



**Impact des éléments mobiles de l'ADN sur l'adaptabilité
et le transfert horizontal chez *Aeromonas salmonicida*
ssp. *salmonicida***

Thèse

Katherine Tanaka

Doctorat en microbiologie
Philosophiæ doctor (Ph. D.)

Québec, Canada

© Katherine Tanaka, 2018

**Impact des éléments mobiles de l'ADN sur l'adaptabilité
et le transfert horizontal chez *Aeromonas salmonicida*
ssp. *salmonicida***

Thèse

Katherine Tanaka

Sous la direction de :

Steve Charette, directeur de recherche
Michel Frenette, codirecteur de recherche

RÉSUMÉ

Les éléments mobiles de l'ADN, des séquences capables de transfert intracellulaire ou extracellulaire, existent dans la plupart des génomes (eucaryotes ou procaryotes). Ce sont des vecteurs importants de variation génétique. Certains portent des gènes accessoires qui procurent un avantage sélectif à leur organisme-hôte. Les séquences d'insertion (IS) sont un cas spécial d'éléments mobiles bactériens. Elles sont exemptes de gènes accessoires, mais autonomes pour leur transposition. Elles peuvent donc causer des variations structurales, ce qui peut perturber les fonctions de la bactérie-hôte. Néanmoins, la plupart des génomes bactériens portent plusieurs copies de différentes familles d'IS. Plusieurs hypothèses, allant du parasitisme des IS à une forme de commensalisme IS-bactérie, ont été formulées pour expliquer la persistance de ces éléments mobiles dans les génomes.

Aeromonas salmonicida ssp. *salmonicida*, l'agent étiologique de la furunculose chez les salmonidés, est un modèle intéressant pour étudier les IS. Son génome possède plusieurs familles d'IS présentes en différents nombres de copies. De plus, des travaux réalisés précédemment ont montré que ces éléments mobiles étaient responsables de variations structurales dans les plasmides de la bactérie et de l'inactivation de certains de ses facteurs de virulence. Bien que les événements mentionnés ci-dessus semblent neutres ou désavantageux, il n'est pas exclu que certains événements impliquant les IS puissent apporter un avantage adaptatif à la bactérie. L'objectif de cette thèse était donc d'explorer les éléments mobiles du génome d'*A. salmonicida* ssp. *salmonicida*, en portant une attention particulière aux IS, afin d'identifier des événements à répercussion bénéfique pour la bactérie.

La richesse d'information et l'accessibilité grandissante du séquençage à haut débit en ont fait une technique de choix pour étudier les IS et autres transposons en association avec un autre élément mobile essentiel à leur transmission intercellulaire, les plasmides. Une première étude a permis de découvrir des variants d'un grand plasmide d'*A. salmonicida* ssp. *salmonicida*, pAsa4, dont la forme de référence porte des gènes de résistance aux antibiotiques, des IS et un transposon. Les variants identifiés lors de l'étude, pAsa4b et pAsa4c, présentent des variations structurales par rapport à la référence. Elles ont des conséquences sur les gènes de résistance portés et sur la capacité de transfert des plasmides.

La génomique comparative faite entre les pAsa4 suggère que les éléments mobiles tiennent un rôle primordial dans la création de certaines variations structurales, tant par leur capacité de transposition ou d'intégration que par leur propension à la recombinaison homologue. L'élargissement de la famille des pAsa4 à trois variants procure d'autres exemples de variabilité attribuable aux éléments mobiles de l'ADN.

