblent avoir pour but d'induire des changements structuraux et fonctionnels de la protéine (Gemayel *et al.*, 2010).

2. Hypothèses

Alors que les isolats d'*A. salmonicida* sous-espèce *salmonicida* ont longtemps été caractérisés comme génétiquement homogènes, la découverte des trois différents éléments génétiques démontre une première forme de diversité insoupçonnée chez cette bactérie. Ainsi, nous proposons que l'utilisation d'une technique RAPD optimisée puisse présenter une sensibilité suffisante pour mener une étude épidémiologique sur une collection d'*A. salmonicida* sous-espèce *salmonicida*.

De plus, sachant que les génomes d'*A. salmonicida* sous-espèce *salmonicida* possèdent de nombreuses familles de MGE, dont les IS sont la famille majoritairement observée, nous supposons que les différences entre les chromosomes des isolats 01-B526 et A449 seront principalement causées par les mouvements de ces derniers.

Parallèlement, nous proposons qu'il serait possible, via des criblages PCR et des analyses bio-informatiques des différents *AsaGEI*, de découvrir de nouveaux éléments génétiques pouvant servir d'indicateur de l'origine géographique des isolats d'*A. salmonicida* sous-espèce *salmonicida*. Ainsi, ces éléments génétiques seront une seconde façon de distinguer les isolats de la collection d'*A. salmonicida* sous-espèce *salmonicida*.

3. Objectifs spécifiques

- 3.1. Mener une étude épidémiologique sur une collection d'*A. salmonicida* sous-espèce *salmonicida* constituée de 139 isolats. Cet objectif englobe la mise au point de la technique RAPD, le criblage de la collection d'isolats et la création d'un dendrogramme.
- 3.2. Faire une analyse comparative des chromosomes des isolats A449 et 01-B526 d'*A. salmonicida* sous-espèce *salmonicida*. Plus précisément, répertorier les différences chromosomiques par l'utilisation de programmes bio-informatique appropriés. Cette analyse aidera à mieux connaître les éléments génétiques responsables de la variabilité chromosomique des isolats d'*A. salmonicida* sous-espèce *salmonicida*.
- 3.3. Analyser les trois éléments génétiques obtenus à la suite des séquençages d'isolats d'*A. salmonicida* sous-espèce *salmonicida*. Cet objectif englobe le criblage PCR de la collection de 139 isolats et les analyses bio-informatique des différents éléments génétiques.

4. Matériels et méthodes

4.1. Isolats bactériens et conditions de croissance

Les isolats d'*A. salmonicida* sous-espèce *salmonicida* qui ont été utilisés lors de cette étude sont présentés à l'annexe A et font partie de la collection de souches et d'isolats possédés par l'équipe du Dr Steve Charette. La liste comprend les informations suivantes pour chaque isolat: numéro dans la collection, nom de l'isolat, espèce de poisson infecté, lieu de l'isolement, référence et les résultats des tests génétiques PCR pour déterminer le type d'*AsaGEI* et la présence du prophage 3. Pour ces deux dernières informations, le lecteur est invité à lire la section Résultats. Les isolats ont été ensemencés sur Furonculose agar (10 g de Bactro-Tryptone, 5 g d'extrait de levure, 1 g de L-tyrosine, 2,5 g de NaCl et 15 g d'agar par litre d'eau distillé) (Hanninen et Hirverikoski, 1997) pour trois jours de croissance à 18°C.

4.2. Extractions d'ADN bactérien

L'ADN bactérien qui a été utilisé pour les amplifications RAPD a été extrait avec la trousse commerciale *ChargeSwitch gDNA Mini Bacteria Kit* (invitrogen, CS11301). Brièvement, cette technique consiste à purifier l'ADN bactérien qui est chargé négativement dans le tampon approprié par le biais de petites billes magnétiques chargées positivement à pH acide (<6,5). Au pH adéquat, l'ADN se liera aux billes magnétiques. Ensuite, la solution a été soumise à un champ magnétique qui attire les billes magnétiques liées à l'ADN. Ainsi, les impuretés peuvent être lavées de l'extraction. Aucune centrifugation n'est nécessaire dans ce protocole. L'ADN a été ensuite élué dans un tampon à pH basique (~8,5), ce qui inverse la charge des billes métalliques qui deviendront chargées négativement. Cette technique requiert un support à microtubes de 1,5 ml ayant une base magnétisée. Lors de cette étude, le *MagnaRack* (invitrogen, CS15000) a été utilisé pour toutes les extractions d'ADN.

La préparation de la collection de lysats d'ADN d'*A. salmonicida* sous-espèce *salmonicida* pour les amplifications PCR traditionnelles a été effectuée comme suit : 1) suspendre les cellules d'une culture bactérienne en milieu gélifié Furonculose provenant d'une sur boîte de Petri dans un eppendorf contenant 1 ml d'eau millipore, puis centrifuger

à 3500 g pendant 5 minutes, 2) enlever le surnageant et resuspendre les bactéries dans 1 ml de solution SWL (50 mM de KCl, 10 mM de Tris-base, 25 mM MgCl₂, 0,45% de Tween 20 et de NP40) (Charette et Cosson, 2004) en vortexant régulièrement, 3) faire lyser les cellules en incubant le tube eppendorf pendant 15 min à 95°C, 4) centrifuger à 16200 g pendant 5 minutes, 5) prélever seulement le surnageant et déposer dans un nouveau tube eppendorf, et finalement 6) les lysats ont été dosés au NanoDrop pour déterminer sa concentration en ADN et elle a été diluée à 100 ng/μL dans du Tris-Hcl 10 mM à pH=8.

4.3. Conception d'amorces RAPD

Afin de concevoir des amorces de neuf nucléotides (nt) avec le plus de sites d'hybridation possibles dans les génomes des isolats 01-B526 et A449 d'*A. salmonicida* sous-espèce *salmonicida*, un programme nommé *fasta_count_kmers*, codé en langage python et écrit par Éric Normandeau (bio-informaticien du RAQ) a été utilisé sur les séquences chromosomiques d'A449 et de 01-B526. Brièvement, le programme fonctionnant dans un environnement Unix coupe chaque génome en séquence de neuf nt. Ensuite, le programme compile le nombre de répétitions pour chaque séquence de neuf nt présente dans chaque chromosome. Les 18 amorces RAPD utilisées lors de cette étude ont été construites par cette méthode et sont répertoriées au Tableau 2.

4.4. RAPD

Des analyses RAPD ont été effectuées pour observer les polymorphismes entre quatre isolats d'*A. salmonicida* sous-espèce *salmonicida*, nommés 01-B526, 09-0167, A449 et HER1085 (Annexe A). Les réactions RAPD ont été effectuées dans un volume de 25 μL et contenaient des concentrations finales de : 1X du tampon «5X Green Go-Taq» (Promega, M791A), 160 μM de dNTPs (Promega, U1240), 3 μM de l'amorce RAPD, 0.03 U de GoTaq (Promega, M3005) et 25 ng d'ADN. Le programme RAPD était le suivant : 2 :30 min à 95°C, 40 cycles de 45 sec à 95°C, 45 sec à 37°C et de 2 :40 min à 72°C, puis se terminait par une élongation finale de 10 min à 72°C. Les patrons de bande suite à la RAPD ont été obtenus par électrophorèse sur gel d'agarose à 1% pendant 1h10 min à 100 volts. Le gel d'agarose a ensuite été immergé dans une solution de bromure d'éthidium (EtBr) à une concentration de 0.5 μg/mL pendant 30 min pour colorer l'ADN,

puis le gel a été immergé dans de l'eau distillée durant 30 min pour enlever le surplus de bromure d'éthidium.

4.5. PCR

Des amplifications PCR ont été effectuées lors des criblages de la collection d'isolats d'*A. salmonicida* sous-espèce *salmonicida* pour déterminer le type d'*AsaGEI* et la présence du prophage 3. La réaction PCR de 20 µL comprend des concentrations finales de : 1X du tampon «5X Green Go-Taq» (Promega, M791A), 160 µM de dNTPs (Promega, U1240), 0.65 µM des amorces 5' et 3', 0.025 U de GoTaq (Promega, M3005) et 100 ng d'ADN. Le programme PCR était le suivant : 2:30 min à 95°C, 30 cycles de 30 sec à 95°C, 30 sec à 55°C et 1 min à 68°C, puis extension finale de 5 min à 68°C. Les amplifications ont été migrées sur gel d'agarose à 0,8% contenant 0,5 µg/ml d'EtBr pendant 35 min à 90 volts. Pour chaque réaction PCR, les contrôles positif et négatif appropriés ont été utilisés. Les analyses ont été faites en duplicata. Les amorces qui ont été utilisées lors de cette étude sont listées à l'annexe B.

4.6. Séquençage et assemblage des génomes

L'ADN génomique total pour neuf souches d'*A. salmonicida* sous-espèce *salmonicida* (JF2267, 01-B522, M15879-11, HER1085, 09-0167, 2009-144K3, 2004-05 MF26, m23067-09 et JF3224) a été extrait avec l'aide de la trousse commerciale *DNeasy Blood and Tissue Kit* (Qiagen, Canada). L'ADN a été séquençé par pyroséquençage à la Plateforme d'Analyse Génomique de l'Institut de Biologie Intégrative et des Systèmes (IBIS; Université Laval) en utilisant différentes grandeurs de fragments et de bibliothèques *mate-pair*. Le protocole détaillé pour construire les bibliothèques *mate-pair* peut être trouvé dans le manuel *Paired End Rapid Library Preparation Method Manual* (Roche, USA). Les lectures de pyroséquençage 454 ont été assemblées *de novo* en utilisant le programme Newbler version 2.5.3 (Margulies *et al.*, 2005) avec les paramètres par défaut.

L'ADN génomique de la souche ATCC 33658 a été extrait comme décrit précédemment (Diamanka *et al.*, 2013). Le séquençage de nouvelle génération de cette souche a été effectué en utilisant un instrument MiSeq de la compagnie Illumina au *Michigan State University Research Technology Support Facility*, et la bibliothèque de séquençage a été préparé d'après le protocole standard du manuel *Illumina TruSeq DNA*

sample prep kit v2. Les différentes lectures ont été assemblées *de novo* avec le logiciel A5 pipeline (Tritt *et al.*, 2012).

4.7. Analyses bio-informatiques

Les résultats des alignements de séquence des chromosomes des isolats 01-B526 et A449 par les outils CONTIGuator (Galardini *et al.*, 2011) et BLASTn (algorithme *megablast*) obtenues sur le site NCBI ont été visualisés par le logiciel ACT (version 13.0.0) (Carver *et al.*, 2005). Les séquences répétées en tandem ont été déterminées par le serveur web Tandem Repeats Finder (Benson, 1999).

Les cadres ouverts de lecture (ORFs) pour chaque *AsaGEI* ont été identifiés par Prodigal (version 1.20) (Hyatt *et al.*, 2010), confirmés par GeneMark (version 2.5) (Besemer *et al.*, 2001) et visualisés sur le logiciel Artemis (version 16.0.0) (Carver *et al.*, 2012). Chaque ORF a été vérifié manuellement à l'aide des analyses BLASTn et BLASTp sur le site NCBI afin de déterminer l'identité de séquence et les fonctions potentielles des produits des ORFs (Altschul *et al.*, 1990). Les bases de données nucléotidiques *non-redundant* (nr/nt) et *whole-genome shotgun* (wgs) ont été utilisées pour les études BLASTn. La base de données protéique *non-redundant protein sequence* (nr) a été utilisée lors des analyses BLASTp. Les séquences annotées des *AsaGEI1a*, *AsaGEI1b* et *AsaGEI2a* ont été déposées sur GenBank sous les numéros d'accèsion KJ626178, KJ626179 et KJ626180, respectivement.

Un programme en langage Perl créé par Antony T. Vincent a été utilisé pour déterminer le pourcentage d'identité entre *AsaGEI1a* et trois régions génomiques provenant des souches *A. hydrophila* RB-AH (Bio-projet GenBank PRJNA253773), *A. jandaei* Riv2 (Bio-projet GenBank PRJNA237126) et *A. salmonicida* sous-espèce *achromogenes* AS03 (Han *et al.*, 2013). Le programme balaye les régions d'intérêt par fenêtre de 350 pb et il utilise l'algorithme heuristique de *fasta36* (Pearson et Lipman, 1988) pour rechercher les identités entre les séquences comparées. Le seuil a été fixé à un minimum de 60% d'identité sur 80% de la longueur. Les ORF des prophages et les ARNt ont été déterminés par les serveurs web Prophage Finder (Bose et Barber, 2006) et tRNAscan-SE (Lowe et Eddy, 1997), respectivement. Le serveur web PipMaker (Schwartz *et al.*, 2000) a été utilisé pour les analyses d'alignement Dotplot. Le programme Easyfig (Sullivan *et al.*, 2011) a été utilisé pour construire les Figures 6A, 7, 8, 9, 12 et 13 représentant les alignements d'éléments génomiques.

Les trois régions chromosomiques utilisées pour les analyses de génomique comparative d'*AsaGEI1a* ont été extraites des génomes d'*A. hydrophila* RB-AH (Bio-projet GenBank PRJNA253773), d'*A. jandaei* Riv2 (Bio-projet GenBank PRJNA237126) et d'*A. salmonicida* subsp. *achromogenes* AS03 (Han *et al.*, 2013). Le génome de la souche *A. salmonicida* sous-espèce *salmonicida* A449 a été utilisé comme référence pour établir les sites d'insertion des *AsaGEI* (numéro d'accèsion GenBank CP000644.1). Les deux prophages putatifs utilisés pour les analyses de génomique comparative avec le prophage 3 d'*A. salmonicida* sous-espèce *salmonicida* ont été extraits des génomes d'*A. hydrophila* 4AK4 (Gao *et al.*, 2013) et d'*A. veronii* B565 (Li *et al.*, 2011).

5. Résultats

5.1. Analyses RAPD

Tout d'abord, pour mener une étude épidémiologique sur les isolats d'*A. salmonicida* sous-espèce *salmonicida*, il est nécessaire d'utiliser un outil sensible pour détecter une gamme de polymorphismes entre les isolats pour ainsi comparer les souches pathogènes. Dans ces travaux de recherche, la technique RAPD a été optimisée en testant plusieurs amorces de 9 nt possédant un nombre élevé de lieux d'hybridation dans le chromosome. Ces amorces ont été prédites par un programme bio-informatique à partir des chromosomes des deux souches auparavant séquencé (A449 et 01-B526). Cette prédiction permet d'utiliser des amorces possédant un nombre optimisé de sites d'appariement. Les amorces sont conçues pour s'hybrider à des séquences présentes fréquemment dans les chromosomes d'*A. salmonicida* sous-espèce *salmonicida*, et ainsi, augmenter la possibilité d'amplification des polymorphismes entre les isolats testés (Tableau 2). Au total, 18 amorces qui sont répertoriées au tableau 2 ont été testées en RAPD sur quatre extraits d'ADN d'*A. salmonicida* sous-espèce *salmonicida*, soit 01-B526, HER1085, A449 et 09-0167 (Annexe A). Ces quatre isolats de la collection d'*A. salmonicida* sous-espèce *salmonicida* ont été sélectionnés pour effectuer les expériences RAPD pour trois raisons : 1) deux isolats provenaient de l'Europe (A449; HER1085) et les deux autres avaient été isolés au Québec (01-B526; 09-0167), 2) chacun des isolats possède soit un génotype différent de GEI illustré à la Figure 2 (Figure 2A pour 01-B526; Figure 2B pour HER1085; Figure 2C pour 09-0167) ou aucun GEI (A449) et 3) les génomes de tous les isolats ont déjà été séquencés et sont disponibles.

Sur les 18 amorces RAPD testées en triplicate biologique, seize présentaient des patrons de bandes homogènes (voir exemple à la Figure 3B) et deux amorces présentaient des patrons de bandes montrant des différences subtiles entre les souches testées. Les amorces JG-RAPD11 et JG-RAPD20 avaient respectivement deux et une bandes qui différaient selon les divers profils d'amplification (Figure 3A et 3C). Donc, pour ces deux oligonucléotides, il y avait une somme de trois sites de polymorphisme qui ont été amplifiés.

Conséquemment aux résultats de criblage des 18 amorces RAPD, cette technique n'est pas assez sensible pour observer une quantité suffisante de polymorphismes pour évaluer la divergence génétique entre les quatre souches testées. Donc, la RAPD ne sera

pas utilisée pour mener une étude épidémiologique sur la collection de 139 isolats d'*A. salmonicida* sous-espèce *salmonicida*.

Tableau 2 : Liste des amorces RAPD testées avec une quantité élevée de site d'hybridation dans les chromosomes d'A449 et de 01-B526 et leurs résultats respective.

Nom de l'amorce RAPD	Séquences 5' → 3'	Nombre d'appariements			Température d'hybridation (en °C)	Observation de profil d'amplification divergent
		génomique		Élément de 51 kb		
		A449	01-B526			
JG-RAPD9	CTGCTGCTG	968	977	10	31.1	Non
JG-RAPD10	CTGGCGCTG	737	741	5	37.2	Non
JG-RAPD11	CTGCTGGCG	674	681	6	37.2	Oui (Figure 3A)
JG-RAPD12	CAGCAGCAG	968	977	10	31.1	Non
JG-RAPD13	CAGCGCCAG	737	741	5	37.2	Non
JG-RAPD14	CGCCAGCAG	674	681	6	37.2	Non
JG-RAPD15	CTGCTGGCC	931	932	4	36.1	Non
JG-RAPD16	GGCCAGCAG	931	932	4	36.1	Non
JG-RAPD17	CCTGCTGGC	847	845	4	36.1	Non
JG-RAPD18	GCCAGCAGG	847	845	4	36.1	Non (Figure 3B)
JG-RAPD19	GCTGCTGGC	803	810	6	37.6	Non
JG-RAPD20	GCCAGCAGC	803	810	6	37.6	Oui (Figure 3C)
JG-RAPD21	GCCCTGCTG	685	692	6	36.1	Non
JG-RAPD22	CAGCAGGGC	685	692	6	36.1	Non
JG-RAPD23	CTGCTCGGC	650	653	3	36.6	Non
JG-RAPD24	GCCGAGCAG	650	653	3	36.6	Non
JG-RAPD25	TGGCGCTGG	739	747	8	39.7	Non
JG-RAPD26	CCAGCGCCA	739	747	8	39.7	Non

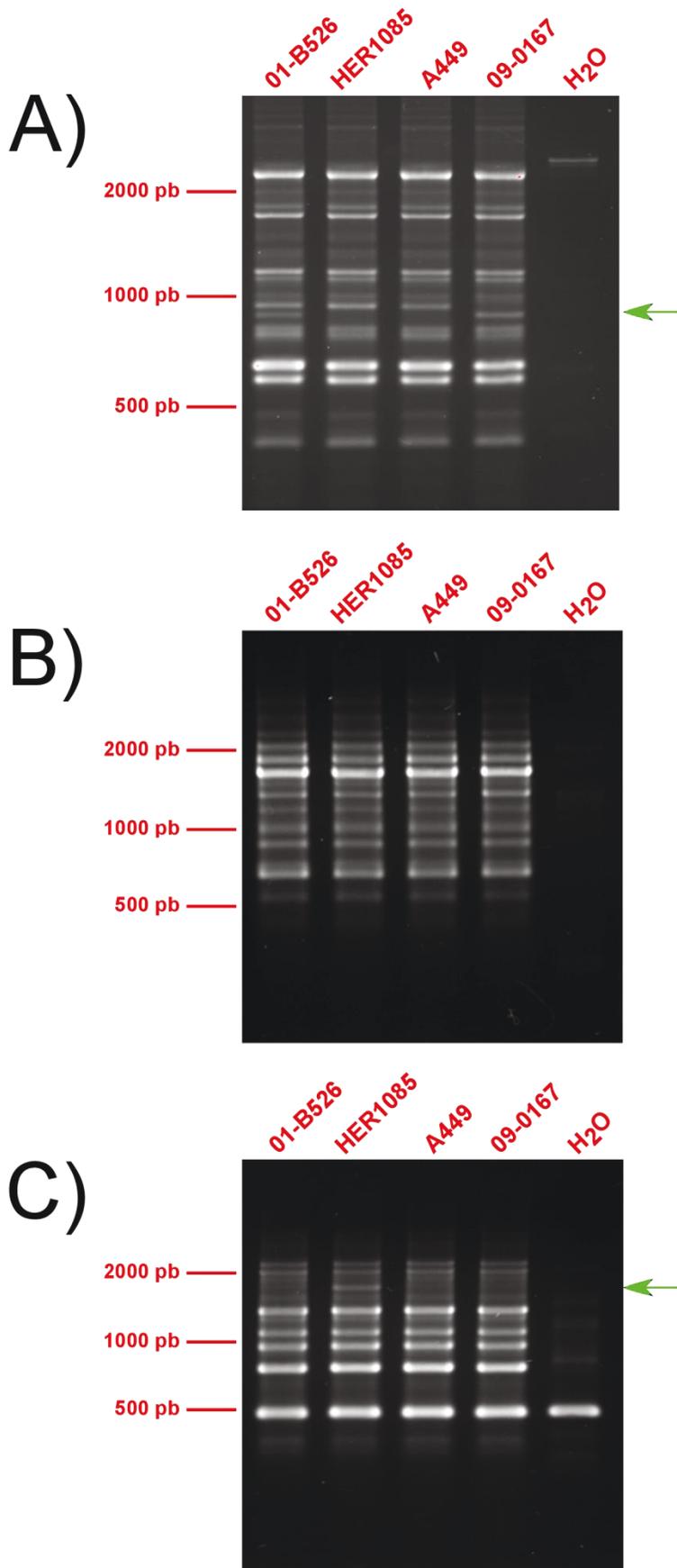


Figure 3 : Résultats de trois RAPD migrés en électrophorèse sur gel d'agarose. Les noms des souches utilisées pour chaque RAPD sont inscrits au-dessus de chaque puits. De l'eau millipore (H₂O) a été utilisée comme contrôle négatif pour chaque expérience. A) Résultat de la RAPD avec l'amorce JG-RAPD11. La flèche verte indique l'emplacement approximatif des deux bandes suggérant un polymorphisme chez les souches testées. B) Résultat de la RAPD avec l'amorce JG-RAPD18. C) Résultat de la RAPD avec l'amorce JG-RAPD20. La flèche verte indique l'emplacement approximatif de la bande suggérant un polymorphisme chez la souche HER1085.

5.2. Étude des variations chromosomiques entre 01-B526 et A449

Alors que les analyses de polymorphisme par la technique RAPD avec des amorces optimisées permettent d'observer seulement des différences subtiles entre les quatre souches testées, il serait important de mieux comprendre les différences chromosomiques parsemant les génomes d'*A. salmonicida* sous-espèce *salmonicida* pour ainsi déterminer des cibles potentielles pour distinguer les isolats.

Dans cette section, les dissemblances chromosomiques entre les souches A449 et 01-B526 ont été répertoriées manuellement. Deux programmes ont été utilisés pour aligner les chromosomes. Les deux outils sont BLASTn du web serveur NCBI et CONTIGuator. Le résultat de chaque alignement a été visualisé par le programme ACT conçu par l'Institut Sanger. Au total, 32 différences ont été observées (Tableau 3). Les caractéristiques de chaque dissemblance ont fait émerger plusieurs tendances. Ainsi, les dissemblances chromosomiques ont été classées en cinq groupes qui sont : mouvements d'IS (n=13), variations du nombre de répétitions dans des séquences TR qui sont soit intra ou inter-géniques (n=12), insertion/délétion d'une séquence dans un gène (n=5), insertion/délétion d'un GEI (n=1) et un site de polymorphisme multi-évènementiel (n=1).

5.2.1. Dissemblances impliquant des IS

Six des treize dissemblances impliquant des IS sont causées par l'insertion dans le chromosome d'A449 de trois ISAS2 (Tableau 3, dissemblances 1, 8 et 32) et de trois ISAS9 (Tableau 3, dissemblances 10, 19 et 31). Dans l'ensemble du génome, trois et une ISAS9 sont distribuées sur le chromosome et le plasmide pAsa4, respectivement. Les quatre ISAS9 sont toutes insérées en 5' du premier gène d'une ISAS5, dans sa séquence non-codante. Contrairement aux quatre ISAS9, l'ISAS2 semble moins restrictive dans ses sites d'insertion. Dans le cas des trois ISAS2 chromosomiques chez A449, une et deux sont insérées dans des régions intra-géniques ou inter-géniques, respectivement. Les séquences DR pour les trois ISAS2 chromosomiques varient entre des motifs de trois ou de quatre nucléotides (CTT; CTTC; GCTT), donc la probabilité de rencontrer le même motif dans un génome est élevée.

Quatre des treize variations causées par des IS sont dues à des ISAS3 insérées dans différents lieux du chromosome. Les quatre ISAS3 sont localisées à l'intérieur de séquences géniques. Trois ISAS3, absentes chez A449, sont insérées chez 01-B526 dans

des gènes codants respectivement pour une *putative transposase (TniA)* (Tableau 3, dissemblance 6), une *DDE endonuclease (ISAS4)* (Tableau 3, dissemblance 11) et pour une protéine *maltose operon periplasmic (Malm)* (Tableau 3, dissemblance 26). La dernière ISAS3 s'est insérée chez A449 dans un gène codant pour une *type-1Aa cytolytic delta-endotoxin* (Tableau 3, dissemblance 16).

Deux des trois dernières dissemblances causées par des IS semblent responsables de deux délétions d'approximativement 4500 et 5500 pb chez 01-B526, une par l'excision d'une ISAS8 (Tableau 3, dissemblance 2) et l'autre par la recombinaison entre deux ISAS3 placées dans la même orientation sur le chromosome (Tableau 3, dissemblance 15), respectivement. La dernière dissemblance aurait été causée par la recombinaison entre deux ISAS4 inversées l'une par rapport à l'autre qui aurait entraîné une inversion chromosomique de 195 kb observable chez A449 (Reith *et al.*, 2008) (Tableau 3, dissemblance 27).

5.2.2. Variations du nombre de répétitions dans des séquences TR

Suite aux analyses bio-informatiques, des dissemblances dans des séquences TR ont été observées en douze sites entre les deux chromosomes. Trois de ces douze variations ont été observées dans trois gènes qui codent respectivement pour une protéine avec des motifs RTX (*asx*) (Tableau 3, dissemblance 4), pour une *acriflavine resistance protein B (acrB)* (Tableau 3, dissemblance 9) et pour une *ABC type transporter ATPases subunit (ASA_2812)* (Tableau 3, dissemblance 21). La séquence TR dans le gène *asx* a une longueur de 336 pb et elle est répétée 6,1 fois chez A449 et 2,1 fois dans le même gène chez 01-B526. Le polymorphisme du gène *acrB* est une séquence TR de 9 pb répétée 6,7 fois chez A449 et 5,7 fois dans le même gène chez 01-B526. La dernière séquence TR identifiée dans le gène appelé *ASA_2812* a une taille de 5 pb et elle est répétée 3,8 fois chez A449 et 4,8 fois chez 01-B526. L'insertion/délétion d'une répétition de 5 pb dans le gène *ASA_2812* (annotation d'A449) a décalé le cadre de lecture (Figure 4A) et ainsi modifié une partie de la séquence protéique (Figure 4B), mais comme les séquences TR sont à la fin du gène, seulement les derniers acides aminés sont modifiés entre la protéine d'A449 et celle de 01-B526.

Neuf des douze polymorphismes causés par des séquences TR sont situés dans des régions inter-géniques (Tableau 3, dissemblances 5, 7, 12, 13, 22, 23, 24, 28 et 30). La longueur et le nombre des répétitions pour ces neuf séquences TR varient entre 6 à 10

pb et de 2 jusqu'à 21, respectivement. Voici deux exemples de séquences TR dans des régions inter-géniques: présence d'une séquence de 6 pb est répétée 7,2 fois chez A449 et 8,2 fois chez 01-B526 (Tableau 3, dissemblance 7) et d'une séquence de 9 pb répétée 32 fois chez A449 et 10 fois chez 01-B526 (Tableau 3, dissemblance 30).

Les douze séquences TR sont dites identique/parfaite. En d'autres termes, le motif nucléotidique est conservé dans toutes ses répétitions pour chacune des séquences TR, donc il n'y a eu aucune érosion génétique (Zhou *et al.*, 2014). Neuf séquences TR sont situées dans des régions inter-géniques, alors que les trois autres séquences TR sont dans des régions intra-géniques.

Théoriquement, les séquences TR chez une bactérie opportuniste possédant un génome > 4 Mb avec un contenu en G+C ~ 60% ont une taille majoritaire entre 5 et 11 pb (Moxon *et al.*, 2006; Treangen *et al.*, 2009). *A. salmonicida* sous-espèce *salmonicida* est un agent pathogène opportuniste qui possède un génome ~ 4,7 Mb avec un contenu en G+C de 58,5% (Reith *et al.*, 2008). La longueur des motifs des douze séquences TR entre A449 et 01-B526 sont de 5, 6, 8, 9, 10 et 336 nucléotides (Tableau 3). On peut constater que la longueur de onze des douze motifs nucléotidiques respecte les observations de Moxon et de Treangen. La quantité totale des séquences TR présentes dans le génome d'*A. salmonicida* sous-espèce *salmonicida* n'a pas été investiguée, donc nous ne savons pas si cette tendance est répandue à l'ensemble des séquences TR génomiques ou seulement à celles démontrant des différences entre les chromosomes d'A449 et 01-B526.

5.2.3. Dissemblances impliquant l'insertion/délétion d'une séquence dans un gène

Cinq des trente-deux dissemblances répertoriées entre les deux chromosomes étaient causées par l'insertion/délétion d'une séquence dans une région intra-génique. Les séquences varient entre 37 et 347 pb et dans toutes les situations, elles sont présentes chez A449, mais absentes chez 01-B526. La longueur des séquences en pb et les gènes associés dans lesquels elles ont été observées sont : 79 pb dans une *pyruvate formate lyase* (Tableau 3, dissemblance 3), 37 pb dans une *diguanylate cyclase* (Tableau 3, dissemblance 17), 347 pb dans une *tryptophanase* (Tableau 3, dissemblance 18), 84 pb dans une *putrescine/spermidine* (Tableau 3, dissemblance 25) et 63 pb dans une *ammonium transporter* (Tableau 3, dissemblance 29). Dans les cinq cas, les insertions/délétions dans les gènes ne provoquent pas un bris de l'ORF (Figure 5).

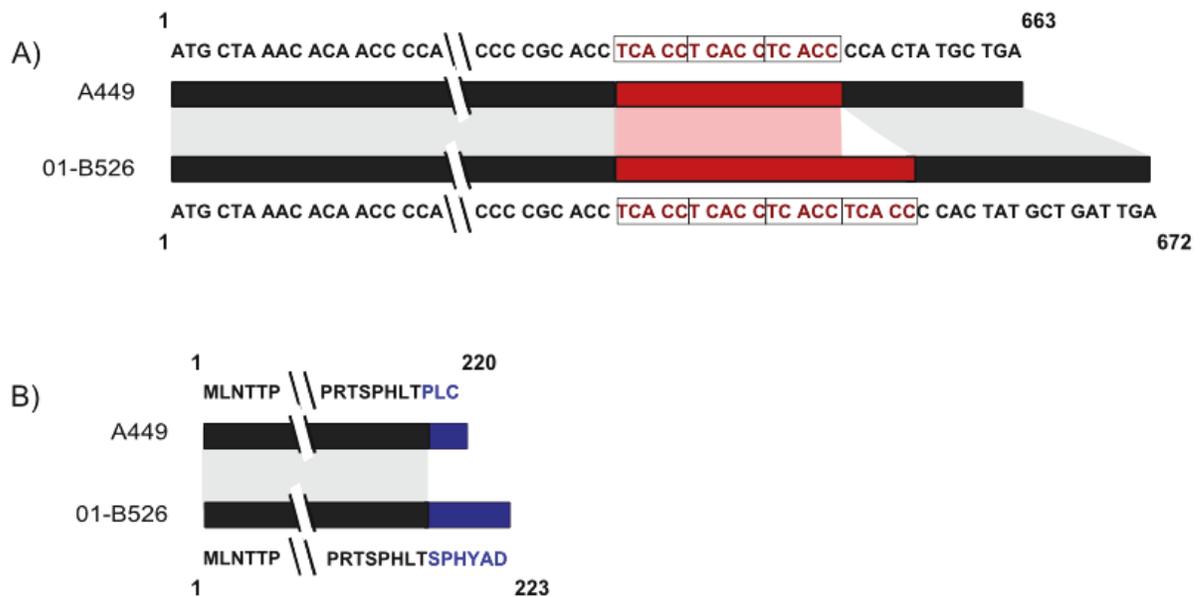


Figure 4 : Représentation des alignements nucléotidique (A) et protéique (B) de l'*ABC-type transporter, ATPase subunit* d'A449 par rapport à celle chez 01-B526. Les sections en gris sont des régions conservées entre les deux séquences nucléotidiques (A) et protéiques (B). La région en rouge représente le polymorphisme des séquences en tandem de 5 pb (entourée des encadrés) répétées 3 fois chez A449 et 4 fois chez 01-B526 (A). La région en bleu de l'alignement protéique (B) représente la section divergente entre les deux protéines *ABC-type transporter* d'A449 et de 01-B526.

5.2.4. Site de polymorphisme multi-évènementiel

Une des trente-deux dissemblances observées entre les deux chromosomes semble issue de plusieurs évènements distincts dans une même fenêtre approximative de 5400 pb (Tableau 3, dissemblance 14). Cette région chez A449, lorsque comparée à la même région dans 01-B526, semble avoir subi quatre évènements différents : une duplication d'une séquence de 6000 pb, deux insertions (une ISAS5 et deux gènes codant pour des protéines hypothétiques) et une délétion de 76 pb dans une région inter-génique. L'ordre chronologique de ces quatre évènements n'a pas été investigué en détail, car cela n'était pas l'objectif de cette recherche.

5.2.5. Différences impliquant l'insertion/délétion d'un GEI

Les analyses des divergences génétiques entre A449 et 01-B526 révèlent que l'insertion de 51 kb est la différence majeure entre les deux chromosomes (Tableau 3, dissemblance 20). L'étude de cet élément génétique et de ses deux variants (Figure 2B et C) pourrait mener à la découverte d'un indicateur qui serait utilisé pour suivre les isolats pathogènes d'*A. salmonicida* sous-espèce *salmonicida* dans le temps et l'espace. Ce GEI sera étudié plus en détail dans la section 5.3.

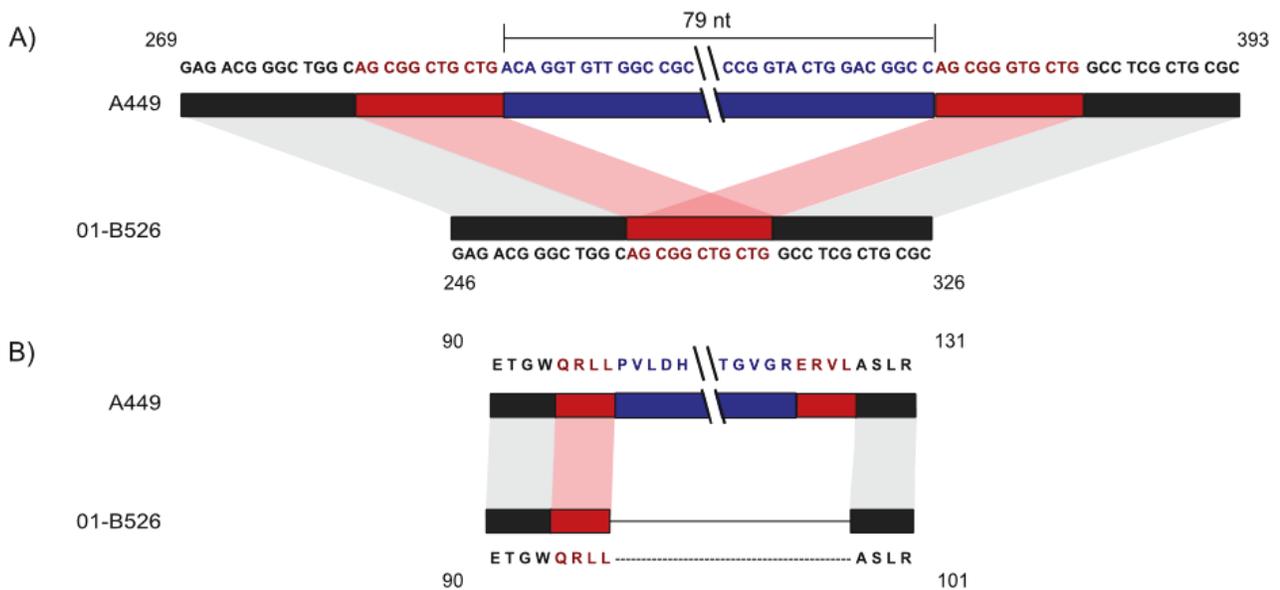


Figure 5 : Représentation d'une section de l'alignement nucléotidique (A) et de l'alignement protéique (B) de la *pyruvate formate lyase* de la souche A449 par rapport à la *pyruvate formate lyase* chez 01-B526. Les régions en rouge représentent le site d'insertion de la séquence de 79 nt qui est représentée en bleu.

Tableau 3 : Listes des dissemblances observées entre le chromosome d'A449, souche qui est utilisée comme référence, par rapport au chromosome de la souche 01-B526.

Numéro de la dissemblance	Position chez A449	Descriptions des dissemblances génétiques
1	59 753 à 61373	Présence d'une ISAS9 dans une région inter-génique (contiguë à une ISAS5)
2	153 123 à 157 615	Présence d'une séquence supplémentaire de ~4500 pb (présence d'une ISAS8 dans la séquence)
3	634 046 à 634 126	Présence d'une séquence supplémentaire de 79 pb dans un gène codant pour une <i>pyruvate formate lyase</i> (ASA_0621). Site d'insertion putatif « AGCGGCTGCTG »
4	846 805 à 848 404	Séquence en tandem de 336 pb répétée 6,1 fois chez A449 et 2,1 fois chez 01-B526 dans des gènes codants pour une protéine avec des motifs RTX
5	874 884 à 874 924	Séquence en tandem de 6 pb « TTGTTT ». Cette répétition est présente 4,3 fois chez A449 et 5,3 fois chez 01-B526 dans des régions inter-géniques
6	1 073 255	Absence d'une ISAS3 qui s'est insérée dans un gène codant pour une <i>TniA transposase</i> (<i>tniA</i>)
7	1 012 677 à 1 027 718	Séquence en tandem de 6 pb « TGCTGA ». Cette répétition est présente 7,2 fois chez A449 et 8,2 fois chez 01-B526 dans des régions inter-géniques
8	1 356 769 à 1 358 395	Présence d'une ISAS9 dans une région inter-génique (contiguë à une ISAS5)
9	1 375 965 à 1 376 023	Séquence en tandem de 9 pb « CAGCAGGAT ». Cette répétition est présente 6,7 fois chez A449 et 5,7 fois chez 01-B526 dans des gènes codants pour une <i>acriflavine resistance protein B</i> (<i>acrD</i>)
10	1 637 717 à 1 638 802	Présence d'une ISAS2 dans un gène codant pour une protéine hypothétique
11	1 680 892	Absence d'une ISAS3 qui s'est intégrée dans une ISAS4
12	2 008 856 à 2 008 708	Séquence en tandem de 9 pb « TCACGCAGG ». Cette répétition est présente 5,9 fois chez A449 et 6,9 fois chez 01-B526 dans des régions inter-géniques
13	2 020 404 à 2 020 453	Séquence en tandem de 8 pb « GGATAACG ». Cette répétition est présente 6,3 fois chez A449 et 9,3 fois chez 01-B526 dans des régions inter-géniques
14	2 141 453 à 2 146 879	Duplication d'une séquence de ~6000 pb, présence d'une ISAS5 et d'une séquence possédant deux CDS encodant des protéines hypothétiques.
15	2 253 680 à 2 259 108	Présence d'une séquence supplémentaire de ~5500 pb entre deux ISAS3 directes
16	2 263 378 à 2 264 603	Présence d'une ISAS3 dans un gène codant pour une <i>cytolytic delta-endotoxin (insecticidal protein)</i> (<i>btt</i>)
17	2 323 382 à 2 323 420	Présence d'une séquence supplémentaire de 37 pb dans un gène codant pour une <i>diguanylate cyclase</i> (ASA_2192). Site d'insertion putatif « CCAGCGCCGC »
18	2 472 094 à 2 472 442	Présence d'une séquence supplémentaire de 347 pb dans un gène codant pour une <i>tryptophanase</i> (<i>tnaA</i>). Site d'insertion putatif « TGACCGACA »
19	2 747 321 à 2 748 405	Insertion d'une ISAS2 dans une région inter-génique
20	2 851 678	Absence d'une séquence correspondante à un élément génétique de 51 kb qui est inséré dans un gène codant pour une protéine

		hypothétique (ASA_2666)
21	3 015 935 à 3 015 949	Séquence en tandem de 5 pb « TCACC ». Cette répétition est présente 3,8 fois chez A449 et 4,8 fois chez 01-B526 dans des gènes codants pour une <i>ABC-type transporter ATPase subunit</i>
22	3 038 306 à 3 038 348	Séquence en tandem de 9 pb « AACGTAAT ». Cette répétition est présente 4,8 fois chez A449 et 6,8 fois chez 01-B526. Dans une région inter-génique
23	3 043 078 à 3 043 123	Séquence en tandem de 10 pb « ATGATGAGCC ». Cette répétition est présente 4,4 fois chez A449 et 3,4 fois chez 01-B526. Dans une région inter-génique
24	3 139 453	Séquence en tandem de 9 pb « CCGTTGTGT ». Cette répétition est présente 4,8 fois chez A449 et 10,2 fois chez 01-B526. Dans une région inter-génique
25	3 257 446 à 3 257 531	Présence d'une séquence supplémentaire de 84 pb dans un gène codant pour une <i>putrescine/spermidine (ASA_3022)</i> . Site d'insertion putatif « CTGGGCTG »
26	3 431 000	Absence d'une ISAS3 qui est insérée dans un gène codant pour une <i>maltose operon periplasmic protein (malM)</i>
27	3 964 442 à 4 157 600	Inversion de 195 kb encadrée par deux ISAS4 inversées
28	4 016 333 à 4 016 354	Séquence en tandem de 8 pb « TTGCCTGA ». Cette répétition est présente 2,6 fois chez A449 et 6,6 fois chez 01-B526 dans des régions inter-géniques de l'inversion
29	4 310 411 à 4 310 475	Présence d'une séquence supplémentaire de 63 pb dans un gène codant pour une <i>ammonium transporter (amt)</i> . Site d'insertion putatif « CGGCGCCCTGGT »
30	4 364 453 à 4 564 543	Séquence en tandem de 9 pb « TTGAGGCAG ». Cette répétition est présente 32,1 fois chez A449 et 11,1 fois chez 01-B526 dans des régions inter-géniques
31	4 573 114 à 4 574 198	Présence d'une ISAS2 dans une région inter-génique
32	4 631 735 à 4 633 358	Présence d'une ISAS9 dans une région inter-génique (contiguë à une ISAS5)

Légende

-  Modification régie par une insertion/délétion ou recombinaison entre des IS
-  Insertion/délétion de séquences entraînant une modification d'un CDS
-  Variations du nombre de répétitions de séquences TR
-  Site de polymorphisme multi-évènementiel
-  Dissemblance causée par l'insertion/délétion d'un GEI

5.2.6. Analyses des sites de polymorphisme dans dix autres isolats d'*A. salmonicida* sous-espèce *salmonicida*.

Les sites génomiques démontrant un polymorphisme de séquence qui ont été répertoriées entre les chromosomes des souches A449 et 01-B526 se révèlent être des cibles potentielles pour évaluer la divergence génétique entre les isolats d'*A. salmonicida* sous-espèce *salmonicida*. Dans l'optique de trouver des loci chromosomiques discriminatoires entre les isolats, le nombre de répétitions de cinq séquences TR et la présence ou l'absence de courtes insertions dans cinq gènes (Tableau 4) ont été déterminés dans dix autres isolats d'*A. salmonicida* sous-espèce *salmonicida* séquencés par le laboratoire Charette. Seulement cinq régions comportant des séquences TR ont été utilisées pour ces analyses et elles ont été choisies au hasard (dissemblances 12, 13, 21, 22 et 28 listées au Tableau 3). Les cinq dissemblances causées par l'insertion de courtes séquences intra-géniques variant entre 37 et 347 pb ont été utilisées pour effectuer les analyses. Ces cinq séquences ont été observées dans les gènes *ASA_0621*, *ASA_2192*, *tnaA*, *ASA_3022* et *amt*, d'après l'annotation de la souche A449 (dissemblances 3, 17, 18, 25 et 29 listées au Tableau 3, respectivement). Les analyses BLASTn ont été effectuées dans les génomes séquencés des isolats JF2267, 01-B522, M15879-11, 09-0167, 2009-144K3, 2004-05 MF26, m23067-09, JF3224, HER1085 et ATCC 33658 (Annexe A).

Tableau 4 : Analyses bio-informatiques par l'outil BLASTn des dissemblances chromosomiques observées entre les souches **A449** et **01-B526** dans les séquences génomiques assemblées en contig de dix autres isolats d'*A. salmonicida* sous-espèce *salmonicida*.

Isolats	Nombre de répétition dans les séquences en tandem des dissemblances indiquées par les numéros référant au Tableau 3					Présence ou absence des courtes insertions dans les gènes référant à l'annotation de la souche A449 (NC_009348)					Origine géographique
	12	13	21	22	28	ASA_0621	ASA_2192	tnaA	ASA_3022	amt	
01-B526	6,9	9,3	4,8	6,8	6,6	Non	Non	Non	Non	Non	Québec
01-B522	6,9	9,3	4,8	6,8	6,6	Non	Oui	Non	Non	Non	Québec
M15879-11	6,9	8,3	4,8	6,8	6,6	Non	Non	Non	Non	Non	Québec
09-0167	6,9	8,3	3,8	4,8	5,6	Oui	Oui	Oui	Non	Non	Québec
2009-144K3	5,9	8,3	3,8	4,8	6,6	Oui	Oui	Oui	Non	Non	Nouveau-Brunswick
2004-05 MF26	8,9	7,3	3,8	4,8	8,6	Oui	Oui	Oui	Non	Non	Nouveau-Brunswick
m23067-09	6,9	8,3	3,8	4,8	5,6	Oui	Oui	Oui	Non	Non	Québec
JF3224	5,9	6,3	3,8	7,8	2,6	Oui	Oui	Oui	Oui	Oui	Suisse
HER1085	9,9	6,3	3,8	7,8	2,6	Oui	Oui	Oui	Oui	Oui	Norvège
ATCC 33568	11,9	6,3	3,8	5,8	2,6	Oui	Oui	Oui	Oui	Oui	Pays de Galles
A449	5,9	6,3	3,8	4,8	2,6	Oui	Oui	Oui	Oui	Oui	France
JF2267	11,9	6,3	3,8	5,8	2,6	Oui	Oui	Oui	Oui	Oui	Suisse

5.3. Caractérisation des *AsaGEI* chez *A. salmonicida*

5.3.1. Caractéristiques générales d'un nouveau GEI dans la souche 01-B526.

L'élément génétique de 51 kb découvert dans la souche d'*A. salmonicida* sous-espèce *salmonicida* 01-B526 possède un contenu en G+C de 57,7%, alors que le chromosome a un contenu en G+C de 58,5%. Il est constitué de 63 ORF encodant des protéines qui ont été prédits par annotation automatique, suivie d'une correction manuelle. Sur les 63 ORF prédits, 19 codent pour des protéines d'origine phagique (30%), dont 12 ont des fonctions connues. L'*orf25* est le seul ORF qui n'est pas d'origine phagique qui code pour une protéine possédant une fonction connue (exo-désoxyribonucléase VIII) (Figure 6A). La liste complète des ORF et de leurs fonctions associées a été placée à l'annexe C.

L'élément génétique de 51 kb est encadré par deux séquences de 19 pb répétées directes (CCATGCCGGGGTAGTAGAT), alors que cette même séquence est seulement présente une fois dans le génome d'A449, qui ne possède pas l'élément. L'absence de l'élément dans le chromosome de la souche de référence A449 suggère qu'il a été acquis par HGT chez 01-B526, un événement impliquant généralement des MGE (Frost *et al.*, 2005). Aucun gène codant pour des protéines structurales de phage n'a été prédit dans l'élément génétique. Sans étude fonctionnelle (activation des phages, microscopie électronique et séquençage), nous ne pouvons affirmer avec certitude qu'il s'agit oui ou non d'un prophage. L'insertion de 51 kb chez 01-B526 possède trois caractéristiques des GEI : il n'est pas compris dans le génome de base d'*A. salmonicida* sous-espèce *salmonicida*, car absent chez certaines souches, telles qu'A449 et JF2267 (résultat non montré), l'élément code pour plusieurs protéines de mobilité d'origines phagiques (intégrase (*orf1*), excisionnase (*orf2*) et deux recombinaisons (*orf14* et *orf24*)) et il est flanqué par des séquences répétées directes (19 pb), indiquant qu'il semble avoir été intégré par transfert horizontal (Darmon et Leach, 2014; Dobrindt *et al.*, 2004). Sur la base de ces éléments, l'insertion de 51 kb a été nommée *Aeromonas salmonicida Genomic Island 1a* (*AsaGEI1a*). À notre connaissance, c'est le premier exemple dans la littérature d'un GEI découvert dans une souche *A. salmonicida* sous-espèce *salmonicida*.

Pour déterminer l'origine potentielle d'*AsaGEI1a*, sa séquence nucléotidique a été comparée par l'outil BLASTn sur le serveur web NCBI contre les bases de données *whole genome shotgun contigs* (*wgs*) contenant des ébauches de séquences génomiques et la base de données *nucleotide collection* (*nr/nt*) contenant des génomes assemblés et

annotés. Suite à la comparaison d'*AsaGE11a* contre les bases de données *nr/nt* et *wgs*, aucune et peu de séquences avec un haut pourcentage d'identité ont été identifiées, respectivement. Au moment de ces analyses, trois souches bactériennes déposées sur la base de données *wgs* possédaient une séquence contiguë présentant des pourcentages élevés d'identité contre *AsaGE11a*, et en ordre décroissant : *A. hydrophila* RB-AH, *A. jandaei* Riv2 et *A. salmonicida* sous-espèce *achromogenes* AS03 (Figure 6B).

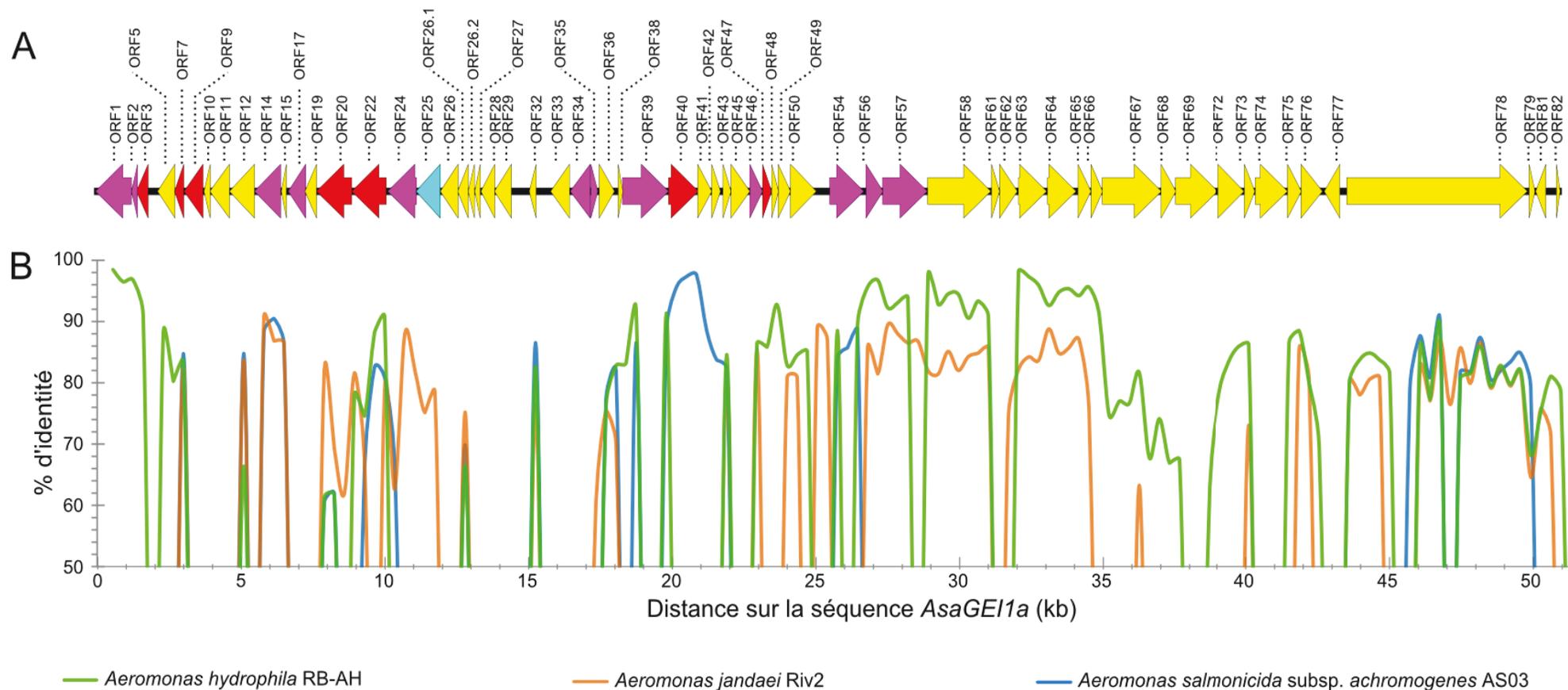


Figure 6 : Caractéristiques génétiques d'*AsaGE1a*. A) Représentation schématique des cadres de lecture ouverts (ORF) d'*AsaGE1a*. Chaque flèche représente un ORF. Les fonctions pour chaque ORF sont indiquées à l'Annexe C. Les flèches de couleur mauve, rouge, bleu et jaune représentent des ORF codants pour des protéines phagiques possédant des fonctions connues, des protéines hypothétiques d'origines phagiques, une protéine avec une fonction connue et des protéines hypothétiques, respectivement. B) Représentation graphique de l'identité de séquence d'*AsaGE1a* avec trois régions génomiques de trois souches appartenant à d'autres espèces ou sous-espèce bactériennes. Le nom des trois souches bactériennes sélectionnées est inscrit en bas du graphique.