Dans un second temps, nous avons approfondi un événement étudié précédemment : la perte du locus du système de sécrétion de type trois (SSTT). Cette région, située sur le grand plasmide pAsa5, peut être perdue lorsque *A. salmonicida* ssp. *salmonicida* est cultivée à 25°C. Dans certains cas, la recombinaison homologue d'IS flanquant les régions touchées avait été mise en cause. Dans le cadre de ce projet, le génome de la souche 01-B526, qui avait été utilisée pour générer des pAsa5 réarrangés, a été séquencé par une combinaison de technologies à courtes et à longues lectures. Un nouveau plasmide, pAsa9, ainsi qu'une nouvelle IS de pAsa5, ISAS5Z, ont été découverts. Ensemble, ces deux éléments ont permis de proposer un nouveau schéma de recombinaison homologue qui s'est avéré exact pour tous les réarrangements de pAsa5 générés lors de l'expérience précédente et considérés comme irréguliers. Ces nouveaux résultats permettent de regrouper tous les scénarios de réarrangement de pAsa5 sous un seul mécanisme, la recombinaison homologue entre IS, en plus d'ajouter une nouvelle entité au plasmidome d'*A. salmonicida* ssp. *salmonicida*.

Plusieurs événements ont donc été ajoutés au catalogue des variations structurales causées par les éléments mobiles de l'ADN d'*A. salmonicida* ssp. *salmonicida*. Dans les derniers chapitres de cette thèse, les conséquences de ces différents événements sont discutées. Certaines variations sur pAsa4 apportent des gènes accessoires qui sont bénéfiques dans un environnement donné. Autrement, l'étude des éléments observés est limitée par le manque de connaissance sur les facteurs environnementaux qui causent ces variations et sur leur impact *in vivo*, chez le poisson. Quoiqu'il en soit, les IS du génome d'*A. salmonicida* ssp. *salmonicida* influencent certainement les relations hôte-pathogène de cette bactérie et participent intimement à son histoire évolutive.

ABSTRACT

Mobile genetic elements are sequences capable of intracellular and extracellular transfer. They exist in the majority of eukaryotic and prokaryotic genomes, and are important vectors of genetic variability. Some mobile genetic elements carry accessory genes that confer selective advantages to their host. Insertion sequences (IS), do not belong to the latter group, as they are exempt of accessory genes, yet they are autonomous in their transposition. ISs are thus important structural variation producers, which may endanger the host's functions and its genomic integrity. However, most bacterial genomes carry ISs. Hypotheses ranging from parasitism to commensalism were formulated to explain how ISs persist in genomes.

A. salmonicida subsp. *salmonicida*, the causative agent of furunculosis in salmonids, is a good model to study ISs. Its genome carries many ISs divided in different families. Moreover, previous work showed that these elements cause structural variations in plasmids and virulence factor inactivations. Even though the events observed to date seemed neutral or detrimental for the bacteria, we cannot exclude that ISs may procure an adaptative advantage in certain cases. This thesis' objective was to explore the *A. salmonicida* subsp. *salmonicida* mobile genetic elements, especially ISs, to identify events with a positive outcome for the bacteria.

ISs and transposons were studied in association with another mobile element accountable for their intercellular transfer, plasmids. High-throughput sequencing was used as a preferred method because of its growing availability and the richness of information it generates. In a first article, variants of a large plasmid that carries antibiotic resistance genes, pAsa4, were discovered. pAsa4b and pAsa4c display structural variations when compared to the reference plasmid that impact their antibiotic resistance genes and their conjugation capabilities. Comparative genomics between the three variants shows that mobile elements, either by transposition or recombination, have an important role in some variations. Adding two new members to the pAsa4 family of plasmids gave further examples of genetic diversity driven by mobile elements.

In a second article, we investigated previous results concerning the type three secretion system (TTSS) loss. This is an essential virulence factor for *A. salmonicida* subsp. *salmonicida*, and its genes are mainly encoded in a single locus of the large plasmid pAsa5. The region can be lost when the bacterium is cultivated at 25°C or higher. In some cases, homologous recombinations between IS copies flanking the TTSS locus have been identified as causing the deletion. In this study, the genome of the strain 01-B526, which was used in a previous study to produce TTSS-negative mutants, was sequenced using Illumina and PacBio technologies. A new plasmid, named pAsa9, and a new pAsa5 IS, ISAS5Z, were discovered. Together, those elements permitted to build a new IS-driven homologous recombination model that was tested in the mutant strains. All TTSS-negative strains generated in previous study and that were not explained by other IS recombination patterns fitted with the new ISAS5 homologous recombination model. These results allowed to regroup all pAsa5 rearrangements under a consistent mechanism: recombination using IS copies as a template.