L'élément génétique découvert dans la bactérie *A. hydrophila* RB-AH (entre 957 062 pb et 1 009 948 pb sur le Contig_1) a une taille de 52,8 kb et une identité de séquence de 92% sur 67% de sa longueur avec *AsaGEI1a*. Un second élément génétique avec 85% d'identité sur 58% de sa longueur contre *AsaGEI1a* a été découvert chez *A. jandaei* Riv2. Ce dernier élément est localisé sur le contig 3 entre 320 688 pb et 373,431 pb. Les éléments chez RB-AH et Riv2 ont beaucoup de régions conservées entre eux qui sont aussi présentes chez *AsaGEI1a* et dont leurs identités de séquence sont de plus de 65% (Figure 7). L'intégrase (*orf1*) et la recombinase (*orf2*) sont conservées entre l'élément chez RB-AH et *AsaGEI1a*. L'élément chez *A. hydrophila* RB-AH est flanqué par la même séquence de 19 pb répétée directe qui a été prédite comme étant le site d'insertion d'*AsaGEI1a*, conséquemment il est présent au même site chromosomique

Le GEI de la souche 01-B526 démontre aussi 90% d'identité de séquence sur 35% de sa longueur avec une région génomique d'*A. salmonicida* sous-espèce *achromogenes* AS03 (entre 138,843 pb et 195,462 pb sur le Contig38) (Han *et al.*, 2013). Trois gènes codants pour des protéines structurales de phage, incluant une protéine *portal*, une protéine structurale de virion et une protéine partielle des fibres de la queue (*tail fiber*), ont été prédits par des analyses bio-informatiques dans la région génomique du Contig38 de la souche AS03, suggérant que ces gènes ont pu être importés en provenance d'un prophage. La région de l'élément chez AS03 codant pour les protéines structurales de phage n'est pas conservée dans *AsaGEI1a* (Figure 8).

Les résultats indiquent que l'élément génétique de 51 kb découvert chez la souche 01-B526 est un GEI et que des éléments avec des pourcentages d'identités élevés ont été observés chez d'autres bactéries du genre *Aeromonas*. Par la suite, la structure des différents variants d'*AsaGEI* présent chez d'autres souches d'*A. salmonicida* sous-espèce *salmonicida* ont été investiguées.

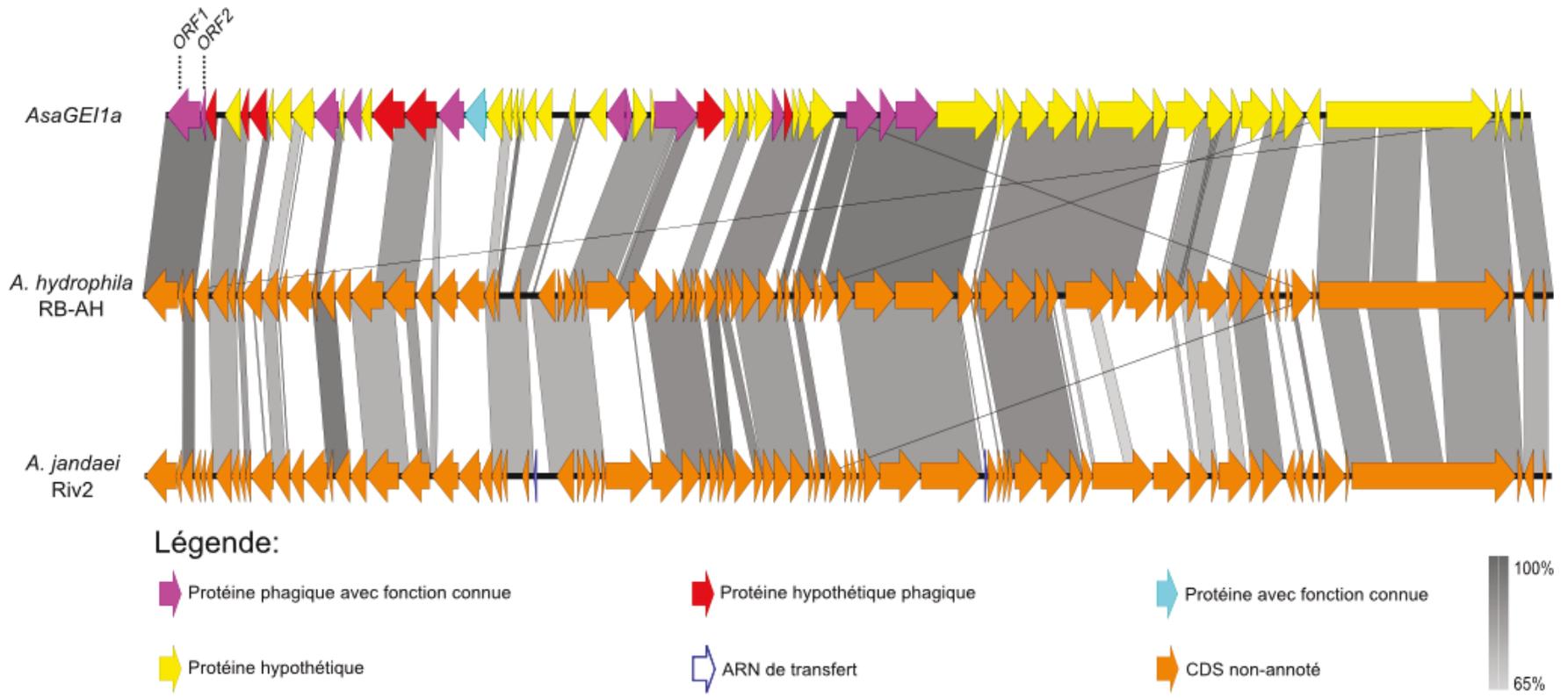


Figure 7 : Comparaison colinéaire d'*AsaGE11a* avec les éléments génétiques des souches d'*A. hydrophila* RB-AH et d'*A. jandaei* Riv2. Les boîtes grises représentent les régions qui partagent une identité de séquence. Le pourcentage d'identité entre les séquences est indiqué par la nuance de gris. Chaque flèche représente un ORF.

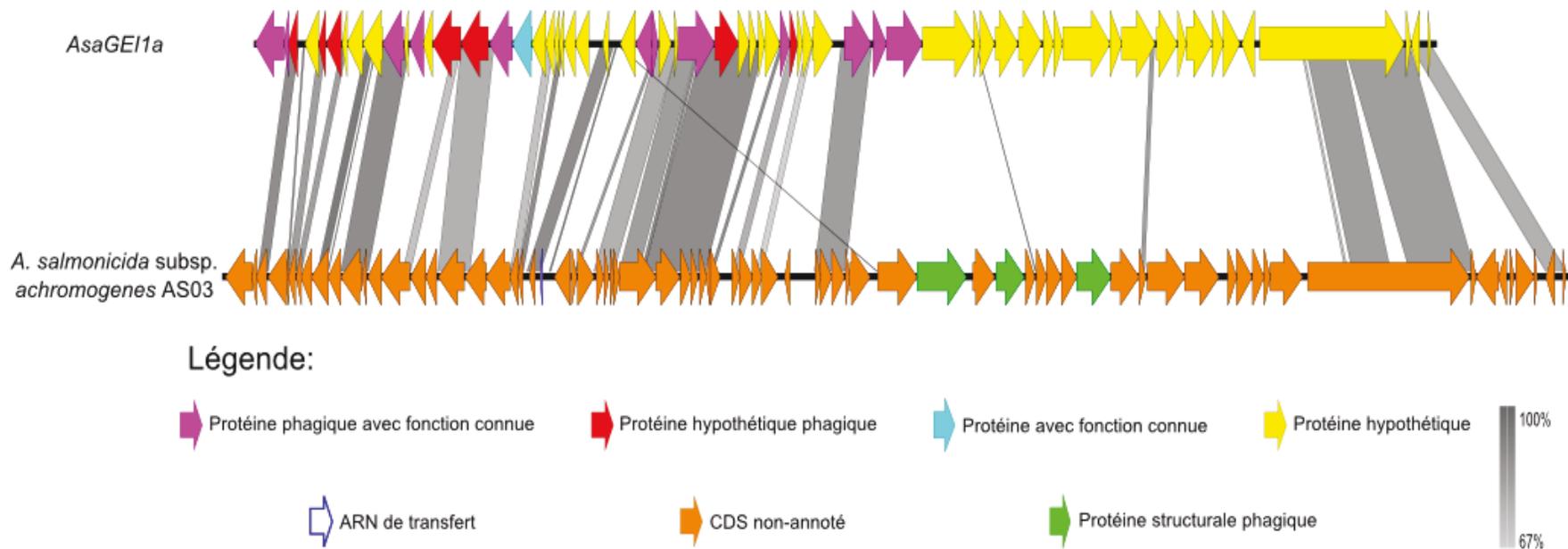


Figure 8 : Comparaison co-linéaire d'*AsaGE11a* avec l'élément génétique de la souche d'*A. salmonicida* sous-espèce *achromogenes* AS03. Les boîtes grises représentent les régions qui partagent une identité de séquence. Le pourcentage d'identité entre les séquences est indiqué par la nuance de gris. Chaque flèche représente un ORF.

5.3.2. Structure d'*AsaGEI1b* et *AsaGEI2a* et leur relation avec *AsaGEI1a*

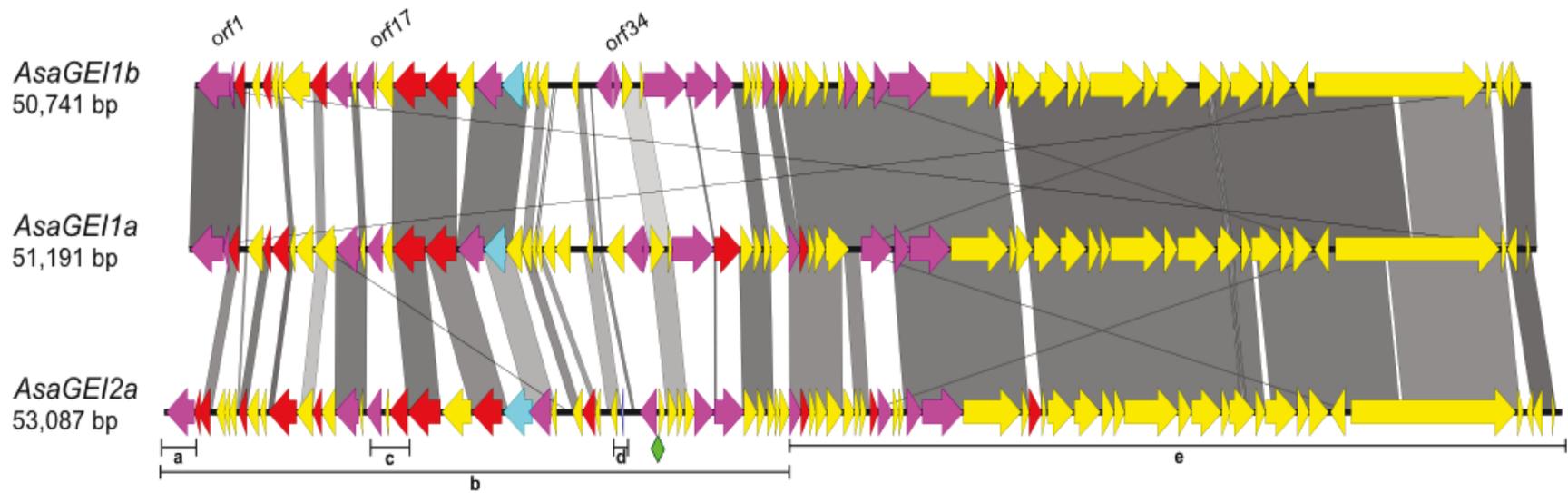
Dans le but d'analyser la structure des autres variants d'*AsaGEI*, les génomes de sept souches additionnelles minutieusement choisies ont été séquencés. Dans les sept souches séquencées, deux démontraient le profil d'amplification indiqué à la Figure 2A (01-B522 et M15879-11), deux avaient le profil d'amplification de la Figure 2B (HER1085 et ATCC 33658) et trois démontraient le profil d'amplification de la Figure 2C (09-0167, m23067-09 et 2009-144K3). Suite au séquençage et à l'assemblage en contig, les séquences correspondantes aux différents GEI ont été extraites et seulement les éléments des souches HER1085 et 09-0167 ont été annotés (un représentant pour chaque groupe). Les GEI dans les souches 01-B522 et M15879-11 se révélaient être aussi des *AsaGEI1a*, ce qui totalise à trois séquences avec celui de la souche 01-B526. Les analyses de séquences démontrent que les trois *AsaGEI1a* ont un pourcentage d'identité de 99% sur 100% de leur longueur (*Query cover*) (annexe D). Le variant d'*AsaGEI1a* observé dans les souches HER1085 et ATCC 33658 a été appelé *AsaGEI1b*, car il partage le même site d'insertion qu'*AsaGEI1a* (voir section 5.3.3). Les analyses de séquences démontrent que les deux *AsaGEI1b* ont un pourcentage d'identité de 99% sur 100% de leur longueur (annexe D). Le second variant d'*AsaGEI1a* observé dans les souches 09-0167, m23067-09 et 2009-144K3 a été appelé *AsaGEI2a*, car il présente un autre site d'insertion que les deux *AsaGEI1* (voir section 5.3.3). Les analyses de séquences démontrent que les trois *AsaGEI2a* ont un pourcentage d'identité de 99% sur 100% de leurs longueur (annexe D).

La représentation schématique de l'alignement des séquences des éléments *AsaGEI1a*, *AsaGEI1b* et *AsaGEI2a*, provenant des génomes des souches 01-B526, HER1085 et 09-0167, respectivement, sont illustrées à la Figure 9. Les listes des ORF annotés pour *AsaGEI1b* et *AsaGEI2a* sont présentées à l'annexe E et l'annexe F, respectivement. La comparaison génomique démontre que les gènes qui codent pour une intégrase (*orf1*) et une excisionase (*orf2*) sont conservés entre *AsaGEI1a* et *AsaGEI1b*, mais que ces deux gènes sont différents chez *AsaGEI2a* (Figure 9-a). L'alignement des trois *AsaGEI* peut être divisé en deux parties. La première partie (Figure 9-b) est ponctuée de plusieurs régions variables (voir l'exemple à la Figure 9-c) encadrées par des régions conservées (voir l'exemple à la Figure 9-d), alors que la deuxième partie correspond à une section qui est conservée entre les trois *AsaGEI* (Figure 9-e). Une autre caractéristique intéressante des éléments *AsaGEI* est qu'on peut observer un changement brusque du

sens de la transcription entre l'*orf34* et l'*orf35* représenté sur la Figure 9 par un losange vert.

Les analyses des régions variables de la section divergente des *AsaGEI* (Figure 9-b) démontrent que les gènes codent pour des protéines avec des fonctions semblables dans certains segments, mais que les séquences géniques sont différentes. Concrètement, les trois *orf17* sont différents dans chacun des *AsaGEI*, mais ils codent tous pour une méthylase phagique. La même observation est vraie pour les *orf34*, lesquels sont différents pour chacun, mais ils codent tout pour une protéine se liant à l'ADN (*DNA-binding protein*). L'annotation d'*AsaGEI2a* révèle aussi la présence d'un ARNt^{Pro} (indiqué en bleu dans la Figure 9). Un alignement de séquence colinéaire entre *AsaGEI2a* et les éléments génétiques découverts chez *A. jandaei* Riv2 et *A. salmonicida* sous-espèce *achromogenes* AS03 démontre que l'ARNt^{Pro} est présent à la même position dans les trois éléments (résultat non montré). L'anticodon de l'ARNt^{Pro} est «TGG».

Globalement, les GEI découverts chez différentes souches d'*A. salmonicida* sous-espèce *salmonicida* ont tous une partie contenant plusieurs variations et une partie conservée. La partie conservée contient majoritairement des gènes codants pour les mêmes protéines hypothétiques. La fonction des différents *AsaGEI* n'a pas pu être prédite. Suite à l'investigation de la structure des GEI, le site d'insertion pour chaque *AsaGEI* a été déterminé à partir des données de séquençage.



Légende:

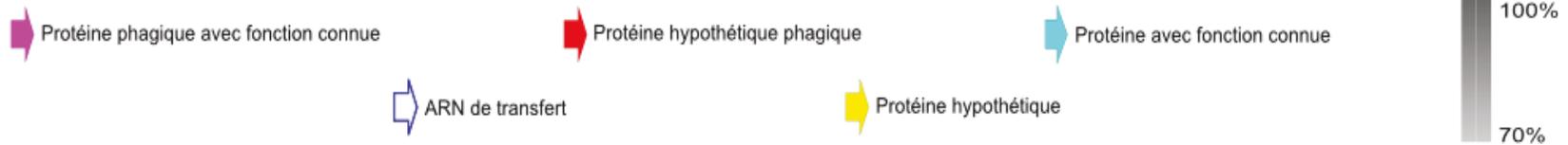


Figure 9 : Comparaison colinéaire des trois *AsaGEI*. Les boîtes grises représentent les régions qui partagent une identité de séquence. Le pourcentage d'identité entre les séquences est indiqué par la nuance de gris. Chaque flèche pleine représente un ORF. Les fonctions pour chaque ORF d'*AsaGEI1a*, d'*AsaGEI1b* et d'*AsaGEI2a* sont indiquées aux Annexes C, E et F, respectivement. Le losange vert indique le site de changement du sens de la transcription. La région mise en évidence par la ligne «a» indique les gènes codant pour une intégrase (*orf1*) et une excisionase (*orf2*) qui sont les mêmes entre *AsaGEI1a* et *AsaGEI1b*, mais dont la région contient deux gènes différents chez *AsaGEI2a*. La partie variable des *AsaGEI* est indiquée par la ligne «b». Les régions variables bordées par des régions conservées sont indiquées par la ligne «c» et la ligne «d», respectivement. La partie conservée entre les trois *AsaGEI* est indiquée par la ligne «e».

5.3.3. Sites d'insertion des *AsaGEI*

Les positions chromosomiques des sites d'insertion des trois *AsaGEI* sont représentées à la Figure 10. Les analyses des séquences génomiques indiquent que les éléments *AsaGEI1a* et *AsaGEI1b* sont localisés au même endroit dans le chromosome (Figure 10A), ce qui confirme les résultats de criblage PCR (Figure 2A et 2B). Ils sont intégrés au début d'un gène codant pour une protéine hypothétique conservée (*ASA_2666* chez A449). Dans la souche 09-0167, *AsaGEI2a* s'est inséré en aval d'un ARNt^{Leu}, à six gènes du prophage 1 (Figure 10B) (Reith *et al.*, 2008). Le site d'insertion d'*AsaGEI2a* comporte 77 nt, dont 59 nt situés à la fin de l'ARNt^{Leu} et 18 nt est en aval de ce même ARNt (**GGGATTTAAAATCCCTCGACGTTTCGCGTCGTGCCGGTTCGATTCCGGCCTCG GGCACCA****TTAAAATCAAAGACTTAC**). L'intégrité de l'ARNt^{Leu} a été conservée suite à l'intégration de l'élément génétique. Les séquences identiques répétées directes de 77 nt qui encadrent *AsaGEI2a* indiquent une duplication du site d'insertion suite à l'intégration de l'élément (Fogg *et al.*, 2014).

Les séquençages des génomes de deux des trois souches possédant *AsaGEI2a* (m23067-09 et 2009-144K3) démontrent la présence d'un nouveau prophage intégré entre l'ARNt^{Leu} et *AsaGEI2a*, nommé prophage 3 (Figure 10C). Suivant la découverte du prophage 3, des paires d'amorces ont été conçues pour déterminer la prévalence du nouveau prophage dans la collection de 139 isolats d'*A. salmonicida* sous-espèce *salmonicida* utilisée dans cette étude. Les résultats du criblage sont inscrits en dessous des boîtes B et C à la Figure 10. Sur 82 isolats possédant un *AsaGEI2a*, 77 ont conjointement le prophage 3. Sur les cinq isolats positifs pour *AsaGEI2a*, mais déficitaires pour le prophage 3, quatre ont été isolés au Québec et un provient d'Europe. Trois de ces quatre isolats québécois proviennent de la même région piscicole (résultat non montré). Le prophage 3 a aussi été identifié dans quatre souches provenant des Grands Lacs ne portant pas *AsaGEI2a*, dont trois semblent posséder un *AsaGEI1a* alors que la dernière souche semble avoir un *AsaGEI1b* (Annexe A). Aucune des souches européennes ne possède le prophage 3, il est majoritairement observé dans des souches portant un *AsaGEI2a* isolées dans l'Est de l'Amérique du Nord.

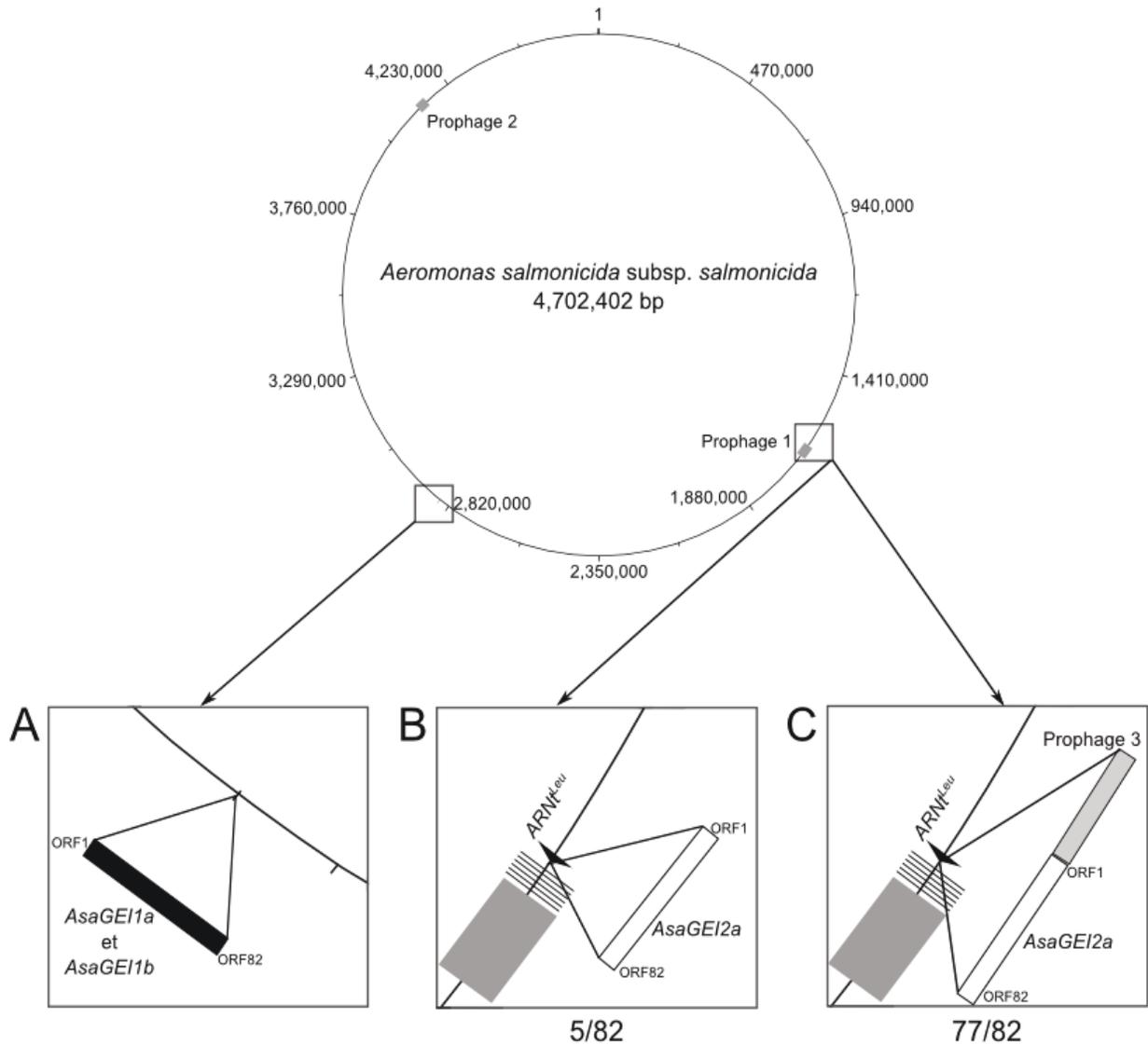


Figure 10 : Représentation schématique des sites d'insertion des trois *AsaGEI* par rapport au génome de la souche d'*A. salmonicida* sous-espèce *salmonicida* A449. Les rectangles noir et blanc représentent *AsaGEI1(a,b)* et *AsaGEI2a*, respectivement. La position relative de l'ORF1 et l'ORF82 sont indiqués pour chaque GEI. A) Lieu d'insertion d'*AsaGEI1a* et *AsaGEI1b*. Site d'insertion d'*AsaGEI2a* qui est associé C) ou non B) avec le prophage 3. Les proportions des souches possédant *AsaGEI2a* et le prophage 3 ou non sont indiquées en bas de chaque carrée. Dans les illustrations en B et C, les six lignes noires représentent les gènes entre le prophage 1 (en gris foncé) et l'ARNt^{Leu} (flèche noire).

5.3.4. Prévalence des différents variants d'*AsaGEI* par criblage PCR

Une approche de génotypage par PCR a été effectuée pour déterminer la prévalence des différents *AsaGEI* dans une collection de 139 isolats d'*A. salmonicida* sous-espèce *salmonicida*. Une grande proportion des isolats provenant du Québec avaient déjà été testés par Mélanie Trudel, alors que peu d'information était disponible sur les isolats européens (souches JF, voir Annexe A). Une partie du criblage a été effectué en utilisant neuf des quinze paires d'amorces discutées dans l'introduction, sept amplifient spécifiquement des ORF (*orf11*, *orf19*, *orf33*, *orf54*, *orf57*, *orf67* et *orf78*) et deux amplifient les sites d'insertion d'*AsaGEI1a* (*ins1* et *ins2*) (Annexe B). Les résultats de génotypage sont inscrits à l'Annexe A à la colonne «type de GEI». La compilation des différents criblages indique que 29 souches ont un signal positif pour les neuf cibles PCR, suggérant qu'elles possèdent toutes *AsaGEI1a* (Figure 11A). 15 des isolats ont le même profil d'amplification négatif que la souche A449 (donc 16 au total), indiquant qu'elles ne possèdent pas d'*AsaGEI* (Figure 11B). Le profil d'amplification pour la souche HER1085 porteuse d'*AsaGEI1b* est indiqué à la Figure 11C. Sur les 139 isolats (incluant la souche HER1085), 12 démontraient le même profil d'amplification qu'*AsaGEI1b* (Figure 11C). 81 isolats démontraient le même profil d'amplification que la souche 09-0167 (82 avec 09-0167), indiquant qu'elles semblent porteuses d'*AsaGEI2a* (Figure 11D).

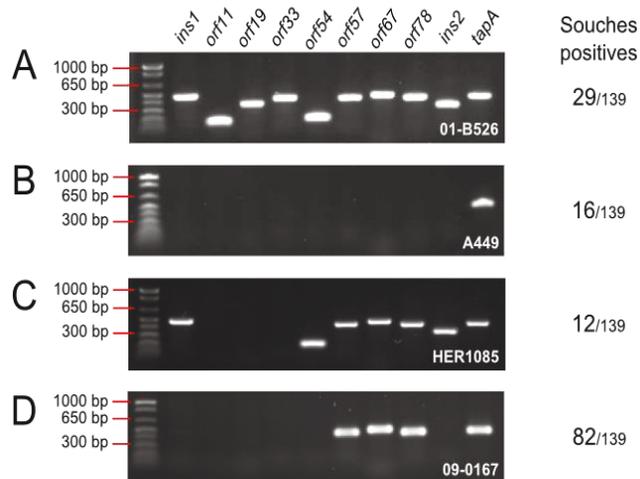


Figure 11 : Génotypage de 139 isolats d'*A. salmonicida* sous-espèce *salmonicida* pour la présence de neuf séquences trouvées chez *AsaGEI1a*. Le gène chromosomique *tapA* été utilisé comme contrôle positif. Analyse d'électrophorèse d'isolats portant *AsaGEI1a* (A), aucun GEI (B), *AsaGEI1b* (C), et *AsaGEI2a* (D). Le nom de chaque isolat qui a été utilisé comme exemple est indiqué en bas de chaque image. Les résultats de criblage pour chaque profil d'amplification sont indiqués à la droite.

5.3.5. Corrélation entre le génotype d'AsaGEI et l'origine géographique des isolats

Les résultats de génotypage des GEI ont été regroupés par rapport à l'origine géographique des isolats d'*A. salmonicida* sous-espèce *salmonicida* (Tableau 5). La provenance de chaque isolat est indiquée à l'Annexe A. Quatre tendances peuvent être observées : (1) 100% (29/29) des isolats qui possèdent *AsaGEI1a* ont été isolés dans la région hydrographique des Grands Lacs et du Saint-Laurent, (2) 87% (71/82) et 12% (10/82) des isolats porteurs d'un *AsaGEI2a* proviennent soit de la région hydrographique des Grands Lacs et du Saint-Laurent ou soit des provinces maritimes de l'Atlantique, respectivement, ce qui signifie que 99% des isolats portant un *AsaGEI2a* proviennent de l'Est de l'Amérique du Nord, (3) 100% (10/10) des isolats provenant des provinces maritimes de l'Atlantique portent un *AsaGEI2a*, et (4) 83% (10/12) des isolats qui portent un *AsaGEI1b* et 88% (14/16) des isolats qui n'ont aucun *AsaGEI* proviennent d'Europe. Cinq isolats ne sont pas inclus dans les quatre tendances décrites ci-haut.

Tableau 5 : Origine géographique des 139 isolats d'*A. salmonicida* sous-espèce *salmonicida* par rapport à leur génotype d'AsaGEI.

	<i>AsaGEI</i>				Total
	1a	2a	1b	sans	
Système hydrographique des Grands Lacs et du Saint-Laurent	29	71	2	2	104
Provinces maritimes de l'Atlantique	0	10	0	0	10
Europe ^a	0	1	10	14	25
Total	29	82	12	16	139

a: Les isolats proviennent de la Suisse (n=12), Norvège (n=5), France (n=4), Royaume-Uni (n=2), Danemark (n=1), et une souche d'un pays inconnu européen (n=1).

5.3.6. Analyse du GEI de la souche JF3224

L'unique isolat européen possédant un *AsaGEI2a*, appelé JF3224, a été séquencé pour confirmer qu'il s'agit bel et bien de ce génotype d'*AsaGEI*. Des résultats de PCR non montrés indiquaient que l'isolat JF3224 a une amplification négative pour l'*orf21*, alors que cet ORF est présent et amplifié chez tous les autres *AsaGEI2a*, ce qui nous indiquait qu'il pouvait s'agir d'un nouveau variant d'*AsaGEI*.

Les résultats de séquençage de l'isolat JF3224 ont été assemblés en contig, puis une PCR a été effectuée pour confirmer l'assemblage du GEI, car il était en deux contigs (résultat non montré). L'élément chez JF3224 a le même site d'insertion et la même orientation d'intégration qu'*AsaGEI2a* (entre le prophage 1 et l'ARNt^{Leu}), ce qui correspondrait à la Figure 10B. Le prophage 3 n'a pas été observé dans le génome de JF3224. Un alignement BLASTn démontre que le GEI de la souche JF3224 a 94% d'identité sur 76% de sa longueur avec *AsaGEI2a*, donc il s'agit bel et bien d'un nouvel *AsaGEI*. Les ORF du GEI ont été prédits et annotés en utilisant les mêmes méthodes bio-informatiques que les trois autres *AsaGEI*. La liste complète des ORF et de leurs fonctions respectives est placée à l'Annexe G.

Le nouvel *AsaGEI* de JF3224 a été aligné avec les trois autres éléments *AsaGEI* auparavant identifiés et le résultat est représenté à la Figure 12. On peut observer que le nouvel *AsaGEI* est divergent de l'*AsaGEI2a*, mais que les deux possèdent la même intégrase (*orf1*) (Figure 12-a) et ARNt^{Pro} (anticodon TGG). Puisque l'élément de JF3224 est différent des trois autres *AsaGEI*, mais qu'il possède la même intégrase et le même site d'insertion que l'élément *AsaGEI2a*, il a été nommé *AsaGEI2b*. Le nouveau GEI peut être aussi séparé en deux parties, soit une partie ponctuée de régions variables encadrées par des séquences conservées (Figure 12-b) et une partie conservée avec les trois autres *AsaGEI* (Figure 12-c). Il y a aussi un changement brusque du sens de la transcription dans *AsaGEI2b* entre l'*orf34* et l'*orf35* (indiqué par un losange vert sur la représentation schématique d'*AsaGEI2a*). En bref, la structure d'*AsaGEI2b* est semblable aux trois autres *AsaGEI*.

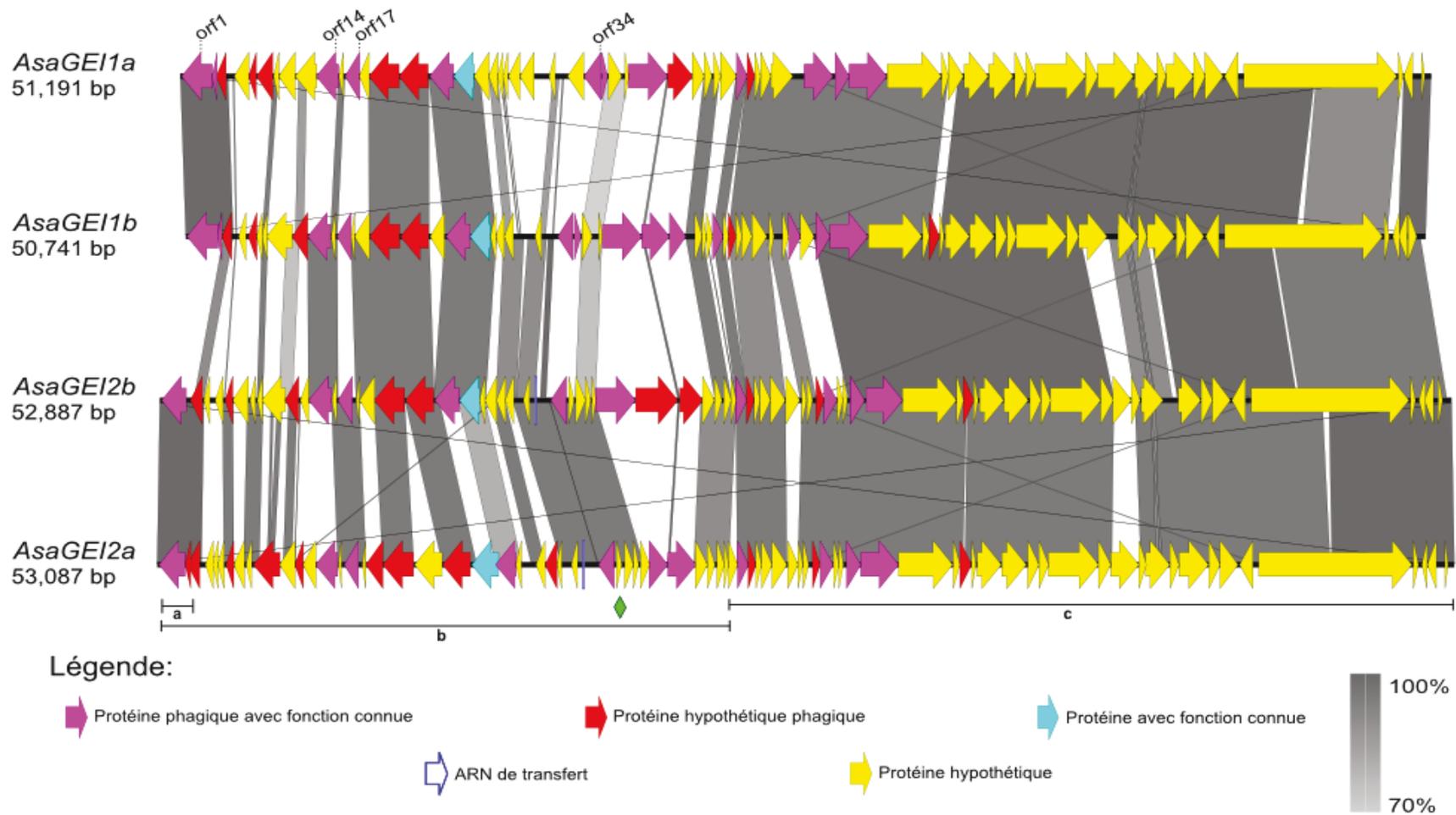


Figure 12 : Comparaison colinéaire des quatre *AsaGEI*. Les boîtes grises représentent les régions qui partagent une identité de séquence. Le pourcentage d'identité entre les séquences est indiqué par la nuance de gris. Chaque flèche pleine représente un ORF. Les fonctions pour chaque ORF d'*AsaGEI1a*, d'*AsaGEI1b*, d'*AsaGEI2a* et d'*AsaGEI2b* sont indiquées aux Annexes C, E, F et G, respectivement. Le losange vert indique le site de changement du sens de la transcription. La région mise en évidence par la ligne «a» indique le gène codant une intégrase (*orf1*) qui est le même entre *AsaGEI2a* et *AsaGEI2b*, mais dont la région contient un nouveau gène d'intégrase chez les *AsaGEI1(a;b)*. La partie variable et la partie conservée des *AsaGEI* sont indiquées par les lignes «b» et «c», respectivement.

5.3.7. Annotation et comparaison génomique du prophage 3

L'annotation révèle que l'élément génétique de 35,3 kb observé entre l'ARNt^{Leu} et *AsaGEI2a* (Figure 10C) possède plusieurs gènes codants pour des protéines structurales phagiques, ce qui indique qu'il s'agit potentiellement d'un prophage. Il a été nommé prophage 3 pour respecter l'appellation des deux autres prophages d'*Aeromonas salmonicida* sous-espèce *salmonicida* déjà observés dans A449 (Reith *et al.*, 2008). La liste complète des ORF prédits dans le prophage 3 et leurs fonctions respectives est placée à l'Annexe H. Un alignement du prophage 3 avec les deux autres prophages connus chez *Aeromonas salmonicida* sous-espèce *salmonicida* (prophage 1 et 2) démontre que le nouveau prophage n'a aucune identité avec les deux autres (résultat non montré). Un alignement BLASTn du prophage 3 contre les séquences déposées sur la base de données publique GenBank *nt/nr* révèle qu'il existe des éléments partageant une longueur et une identité de plus de 46% et 77%, respectivement, découverts chez deux autres bactéries du genre *Aeromonas*. Pour cette étude de génomique comparative, les deux séquences avec des pourcentages de longueur et d'identité élevés ont été extraites pour les analyses. Les séquences proviennent des génomes d'*A. veronii* B565 (Li *et al.*, 2011) et d'*A. hydrophila* 4AK4 (Gao *et al.*, 2013). L'alignement du prophage 3 avec les deux autres éléments est représenté à la Figure 13. Plusieurs régions sont conservées entre le prophage 3 et les deux autres éléments. Les deux autres éléments génétiques semblent être aussi des prophages, car plusieurs ORF encodant pour des protéines de structures phagiques sont conservés entre les trois séquences.

Un fait intéressant est que deux régions du prophage 3 indiquées par les lignes rouges sur la Figure 13 ont un pourcentage d'identité de plus de 60% avec des régions génomiques observées chez d'autres souches d'ordres bactériens variés, soit, par exemple, des bactéries de l'ordre des *Aeromonadales*, des *Enterobacteriales* et des *Pseudomonadales*. Les trois gènes contenus dans ces deux régions codent pour une protéine de structure phagique (*phage tail sheath*) et pour deux protéines responsables de l'encapsidation de l'ADN virale (*terminase* et une protéine *phage portal vertex-like*) (Black et Rao, 2012; Black, 1995). De façon générale, les protéines de structures et les protéines responsables de l'encapsidation de l'ADN sont caractérisées pour être conservées chez les phages (Martel et Moineau, 2014).

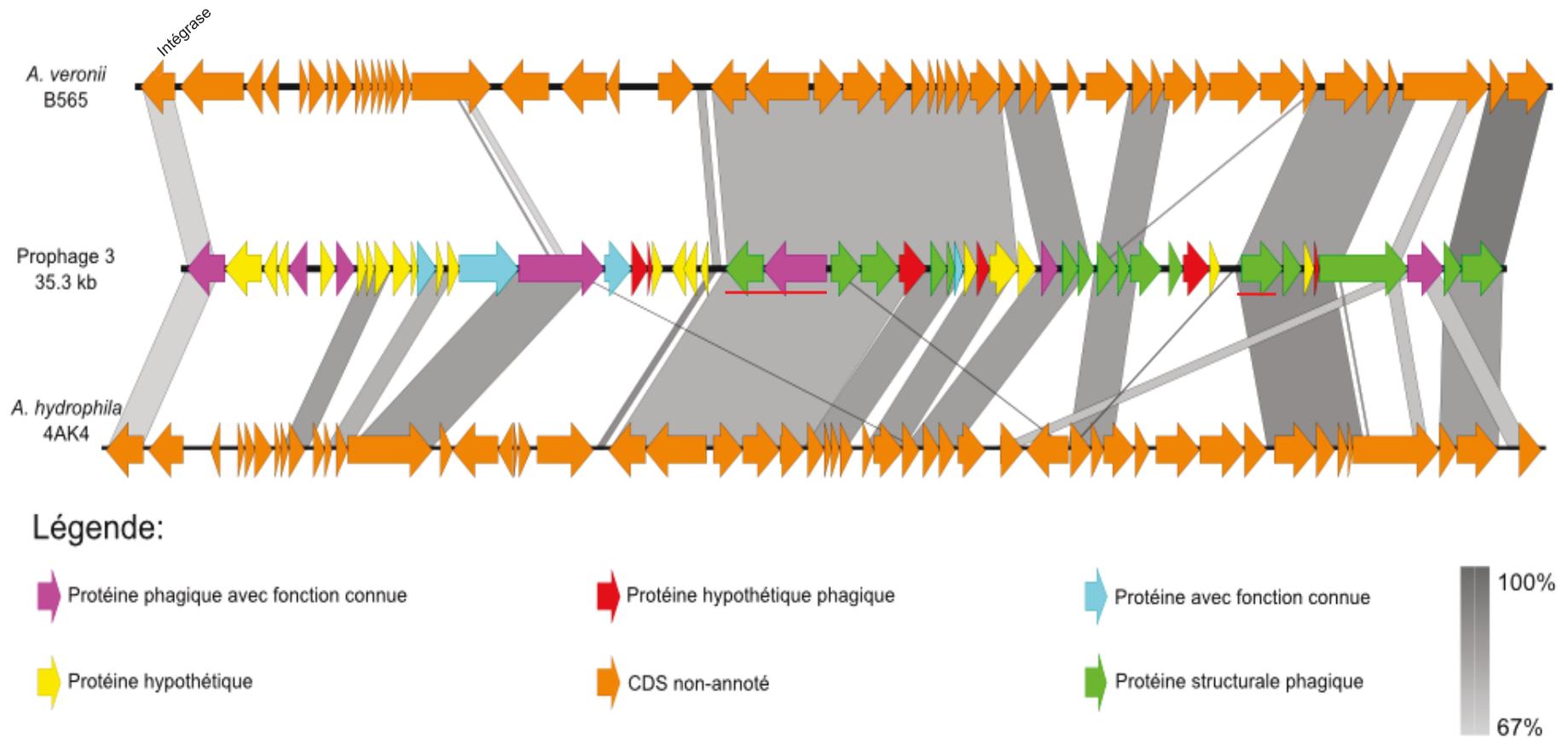


Figure 13 : Comparaison colinéaire du prophage 3 avec les éléments génétiques découvert chez les souches d'*A. veronii* B565 et d'*A. hydrophila* 4AK4. Les boîtes grises représentent les régions qui partagent une identité de séquence. Le pourcentage d'identité entre les séquences est indiqué par la nuance de gris. Chaque flèche pleine représente un ORF. Les fonctions pour chaque ORF du prophage 3 sont indiquées à l'Annexe H. Les lignes rouges représentent les deux régions conservées chez d'autres souches bactériennes de l'ordre des *Aeromonadales*, des *Enterobacteriales* et des *Pseudomonadales*

6. Discussion

Face à la croissance constante de la population mondiale et du déclin des stocks de poisson à l'état naturel, l'aquaculture est la solution pour répondre à la demande grandissante de poisson et ainsi assurer l'approvisionnement pour les prochaines générations. Au Québec, les aquaculteurs élèvent principalement des poissons de la famille des Salmonidés, mais plus précisément, l'omble de fontaine (*Salvelinus fontinalis*). Les techniques d'élevage intensives de l'omble de fontaine augmentent sa susceptibilité aux infections par des agents pathogènes, tel *A. salmonicida* sous-espèce *salmonicida*, responsable de la furunculose dite typique. En vue de combattre cette infection bactérienne, une meilleure compréhension de cet agent pathogène est nécessaire. Dans ces travaux de recherche, plusieurs éventualités ont été envisagées dans le but de mettre au point des outils pour mener des études épidémiologiques sur les isolats responsables d'éclosion de furunculose dans les aquacultures. Les études épidémiologiques antérieures sur *A. salmonicida* sous-espèce *salmonicida* démontraient que les isolats semblaient génétiquement homogènes (Belland et Trust, 1988; García et al., 2000; Hänninen et al., 1995; McCormick et al., 1990; Miyata et al., 1995; O'hici et al., 2000). Les travaux de recherche présentés dans ce document mettent en évidence les différences génétiques observées entre les isolats d'*A. salmonicida* subsp. *salmonicida* et démontrent une spécificité des éléments *AsaGEI* d'après l'origine géographique des isolats (Emond-Rheault et al., 2015).

6.1. RAPD

Dans le but de mener une étude épidémiologique sur une collection de 139 isolats d'*A. salmonicida* sous-espèce *salmonicida*, la technique RAPD a été optimisée en testant 18 amorces de 9 nt (Tableau 2). Les séquences des amorces ont été choisies d'après le nombre de sites théoriques d'appariement dans les chromosomes des souches A449 et 01-B526. Les extractions d'ADN pour les expériences RAPD ont été effectuées avec une trousse commerciale pour deux raisons. Tout d'abord, l'utilisation d'une trousse commerciale se démontre plus rapide qu'une extraction d'ADN traditionnelle (phénol/chloroforme), ce qui aurait accéléré l'étude si elle avait été appliquée aux 139 isolats de la collection d'*A. salmonicida* sous-espèce *salmonicida* du laboratoire. Ensuite, la trousse a été utilisée pour éviter une perte de signal pour certaines bandes, un

phénomène observable lors de l'emploi pour la RAPD de lysats d'ADN créés par une technique de lyse par ébullition généralement utilisée pour les amplifications PCR traditionnelles (Madico et Akopyants, 1995).

Les résultats de cette étude démontrent que peu de polymorphismes ont été observés par la technique RAPD entre les quatre isolats testés (Figure 3). Seulement deux amorces sur les dix-huit utilisées ont amplifié des patrons de bandes présentant des différences (Figure 3A et C). Un total de trois bandes variaient avec les deux amorces, donc les quatre souches peuvent être considérées comme génétiquement homogènes telles que proposées par une étude ayant établi ce type de paramètres (Tenover *et al.*, 1995).

La RAPD peut être caractérisée comme étant une technique à faible résolution, car elle ne permet de déceler qu'une quantité limitée de polymorphismes présents entre les différents génomes. Certaines des études épidémiologiques antérieures sur *A. salmonicida* sous-espèce *salmonicida* avaient démontré aussi la présence de polymorphismes entre les différents échantillons, mais il n'y avait pas assez de variations pour affirmer que les isolats étaient génétiquement hétérogènes (McCormick *et al.*, 1990; Hänninen *et al.*, 1995; García *et al.*, 2000; O'hici *et al.*, 2000). Les techniques moléculaires qui ont été utilisées lors des études précédentes étaient la RLFP (McCormick *et al.*, 1990), la RAPD (Hänninen *et al.*, 1995; O'hici *et al.*, 2000) et la PFGE (García *et al.*, 2000; O'hici *et al.*, 2000). Les trois techniques se sont démontrées inefficaces pour distinguer entre les différents isolats testés.

Dans cette présente étude, même si la RAPD a été optimisée en utilisant des amorces à hauts sites d'hybridation, cette technique ne demeure pas assez sensible pour faire la distinction entre les quatre isolats d'*A. salmonicida* sous-espèce *salmonicida*, car la RAPD n'amplifie pas spécifiquement les sites de polymorphismes qui semblent peu nombreux entre les génomes des isolats. Bien que la RAPD a déjà été utilisée pour distinguer entre des isolats de certaines espèces bactériennes, telles *Mycoplasma hyopneumoniae* (Artiushin et Minion, 1996), *Salmonella enteritidis* (Lin *et al.*, 1996) et *Pseudomonas aeruginosa* (Mahenthiralingam *et al.*, 1996; Ouellet *et al.*, 2015), son utilisation dans le cadre d'une étude épidémiologique sur *A. salmonicida* sous-espèce *salmonicida* n'est pas envisageable. Dans l'objectif de mettre au point des techniques d'identification adaptée aux isolats d'*A. salmonicida* sous-espèce *salmonicida*, les analyses des différences chromosomiques entre les souches A449 et 01-B526 se révèlent une alternative plus prometteuse.

6.2. Différences chromosomiques entre A449 et 01-B526

La plasticité génomique chez les bactéries est un mécanisme complexe et dynamique qui permet l'évolution de leur matériel génétique et ainsi leur permet de s'adapter à un environnement en perpétuel changement. Plusieurs éléments ont clairement été identifiés dans la littérature comme pouvant causer la plasticité génomique, tels les MGE et les séquences TR (Darmon et Leach, 2014; Zhou et al., 2014).

Pour tenter de déduire les effets des différentes pressions évolutives qui ont affecté les génomes des deux souches d'*A. salmonicida* sous-espèce *salmonicida* A449 (Reith *et al.*, 2008) et 01-B526 (Charette *et al.*, 2012), leurs chromosomes ont été alignés et les types et les lieux des différences génétiques ont été déterminés manuellement. La souche A449 a été isolée en France, alors que la souche 01-B526 provient du Québec. Étant deux souches bactériennes de la même espèce, mais provenant de deux continents, il était soupçonné que chacun des génomes ait divergé pour ainsi s'adapter à leur environnement.

Suite aux analyses, trente-deux polymorphismes chromosomiques ont été répertoriés entre A449 et 01-B526 (Tableau 3). Les différences ont été classées en cinq catégories : 1) les modifications causées par des IS (13/32), 2) les variations du nombre de répétitions dans des séquences TR (12/32) (voir exemple Figure 4), 3) les insertions/délétions de séquences dans des gènes (5/32) (voir exemple Figure 5), 4) l'insertion/délétion d'un GEI (1/32) et 5) une région possédant plusieurs différences qui sont issues de plusieurs évènements (1/32).