In this thesis, more structural variation events involving ISs were uncovered. In the last chapters, their involvement in *A. salmonicida* subsp. *salmonicida* fitness *in vitro* and in the fish host is discussed. Some variations in pAsa4 bring new resistance genes, which are beneficial for the bacteria in an aquaculture context. However, for most events, we lack situational information to properly conclude their fitness impact. Regardless, ISs of the *A. salmonicida* subsp. *salmonicida* genome certainly influence host-pathogen relations and participate in the bacteria evolutive history.

TABLE DES MATIÈRES

Résumé	iii
Abstract	v
Table des matières	vii
Liste des Tableaux	x
Liste des Figures	xi
Liste des Fichiers Additionnels	xii
Liste des Abréviations	xiii
Remerciements	xvii
Avant-propos	xix
Chapitre 1 Introduction	1
1.1 La génomique, un nouveau paradigme de l'étude du vivant	1
1.1.1 La génomique bactérienne	2
1.2 Le séquençage à haut débit	2
1.2.1 Comparaison du séquençage à haut débit avec la méthode Sanger	4
1.2.2 Préparation de l'ADN	4
1.2.3 Le pyroséquençage	5
1.2.4 Séquençage par synthèse de type Illumina	6
1.2.5 Séquençage de molécule unique de type PacBio	7
1.2.6 Assemblage et analyse des séquences	8
1.3 Les éléments mobiles de l'ADN	10
1.3.1 Les transposons	11
1.3.2 Les intégrons	12
1.3.2.1 Étude de cas : le transposon Tn21	14
1.3.3 Les îlots génomiques	14
1.3.4 Les plasmides	15
1.3.5 Les séquences d'insertion	16
1.3.5.1 Les transposases	17
1.3.5.2 Réplication en cercle roulant : le cas particulier d'IS91 et les ISCR	18
1.3.5.3 Impact des IS sur les gènes et les génomes	18
1.3.5.4 La théorie de réduction des génomes	20
1.3.5.5 Paradigme de la coévolution IS-bactéries	20
1.4 L'organisme à l'étude : <i>Aeromonas salmonicida</i> ssp. <i>salmonicida</i>	22
1.4.1 La problématique de la furunculose	24
1.4.2 Les facteurs de virulence d' <i>A. salmonicida</i> ssp. <i>salmonicida</i>	26
1.4.2.1 Toxines extracellulaires	27
1.4.2.2 La couche de surface A et les lipopolysaccharides	27
1.4.2.3 Autres aspects en lien avec la membrane	28
1.4.2.4 Métabolisme dans l'hôte	29
1.4.2.5 Le système de sécrétion de type trois	29
1.4.3 Le génome d' <i>A. salmonicida</i> ssp. <i>salmonicida</i> est riche en information	32
1.4.4 Le génome de référence d' <i>A. salmonicida</i> ssp. <i>salmonicida</i>	33
1.4.5 Le plasmidome d' <i>A. salmonicida</i> ssp. <i>salmonicida</i>	34
1.4.5.1 Les petits plasmides à haut nombre de copies	34
1.4.5.2 Les plasmides portant des gènes de résistance	35
1.4.5.2.1 Le plasmide pAsa8	36
1.4.5.2.2 pRAS1, pAr-32 et les pASOT : des plasmides apparentés ?	36
1.4.5.2.3 Autres plasmides de la suite pRAS	37
1.4.5.2.4 Une accumulation de résistances : pAb5S9b et pSN254b	37

1.4.5.2.5 pAsa4, un plasmide spécifique d' <i>A. salmonicida</i> ssp. <i>salmonicida</i>	38
1.4.5.2.6 Un autre grand plasmide d'A449, pAsa5	39
1.4.6 Les séquences d'insertion chez <i>A. salmonicida</i> ssp. <i>salmonicida</i>	42
1.4.6.1 Variations cryptiques dans les plasmides	44
1.4.6.2 Transposition des IS <i>AS1</i> et IS <i>AS2</i> dans le gène <i>vapA</i> et <i>abcA</i>	44
1.4.6.3 Perte du SSTT et de la virulence	45
1.5 Énoncé de la problématique de recherche	47
Chapitre 2 Les conséquences sur la résistance aux antibiotiques de l'architecture en mosaïque des plasmides de la famille de pAsa4	50
Résumé	50
The mosaic architecture of <i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i> pAsa4 plasmid and its consequences on antibiotic resistance	51
Abstract	52
Introduction	53
Material and Methods	54
Bacterial isolates, growth conditions, antibiotic resistance profiles, and conjugation assays	54
DNA extraction and sequencing	55
Sequence analysis	55
PCR analyses	56
Results	57
Complete sequences of the new pAsa4 variants	57
Insertion sequences	61
Detailed plasmid architecture	61
Comparative analysis of the pAsa4 architecture	63
Variant genotyping and antibiotic resistances	65
Discussion	67
Conclusion and Perspectives	69
Acknowledgments	70
Relevant supporting information	71
References	72
Chapitre 3 Un nouveau patron de recombinaison des IS du plasmide pAsa5 chez <i>Aeromonas salmonicida</i> ssp. <i>salmonicida</i> 01-B526	76
Résumé	76
Plasmid composition in <i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i> 01-B526 unravels unsuspected type three secretion system loss patterns	77
Abstract	78
Background	79
Results	82
Characterization of pAsa9	82
Comparison between pAsa9 and pAsa5	83
Genotyping unexplained type 1 loss profile and pAsa9 occurrence	85
01-B516 and its assigned daughter strains	89
pAsa9 prevalence	91
Plasmid fusion in 01-B526-R19	91
Discussion	93
Conclusions	95
Methods	96
Bacterial isolates and growth conditions	96
DNA extraction and sequencing	96
Sequence analysis	97
PCR analyses	97

Plasmid visualization by agarose gel electrophoresis	98
Abbreviations	98
Acknowledgments	98
Relevant supporting information.....	99
References	101
Chapitre 4 Discussion.....	106
4.1 Rappel de l’hypothèse de départ et des résultats.....	106
4.2 Les variants du plasmide pAsa4 confèrent différentes résistances aux antibiotiques	106
4.2.1 Retour sur les variations structurales de pAsa4 causées par les IS	107
4.2.1.1 L’inversion dans pAsa4.....	109
4.2.1.2 Insertion de différentes familles d’IS dans ISAS5.....	111
4.2.1.3 L’ISCR et l’insertion de <i>catA1</i>	112
4.2.2 Distribution des IS entre les pAsa4, les autres plasmides et le chromosome.....	113
4.2.3 Implication des autres éléments mobiles sur pAsa4.....	116
4.2.4 L’évolution du plasmide pAsa4 et les liens évolutifs avec d’autres plasmides	117
4.2.4.1 Liens évolutifs entre les plasmides de la famille de pAsa4 et ceux du groupe d’incompatibilité IncA/C.....	117
4.2.5 Les plasmides de la famille de pAsa4, un exemple de coévolution entre plasmide et hôte ?	118
4.2.6 Apport de pAsa4 à la plasticité génomique d’ <i>A. salmonicida</i> ssp. <i>salmonicida</i>	119
4.3 Nouvelles cibles de recombinaison et nouveau plasmide pour les <i>A. salmonicida</i> ssp. <i>salmonicida</i> québécoises	119
4.3.1 Retour sur les mécanismes de délétion du SSTT	120
4.3.1.1 Trouver la recombinase responsable de la délétion.....	123
4.3.1.2 Le stress, déclencheur de la perte du SSTT ?.....	123
4.3.2 Conséquence d’une perte du SSTT pour <i>A. salmonicida</i> ssp. <i>salmonicida</i>	124
4.3.3 Liens évolutifs entre pAsa5, pAsa9, et l’origine du locus du SSTT	127
Chapitre 5 Conclusion	128
5.1 Le passé, présent et futur d’ <i>A. salmonicida</i> ssp. <i>salmonicida</i> et de ses IS.....	128
Bibliographie	130
Annexe 1 Méthode employée pour l’identification des IS dans les génomes d’ <i>A. salmonicida</i> ssp. <i>salmonicida</i>	150
Mise en contexte.....	150
Méthode.....	150
Analyse.....	151