6.2.1. Modifications impliquant des IS

L'activité de certaines IS chez *A. salmonicida* sous-espèce *salmonicida* a déjà été observée suite à une incubation prolongée des cultures à des températures élevées pour les ISAS1, ISAS2 et ISAS11 (Gustafson *et al.*, 1994; Tanaka *et al.*, 2012) ou lors d'analyses bio-informatiques du plasmide pAsal1 (ISAS5) (Trudel *et al.*, 2013) ou du chromosome d'A449 (ISAS4) (Reith *et al.*, 2008). Les analyses effectuées dans le cadre de cette étude démontrent qu'il y aurait eu activation de certaines des IS de type ISAS2, ISAS3, ISAS4, ISAS8 et ISAS9 entre les génomes d'A449 et de 01-B526, dont trois groupes (ISAS3, ISAS8 et ISAS9) dont l'activité n'avait encore jamais été observée jusqu'à maintenant dans la littérature. Certaines des ISAS2, ISAS3, ISAS4, ISAS8 et

ISAS9 semblent avoir mené soit à des mouvements d'IS ou à des recombinaisons entre deux IS du même groupe.

Les analyses révèlent que deux groupes d'IS, appelés ISAS2 et ISAS9, sont présents dans le génome d'A449, mais absent chez 01-B526, à raison de trois copies pour chacune des IS dans le chromosome et de deux et une copies sur le plasmide de grande taille pAsa4, respectivement (résultat non montré). Un fait intéressant est que la souche 01-B526 ne possède ni le plasmide pAsa4, ni les ISAS2 et ISAS9. On pourrait proposer l'hypothèse que le plasmide pAsa4 a été le vecteur de propagation de ces deux IS chez A449. Donc, les trois ISAS2 et trois ISAS9 observées présentement dans le chromosome d'A449 étaient potentiellement au préalable sur le plasmide pAsa4, mais suite à l'acquisition de pAsa4 par A449, elles auraient transposées au fil des générations dans le chromosome.

Lors d'une incubation prolongée d'A449 à une température inadéquate, l'ISAS2 peut se transposer dans le gène *vapA* (Gustafson *et al.*, 1994). La souche A449 a été isolée en France en 1979 et la séquence de son génome a été publiée en 2008. Entre temps, les conditions dans lesquelles elle a été cultivée lors des années avant d'être séquencée sont méconnues. Si la souche A449 a été incubée à des températures égales ou supérieures à 25°C de façon prolongée, ceci aurait pu entraîner une activation des IS et ainsi entraîner des dissemblances génomiques, comme démontrent les études sur les ISAS1, ISAS2 et ISAS11 (Gustafson *et al.*, 1994; Tanaka *et al.*, 2012). Puisque le Bergey's (Brenner *et al.*, 2005) propose avec erreur (Daher *et al.*, 2011) qu'*A. salmonicida* sous-espèce *salmonicida* puisse être cultivé à 25°C, il est fort probable que la souche A449 ait été cultivée de façon inadéquate dans les décennies séparant son isolement du séquençage de son génome. Cependant, on ne peut affirmer hors de tout doute que la présence des ISAS2 et ISAS9 dans le chromosome d'A449 est due à des mouvements de ces IS du plasmide pAsa4 vers le chromosome suite à la culture inadéquate de cette souche, mais cette éventualité peut être envisagée.

Les analyses des variations causées par ces deux types d'IS démontrent qu'une ISAS2 est présente dans un gène (Tableau 3, dissemblance 10), alors que dans les cinq autres IS sont intégrées dans des séquences non-codantes (Tableau 3, dissemblances 1,8, 19, 31 et 32). Il n'y a pas de conséquences claires sur l'insertion des trois ISAS9 dans le génome d'A449, car elles sont toutes intégrées dans des séquences non-codantes de l'ISAS5, en 5' du premier gène. D'un autre côté, les ISAS2 semblent plus incommodes pour le chromosome d'A449. Une ISAS2 a interrompu un ORF correspondant à un gène

codant pour une protéine hypothétique et il l'a scindé en deux (*ASA_1563* et *ASA_1565*), donc ce gène est considéré comme inactif. Les deux autres ISAS2 se sont intégrés dans des régions inter-géniques, donc elles risquent d'affecter la transcription des gènes placés en 3'. Pour chaque situation, un et deux gènes sont placés en 3' d'un ISAS2. Les gènes codent soit pour une protéine hypothétique (*ASA_2570*) (Tableau 3, dissemblance 19), une *GntR family transcriptional regulator* (*ASA_4274*) ou une *PTS system alpha-glucoside-specific EIICB component* (*ASA_4276*) (Tableau 3, dissemblance 31).

Douze et treize ISAS3, aussi appelées ISAS7 dans le plasmide pAsa5 (numéro GenBank NC_009348) (Tableau 1), sont présentes dans le chromosome d'A449 et dans celui de 01-B526, respectivement. Deux autres ISAS3 complètes et une partielle interrompue par une séquence d'à peu près 7600 pb sont situées sur le plasmide de grande taille pAsa5 d'A449 (numéro GenBank CP00646). Le nombre d'ISAS3 présent sur le plasmide pAsa5 de 01-B526 n'est pas encore connu, car ce plasmide n'a pas été terminé et assemblé de façon définitive. Cependant, il serait intéressant d'en connaître le nombre, car la séquence de ce plasmide pourrait être la source de la treizième ISAS3 située dans le chromosome de 01-B526. Quatre différences chromosomiques sont causées par l'intégration de l'ISAS3 dans une séquence génique (Tableau 3, dissemblances 6, 11, 16 et 26). Trois ISAS3 interrompent spécifiquement des gènes chez 01-B526 et une ISAS3 est présente uniquement dans un gène chez A449. L'intégration d'une ISAS3 chez A449 dans une séquence génique codant pour une *cytolytic delta-endotoxin* interrompt son cadre de lecture et devrait probablement inactiver ainsi le gène. Pour confirmer l'inactivation des gènes interrompus par les IS, des analyses de transcriptomique pourraient être effectuées. Cependant, la caractérisation des gènes inactivés par l'intégration d'une IS n'était pas l'objectif de cette étude. Chez 01-B526, trois ISAS3 se sont insérés dans les gènes codant pour une *TniA transposase (tniA)*, ISAS4 et pour une *maltose operon periplasmic protein (malM)*. L'alignement chromosomique entre les deux souches démontre aussi la présence d'une séquence supplémentaire de 5500 pb dans A449 encadrée par deux ISAS3 en orientation directe. À la position où devrait se situer la séquence de 5500 pb encadrée par deux ISAS3 chez 01-B526, on peut observer seulement une ISAS3 (Tableau 3, dissemblance 15). Donc, il est probable qu'une recombinaison entre les deux ISAS3 chez 01-B526 ait ainsi mené à l'excision d'une des ISAS3 et entraîné l'élimination de la séquence inter-IS de 5500 pb avec elle (résultat non montré). Cependant, l'excision de la séquence de 5500 pb et d'une des deux ISAS3 aurait pu se produire en deux événements distincts.

Dans l'une des autres différences observées, l'interaction entre deux ISAS4 inversées sur le chromosome aurait entraîné une inversion chromosomique présente dans A449, mais pas chez 01-B526 (Tableau 3, dissemblance 27). Cette inversion chromosomique de 195 kb chez A449 avait déjà été observée lors de l'analyse du biais en GC par Reith et ses collaborateurs en 2008 (Reith *et al.*, 2008). Nous avons cependant observé que le prophage 2 est positionné dans l'inversion de 195 kb, donc, le prophage 2 est dans des orientations contraires sur les chromosomes des souches A449 et 01-B526.

La dernière différence causée par le mouvement d'une IS semble due à l'excision de l'ISAS8 dans le chromosome de 01-B526 qui aurait entraîné des séquences en aval et en amont avec elle (Tableau 3, dissemblance 2). La séquence perdue dans 01-B526 totalise ~4500 pb (en comptant l'ISAS8 qui fait 754 pb). L'alignement chromosomique entre les deux souches révèle l'absence de la séquence de 4500 pb chez 01-B526. L'annotation du génome d'A449 démontre la présence d'une ISAS8 dans le 4500 pb manquant chez 01-B526. Le motif putatif du site de recombinaison ayant causé l'excision est CCGGTC. Ce motif encadre la séquence supplémentaire présente chez A449, mais est seulement présent une fois chez 01-B526. Le deuxième gène de l'ISAS8 chez A449, qui est appelé *insB* (locus tag : ASA_0144), est annoté comme un pseudogène. Des mutations ponctuelles ont créé deux codons-stops dans ce gène tronquant la protéine lors de la traduction, donc cette ISAS8 est probablement devenue inactive chez A449.

Les IS sont connues dans la littérature comme l'une des causes de l'instabilité génomique chez les bactéries (Darmon et Leach, 2014). Les divergences chromosomiques entre A449 et 01-B526 sont avant tout causées par le mouvement des IS (13 fois sur 32 dissemblances chromosomiques). Brièvement, des mouvements d'IS ont mené à une inversion chromosomique de 195 kb chez A449, deux délétions de séquences chez 01-B526, et à des bris d'ORF et de régions inter-géniques dans les deux chromosomes.

6.2.2. Variations du nombre de répétitions dans des séquences TR

Les conséquences de la variation des régions TR dépendent du lieu où elles sont situées. Si la séquence TR est placée dans une région inter-génique ou dans une région intra-génique, les effets sont variés (Zhou *et al.*, 2014). Dans le cadre de cette recherche, aucune étude fonctionnelle sur les régions TR n'a été effectuée puisque l'intérêt premier de l'étude était de trouver des marqueurs génétiques permettant de distinguer les souches

d'*A. salmonicida* sous-espèce *salmonicida*. On constate par contre une corrélation entre le nombre des répétitions dans une séquence TR, l'origine géographique de l'isolat et le type d'*AsaGEI* qu'il possède (Tableau 4). Par exemple, le motif de la séquence TR #28 est répété 2,6 dans toutes les souches analysées provenant d'Europe, alors que ce même motif est répété entre 5,6 et 8,6 fois dans les souches étudiées provenant d'Amérique du Nord. Un deuxième exemple : les souches possédant *AsaGEI1a* ont le même nombre de répétitions pour les séquences TR #13, #22 et #28 (Tableau 4).

Pour l'instant, seulement cinq des douze séquences TR ont été analysées dans dix autres génomes d'isolats d'*A. salmonicida* sous-espèce *salmonicida* séquencés par le laboratoire Charette (Tableau 4). Mais les tendances de spécificité du nombre de répétitions des motifs dans les séquences TR d'après l'origine géographique et du type d'*AsaGEI* sont déjà présentes. Les cinq isolats provenant de l'Europe ont le même nombre de répétitions pour trois des cinq séquences TR (#13, #21 et #28). Il ne semble pas avoir de concordance entre le nombre des répétitions et le génotype d'*AsaGEI* pour les isolats européens, alors que deux groupes sont observables entre les sept isolats d'Amérique du Nord, selon qu'ils possèdent soit un *AsaGEI1a* ou *AsaGEI2a* (Tableau 4).

L'utilisation des séquences VNTR pour définir les relations évolutives entre des isolats de la même espèce bactérienne est appelé *multi-locus VNTR analysis* (MLVA). L'utilisation de la technique MLVA a été proposée pour de nombreux agents pathogènes, comme pour *Escherichia coli* O157:H7 (Lindstedt *et al.*, 2004), *Pseudomonas aeruginosa* (Onteniente *et al.*, 2003), *Yersinia pestis* (Pourcel *et al.*, 2004), *Bacillus anthracis* (Flèche *et al.*, 2001), *Mycobacterium tuberculosis* (Flèche *et al.*, 2002), *Legionella pneumophila* (Pourcel *et al.*, 2003) et *Salmonella enterica* sous-espèce *enterica* (Ramisse *et al.*, 2004). La technique MLVA n'a jamais encore été utilisée dans la littérature pour définir les liens de parenté entre des isolats d'*A. salmonicida* sous-espèce *salmonicida*, mais cette technique semble prometteuse suite aux analyses préliminaires de cinq séquences VNTR entre douze génomes d'isolats de cette bactérie.

Finalement, elle pourrait être utilisée pour faire la distinction entre les isolats provenant soit d'Europe ou d'Amérique du Nord et même faire la différence entre les groupes d'isolats d'Amérique du Nord (*AsaGEI1a* ou *AsaGEI2a*) (Tableau 4).

6.2.3. Différences impliquant l'insertion/délétion d'une séquence dans un gène

Le troisième groupe représentant cinq différences chromosomiques est constitué de séquences insérées/excisées dans des gènes (Tableau 3, dissemblances 3, 17, 18, 25 et 29). Les caractéristiques générales de ces séquences sont que leurs longueurs varient entre 37 et 347 pb et qu'elles sont encadrées par des régions DR (site d'insertion dupliqué).

Une recherche approfondie pour déterminer la nature des cinq séquences intra-géniques présentes chez A449, mais absentes chez 01-B526, révélait qu'elles possèdent plusieurs caractéristiques de la famille des éléments miniatures transposables avec région répétée inversée, ou en anglais, *miniature inverted-repeat transposable-element* (MITE) (Darmon et Leach, 2014). Mais, il manque une caractéristique essentielle aux séquences chez A449 pour faire partie de la famille des MITE : elles ne sont pas encadrées de séquences répétées inversées (IR). Donc l'origine et la nature des cinq différences causées par l'insertion/délétion de séquence dans des gènes n'ont pas encore été déterminées pour le moment faute d'avoir trouvé des situations équivalentes dans la littérature.

En complément, des analyses déterminant la présence ou l'absence des cinq séquences intra-géniques dans les dix autres génomes d'isolats d'*A. salmonicida* sous-espèce *salmonicida* ont été effectuées. Les résultats révèlent qu'il y a plusieurs corrélations entre la présence/absence de ces séquences dans les gènes correspondants, l'origine géographique de l'isolat et le type d'*AsaGEI* présent dans le génome de l'isolat (Tableau 4). Quatre des dix isolats analysés proviennent de l'Europe et ils démontrent tous la même tendance: les quatre possèdent les cinq séquences intra-géniques, tout comme A449. Dans les quatre isolats, deux possèdent *AsaGEI1b* (HER1085 et ATCC 33586), un porte *AsaGEI2b* (JF3224) et le dernier ne possède aucun GEI (JF2267). Six des dix isolats proviennent de l'Est de l'Amérique du Nord (Québec et Nouveau-Brunswick). Trois groupes sont distinctifs au regard des analyses de présence/absence des cinq séquences intra-géniques sur les six isolats d'Amérique du Nord. Le premier groupe est constitué des isolats 09-0167, 2009-144K3, 2004-05 MF26 et m23067-09 qui portent un *AsaGEI2a*. Les quatre isolats possèdent trois des cinq séquences insérées dans les gènes *ASA_0621*, *ASA_2192* et *tnaA*. Le deuxième groupe est constitué de l'isolat m15789-11 possédant l'élément *AsaGEI1a*. Cet isolat ne possède aucune des séquences intra-géniques d'A449, donc il démontre le même résultat que 01-B526. Le troisième groupe est constitué de l'isolat 01-B522 qui porte aussi *AsaGEI1a*. Quant à lui,

01-B522 possède seulement la séquence insérée dans le gène *ASA_2192*. Une des différences entre 01-B522 et les deux autres isolats possédant aussi *AsaGEI1a* est qu'elle porte le plasmide de grande taille pAsa4, absent chez 01-B526 et m15789-11.

Finalement, la présence/absence des cinq séquences insérées dans les gènes *ASA_0621*, *ASA_2192*, *tnaA*, *ASA_3022* et *amt* est révélatrice de l'origine géographique des isolats et du génotype d'*AsaGEI* que les isolats possèdent. Ces cinq séquences pourraient être utilisées comme indicateur de la provenance des différents isolats. De plus, les résultats de ces analyses pourraient être couplés à une technique de MLVA pour ainsi déterminer plus précisément l'origine des souches d'*A. salmonicida* sous-espèce *salmonicida* responsables d'éclosion de furonculose dans les aquacultures

6.3. Caractérisation des *AsaGEI* chez *Aeromonas salmonicida*

Dans la littérature, la grande plasticité des GEI a souvent été observée et cette particularité entraîne un changement constant des gènes les constituant. L'origine et la nature des gènes qui sont intégrés dans les GEI au fil de l'évolution ne sont pas encore tout à fait comprises. Le type de GEI dépend de la fonction des protéines pour lesquelles ses gènes encodent. Par exemple, un GEI possédant des gènes codant soit pour une voie métabolique ou pour des facteurs de virulence sera appelé îlot métabolique ou îlot de pathogénicité (PAI), respectivement (Darmon et Leach, 2014). Les GEI augmentent généralement le *fitness* des bactéries qui les portent. En d'autres termes, les GEI permettent aux bactéries de peupler de nouveaux environnements et favorisent leur survie et leur compétitivité (Dobrindt *et al.*, 2004).

La comparaison chromosomique des souches A449 et 01-B526 d'*A. salmonicida* sous-espèce *salmonicida* a entraîné la découverte d'un GEI. Ce type de MGE n'avait encore jamais été observé chez cette espèce bactérienne. Ce nouveau GEI, qui a été appelé *AsaGEI1a* pour *Aeromonas salmonicida Genomic Island 1a*, est la différence majeure entre les chromosomes des souches A449 et 01-B526. De plus, deux variants d'*AsaGEI1a* avaient été détectés par PCR dans d'autres isolats d'*A. salmonicida* sous-espèce *salmonicida* (Figure 2). Ces deux particularités lui ont accordé la majeure section de cette étude.

6.3.1. Relation évolutive entre les différents éléments AsaGEI ou reliée aux AsaGEI

L'alignement d'*AsaGEI1a* de la souche 01-B526 avec les deux autres éléments provenant aussi de génomes de bactéries du genre *Aeromonas* (*A. hydrophila* RB-AH et *A. jandaei* Riv2) (Figure 7) révèle une grande plasticité entre les trois éléments. Tandis que les trois éléments partagent des pourcentages d'identité et de longueur élevés et qu'ils possèdent une structure similaire, nous proposons qu'ils appartiennent à la même famille de GEI, mais ils auraient évolué dans des environnements différents pour ainsi devenir trois éléments distincts. L'élément génétique de la souche RB-AH présente le même contexte chromosomique et encadré par le même 19 pb que l'élément *AsaGEI1a* de 01-B526 (résultat non montré). Ces observations nous indiquent que ces deux éléments génétiques sont probablement plus proches évolutivement entre eux qu'avec l'élément de l'isolat Riv2, qui est intégré à un autre site chromosomique (résultat non montré). La découverte des éléments génétiques chez *A. hydrophila* RB-AH et *A. jandaei* Riv2 indique que cette famille de GEI a été transférée à d'autres bactéries du genre *Aeromonas* et qu'il est probablement assez répandu dans ce genre bactérien. Pour l'instant, les résultats de séquençage de génomes bactériens présents sur les bases de données publiques sont limités, une augmentation de la disponibilité de ces données pourrait accroître la découverte de nouveaux variants d'éléments génétiques semblables à *AsaGEI1a* chez d'autres bactéries du genre *Aeromonas*.

Les analyses BLASTn sur l'élément *AsaGEI1a* contre la base de données *wgs* révélaient aussi qu'il partage une longueur de 35% avec une identité de séquence de 90% avec un élément génétique de 58 kb chez *A. salmonicida* sous-espèce *achromogenes* AS03 (Figure 8). L'annotation de l'élément chez AS03 démontre la présence de trois gènes codants pour des protéines structurales phagiques, donc l'élément chez AS03 pourrait être un prophage putatif. En plus du lien de parenté avec un prophage putatif, *AsaGEI1a* est constitué à 30% (19/63) d'ORF d'origines phagiques, mais aucun gène codant pour des protéines de structure de phage n'a été identifié. La région codante pour les protéines de structures phagiques chez AS03 n'est pas conservée dans *AsaGEI1a* (Figure 8). De plus, les éléments chez l'isolat *A. jandaei* Riv2 et chez *A. salmonicida* sous-espèce *achromogenes* AS03 ont la même intégrase et sont tous deux intégrés dans le même lieu génomique (résultat non montré). Sachant ces informations, il est proposé que l'élément *AsaGEI1a* et les éléments chez RB-AH et Riv2 dérivent d'un même prophage ancestral qui aurait fortement évolué pour perdre de nombreux éléments phagiques.

À part *AsaGEI1a*, trois autres éléments *AsaGEI* ont été identifiés dans des génomes d'isolats d'*A. salmonicida* sous-espèce *salmonicida*, séquencés et annotés (*1b;2a;2b*). Les listes des fonctions des gènes pour les *AsaGEI1a*, *AsaGEI1b*, *AsaGEI2a* et *AsaGEI2b* sont placées à l'Annexe C, E, F et G, respectivement. L'alignement des quatre *AsaGEI* est présenté à la Figure 12. On peut constater que les quatre éléments ont une structure équivalente. Ils ont aussi en commun une région variable (Figure 12-b) et une région conservée (Figure 12-c). Dans la région variable, les gènes appelés *orf17* dans chacun des *AsaGEI*, qui sont des gènes différents chez *AsaGEI1a*, *1b* et *2a*, codent dans les quatre *AsaGEI* pour une protéine présentant une fonction potentielle de DAM méthylase. La même observation a été faite pour trois autres gènes dans la région variable, appelés dans chacun des *AsaGEI* *orf1*, *orf14* et *orf34*. Ils codent respectivement pour des protéines possédant la fonction d'intégrase, de recombinaise et de *DNA-binding*. Alors que plusieurs gènes de la région variable (Figure 12-b) sont différents entre les quatre *AsaGEI*, les fonctions des protéines encodées par certains de ces gènes sont conservées entre chaque élément *AsaGEI*. En d'autres mots, la région variable présente une certaine constance en fin de compte, car les fonctions des protéines encodées par certains gènes sont conservées.

Trois intégrases possédant chacune un site d'insertion spécifique ont été identifiées pour les différents éléments *AsaGEI* (*1a;1b;2a;2b*) ou reliées aux *AsaGEI* (Riv2, RB-AH et AS03). Les intégrases sont déjà connues dans la littérature pour leurs spécificités face aux motifs nucléotidiques qu'elles reconnaissent (Fogg *et al.*, 2014). On peut observer à la Figure 14 la représentation de la relation évolutive entre les quatre *AsaGEI* et les trois éléments reliés aux *AsaGEI* d'après leurs sites d'insertion (**A**, **B** et **C**) et les pourcentages de longueur partagée (*query cover*) entre les différents éléments possédant le même site d'insertion.

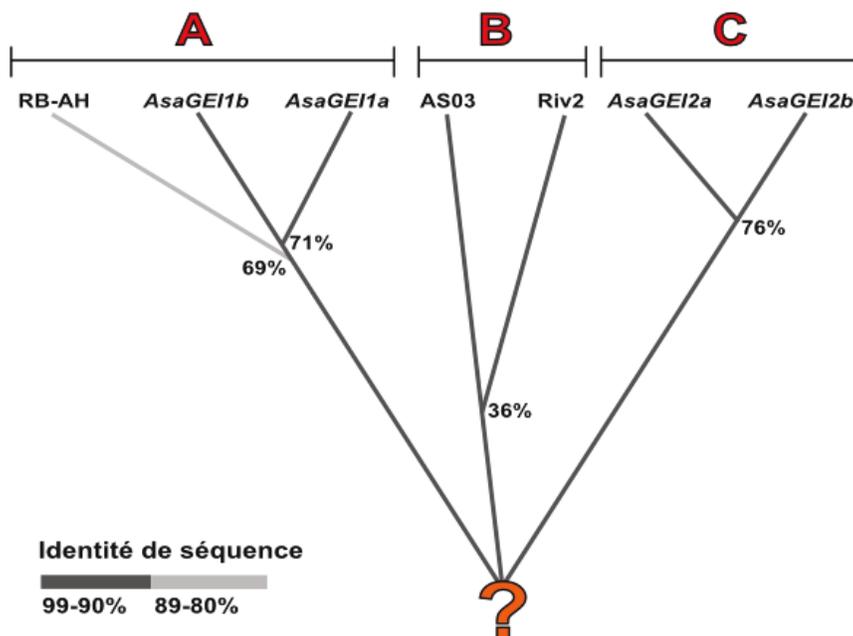


Figure 14 : Relation évolutive entre les quatre *AsaGEI* et les trois éléments apparentés aux *AsaGEI* d'après leurs sites d'insertion (**A**, **B** et **C**). Les pourcentages de longueur partagée entre chaque élément possédant le même site d'insertion (ou intégrase) sont indiqués aux nœuds. Chaque branche représente une valeur de longueur partagée de 100%. En **A**, les pourcentages indiqués aux nœuds de RB-AH et *AsaGEI1a* ont été déterminés par rapport à *AsaGEI1b*, car ce dernier élément possède les pourcentages de longueur partagée les plus élevés avec RB-AH et *AsaGEI1a* (résultat non montré).

6.3.2. Relation entre l'origine géographique des isolats, le génotype d'*AsaGEI* et la présence/absence du prophage 3

Antérieurement à la découverte des différents variants d'*AsaGEI1a* dans d'autres isolats d'*A. salmonicida* sous-espèce *salmonicida* (Figure 2), des criblages PCR ont été effectués pour déterminer les génotypes d'*AsaGEI* dans les autres isolats de la collection Charette (Figure 11). Les résultats du génotypage pour la présence des ORF d'*AsaGEI1a* sont inscrits dans le Tableau 5. On peut observer dans ce Tableau que cinq isolats ne suivent pas les tendances générales des résultats de génotypage. Quatre des cinq isolats proviennent du Système hydrographique des Grands Lacs et du Saint-Laurent, parmi lesquels deux portent un *AsaGEI1b* et deux n'en porte aucun. Le dernier isolat provient de l'Europe et porte un *AsaGEI2a*.

Pour confirmer le résultat de génotypage de l'isolat *AsaGEI2a* européen (JF3224), son génome a été séquencé. Les données de séquençage pour cet isolat indiquent qu'il

est en réalité porteur d'un nouveau variant d'*AsaGEI2*, appelé *AsaGEI2b*. Le nouvel élément possède le même site d'insertion qu'*AsaGEI2a* (Figure 10B). L'alignement génomique des quatre éléments *AsaGEI* est représenté à la Figure 12.

Pour déterminer si les deux isolats d'Amérique du Nord portent bel et bien un *AsaGEI1b*, des PCR ont été effectuées visant spécifiquement ce GEI et les résultats indiquaient que les deux éléments génétiques n'étaient finalement pas des *AsaGEI1b* (résultat non montré). Des expériences supplémentaires seront nécessaires pour déterminer le génotype des deux isolats d'Amérique du Nord qui reste pour le moment indéterminé. Il n'est pas exclu que ces isolats portent une nouvelle forme d'*AsaGEI*.

Tableau 6 : Origine géographique des 139 isolats d'*A. salmonicida* sous-espèce *salmonicida* par rapport à leur génotype d'*AsaGEI* (Tableau 5 corrigé).

	<i>AsaGEI</i>						Total
	<i>1a</i>	<i>2a</i>	?	<i>1b</i>	<i>2b</i>	sans	
Système hydrographique des Grands Lacs et du Saint-Laurent	29	71	2	0	0	2	104
Provinces maritimes de l'Atlantique	0	10	0	0	0	0	10
Europe ^a	0	0	0	10	1	14	25
Total	29	81	2	10	1	16	139

À la lumière de ces nouveaux résultats qui sont inscrits dans le Tableau 6 et illustrés à la Figure 15, chaque *AsaGEI* est spécifique à une région hydrographique : *AsaGEI1a* et *AsaGEI2a* sont observés seulement dans l'Est de l'Amérique du Nord, alors qu'*AsaGEI1b* et *AsaGEI2b* sont spécifiques à l'Europe. Une majorité des isolats sans GEI proviennent de l'Europe. Ainsi, les génotypes des éléments *AsaGEI* sont indicatifs de l'origine géographique des isolats (Amérique du Nord ou Europe).

Une réflexion est proposée : *AsaGEI1a* et *AsaGEI1b* sont deux éléments différents (Figure 12), ils partagent des pourcentages de longueur et d'identité élevés l'un avec l'autre (Annexe D), ils ont le même site d'insertion (Figure 10A) et ils sont spécifiques chacun à un continent différent (Figure 15). Donc, il est proposé qu'*AsaGEI1a* et *AsaGEI1b* ont un ancêtre commun récent et qu'ils auraient évolué indépendamment l'un par rapport à l'autre dans deux environnements pour ainsi divergés et devenir deux éléments *AsaGEI* distincts (Figure 14). La même hypothèse peut être proposée pour la

relation évolutive entre *AsaGEI2a* et *AsaGEI2b*, qui sont aussi deux éléments différents (Figure 15) avec des pourcentages de longueur et d'identité l'un avec l'autre de 76% et 94%, respectivement.

Suite aux séquençages d'isolats possédant *AsaGEI2a* (2009-144K3 et m23067-09), un nouveau prophage a été découvert (prophage 3). Ce nouveau prophage a des pourcentages de longueur et d'identité élevés avec deux autres éléments dans les génomes d'*A. veronii* B565 et d'*A. hydrophila* 4AK4 (Figure 13), dont les deux semblent aussi être des prophages. Il est proposé que les trois prophages alignés à la Figure 13 soient de la même famille. De plus, le prophage 3 est spécifique à l'Amérique du Nord, car exclusivement observé chez les isolats d'*A. salmonicida* sous-espèce *salmonicida* provenant de ce continent. Le prophage 3 est fortement lié à *AsaGEI2a*, car 95% des isolats portant le prophage possèdent aussi *AsaGEI2a* (77/81) (Annexe A).

Finalement, il existe des éléments semblables aux quatre *AsaGEI* d'*A. salmonicida* sous-espèce *salmonicida* chez deux autres espèces bactériennes du genre *Aeromonas*. De plus, les six éléments ont des liens de parenté avec un prophage putatif chez *A. salmonicida* sous-espèce *achromogenes* AS03. Il a été observé aussi que chacun des *AsaGEI* est spécifique à un continent (Figure 15), tout comme le prophage 3, qui est seulement présent en Amérique du Nord et majoritairement chez les isolats *AsaGEI2a*. Ces deux types de MGE pouvant être utilisés comme indicateur de la provenance des isolats étaient les éléments recherchés au début de cette étude.



Figure 15 : Représentation de la provenance des 139 isolats d'*Aeromonas salmonicida* sous-espèce *salmonicida* de la collection Charette listés à l'Annexe A utilisés pour ces travaux de recherche. Les chiffres en bleu indiquent le nombre d'isolats d'*A. salmonicida* sous-espèce *salmonicida* provenant de chaque pays. Les chiffres en noir indiquent le nombre d'isolats présentant chaque des génotypes d'*AsaGEI*.

6.4 Éléments génétiques responsables de la variabilité chromosomique chez *A. salmonicida* sous-espèce *salmonicida*

Les chromosomes des isolats d'*A. salmonicida* sous-espèce *salmonicida* font face à plusieurs familles d'éléments génétiques modulant la séquence et la structure de leur génome. Dans le cadre de cette étude, plusieurs familles d'éléments créant des dissemblances entre les différentes souches analysées ont été identifiées. Les principales familles d'éléments génétiques sont les IS, les séquences TR, les courtes séquences intra-géniques encadrées par une région DR (famille indéterminée pour l'instant), les GEI et les prophages.

Suite à l'alignement chromosomique d'A449 et de 01-B526, les IS, les séquences TR et les courtes séquences intra-géniques sont la cause de 94% des dissemblances (30/32) observées entre les deux isolats. Les IS et les séquences TR sont responsables à elles seules de 78% des dissemblances (25/32). Ces deux familles d'éléments génétiques sont bien connues dans la littérature comme l'une des causes de l'instabilité génomique chez les bactéries (Darmon et Leach, 2014; Zhou et al., 2014). Le troisième groupe englobe les courtes séquences intra-géniques. Elles représentent 16% des dissemblances (5/32) entre les deux souches, mais leurs mécanismes sont pour le moment inconnus. Ces trois types d'éléments génétiques de petite taille ont créé des dissemblances entre les génomes de souches d'*A. salmonicida* sous-espèce *salmonicida* (A449 et 01-B526) et les font ainsi diverger.

Cinq éléments génétiques de grandes tailles (> 35 kb) ont été découverts entre des isolats d'*A. salmonicida* sous-espèce *salmonicida*. Les éléments observés dans les chromosomes sont les quatre variants d'un GEI appelé *AsaGEI* (*AsaGEI1a*, *AsaGEI1b*, *AsaGEI2a* et *AsaGEI2b*) et un prophage (prophage 3). Les GEI confèrent généralement à la bactérie certains avantages, car ils peuvent coder pour des protéines augmentant le *fitness* (Darmon et Leach, 2014; Juhas et al., 2009). Les prophages peuvent coder quant à eux pour des facteurs de virulence et ainsi augmenter la pathogénicité de leurs hôtes (Cerdeno-Tarraga, 2003; Plunkett III et al., 1999; Waldor et Mekalanos, 1996). Finalement, ces deux familles de MGE créent une mutation à leur site d'insertion, mais ils apportent aux isolats d'*A. salmonicida* sous-espèce *salmonicida* de nouveaux gènes, ce qui pourrait potentiellement augmenter leur *fitness*.

7. Conclusion

Les isolats d'*A. salmonicida* sous-espèce *salmonicida* ont longtemps été considérés comme étant génétiquement homogènes (Belland et Trust, 1988; García et al., 2000; Hänninen et al., 1995; McCormick et al., 1990; Miyata et al., 1995; O'hici et al., 2000; Studer et al., 2013). Dans le cadre de cette étude, plusieurs éventualités ont été explorées dans l'optique de mettre en évidence une méthode pour distinguer des divergences génomiques entre les différents isolats d'*A. salmonicida* sous-espèce *salmonicida* qui seraient utiles à un suivi épidémiologique pour ainsi comprendre et traquer le déplacement de souches pathogènes responsable d'éclosion de furunculose.

Dans un premier temps, la technique de RAPD a été optimisée sur quatre isolats d'*A. salmonicida* sous-espèce *salmonicida* en testant dix-huit amorces de 9 nt à hauts sites d'hybridation (Tableau 2). Deux des dix-huit amorces utilisées pour la RAPD amplifiaient des sites de polymorphismes. Brièvement, les amorces JG-RAPD11 et JG-RAPD20 amplifiaient chacun soit deux et un sites polymorphes, respectivement (Figure 3A et C). Même en utilisant des amorces ayant un nombre élevé de sites d'hybridation, ce qui augmente la probabilité d'amplifier des sites de polymorphisme, la technique RAPD ne s'est pas montrée assez sensible pour distinguer entre les différents isolats.

Pour mettre au point une technique adaptée à *A. salmonicida* sous-espèce *salmonicida*, il est apparu nécessaire d'identifier les différentes dissemblances distinguant les deux chromosomes. Ainsi, un alignement chromosomique d'A449 et de 01-B526 a été effectué dans l'optique de mettre en évidence l'éventail des différences génétiques entre les deux souches. Trente-et-une dissemblances et un site de polymorphisme multi-événementiel ont été identifiés (Tableau 3). Le site possédant plusieurs variations n'a pas été analysé en détail dans le cadre de cette étude. Les trente-et-une dissemblances peuvent être regroupées en quatre familles d'après les agents causaux : mouvement d'IS (13), séquences TR (12) (Figure 4), insertion/délétion de séquences intra-géniques (5) (Figure 5) et insertion/délétion d'un GEI. Subséquemment, des analyses du nombre de répétitions dans cinq séquences TR et de la présence/absence de cinq séquences intra-géniques ont été effectuées dans dix autres génomes d'*A. salmonicida* sous-espèce *salmonicida* (Tableau 4). Les résultats démontrent l'existence de plusieurs sites polymorphes dans les génomes d'*A. salmonicida* sous-espèce *salmonicida*. Les cinq isolats européens ont des caractéristiques similaires, alors que deux groupes sont

observables en Amérique du Nord, dépendant si les isolats possèdent soit un *AsaGEI1a* ou un *AsaGEI2a*.

En dernier, une analyse en génomique comparative a été effectuée sur quatre variants d'une même famille de GEI (*AsaGEI*) observé dans la majorité des isolats d'*A. salmonicida* sous-espèce *salmonicida* étudiés. La présence d'un des quatre variants d'*AsaGEI* chez un isolat se révèle fortement corrélée avec son origine géographique. Les génotypes *AsaGEI1a* et *AsaGEI2a* sont seulement observés dans les isolats d'Amérique du Nord, alors que les variants *AsaGEI1b* et *AsaGEI2b* sont seulement présents dans les isolats européens. Par ce fait, la détection du génotype d'*AsaGEI* pourrait être utilisée comme indicateur de provenance de souches pathogènes responsable d'éclosion de furunculose.

En conclusion, plusieurs sites chromosomiques présentant des polymorphismes ont été découverts et répertoriés entre deux génomes d'*A. salmonicida* sous-espèce *salmonicida*. Dix de ces variations génétiques (cinq séquences TR et cinq insertions intra-géniques) ont été analysées dans un total de douze génomes d'*A. salmonicida* sous-espèce *salmonicida* (Tableau 4). Les analyses de ces dix dissemblances semblent distinguer entre les douze isolats à deux niveaux. Tout d'abord, les isolats possédant la même origine géographique regroupent ensembles (Europe et de l'Amérique du Nord). Ensuite, on peut distinguer entre les isolats provenant d'Amérique du Nord portant soit un *AsaGEI1a* ou *AsaGEI2a*. Ces deux familles de dissemblances génétiques représentent une plus grande diversité chromosomique que les *AsaGEI* dans les douze génomes testés, donc il serait préférable d'utiliser ces éléments pour définir les liens de parenté entre les isolats d'*A. salmonicida* sous-espèce *salmonicida* de la collection Charette, ce qui pourrait mener à une meilleure distinction des différents sous-groupes. Subséquemment, pour faire suite aux analyses des séquences chromosomiques, une méthode ciblant spécifiquement les variations chromosomiques pourrait être mise au point dans le but de mener des études épidémiologiques sur les isolats d'*A. salmonicida* sous-espèce *salmonicida* responsables d'éclosion de furunculose. De plus, les isolats de cette bactérie ne sont pas génétiquement homogènes, car plusieurs sites polymorphiques et éléments génétiques de grandes tailles (GEI et prophage) ont été découverts entre diverses souches.

8. Perspectives

Dans cette section, plusieurs idées pour des expériences futures sont proposées :

- 1) Historiquement, le laboratoire Charette déterminait le génotype d'*AsaGEI* chez les nouveaux isolats d'*A. salmonicida* sous-espèce *salmonicida* en utilisant 15 paires d'amorces ciblant chacun un gène et les deux sites d'insertion dans *AsaGEI1a* (*ins1*, *orf11*, *orf19*, *orf33*, *orf45*, *orf54*, *orf55*, *orf57*, *orf58*, *orf64*, *orf66*, *orf67*, *orf72*, *orf78* et *ins2*). Cependant, l'utilisation des 15 cibles PCR crée des erreurs de génotypage. Par exemple, il était soupçonné à un moment de l'étude que l'isolat européen JF3224 possédait *AsaGEI2a*, car il démontrait le même profil d'amplification que la souche 09-0167, qui porte un *AsaGEI2a* confirmé suite au séquençage. Pour contrer ces erreurs, le génotypage des nouveaux isolats pourraient être faits en deux temps :
 - a) Tout d'abord, effectuer un premier criblage en utilisant seulement sept cibles pour *AsaGEI1a* (*ins1*, *orf19*, *orf33*, *orf57*, *orf67*, *orf78* et *ins2*). À ce moment, il serait déjà possible d'avoir une idée du génotype approximatif que possède l'isolat, soit *AsaGEI1* ou *AsaGEI2*.
 - b) Ensuite, utiliser les amorces spécifiques pour chaque *AsaGEI* (Annexe B). Trois paires d'amorces ont été créées ciblant spécifiquement trois régions variables dans chacun des quatre génotypes d'*AsaGEI*. Les trois paires d'amorces ciblant *AsaGEI1b* ont été testées et elles amplifient spécifiquement les trois régions (résultat non montré). Brièvement, suite à la détermination du génotype *AsaGEI1*, il suffit d'utiliser les six paires d'amorces pour distinguer les deux sous-génotypes, soit *AsaGEI1a* et *AsaGEI1b*. Suivre le même raisonnement pour les isolats possédant le génotype *AsaGEI2* suite au premier criblage en (a).
- 2) Ultérieurement, il sera nécessaire de terminer les analyses bio-informatiques du nombre de répétitions pour chaque séquence TR dans les dix ébauches de génomes d'*A. salmonicida* sous-espèce *salmonicida*, mais aussi dans d'autres génomes dont la séquence est disponible. Ensuite, les résultats devront être analysés avec l'aide d'outils statistiques pour déterminer si les marqueurs sont significatifs. En dernier, des analyses de parenté génétique pourront être effectuées avec les marqueurs qui sont significatifs.

Bibliographie

- Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Artiushin S & Minion FC (1996) Arbitrarily primed PCR analysis of *Mycoplasma hyopneumoniae* field isolates demonstrates genetic heterogeneity. *Int J Syst Bacteriol* **46**: 324–328.
- Atienzar FA & Jha AN (2006) The random amplified polymorphic DNA (RAPD) assay and related techniques applied to genotoxicity and carcinogenesis studies: a critical review. *Mutat Res* **613**: 76–102.
- Attia AS & Hansen EJ (2006) A conserved tetranucleotide repeat is necessary for wild-type expression of the *Moraxella catarrhalis* UspA2 protein. *J Bacteriol* **188**: 7840–7852.
- Beilstein F & Dreiseikelmann B (2008) Temperate bacteriophage PhiO18P from an *Aeromonas media* isolate: characterization and complete genome sequence. *Virology* **373**: 25–29.
- Belland RJ & Trust TJ (1988) DNA:DNA reassociation analysis of *Aeromonas salmonicida*. *J Gen Microbiol* **134**: 307–315.
- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* **27**: 573–580.
- Bernoth E-M, Ellis AE, Midtlyng PJ, Olivier G & Smith P (1997) *Furunculosis: multidisciplinary fish disease research*. Academic Press, New-York.
- Besemer J, Lomsadze A & Borodovsky M (2001) GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* **29**: 2607–2618.
- Black LW (1995) DNA packaging and cutting by phage terminases: control in phage T4 by a synaptic mechanism. *BioEssays* **17**: 1025–1030.
- Black LW & Rao VB (2012) *Structure, assembly, and DNA packaging of the bacteriophage T4 head*. 1st ed. Elsevier Inc.
- Bose M & Barber RD (2006) Prophage Finder: a prophage loci prediction tool for prokaryotic genome sequences. *In Silico Biol* **6**: 223–227.
- Boyd J, Williams J, Curtis B, Kozera C, Singh R & Reith M (2003) Three small, cryptic plasmids from *Aeromonas salmonicida* subsp. *salmonicida* A449. *Plasmid* **50**: 131–144.

- Brenner DJ, Krieg NR, Staley JR & Garrity GM (2005) *Bergey's Manual of Systematic Bacteriology. Volume Two: The Proteobacteria, Part B: The Gammaproteobacteria*. 2nd ed. Garrity GM, ed. Springer, New-York.
- Bru H & Hendrix RW (2002) Phage Genomics : Small Is Beautiful. **108**: 13–16.
- Burr SE & Frey J (2007) Analysis of type III effector genes in typical and atypical *Aeromonas salmonicida*. *J Fish Dis* **30**: 711–714.
- Canchaya C, Fournous G & Brüssow H (2004) The impact of prophages on bacterial chromosomes. *Mol Microbiol* **53**: 9–18.
- Carson SDB, Stone B, Beucher M, Fu J & Sparling PF (2002) Phase variation of the gonococcal siderophore receptor FetA. *Mol Microbiol* **36**: 585–593.
- Carver T, Harris SR, Berriman M, Parkhill J & McQuillan J a (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28**: 464–469.
- Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG & Parkhill J (2005) ACT: the Artemis Comparison Tool. *Bioinformatics* **21**: 3422–3423.
- Casjens S, Palmer N, Vugt A Van, *et al.* (2000) A bacterial genome in flux : the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. *Mol Microbiol* **35**: 490–516.
- Cerdeno-Tarraga a. M (2003) The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. *Nucleic Acids Res* **31**: 6516–6523.
- Chai B, Wang H & Chen X (2012) Draft genome sequence of high-melanin-yielding *Aeromonas media* strain WS. *J Bacteriol* **194**: 6693–6694.
- Charette SJ & Cosson P (2004) Preparation of genomic DNA from *Dictyostelium discoideum* for PCR analysis. *Biotechniques* **36**: 574–575.
- Charette SJ, Brochu F, Boyle B, Filion G, Tanaka KH & Derome N (2012) Draft genome sequence of the virulent strain 01-B526 of the fish pathogen *Aeromonas salmonicida*. *J Bacteriol* **194**: 722–723.
- Chen P-L, Wu C-J, Tsai P-J, *et al.* (2014) Virulence diversity among bacteremic *Aeromonas* isolates: *ex vivo*, animal, and clinical evidences. *PLoS One* **9**: 1–7.
- Comeau AM, Tremblay D, Moineau S, Rattei T, Kushkina AI, Tovkach FI, Krisch HM & Ackermann H-W (2012) Phage morphology recapitulates phylogeny: the comparative genomics of a new group of myoviruses. *PLoS One* **7**: e40102.
- Dacanay A, Knickle L, Solanky KS, Boyd JM, Walter JA, Brown LL, Johnson SC & Reith M (2006) Contribution of the type III secretion system (TTSS) to virulence of *Aeromonas salmonicida* subsp. *salmonicida*. *Microbiology* **152**: 1847–1856.

- Daher RK, Filion G, Tan SGE, Dallaire-Dufresne S, Paquet VE & Charette SJ (2011) Alteration of virulence factors and rearrangement of pAsa5 plasmid caused by the growth of *Aeromonas salmonicida* in stressful conditions. *Vet Microbiol* **152**: 353–360.
- Dalsgaard I, Nielsen B & Larsen JL (1994) Characterization of *Aeromonas salmonicida* subsp. *salmonicida*: a comparative study of strains of different geographic origin. *J Appl Bacteriol* **77**: 21–30.
- Darmon E & Leach DRF (2014) Bacterial genome instability. *Microbiol Mol Biol Rev* **78**: 1–39.
- Dautremepuits C, Fortier M, Croisetiere S, Belhumeur P & Fournier M (2006) Modulation of juvenile brook trout (*Salvelinus fontinalis*) cellular immune system after *Aeromonas salmonicida* challenge. *Vet Immunol Immunopathol* **110**: 27–36.
- Diamanka A, Loch TP, Cipriano RC & Faisal M (2013) Polyphasic characterization of *Aeromonas salmonicida* isolates recovered from salmonid and non-salmonid fish. *J Fish Dis* **36**: 949–963.
- Dobrindt U, Hochhut B, Hentschel U & Hacker J (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2**: 414–424.
- Domde P, Kapley A & Purohit HJ (2007) Impact of bioaugmentation with a consortium of bacteria on the remediation of wastewater-containing hydrocarbons. *Environ Sci Pollut Res Int* **14**: 7–11.
- Dworkin M, Falkow S, Rosenberg E, Schleifer K-H & Stackebrandt E (2006) *The Prokaryotes. Proteobacteria: Gamma Subclass*. Third edit. Springer Science & Business Media <https://books.google.com/books?id=sQ1To-cNyMkC&pgis=1> (Accessed March 30, 2015).
- Ebanks RO, Knickle LC, Goguen M, Boyd JM, Pinto DM, Reith M & Ross NW (2006) Expression of and secretion through the *Aeromonas salmonicida* type III secretion system. *Microbiology* **152**: 1275–1286.
- Emond-Rheault J-G, Vincent AT, Trudel M V., et al. (2015) Variants of a genomic island in *Aeromonas salmonicida* subsp. *salmonicida* link isolates with their geographical origins. *Vet Microbiol* **175**: 68–76.
- Van der Ende a., Hopman CTP & Dankert J (2000) Multiple mechanisms of phase variation of PorA in *Neisseria meningitidis*. *Infect Immun* **68**: 6685–6690.
- Fehr D, Casanova C, Liverman A, Blazkova H, Orth K, Dobbelaere D, Frey J & Burr SE (2006) AopP, a type III effector protein of *Aeromonas salmonicida*, inhibits the NF-kappaB signalling pathway. *Microbiology* **152**: 2809–2818.
- Figueras MJ (2005) Clinical relevance of *Aeromonas* sM503. *Rev Med Microbiol* **16**: 145–153.

- Flèche P Le, Hauck Y, Onteniente L, Prieur A, Denoeud F, Ramisse V, Sylvestre P, Benson G, Ramisse F & Vergnaud G (2001) A tandem repeats database for bacterial genomes : application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. *BMC Microbiol* **1**.
- Flèche P Le, Fabre M, Denoeud F, Koeck J & Vergnaud G (2002) High resolution , on-line identification of strains from the *Mycobacterium tuberculosis* complex based on tandem repeat typing. *BMC Microbiol* **2**: 1–12.
- Fogg PCM, Colloms S, Rosser S, Stark M & Smith MCM (2014) New applications for phage integrases. *J Mol Biol* **426**: 2703–2716.
- Food and Agriculture Organization of the United Nations (2014) *The State of World Fisheries and Aquaculture 2014*.
- Fortier L-C & Sekulovic O (2013) Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence* **4**: 354–365.
- Frost LS, Leplae R, Summers AO & Toussaint A (2005) Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3**: 722–732.
- Galardini M, Biondi EG, Bazzicalupo M & Mengoni A (2011) CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* **6**: 1–5.
- Gao X, Jian J, Li W-J, Yang Y-C, Shen X-W, Sun Z-R, Wu Q & Chen G-Q (2013) Genomic study of polyhydroxyalkanoates producing *Aeromonas hydrophila* 4AK4. *Appl Microbiol Biotechnol* **97**: 9099–9109.
- García JA, Larsen JL, Dalsgaard I & Pedersen K (2000) Pulsed-field gel electrophoresis analysis of *Aeromonas salmonicida* ssp. *salmonicida*. *FEMS Microbiol Lett* **190**: 163–166.
- Garcia-Russell N, Elrod B & Dominguez K (2009) Stress-induced prophage DNA replication in *Salmonella enterica* serovar Typhimurium. *Infect Genet Evol* **9**: 889–895.
- Gemayel R, Vinces MD, Legendre M & Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu Rev Genet* **44**: 445–477.
- Gustafson CE, Chu S & Trust TJ (1994) Mutagenesis of the paracrystalline surface protein array of *Aeromonas salmonicida* by endogenous insertion elements. *J Mol Biol* **237**: 452–463.
- Hacker J & Carniel E (2001) Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Rep* **2**: 376–381.
- Han JE, Kim H, Shin P, Jun W, Chai Y & Park C (2013) Draft genome sequence of *Aeromonas salmonicida* subsp . *achromogenes* AS03 , an atypical strain isolated

- from Crucian Carp (*Carassius carassius*) in the Republic of Korea. *Genome Announc* **1**: 2007–2008.
- Hanninen M & Hirverikoski V (1997) Microbiology molecular and phenotypic methods for the characterization of atypical *Aeromonas salmonicida*. *Vet Microbiol* **56**: 147–158.
- Hänninen ML, Ridell J & Hirvelä-Koski V (1995) Phenotypic and molecular characteristics of *Aeromonas salmonicida* subsp. *salmonicida* isolated in southern and northern Finland. *J Appl Bacteriol* **79**: 12–21.
- Hedrick RP (1996) Movements of pathogens with the international trade of live fish : problems and solutions. *Rev Sci Tech* **15**: 523–531.
- Herschleb J, Ananiev G & Schwartz DC (2007) Pulsed-field gel electrophoresis. *Nat Protoc* **2**: 677–684.
- Ho Sui SJ, Fedynak A, Hsiao WWL, Langille MGI & Brinkman FSL (2009) The association of virulence factors with genomic islands. *PLoS One* **4**: 1–11.
- Hossain MJ, Waldbieser GC, Sun D, *et al.* (2013) Implication of lateral genetic transfer in the emergence of *Aeromonas hydrophila* isolates of epidemic outbreaks in channel catfish. *PLoS One* **8**: e80943.
- Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW & Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.
- Jackson RW, Vinatzer B, Arnold DL, Dorus S & Murillo J (2011) The influence of the accessory genome on bacterial pathogen evolution. *Mob Genet Elements* **1**: 55–65.
- Janda JM & Abbott SL (2010) The genus *Aeromonas*: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev* **23**: 35–73.
- Janulczyk R, Masignani V, Maione D, Tettelin H, Grandi G & Telford JL (2010) Simple sequence repeats and genome plasticity in *Streptococcus agalactiae*. *J Bacteriol* **192**: 3990–4000.
- Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW & Crook DW (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* **33**: 376–393.
- Jun JW, Kim HJ, Yun SK, Chai JY & Park SC (2014) *Genomic structure of the Aeromonas bacteriophage pAh6-C and its comparative genomic analysis.*
- Kassai-Jäger E, Ortutay C, Tóth G, Vellai T & Gáspári Z (2008) Distribution and evolution of short tandem repeats in closely related bacterial genomes. *Gene* **410**: 18–25.
- Kim JH, Son JS, Choi YJ, Choresca CH, Shin SP, Han JE, Jun JW & Park SC (2012) Complete genome sequence and characterization of a broad-host range T4-like

- bacteriophage phiAS5 infecting *Aeromonas salmonicida* subsp. *salmonicida*. *Vet Microbiol* **157**: 164–171.
- Kim JH, Son JS, Choresca CH, Shin SP, Han JE, Jun JW, Kang D-H, Oh C, Heo S-J & Park SC (2012) Complete genome sequence of bacteriophage phiAS7, a T7-like virus that infects *Aeromonas salmonicida* subsp. *salmonicida*. *J Virol* **86**: 2894–2895.
- Kim JH, Son JS, Choi YJ, Choresca CH, Shin SP, Han JE, Jun JW & Park SC (2012) Complete genomic sequence of a T4-like bacteriophage, phiAS4, infecting *Aeromonas salmonicida* subsp. *salmonicida*. *Arch Virol* **157**: 391–395.
- Lafaille A (2010) *Rapport des activités en Ichtyopathologie 1 er avril 2009 au 31 mars 2010*. Service de diagnostic en Ichtyopathologie, Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe.
- Lafaille A (2011) *Rapport des activités en Ichtyopathologie 1 er avril 2010 au 31 mars 2011*. Service de diagnostic en Ichtyopathologie, Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe.
- Lafaille A (2012) *Rapport des activités en Ichtyopathologie 1 er avril 2011 au 31 décembre 2011*. Service de diagnostic en Ichtyopathologie, Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe.
- Lafferty KD, Harvell CD, Conrad JM, Friedman CS, Kent ML, Kuris AM, Powell EN, Rondeau D & Saksida SM (2015) Infectious diseases affect marine fisheries and aquaculture economics. *Ann Rev Mar Sci* **7**: 1–26.
- Li Y, Liu Y, Zhou Z, Huang H, Ren Y, Zhang Y, Li G, Zhou Z & Wang L (2011) Complete genome sequence of *Aeromonas veronii* strain B565. *J Bacteriol* **193**: 3389–3390.
- Lima WC, Paquola ACM, Varani AM, Van Sluys M-A & Menck CFM (2008) Laterally transferred genomic islands in Xanthomonadales related to pathogenicity and primary metabolism. *FEMS Microbiol Lett* **281**: 87–97.
- Lin AW, Usera MA, Barrett TJ, Goldsby RA, Lin AW, Usera MA & Barrett TJ (1996) Application of random amplified polymorphic DNA analysis to differentiate strains of *Salmonella enteritidis*. *J Clin Microbiol* **34**: 870–876.
- Lindstedt B-A, Vardund T & Kapperud G (2004) Multiple-locus variable-number tandem-repeats analysis of *Escherichia coli* O157 using PCR multiplexing and multi-colored capillary electrophoresis. *J Microbiol Methods* **58**: 213–222.
- Lowe TM & Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**: 955–964.
- Lund V, Jenssen LM & Wesmajervi MS (2002) Assessment of genetic variability and relatedness among atypical *Aeromonas salmonicida* from marine fishes, using AFLP-fingerprinting. *Dis Aquat Organ* **50**: 119–126.