LISTE DES TABLEAUX

Tableau 1.1 Comparaison des techniques de séquençage utilisées au cours de la thèse	8
Tableau 1.2 Le plasmidome d' <i>A. salmonicida</i> ssp. <i>salmonicida</i>	40
Tableau 1.3 Séquences d'insertion retrouvées dans le génome d' <i>A. salmonicida</i> ssp. <i>salmonicida</i> A449.....	43
Table 2.1 <i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i> strains bearing pAsa4 variants	58
Table 3.1 <i>A. salmonicida</i> subsp. <i>salmonicida</i> isolates used in this study.....	87
Tableau 4.1 Nombre ou présence d'IS dans les génomes dans une sélection de génomes d' <i>A. salmonicida</i> ssp. <i>salmonicida</i>	115

LISTE DES FIGURES

Figure 1.1 Différents types de transposons	12
Figure 1.2 Schéma d'un intégron de classe 1	13
Figure 1.3 Deux exemples de perturbation impliquant les IS	19
Figure 1.4 Le système de sécrétion (SSTT) d' <i>A. salmonicida</i> ssp. <i>salmonicida</i>	32
Figure 1.5 Résumé des études sur la délétion du locus du SSTT par recombinaison homologue	46
Figure 2.1 Nucleotide alignment of three plasmid variants: pAsa4, pAsa4b, and pAsa4c ..	60
Figure 2.2 In2 comparison between pAsa4b and the other variants.....	63
Figure 2.3 Clustering tBlastn results for pAsa4b	64
Figure 2.4 Nucleotide alignment between pAsa4b and pSN254b.....	65
Figure 3.1 Alignment of 01-B526 pAsa5 and pAsa9, and proposed recombination patterns for 01-B526 pAsa5	81
Figure 3.2 pAsa9 map.....	84
Figure 3.3 Large plasmids from parental and derived strains	90
Figure 3.4 Fusion of the 01-B526-R19 pAsa5 plasmid with pAsa9 through their ISAS5....	92
Figure 4.1 Mise en évidence des variations structurales entre les plasmides pAsa4.....	108
Figure 4.2 Les séquences répétées du pAsa5 de 01-B526.....	121
Figure 4.3 Courbes de croissance d' <i>A. salmonicida</i> ssp. <i>salmonicida</i> 01-B526 et de ses réarrangés.....	125

LISTE DES FICHIERS ADDITIONNELS

Additional File 2.1 Substitution rate between pAsa4c (JF2267) and pAsa4b (01-B522) using 1-kbp windows.....	71
Additional File 3.1 Circular map of pAsa5 of 01-B526 strain with G+C skew	99
Additional File 3.2 Alignment between pAsa5 and pAsa9 from 01-B526 with pAsa6.	100