- Madico G & Akopyants NS (1995) Arbitrarily primed PCR DNA fingerprinting of *Escherichia coli* O157 : H7 strains by using templates from boiled cultures . *J Clin Microbiol* **33**: 1534–1536.
- Mahenthiralingam E, Campbell ME, Foster J, Lam JS, Mahenthiralingam E & Campbell ME (1996) Random amplified polymorphic DNA typing of *Pseudomonas aeruginosa* isolates recovered from patients with cystic fibrosis. *J Clin Microbiol* **34**: 1129–1135.
- Mahillon J & Chandler M (1998) Insertion sequences. *Microbiol Mol Biol Rev* **62**: 725–774.
- Makarova KS, Wolf YI, Snir S & Koonin E V (2011) Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J Bacteriol* **193**: 6039–6056.
- MAPAQ (2012) Aquaculture. 10 octobre 2014.
- MAPAQ (2013) Choix de l'espèce. 10 octobre 2014.
- Margulies M, Egholm M, Altman WE, *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–380.
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1464427&tool=pmcentrez&rendertype=abstract> (Accessed April 28, 2014).
- Martel B & Moineau S (2014) CRISPR-Cas: an efficient tool for genome engineering of virulent bacteriophages. *Nucleic Acids Res* **42**: 9504–9513.
- Martin-Carnahan A & Joseph SW (2005) Aeromonadales. *Bergey's Manual® of Systematic Bacteriology - Volume 2*, (Garrity GM, ed), pp. 556–587. Springer, New-York.
- McCormick W a, Stevenson RM & MacInnes JI (1990) Restriction endonuclease fingerprinting analysis of Canadian isolates of *Aeromonas salmonicida*. *Can J Microbiol* **36**: 24–32.
- Merino S, Gavín R, Vilches S, Shaw JG & Tomás JM (2003) A colonization factor (production of lateral flagella) of mesophilic *Aeromonas* spp. is inactive in *Aeromonas salmonicida* strains. *Appl Environ Microbiol* **69**: 663–667.
- Metruccio MME, Pigozzi E, Roncarati D, Berlanda Scorza F, Norais N, Hill S a, Scarlato V & Delany I (2009) A novel phase variation mechanism in the meningococcus driven by a ligand-responsive repressor and differential spacing of distal promoter elements. *PLoS Pathog* **5**: e1000710.
- Miyata M, Aoki T, Inglis V, Yoshida T & Endo M (1995) RAPD analysis of *Aeromonas salmonicida* and *Aeromonas hydrophila* . *J Appl Bacteriol* **79**: 181–185.
- Morgan UM, Constantine CC, Greene WK & Thompson RCA (1993) RAPD (random amplified polymorphic DNA) analysis of *Giardia* DNA and correlation with isoenzyme data. *Trans R Soc Trop Med Hyg* **87**: 702–705.

- Morin R (2010) L'aquicole. *Le Bull l'Association des aquaculteurs du Québec* **15**.
- Moxon R, Bayliss C & Hood D (2006) Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu Rev Genet* **40**: 307–333.
- Mrázek J (2006) Analysis of distribution indicates diverse functions of simple sequence repeats in *Mycoplasma* genomes. *Mol Biol Evol* **23**: 1370–1385.
- Mrázek J, Guo X & Shah A (2007) Simple sequence repeats in prokaryotic genomes. *PNAS* **104**: 8472–8477.
- O'hici B, Olivier G & Powell R (2000) Genetic diversity of the fish pathogen *Aeromonas salmonicida* demonstrated by random amplified polymorphic DNA and pulsed-field gel electrophoresis analyses. *Dis Aquat Organ* **39**: 109–119.
- Onteniente L, Brisse S, Tassios PT & Vergnaud G (2003) Evaluation of the polymorphisms associated with tandem repeats for *Pseudomonas aeruginosa* strain typing. *J Clin Microbiol* **41**: 4991–4997.
- Ouellet MM, Leduc A, Nadeau C, Barbeau J, Charette SJ, Christaki E & General AP (2015) *Pseudomonas aeruginosa* isolates from dental unit waterlines can be divided in two distinct groups , including one displaying phenotypes similar to isolates from cystic fibrosis patients. *Front Microbiol* **5**: 1–11.
- Pearson WR & Lipman DJ (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* **85**: 2444–2448.
- Petrov VM, Ratnayaka S, Nolan JM, Miller ES & Karam JD (2010) Genomes of the T4-related bacteriophages as windows on microbial genome evolution. *Virology* **7**: 292.
- Plunkett III G, Rose DJ, Durfee TJ & Blattner FR (1999) Sequence of shiga toxin 2 phage 933W from *Escherichia coli* O157 : H7 : shiga toxin as phage late-gene product. *Am Soc Microbiol* **181**: 1767–1778.
- Pourcel C, Vidgop Y, Ramisse F, Vergnaud G & Tram C (2003) Characterization of a tandem repeat polymorphism in *Legionella pneumophila* and its use for genotyping. *J Clin Microbiol* **41**: 1819–1826.
- Pourcel C, Neubauer H, Ramisse F & Vergnaud G (2004) Tandem repeats analysis for the high resolution phylogenetic analysis of *Yersinia pestis*. *BMC Microbiol* **4**: 1–9.
- Power PM, Sweetman W a, Gallacher NJ, Woodhall MR, Kumar G a, Moxon ER & Hood DW (2009) Simple sequence repeats in *Haemophilus influenzae*. *Infect Genet Evol* **9**: 216–228.
- Pridgeon JW, Zhang D & Zhang L (2014a) Complete genome sequence of a moderately virulent *Aeromonas hydrophila* strain , pc104A , isolated from soil of a catfish pond in West Alabama. *Genome Announc* **2**: 94–95.

- Pridgeon JW, Zhang D & Zhang L (2014b) Complete genome sequence of the highly virulent *Aeromonas hydrophila* AL09-71 isolated from diseased channel catfish in West Alabama. *Genome Announc* **2**: 9–10.
- Que F, Wu S & Huang R (2013) *Salmonella* pathogenicity island 1(SPI-1) at work. *Curr Microbiol* **66**: 582–587.
- Ramisse V, Houssu P, Hernandez E, Denoeud F, Lisanti O, Cavallo J, Vergnaud G & Mande S (2004) Variable number of tandem repeats in *Salmonella enterica* subsp . *enterica* for typing purposes. *J Clin Microbiol* **42**: 5722–5730.
- Ravatt R, Studer S, Springael D, Zehnder AJ & van der Meer JR (1998) Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from *Pseudomonas* sp. strain B13. *J Bacteriol* **180**: 4360–4369.
- Reith ME, Singh RK, Curtis B, *et al.* (2008) The genome of *Aeromonas salmonicida* subsp. *salmonicida* A449: insights into the evolution of a fish pathogen. *BMC Genomics* **9**: 427.
- Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R & Miller W (2000) PipMaker--a web server for aligning two genomic DNA sequences. *Genome Res* **10**: 577–586.
- Seshadri R, Joseph SW, Chopra AK, *et al.* (2006) Genome sequence of *Aeromonas hydrophila* ATCC 7966T: jack of all trades. *J Bacteriol* **188**: 8272–8282.
- Shen C, Liu Y & Lu C (2012) Complete genome sequence of *Aeromonas hydrophila* phage CC2. *J Virol* **86**: 10900.
- Sherwood LM, Willey JM & Woolverton CJ (2010) *Microbiologie*. De Boeck.
- Shin EK, Seo YS, Han JH & Hahn TW (2007) Diversity of swine *Bordetella bronchiseptica* isolates evaluated by RAPD analysis and PFGE. *J Vet Sci* **8**: 65–73.
- Studer N, Frey J & Vanden Bergh P (2013) Clustering subspecies of *Aeromonas salmonicida* using IS630 typing. *BMC Microbiol* **13**: 13–36.
- Sullivan JT, Trzebiatowski JR, Cruickshank RW, *et al.* (2002) Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J Bacteriol* **184**: 3086–3095.
- Sullivan MJ, Petty NK & Beatson S a (2011) Easyfig: a genome comparison visualizer. *Bioinformatics* **27**: 1009–1010.
- Sung JY, Koo SH, Cho HH & Kwon KC (2012) AbaR7, a genomic resistance island found in multidrug-resistant *Acinetobacter baumannii* isolates in Daejeon, Korea. *Ann Lab Med* **32**: 324–330.

- Tanaka KH, Dallaire-Dufresne S, Daher RK, Frenette M & Charette SJ (2012) An insertion sequence-dependent plasmid rearrangement in *Aeromonas salmonicida* causes the loss of the type three secretion system. *PLoS One* **7**: e33725.
- Tenover FC, Arbeit RD, Goering R V, Mickelsen PA, Murray BE, Persing DH & Swaminathan B (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis : criteria for bacterial strain typing . *J Clin Microbiol* **33**: 2233–2239.
- Tkedar HC, Waldbieser GC, Karsi A, *et al.* (2013) Complete genome sequence of a channel catfish epidemic isolate, *Aeromonas hydrophila* strain ML09-119. *Genome Announc* **1**: 1–2.
- Treangen TJ, Abraham A-L, Touchon M & Rocha EPC (2009) Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS Microbiol Rev* **33**: 539–571.
- Tritt A, Eisen J a., Facciotti MT & Darling AE (2012) An integrated pipeline for de novo assembly of microbial genomes. *PLoS One* **7**: e42304.
- Trudel M V, Tanaka KH, Filion G, Daher RK, Frenette M & Charette SJ (2013) Insertion sequence AS5 (ISAS5) is involved in the genomic plasticity of *Aeromonas salmonicida*. *Mob Genet Elements* **5**: 1–7.
- Uhland FC (2004) Maladies communes dans les piscicultures du Québec et interventions vétérinaires. *Le médecin vétérinaire du Québec* **34**: 268–274.
- Uhland FC, Péloquin R & Morin R (1997) *SANTÉ*. Gouvernement du Québec.
- Waldor MK & Mekalanos JJ (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**: 1910–1914.
- Wiklund T & Dalsgaard I (1998) Occurrence and significance of atypical *Aeromonas salmonicida* in non-salmonid and salmonid fish species: a review. *Dis Aquat Organ* **32**: 49–69.
- Williams JG, Kubelik a R, Livak KJ, Rafalski J a & Tingey S V (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* **18**: 6531–6535.
- Yang J, Wang J, Chen L, Yu J, Dong J, Yao Z-J, Shen Y, Jin Q & Chen R (2003) Identification and characterization of simple sequence repeats in the genomes of *Shigella* species. *Gene* **322**: 85–92.
- Zhao S, Poulin L, Rodriguez-r LM, *et al.* (2012) Development of a variable number of tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv . *oryzicola*. *Phytopathology* **3**: 948–956.
- Zhou K, Aertsen A & Michiels CW (2014) The role of variable DNA tandem repeats in bacterial adaptation. *FEMS Microbiol Rev* **38**: 119–141.

Annexe A : Souches d'*A. salmonicida* sous-espèce *salmonicida* utilisées dans cette étude

Num	Isolats	Origine	Source et/ou référence	Type de GEI	Prophage 3
1	01-B522	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
2	01-B526	Omble de fontaine (QC,Canada)	FMVUM (Dautremepuits <i>et al.</i> , 2006)	AsaGEI1a	Non
3	01-B516	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
4	HER1108	Truite (Denmark)	FHRC	AsaGEI1b	Non
5	HER1104	Truite (France)	FHRC	sans	Non
6	HER1085	Truite (Norvège)	FHRC	AsaGEI1b	Non
7	HER1084	Truite (France)	FHRC	sans	Non
8	A449	Truite brune (France)	Dacanay <i>et al.</i> , 2006	sans	Non
9	07-9324	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
10	07-7817	* (QC,Canada)	FMVUM	AsaGEI1a	Non
11	07-7346	Saumon de l'Atlantique (QC,Canada)	FMVUM	AsaGEI2a	Non
12	07-5957	Saumon de l'Atlantique (QC,Canada)	FMVUM	AsaGEI2a	Non
13	08-2647	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
14	09-0167	Saumon de l'Atlantique (QC,Canada)	FMVUM	AsaGEI2a	Non
15	07-7287	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
16	08-2783	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
17	08-4188	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
18	5093-3	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
19	5093-4	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
20	5490-1	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
21	5490-2	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
22	5490-4	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
23	5490-5	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
24	5704-1	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
25	5704-2	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
26	5704-4	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
27	5704-5	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
28	5704-6	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
29	5704-3	Omble de fontaine (QC,Canada)	MAPAQ	AsaGEI2a	Oui
30	m17524-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
31	m14349-09	Saumon de l'Atlantique (QC,Canada)	FMVUM	AsaGEI1a	Non
32	m23281-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
33	m23067-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
34	m19438-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
35	m16583-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
36	m14231-09	Saumon de l'Atlantique (QC,Canada)	FMVUM	AsaGEI2a	Oui
37	m11743-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui

38	m11431-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
39	m10419-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
40	m9906-09	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
41	m9954-10	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
42	m8029-10	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
43	m11603-10	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
44	m6363-10	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
45	m9221-10	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
46	2009-178 K9	Saumon de l'Atlantique (NB,Canada)	FOC	AsaGEI2a	Oui
47	2009-157 K5	Omble de fontaine (NB,Canada)	FOC	AsaGEI2a	Oui
48	2010-47 K18	Omble de fontaine (NB,Canada)	FOC	AsaGEI2a	Oui
49	2004-05 MF26	* (NB,Canada)	FOC	AsaGEI2a	Oui
50	2004-68 K52	Saumon de l'Atlantique (NE,Canada)	FOC	AsaGEI2a	Oui
51	2009-195 K29	Omble de fontaine (NB,Canada)	FOC	AsaGEI2a	Oui
52	2005-70	* (NB,Canada)	FOC	AsaGEI2a	Oui
53	2009-144 K3	Omble de fontaine (NB,Canada)	FOC	AsaGEI2a	Oui
54	2005-175 K2	Omble de fontaine (NB,Canada)	FOC	AsaGEI2a	Oui
55	2004-208	* (NB,Canada)	FOC	AsaGEI2a	Oui
56	M10935-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
57	M15448-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
58	M16474-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
59	M19878-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
60	M11500-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
61	M16486-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
62	M13460-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
63	M13729-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
64	M14481-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
65	M15879-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
66	M17739-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
67	M13732-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
68	M17053-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
69	M15878-11	Truite arc-en-ciel (QC, Canada)	FMVUM	AsaGEI2a	Oui
70	M13182-11	Saumon de l'Atlantique (QC,Canada)	FMVUM	AsaGEI2a	Oui
71	M17735-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
72	M15576-11	Truite brune (QC, Canada)	FMVUM	AsaGEI2a	Non
73	M15469-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
74	M22710-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
75	M13764-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI2a	Oui
76	M18076-11	Lompe (QC, Canada)	FMVUM	sans	Non
77	M23911-11	Omble de fontaine (QC,Canada)	FMVUM	AsaGEI1a	Non
78	JF2267	Omble chevalier (Suisse)	Studer <i>et al.</i> , 2013	sans	Non
79	RS 1458	Truite arc-en-ciel (Ganaraska River, ON, Canada)	RS, University of Guelph	AsaGEI2a	Oui
80	RS 1706	Saumon chinook (Credit River, ON, Canada)	RS, University of Guelph	AsaGEI2a	Oui
81	RS 1835	Saumon coho (Credit River, ON, Canada)	RS, University of Guelph	AsaGEI2a	Oui
82	RS 530	Saumon de l'Atlantique (Pays de Galles, Royaume-Uni)	RS, University of Guelph	AsaGEI1b	Non

	(ATCC33658)				
83	RS 534	* (France)	RS, University of Guelph	sans	Non
84	RS 1752	Perche-soleil (Mississippi Lake, ON, Canada)	RS, University of Guelph	Indéterminé	Non
85	RS 1705	Ombre de fontaine (Tarentorus Fish Culture Station, Sault Ste-Marie, ON, Canada)	RS, University of Guelph	AsaGEI1a	Non
86	RS 1744	Saumon coho (Credit River, ON, Canada)	RS, University of Guelph	AsaGEI2a	Oui
87	M12357-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI1a	Non
88	M21375-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
89	M16237-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
90	M12976-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
91	M22895-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
92	M10745-12	Truite brune (QC, Canada)	FMVUM	AsaGEI1a	Non
93	M9754-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
94	M17930-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI1a	Non
95	M12418-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
96	M21368-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
97	M13050-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
98	M24783-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
99	M13566-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI1a	Non
100	M14404-12	Saumon de l'Atlantique (QC, Canada)	FMVUM	AsaGEI2a	Oui
101	M16671-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI2a	Oui
102	M16042-12	Ombre de fontaine (QC, Canada)	FMVUM	AsaGEI1a	Non
103	JF2506	Saumon de l'Atlantique (Norvège)	Studer <i>et al.</i> , 2013	sans	Non
104	JF2507	Saumon de l'Atlantique (Ecosse, Royaume-Uni)	Studer <i>et al.</i> , 2013	sans	Non
105	JF2510	Saumon de l'Atlantique (Norvège)	Studer <i>et al.</i> , 2013	sans	Non
106	JF3224	Truite brune (Suisse)	Studer <i>et al.</i> , 2013	AsaGEI2b	Non
107	JF3327	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	AsaGEI1b	Non
108	JF3496	Truite brune (Suisse)	Studer <i>et al.</i> , 2013	sans	Non
109	JF3517	Turbot (Norvège)	Studer <i>et al.</i> , 2013	sans	Non
110	JF3518	Turbot (Norvège)	Studer <i>et al.</i> , 2013	sans	Non
111	JF3844	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	sans	Non
112	JF2869	Ombre chevalier (pays européen inconnu)	Studer <i>et al.</i> , 2013	AsaGEI1b	Non
113	JF3223	Poisson blanc (Suisse)	Studer <i>et al.</i> , 2013	AsaGEI1b	Non
114	JF3519	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	sans	Non
115	JF3791	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	sans	Non
116	JF4111	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	AsaGEI1b	Non
117	JF4112	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	AsaGEI1b	Non
118	JF4113	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	AsaGEI1b	Non
119	JF4114	Ombre chevalier (Suisse)	Studer <i>et al.</i> , 2013	AsaGEI1b	Non
120	As4	Saumon chinook (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI1a	Oui
121	As9	Saumon chinook (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
122	As11	Saumon chinook (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
123	As13	Saumon chinook (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
124	As15	Saumon chinook (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI1a	Oui
125	As18	Truite arc-en-ciel (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI1a	Oui

126	As25	Saumon coho (Plate River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
127	As33	Saumon coho (Plate River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
128	As36	Saumon de l'Atlantique (Plate River State Fish Hatchery, Michigan, USA)	Diamanka <i>et al.</i> , 2013	Indéterminé	Oui
129	As37	Perche jaune (Newaygo County, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
130	As7	Saumon chinook (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
131	As19	Saumon chinook (Little Manistee River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI1a	Non
132	As28	Saumon coho (Plate River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI1a	Non
133	As29	Saumon coho (Plate River Weir, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
134	As40	Saumon de l'Atlantique (St Mary's River, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
135	As42	Omble de fontaine (Alcona County, Michigan, USA)	Diamanka <i>et al.</i> , 2013	sans	Non
136	As44	Omble de fontaine (Gogebic County, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
137	As48	Saumon de l'Atlantique (St Mary's River, Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
138	050906-1 4Les	*(Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui
139	060921-1 13As	*(Michigan, USA)	Diamanka <i>et al.</i> , 2013	AsaGEI2a	Oui

*: Informations non disponible ou non traçable

FMVUM: Faculté de médecine vétérinaire, Université de Montréal (Québec, Canada)

FHRC: Félix d'Hérelle Reference Center, Département de biochimie, de microbiologie et de bio-informatique, Université Laval (Québec, Canada).