LISTE DES ABRÉVIATIONS

A	adénine
ADN/DNA	acide désoxyribonucléique/ <i>deoxyribonucleic acid</i>
ADP	adénosine diphosphate
ATP	adénosine triphosphate
ARN/RNA	acide ribonucléique/ <i>ribonucleic acid</i>
ATCC	<i>American Type Culture Collection</i>
ATP	adénosine triphosphate
C	cytosine
Cas	CRISPR-associated system
CC BY	<i>creative commons attribution</i>
CDS	<i>coding sequence</i>
CRISPR	<i>Clustered Regularly Interspaced Short palindromic repeats</i>
CRSNG	Conseil de recherche en sciences naturelles et en génie
Dtr	<i>DNA transfer and replication</i>
EDTA	éthylène diamine tétra-acétique
FAO	L'Organisation des Nations Unies pour l'alimentation et l'agriculture (<i>Food and Agriculture Organization</i>)
FRQNT	Fonds de recherche Nature et technologie
G	guanine
Gb	gigabases
GCAT	glycérophospholipide-cholestérol acyltransférase
GREB	Groupe de Recherche en Écologie Buccale
GTP	guanosine triphosphate
IBIS	Institut de Biologie intégrative et des Systèmes
ICE	<i>integrative and conjugative elements</i>
In	intégron
IR	répétitions inversées (<i>inverted repeats</i>)
IS	séquence d'insertion (<i>insertion sequence</i>)
ISCR	<i>insertion sequence common region</i> (un type d'IS)
kb	kilobases
kpb/kbs	kilopaires de bases/ <i>kilobase pairs</i>
LB	<i>lysogeny broth</i>
LPS	lipopolysaccharide
Mb	mégabases
Mpf	<i>mating pair formation</i>
NCBI	<i>National Center for Biotechnology Information</i>
nr/nt	non-redondant/nucléotides (nom d'une base de données)
ORF	<i>open reading frame</i>
pb/bp	paires de bases/ <i>base pair</i>
PCR	réaction de polymérisation en chaîne (<i>polymerase chain reaction</i>)

pv.	pathovar
RAQ	Ressources Aquatiques Québec
rpm	rotations par minute
SDS	<i>sodium dodecyl sulfate</i>
SMRT	<i>single molecule real-time</i>
sp.	<i>species</i>
ssp./subsp.	sous-espèce/ <i>subspecies</i>
SSTT/TTSS	système de sécrétion de type trois/ <i>type three secretion system</i>
SXT	sulfaméthoxazole-triméthoprim
T	thymine
Tn	transposon
TSA	<i>tryptic soy agar</i>
TSB	<i>tryptic soy broth</i>
UV	ultraviolet
wgs	<i>whole genome shotgun</i> (nom d'une base de données)

À mamie Jeanine, le cœur et l'âme d'une grande famille

*Lorsque vous avez éliminé l'impossible, ce qui reste,
si improbable soit-il, est nécessairement la vérité.*

Sir Arthur Conan Doyle, *Le Signe des Quatre*

REMERCIEMENTS

Le projet qui vous est présenté ici s'est tissé au fil de sept ans de collaboration, de travail commun et d'amitié. C'est l'histoire de beaucoup d'espoir et d'ambitions qui sont partagés entre des personnes extraordinaires. Maintenant que mon projet est prêt à prendre une vie propre, j'aimerais profiter de cet instant pour exprimer de sincères remerciements.

D'abord, rien de tout cela n'aurait pu exister sans sources de financement. De nombreux organismes subventionnaires ont appuyé à un moment ou à un autre le pur potentiel de quelque chose de plus grand. Merci au CRSNG et au FRQNT et à l'Université Laval d'avoir cru en mes capacités de jeune chercheure. Au cours de mon cheminement, j'ai aussi eu la chance d'obtenir du financement, mais surtout de côtoyer les gens extraordinaires du RAQ. Merci à toute cette famille élargie, et plus particulièrement à Céline Audet.

Je lève ensuite mon chapeau à mon comité d'encadrement. Merci à Claude Lemieux et à Nicolas Derome d'avoir accepté mon invitation, il y a cinq ans, et d'avoir suivi mon cheminement depuis. Nos conversations ont été des plus enrichissantes. Merci au professeur Vincent Burrus, qui se joint au jury d'évaluation de la thèse. Merci à tous au département de biochimie, de microbiologie et de bio-informatique d'avoir veillé sur ma vie administrative. Vous gardez cette structure vivante et humaine.