MAPAQ: Ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec

FOC: Aquatic Animal Health Department, Fisheries and Oceans Canada (NB, Canada)

NB: Nouveau-Brunswick

NE: Nouvelle-Écosse

QC: Québec

RS : Roselynn M.W. Stevenson

Annexe B : Amorces utilisées dans cette étude

Nom	Séquence (5' → 3')	Référence
Détection d' <i>AsaGEI1a</i>		
MT21-Ins1-F1	AAATACCTTGTGGCGGGCAAAGTC	(Emond-Rheault <i>et al.</i> , 2015)
MT23-Ins1-R1	ATTTGACATCGCACGGAAACCTCG	(Emond-Rheault <i>et al.</i> , 2015)
MT1-orf11-F1	GCGCTGGCCAACATATTCATCAGT	(Emond-Rheault <i>et al.</i> , 2015)
MT4-orf11-R2	AGCTGTTTAGCCACCCAGGTATGT	(Emond-Rheault <i>et al.</i> , 2015)
MT5-orf19-F1	TGAACTTCGTCACCGACGACAACA	(Emond-Rheault <i>et al.</i> , 2015)
MT7-orf19-R1	TCAATCAGAATTGCCATTGGCGGG	(Emond-Rheault <i>et al.</i> , 2015)
MT10-orf33-F2	ATGTGCTGTTGTTGGCCGTACTION	(Emond-Rheault <i>et al.</i> , 2015)
MT12-orf33-R2	GTTTGGTCTGAAATGTGAGCGCGA	(Emond-Rheault <i>et al.</i> , 2015)
MT13-orf45-F2	ACATGGCTTGGGATTAACGATGC	Cette étude
MT15orf45-R2	ACAGGGATGAAGGCACAGTTGTA	Cette étude
MT2-orf54-F2	TGCCTATGTGCAACTCTCCCACAA	(Emond-Rheault <i>et al.</i> , 2015)
MT4-orf54-R2	CCTTGCACTTCGATTGAGGC	(Emond-Rheault <i>et al.</i> , 2015)
MT5-orf55-F1	AAGAACAATCGCCCAAGCGCAAG	Cette étude
MT7-orf55-R1	AGTTATAGGCTCGCTCACAGGCTA	Cette étude
MT9-orf57-F1	TGCGTCATGTTTCATGCTGGATGTG	(Emond-Rheault <i>et al.</i> , 2015)
MT11-orf57-R1	ATCCGGTGAGTTGGTGTCCATGAT	(Emond-Rheault <i>et al.</i> , 2015)
MT13-orf58-F1	GTCCATGCACATCGACAATGCCAA	Cette étude
MT15orf58-R1	TCTTGACCTCGGCCATGACTION	Cette étude
MT17-orf64-F1	CCAAGAACAAGGGTCAAACCGTCA	Cette étude
MT20-orf64-R2	ACGAAGCGGAACCTTCCACAGAT	Cette étude
MT22-orf66-F2	ATCCAGCTGCTGCGTAATGATTGG	Cette étude
MT24-orf66-R2	GCTTCACCAGCAACCCGTAATAGA	Cette étude
MT18-ORF67-F2	CTGACGCATCTTGCTGTCACCATT	(Emond-Rheault <i>et al.</i> , 2015)
MT19-ORF67-R1	ATCGGTGCATTGGTGTACCCGTAA	(Emond-Rheault <i>et al.</i> , 2015)
MT30-orf72-F2	AAGACGGTCTGGAACCTGGCTGAT	Cette étude
MT31-orf72-R1	CGTTGCCATTGACATAGCCACCAA	Cette étude
MT30-orf78-F2	AAGACGGTCTGGAACCTGGCTGAT	(Emond-Rheault <i>et al.</i> , 2015)
MT31-orf78-R1	CGTTGCCATTGACATAGCCACCAA	(Emond-Rheault <i>et al.</i> , 2015)
MT25-ins2-F1	GTTCCGAGGTGGAATGTTGGTTT	(Emond-Rheault <i>et al.</i> , 2015)
MT27-ins2-R1	AGCGTTCACCCATATCTCCACT	(Emond-Rheault <i>et al.</i> , 2015)
Site d'insertion d' <i>AsaGEI2a</i>		
JG-ins1_Pro1-FWD2	CGTGCTCACCTTCTGGTCCATGTT	(Emond-Rheault <i>et al.</i> , 2015)
JG-ins1_Pro1-REV2	CTGTCGCCACTTGGCCAAAACAA	(Emond-Rheault <i>et al.</i> , 2015)
JG-Pro3_Asa3a-FWD	GTGCCAGCAGGAGAGATATTGCCT	(Emond-Rheault <i>et al.</i> , 2015)
JG-Pro3_Asa3a-REV	GGAAGTAGCCGCATCAGCATCAA	(Emond-Rheault <i>et al.</i> , 2015)
JG-ins1_AsaGEI2a-FWD2	CTCAACCCGCTGACCATCAACAAA	(Emond-Rheault <i>et al.</i> , 2015)
JG-ins1_AsaGEI2a-REV2	GGGGCAAGTGCAGGTATATGACAG	(Emond-Rheault <i>et al.</i> , 2015)
JG-orf21_AsaGEI2a-FWD	AGTTCGTCGCGCATTCTTCAA	(Emond-Rheault <i>et al.</i> , 2015)
JG-orf21_AsaGEI2a-REV	GGGTGGACAGAACAGGAATCGTT	(Emond-Rheault <i>et al.</i> , 2015)
JG-ins2_AsaGEI2a-FWD	CGCCGGTACGGAGGCTGTAAAAT	(Emond-Rheault <i>et al.</i> , 2015)
JG-ins2_AsaGEI2a-REV	CCTTTTCTGCCGATGTCAGGCTTT	(Emond-Rheault <i>et al.</i> , 2015)
Contrôle positif		
TapA-F	ACATGAAGAAGCAATCAGGC	(Ebanks <i>et al.</i> , 2006)
TapA-R	AGAGGTCATGCGTTAGCAG	(Ebanks <i>et al.</i> , 2006)
Région spécifique chez <i>AsaGEI1a</i>		
AsaGEI1a_orf5_For	CGCAGTAGTGATAATGCCGGTCAA	Cette étude
AsaGEI1a_orf5_Rev	CCGCATCGAGAAGAACTGAGCAA	Cette étude
AsaGEI1a_orf11_For	CAAAACACCAGCATGCAGCCTTGG	Cette étude
AsaGEI1a_orf11_Rev	CGGGAAAAGACCTGCTTGTGTCC	Cette étude
AsaGEI1a_orf39_For	CGAAATCGCATTCAATCAGACCC	Cette étude
AsaGEI1a_orf39_Rev	GCGCTTACGTTCGAAAATCGACGG	Cette étude
Région spécifique chez <i>AsaGEI1b</i>		
AsaGEI1b_orf5_For	GCGAAATTGATCAGCGCGGTTGA	Cette étude
AsaGEI1b_orf5_Rev	GCCGAACGCAAGTAACTATCGACG	Cette étude

AsaGEI1b_orf11_For	CGTGATATGGGCGGCAAAGATCTT	Cette étude
AsaGEI1b_orf11_Rev	GTTGCGGGAGGTAACCCAGATAAT	Cette étude
AsaGEI1b_orf39_For	GTGGTTCGCCGAGATTGAGCCATT	Cette étude
AsaGEI1b_orf39_Rev	GGCGGATTCTCTCGTACCAATTCA	Cette étude
Région spécifique chez <i>AsaGEI2a</i>		
AsaGEI2a_orf5_For	CCCTTTCTTACCGCAATGTTGGCA	Cette étude
AsaGEI2a_orf5_Rev	GCATAACGTGCGACATGGGATGTT	Cette étude
AsaGEI2a_orf10.2_For	GCGTGCTCATGGAGGTTATTGAAC	Cette étude
AsaGEI2a_orf10.2_Rev	GCTCGACATGGTTTCGCTGTCAGA	Cette étude
AsaGEI2a_orf39_For	GTCCTTGATAGGTATGCGGCTCA	Cette étude
AsaGEI2a_orf39_Rev	GTCCTTGATAGGTATGCGGCTCA	Cette étude
Région spécifique chez <i>AsaGEI2b</i>		
AsaGEI2b_orf5.2_For	GACTTATGGGTAACCAGCTCCCAC	Cette étude
AsaGEI2b_orf5.2_Rev	CGGCATTAACAGGAGGCAGAGATT	Cette étude
AsaGEI2b_orf10.3_For	CGGCACTTCACTGGTAATCCGCTT	Cette étude
AsaGEI2b_orf10.3_Rev	GGCGAACGATAGCGTGATTCCA	Cette étude
AsaGEI2b_orf39_For	CCCAGCATTTTCCTCACCCAGGAA	Cette étude
AsaGEI2b_orf39_Rev	GCAACCATGGTGGATACTGCGGAA	Cette étude

Annexe C : Fonction et position des ORF dans *AsaGEI1a*

Gène	Fonction	Position	Brin
ORF1	Intégrase	50..1303	-1
ORF2	Putative phage excisionnase	1303..1509	-1
ORF3	protéine hypothétique phagique	1473..1883	-1
ORF5	protéine hypothétique	2210..2809	-1
ORF7	protéine hypothétique phagique	2806..3111	-1
ORF9	protéine hypothétique phagique	3128..3778	-1
ORF10	protéine hypothétique	3821..4039	-1
ORF11	protéine hypothétique	4026..4703	-1
ORF12	protéine hypothétique	4706..5572	-1
ORF14	Phage recombinase	5569..6480	-1
ORF15	protéine hypothétique	6503..6682	-1
ORF17	putative Dam methylase	6741..7343	-1
ORF19	protéine hypothétique	7327..7728	-1
ORF20	protéine hypothétique phagique	7725..8954	-1
ORF22	protéine hypothétique phagique	8954..10144	-1
ORF24	phage recombination protein Bet	10201..11184	-1
ORF25	exodeoxyribonuclease VIII	11181..12026	-1
ORF26	protéine hypothétique	12023..12634	-1
ORF26.1	protéine hypothétique	12631..12996	-1
ORF26.2	protéine hypothétique	12993..13211	-1
ORF27	protéine hypothétique	13214..13405	-1
ORF28	protéine hypothétique	13429..13905	-1
ORF29	protéine hypothétique	13902..14489	-1
ORF32	protéine hypothétique	15103..15348	-1
ORF33	protéine hypothétique	15857..16510	-1
ORF34	DNA-binding protein	16526..17224	-1
ORF35	putative DNA-binding transcriptional regulator	17251..17493	1
ORF36	protéine hypothétique	17535..18071	1
ORF38	protéine hypothétique	18198..18353	1
ORF39	type II DNA modification methyltransferase	18346..19950	1
ORF40	protéine hypothétique phagique	19947..20981	1
ORF41	protéine hypothétique	20959..21447	1
ORF42	protéine hypothétique	21450..21752	1
ORF43	protéine hypothétique	21836..22117	1
ORF45	protéine hypothétique	22114..22773	1
ORF46	NinB protein	22770..23207	1
ORF47	protéine hypothétique phagique	23204..23521	1
ORF48	protéine hypothétique	23518..23757	1
ORF49	protéine hypothétique	23754..24176	1
ORF50	protéine hypothétique	24173..25069	1
ORF54	phage lysozyme	25542..26717	1
ORF56	putative small subunit bacteriophage terminase	26799..27386	1
ORF57	phage terminase-like family (TerL)	27376..28929	1

ORF58	protéine hypothétique	28934..31138	1
ORF61	protéine hypothétique	31161..31436	1
ORF62	protéine hypothétique	31433..32053	1
ORF63	protéine hypothétique	32088..33074	1
ORF64	protéine hypothétique	33092..34093	1
ORF65	protéine hypothétique	34154..34609	1
ORF66	protéine hypothétique	34612..35007	1
ORF67	protéine hypothétique	35001..37052	1
ORF68	protéine hypothétique	37045..37557	1
ORF69	protéine hypothétique	37550..39007	1
ORF72	protéine hypothétique	39060..39965	1
ORF73	protéine hypothétique	39975..40346	1
ORF74	protéine hypothétique	40349..41479	1
ORF75	protéine hypothétique	41470..41949	1
ORF76	protéine hypothétique	41946..42680	1
ORF77	protéine hypothétique	42742..43287	-1
ORF78	protéine hypothétique	43530..49784	1
ORF79	protéine hypothétique	49855..50058	1
ORF81	protéine hypothétique	50075..50437	-1
ORF82	protéine hypothétique	50803..50919	1

Annexe D : Matrice carrée des comparaisons génomique des trois *AsaGEI* chez les huit souches d'*A. salmonicida* sous-espèce *salmonicida*

		AsaGEI1a (51191 pb)			AsaGEI1b (50741 pb)		AsaGEI2a (53087 pb)		
		01-B526	01-B522	M15879-11	HER1085	ATCC 33658	09-0167	2009-144K3	m23067-09
AsaGEI1a	01-B526		100%/99%	100%/99%	73%/98%	73%/98%	70%/96%	70%/96%	70%/96%
	01-B522	100%/99%		100%/99%	73%/98%	73%/98%	70%/96%	70%/96%	70%/96%
	M15879-11	100%/99%	100%/99%		73%/97%	73%/97%	70%/95%	70%/95%	70%/95%
AsaGEI1b	HER1085	74%/98%	74%/98%	74%/97%		100%/99%	71%/95%	71%/95%	71%/95%
	ATCC 33658	74%/98%	74%/98%	74%/97%	100%/99%		71%/95%	71%/95%	71%/95%
AsaGEI2a	09-0167	67%/96%	67%/96%	67%/95%	68%/95%	68%/95%		100%/99%	100%/99%
	2009-144K3	67%/96%	67%/96%	67%/95%	68%/95%	68%/95%	100%/99%		100%/99%
	m23067-09	67%/96%	67%/96%	67%/95%	68%/95%	68%/95%	100%/99%	100%/99%	

% de longueur / % d'identité

Annexe E : Fonction et position des ORF dans *AsaGEI1b*

Gène	Fonction	Position	Brin
ORF1	Intégrase	50..1330	-1
ORF2	Putative phage excisionnase	1303..1476	-1
ORF3	protéine hypothétique phagique	1473..1883	-1
ORF4	protéine hypothétique	2167..2460	-1
ORF5	protéine hypothétique phagique	2563..2910	-1
ORF6	protéine hypothétique	2913..3128	-1
ORF10	protéine hypothétique	3122..3340	-1
ORF11	protéine hypothétique	3304..4365	-1
ORF12	protéine hypothétique phagique	4362..4994	-1
ORF14	phage exonucléase	4991..5929	-1
ORF15	protéine hypothétique	5952..6131	-1
ORF17	phage N-6-adénine-méthyltransférase	6191..6769	-1
ORF18	protéine hypothétique	6769..6909	-1
ORF19	protéine hypothétique	6906..7520	-1
ORF20	protéine hypothétique phagique	7517..8746	-1
ORF22	protéine hypothétique phagique	8793..9935	-1
ORF23	protéine hypothétique	9999..10586	-1
ORF24	phage recombination (Bet)	10650..11633	-1
ORF25	exodéoxyribonucléase VIII	11630..12475	-1
ORF26	protéine hypothétique	12472..12690	-1
ORF26.4	protéine hypothétique	12687..13043	-1
ORF26.1	protéine hypothétique	13040..13405	-1
ORF32	protéine hypothétique	14305..14556	-1
ORF34	phage répresseur homologue	15228..15842	-1
ORF35	putative phage régulatory protéine	15900..16184	1
ORF36	protéine hypothétique	16223..16642	1
ORF38	protéine hypothétique	16902..17054	1
ORF39	phage N-6-adnine-méthyltransférase	17051..18664	1
ORF40	phage réplication protéine	18661..19803	1
ORF41	phage réplication protéine P	19803..20474	1
ORF43	protéine hypothétique	20847..21143	1
ORF44	protéine hypothétique	21140..21313	1
ORF45	protéine hypothétique	21316..21567	1
ORF46	protéine NinB phagique	21564..22004	1
ORF46.1	protéine hypothétique	22001..22210	1
ORF47	protéine hypothétique phagique	22207..22524	1
ORF48	protéine hypothétique	22521..22760	1
ORF49	protéine hypothétique	22757..23179	1
ORF50	protéine hypothétique	23176..23787	1
ORF51	protéine hypothétique	23876..24070	1
ORF52	protéine hypothétique	24432..24674	1
ORF54	Lysozyme phagique	24671..25162	1
ORF55	protéine hypothétique	25162..25716	1
ORF56	putative small subunit bacteriophage	25798..26385	1

	terminase		
ORF57	phage terminase-like family (TerL)	26369..27928	1
ORF58	protéine hypothétique	27933..30137	1
ORF61	protéine hypothétique	30160..30435	1
ORF62.1	protéine hypothétique phagique	30432..30854	1
ORF62	protéine hypothétique	30862..31032	1
ORF63	protéine hypothétique	31087..32073	1
ORF64	protéine hypothétique	32091..33092	1
ORF65	protéine hypothétique	33153..33608	1
ORF66	protéine hypothétique	33611..34006	1
ORF67	protéine hypothétique	34000..36051	1
ORF68	protéine hypothétique	36044..36556	1
ORF69	protéine hypothétique	36549..37703	1
ORF72	protéine hypothétique	38147..38962	1
ORF73	protéine hypothétique	38972..39343	1
ORF74	protéine hypothétique	39346..40461	1
ORF75	protéine hypothétique	40533..40940	1
ORF76	protéine hypothétique	40937..41671	1
ORF77	protéine hypothétique	41733..42278	-1
ORF78	protéine hypothétique	42521..48979	1
ORF79	protéine hypothétique	49050..49253	1
ORF80	protéine hypothétique	49401..49655	-1
ORF81	protéine hypothétique	49624..49986	-1
ORF82	protéine hypothétique	50025..50375	1

Annexe F : Fonction et position des ORF dans *AsaGEI2a*

Gène	Fonction	Position	Brin
ORF1	phage integrase	79..1155	-1
ORF2	protéine hypothétique phagique	1155..1376	-1
ORF3	protéine hypothétique phagique	1258..1746	-1
ORF4	protéine hypothétique	1983..2297	-1
ORF5	protéine hypothétique	2278..2499	-1
ORF6	protéine hypothétique	2489..2749	-1
ORF7	protéine hypothétique phagique	2830..3135	-1
ORF9	protéine hypothétique	3138..3572	-1
ORF9.1	protéine hypothétique	3646..3873	-1
ORF10	protéine hypothétique phagique	3941..5023	-1
ORF12	protéine hypothétique	5020..5646	-1
ORF12.1	protéine hypothétique phagique	5643..5978	-1
ORF13	protéine hypothétique	5975..6490	-1
ORF14	phage recombinase	6487..7398	-1
ORF15	protéine hypothétique	7423..7545	-1
ORF17	adenine phage methylase	7662..8246	-1
ORF18	protéine hypothétique	8407..8547	-1
ORF19	protéine hypothétique phagique	8544..9245	-1
ORF20	protéine hypothétique phagique	9242..10495	-1
ORF21	protéine hypothétique	10542..11660	-1
ORF22	protéine hypothétique phagique	11683..12825	-1
ORF24	recombinase (RecT)	12877..13977	-1
ORF25	phage exodeoxyribonuclease VIII	13851..14693	-1
ORF26	protéine hypothétique	14690..14908	-1
ORF26.1	protéine hypothétique	15515..15877	-1
ORF28	protéine hypothétique phagique	15874..16368	-1
ORF29	protéine hypothétique	16371..16562	-1
ORF32	protéine hypothétique	16961..17206	-1
ORF34	phage DNA-binding protein	18070..18708	-1
ORF35	protéine hypothétique	18803..19021	1
ORF36	protéine hypothétique	19135..19491	1
ORF37	protéine hypothétique	19496..19708	1
ORF38	protéine hypothétique	19783..20154	1
ORF39	phage phosphoadenosine phosphosulfate reductase	20151..20906	1
ORF40	replication phage protein	20903..22063	1
ORF41	protéine hypothétique	22041..22526	1
ORF42	protéine hypothétique	22529..22831	1
ORF43	protéine hypothétique	22915..23211	1
ORF44	protéine hypothétique	23208..23387	1
ORF45	protéine hypothétique	23384..23758	1
ORF46	NinB phage protein	23755..24198	1
ORF47	protéine hypothétique phagique	24195..24512	1
ORF48	protéine hypothétique	24509..24748	1

ORF49	protéine hypothétique	24745..25167	1
ORF50	protéine hypothétique	25164..25775	1
ORF50.1	protéine hypothétique	25821..26234	1
ORF51	protéine hypothétique	26231..26425	1
ORF51.1	protéine hypothétique	26444..26671	1
ORF52	protéine hypothétique phagique	26821..27135	1
ORF54	phage lysosyme protéine	27132..27680	1
ORF55.1	protéine hypothétique	27688..27912	1
ORF55.2	protéine hypothétique	27909..28091	1
ORF56	putative small subunit bacteriophage terminase	28218..28796	1
ORF57	phage terminase-like family (TerL)	28796..30349	1
ORF58	protéine hypothétique	30354..32558	1
ORF61	protéine hypothétique	32581..32856	1
ORF62.1	protéine hypothétique phagique	32853..33335	1
ORF62	protéine hypothétique	33343..33513	1
ORF63	protéine hypothétique	33569..34555	1
ORF64	protéine hypothétique	34573..35574	1
ORF65	protéine hypothétique	35635..36090	1
ORF66	protéine hypothétique	36093..36488	1
ORF67	protéine hypothétique	36482..38533	1
ORF68	protéine hypothétique	38526..39038	1
ORF69	protéine hypothétique	39031..40185	1
ORF71	protéine hypothétique	40187..40486	1
ORF72	protéine hypothétique	40455..41444	1
ORF73	protéine hypothétique	41454..41825	1
ORF75	protéine hypothétique	43049..43510	1
ORF76	protéine hypothétique	43507..44241	1
ORF77	protéine hypothétique	44303..44848	-1
ORF78	protéine hypothétique	45091..51351	1
ORF79	protéine hypothétique	51414..51617	1
ORF80	protéine hypothétique	51765..51983	-1
ORF81	protéine hypothétique	51988..52350	-1
ORF82	protéine hypothétique	52716..52832	1

Annexe G : Fonction et position des ORF dans *AsaGEI2b*

Gène	Fonction	Position	Brin
ORF82	protéine hypothétique	376..543	-1
ORF81	protéine hypothétique	738..1100	1
ORF80	protéine hypothétique	1069..1323	1
ORF79	protéine hypothétique	1471..1674	-1
ORF78	protéine hypothétique	1737..8201	-1
ORF77	protéine hypothétique	8444..8989	1
ORF76	protéine hypothétique	9051..9785	-1
ORF75	protéine hypothétique	9782..10243	-1
ORF74	protéine hypothétique	10261..11181	-1
ORF72	protéine hypothétique	11812..12684	-1
ORF71	protéine hypothétique	12737..13099	-1
ORF69.2	protéine hypothétique	13099..13854	-1
ORF68	protéine hypothétique	13847..14359	-1
ORF67	protéine hypothétique	14352..16403	-1
ORF66	protéine hypothétique	16397..16792	-1
ORF65	protéine hypothétique	16795..17250	-1
ORF64	protéine hypothétique	17312..18313	-1
ORF63	protéine hypothétique	18331..19317	-1
ORF62.2	protéine hypothétique	19373..19543	-1
ORF62.1	protéine hypothétique	19551..19973	-1
ORF61	protéine hypothétique	19970..20245	-1
ORF58	protéine hypothétique	20262..22466	-1
ORF57	phage terminase-like family (TerL)	22471..23958	-1
ORF56	putative small subunit bacteriophage terminase	24014..24601	-1
ORF55.3	protéine hypothétique	24728..24910	-1
ORF55.2	protéine hypothétique	24907..25131	-1
ORF55.1	phage lysozyme protéine	25139..25687	-1
ORF54	phage membrane protéine	25684..25998	-1
ORF53	protéine hypothétique	26149..26373	-1
ORF52	protéine hypothétique	26395..26589	-1
ORF50.2	protéine hypothétique	26675..27250	-1
ORF50	protéine hypothétique	27265..27876	-1
ORF49	protéine hypothétique	27873..28295	-1
ORF48	hypthetical protein	28292..28531	-1
ORF47	protéine hypothétique	28528..28845	-1
ORF46	NinB	28842..29285	-1
ORF44	protéine hypothétique	29288..29485	-1
ORF43	protéine hypothétique	29482..29778	-1
ORF42	protéine hypothétique	29856..30158	-1
ORF41	protéine hypothétique	30161..30649	-1
ORF40.2	protéine hypothétique phagique	30627..31592	-1

ORF40	protéine hypothétique phagique	31658..33394	-1
ORF39	type II DNA modification methyltransferase	33391..35028	-1
ORF38	protéine hypothétique	35021..35176	-1
ORF37	protéine hypothétique	35227..35439	-1
ORF36	protéine hypothétique	35444..35833	-1
ORF35	DNA-binding	35914..36132	-1
ORF34	DNA-binding protein	36227..36865	1
ORF32	protéine hypothétique	37736..37981	1
ORF28	protéine hypothétique	38380..38724	1
ORF27	protéine hypothétique	38714..39079	1
ORF26.2	protéine hypothétique	39076..39558	1
ORF26	protéine hypothétique	39555..39773	1
ORF25	exodeoxyribonuclease VIII	39770..40615	1
ORF24	recombinase	40612..41607	1
ORF23	protéine hypothétique	41663..42805	1
ORF20	protéine hypothétique	42852..44081	1
ORF19	protéine hypothétique	44078..44692	1
ORF18	protéine hypothétique	44689..44829	1
ORF17	adenine phage methylase	44990..45574	1
ORF15	protéine hypothétique	45634..45813	1
ORF14	recombinase	45836..46774	1
ORF11.2	protéine hypothétique	46771..47115	1
ORF11	protéine hypothétique	47187..47726	1
ORF10.3	protéine hypothétique	47729..48694	1
ORF10	protéine hypothétique	48658..48876	1
ORF10.1	hypthetical protein	48944..49171	1
ORF8.2	protéine hypothétique	49244..49882	1
ORF7	protéine hypothétique	49886..50191	1
ORF5.2	protéine hypothétique	50278..50595	1
ORF4.2	protéine hypothétique	50816..51055	1
ORF3	protéine hypothétique	51121..51609	1
ORF1	Phage integrase	51784..52788	1

Annexe H : Fonction et position des ORF dans le prophage 3

Produit	Position	Brin	Note
tail protéine	115..1275	-1	D domain
tail assembly protéine	1275..1766	-1	
Rha family transcriptional regulator	1769..2797	-1	phage_pRha domain
tail protéine	2784..5285	-1	
P2 GpE family protéine	5282..5413	-1	Phage_P2_GpE domain
protéine hypothétique	5422..5706	-1	FluMu_gp41 domain
major tail tube protéine	5773..6291	-1	Phage_tube domain
tail sheath protéine	6301..7479	-1	FI domain
protéine hypothétique	8072..8365	-1	
protéine hypothétique phagique	8362..9117	-1	COG5301 domain; Phage-related tail fibre
phage tail protéine I	9125..9514	-1	Tail_P2_I domain
baseplate assembly protéine	9712..10614	-1	J domain
baseplate assembly protéine	10611..10970	-1	W domain
phage baseplate assembly protéine	10967..11539	-1	phage_P2_V domain
Phage virion morphoproductsis family	11618..12076	-1	Phage_tail_S domain
tail fiber protéine	12058..12528	-1	P2_Phage_GpR domain
LysB	12642..13109	-1	phage_LysB domain
protéine hypothétique	13254..13766	-1	
protéine hypothétique	13756..14577	-1	
protéine hypothétique phagique	14574..14915	-1	DUF754 domain
protéine hypothétique	14918..15286	-1	membrane protéine
conjugal transfer protéine TraR	15302..15559	-1	zf-dskA_traR domain
tail protéine X	15550..15753	-1	
Phage head completion protéine	15753..16229	-1	Phage_GPL domain
terminase	16340..17113	-1	Phage_term_smal domain
capsid protéine	17124..18182	-1	Phage_cap_P2 domain
capsid protéine	18195..19028	-1	Phage_GPO domain
oxidoreductase	19177..20943	1	
Portal vertex-like protéine	20943..22016	1	
protéine hypothétique	22495..22716	1	
protéine hypothétique	22831..23196	1	HTH_XRE
protéine hypothétique	23138..23515	1	
protéine hypothétique	23777..24067	-1	
bacteriophage protéine gp46	24077..24199	-1	
phage protéine	24196..24651	-1	DUF3850 domain
DNA methylase	24648..25412	-1	N6_N4_Mtase domain
replication initiation protéine	25423..27834	-1	Phage_GPA domain
type II restriction endonuclease subunit M	27831..29504	-1	Dcm domain
protéine hypothétique	29501..29824	-1	
protéine hypothétique	29947..30153	-1	

DNA methyltransferase	30150..30686	-1	AdoMet_MTases domain
protéine hypothétique	30683..30859	-1	
hypthetical protéine	30856..31377	-1	
protéine hypothétique	31441..31884	-1	
protéine hypothétique	31900..32121	-1	
protéine hypothétique	32118..32378	-1	Ogr_Delta
Phage regulatory protéine CII	32447..32965	-1	Phage_CP76
protéine hypothétique	32999..33412	-1	
putative cl repressor	33806..34324	1	S24_LexA-like domain
protéine hypothétique	34338..34616	1	Excalibur domain; cold-shock protéine like
protéine hypothétique	34643..35026	1	
protéine hypothétique	35083..36120	1	UPF0547
integrase	36117..37175	1	PHA02601