À mon codirecteur, Michel Frenette, j'adresse mes remerciements les plus sincères. Merci d'avoir offert ton soutien, avec cœur, à la science, mais aussi à la scientifique, d'avoir offert une compréhension qui transcende les mots. Et à mon directeur, Steve Charette, simplement, milles mercis. Pour avoir appliqué maintes et maintes fois ton raisonnement de pédagogue empreint de compassion. Pour tous les projets que nous avons commencés, ceux qui sont terminés, et ceux qui se reposent. Pour avoir tenu la lampe et montré le chemin dans les moments les plus sombres.

Sept ans dans un laboratoire, c'est aussi beaucoup de moments passés avec des gens fantastiques. C'est beaucoup de bienvenues et autant d'au revoir. À tous mes collègues et amis des laboratoires Charette, Lagüe, Shi, Tanguay, Levesque et Landry, à tous les gens fantastiques ainsi qu'à l'IBIS et sa plateforme de séquençage, je vous dis merci. Plus particulièrement, merci à Cynthia, Alicia, Alix, William, Myriam et Mélanie pour

l'ambiance du bureau. À Alex, le combattant des huiles essentielles. À Perrine et à Marie-Stéphanie, pour vos sourires sympathiques. Au fan-club d'*Asal*, notre bibitte compliquée préférée : Gabriel, Sabrina, Antony, Jean-Guillaume, Marcin, François. Vous êtes extraordinaires dans ce que vous faites, mais surtout qui vous êtes. J'aimerais aussi envoyer de gros mercis à celles qui sont plus loin, mes *senpai* de labo et de projet : Rana, Geneviève et Stéphanie. J'espère être digne de vous quand viendra le temps de passer le flambeau.

Mes derniers remerciements à l'équipe Charette vont à Valérie Paquet, maman inconditionnelle de labo. Valérie, tu m'as donné, tout au long de ces sept ans, l'opinion lucide et honnête d'une amie. Tu fais partie des personnes qui m'ont donné cette dernière poussée dans le dernier droit. Ta présence au quotidien me manquera.

J'aimerais aussi remercier ma famille, toutes mes familles pour leur soutien spécial et pour avoir contribué à faire qui je suis. Au Gosselin, à ma mère, Carole, et à ma grand-mère Monique, merci d'avoir encouragé cette petite dérive de l'écriture créative. Aux Noël, et particulièrement à Guylaine : tout a commencé quand vous nous avez accueillis dans la famille et n'a cessé de croître depuis. Merci pour toutes ces formes de support. À mon père, Thomas, merci pour toutes les discussions des dernières années, et pour m'avoir appris la valeur de la persévérance. Ta fierté est un de mes plus grands trésors. Merci à ma sœur Laurie-Ann. Tu dis ne pas partager ma passion pour l'*entique*, mais tu as quand même poussé plus fort que tout pour que je rédige. L'ultime récompense aura été cette bulle dans le temps, au Japon.

À ma grosse *gang* de Maillets, j'espère que vous savez à quel point vous êtes spéciaux. Non, dans le bon sens du terme. Ça prend une amitié en or pour accueillir de nouvelles personnes dans un groupe d'amis du secondaire. Mais ça prend un amour véritable et un respect profond pour encourager les ambitions de tous.

Enfin, mes derniers mots vont à mon partenaire des neuf dernières années, mon âme sœur, mon amour, Jason. Tu as appuyé inconditionnellement mes rêves et mon projet, conscient de tous les sacrifices que ça aura engendré. Tu as attendu que je franchisse cette étape de ma vie avec une patiente héroïque. Même au plus fort de la nuit, tu es toujours resté à mes côtés. Merci d'être une force inébranlable, un complice des bons et des moins bons moments, et un flot d'amour continu.

AVANT-PROPOS

Les principaux résultats de cette thèse ont fait l'objet de publications dans des journaux avec comité de lecture. Ils seront donc présentés sous la forme de manuscrits scientifiques en deux chapitres, les chapitres 2 et 3. Les deux articles, dont je suis la première auteure, sont reproduits avec l'autorisation des coauteurs et dans le respect des licences de diffusion. Ces deux chapitres sont précédés d'une introduction détaillée et suivis par une discussion générale.

Chapitre 2 : Les conséquences sur la résistance aux antibiotiques de l'architecture en mosaïque des plasmides de la famille de pAsa4. Ce chapitre reproduit l'article *The mosaic architecture of Aeromonas salmonicida subsp. salmonicida pAsa4 plasmid and its consequences on antibiotic resistance* qui a été publié dans le journal *PeerJ* le 27 octobre 2016. L'article original est accessible en ligne en suivant le permalien <http://dx.doi.org/107717/peerJ.2595>. Les auteurs de cet article sont Katherine H. Tanaka, Antony T. Vincent, Mélanie V. Trudel, Valérie E. Paquet, Michel Frenette et Steve J. Charette. J'ai participé à la conception des expériences, à leur réalisation, à la conception des moyens d'analyse, à l'analyse des résultats, à la conception et à la préparation des tableaux et des figures. J'ai rédigé l'ensemble de la première version du manuscrit et participé au polissage de la version finale. La participation de tous les coauteurs est explicitée dans l'article original.

Le corps du manuscrit, les tableaux et les figures sont reproduits dans leur intégralité. Les numéros de figures et de tableaux ont été modifiés pour respecter la facture du texte. Seule la Figure supplémentaire 1 est incluse dans la thèse. Le reste du matériel supplémentaire ainsi que les déclarations relatives à l'éthique et au financement sont accessibles dans l'article en ligne, en suivant le lien précédemment cité. Ce manuscrit est reproduit ici conformément à sa licence *Creative Commons Attribution 4.0 International* (<http://creativecommons.org/licenses/by/4.0>) qui en permet le partage libre moyennant attribution.

Chapitre 3 : Un nouveau patron de recombinaison des IS du plasmide pAsa5 chez *Aeromonas salmonicida* ssp. *salmonicida* 01-B526. Ce chapitre reproduit l'article *Plasmid composition in Aeromonas salmonicida subsp. salmonicida 01-B526 unravels unsuspected type three secretion system loss patterns* qui a été publié dans le journal *BMC Genomics* le 12 juillet 2017. L'article est accessible en ligne en suivant le permalien <http://dx.doi.org/10.1186/s12864-017-3921-1>. Les auteurs de cet article sont Katherine H. Tanaka, Antony T. Vincent, Jean-Guillaume Emond-Rheault, Marcin Adamczuk, Michel Frenette et Steve J. Charette. J'ai participé à la conception et à la réalisation des expériences, à l'analyse des résultats, à la conception et à la création des figures. J'ai rédigé l'ensemble de la première version du manuscrit et participé au polissage de la version finale. La participation de tous les coauteurs est explicitée dans l'article original.

Le corps du manuscrit, les tableaux et les figures sont reproduits dans leur intégralité. Les numéros de figures et de tableaux ont été modifiés pour respecter la facture du texte. Seuls les Fichiers additionnels 1 et 2 sont inclus dans la thèse. Le reste du matériel supplémentaire ainsi que les déclarations relatives à l'éthique et au financement sont accessibles dans l'article en ligne, en suivant le lien précédemment cité. Ce manuscrit est reproduit ici conformément à sa licence *Creative Commons Attribution 4.0 International*.

Chapitre 1 INTRODUCTION

1.1 LA GÉNOMIQUE, UN NOUVEAU PARADIGME DE L'ÉTUDE DU VIVANT

Comprendre le fonctionnement du vivant : voici le grand défi des sciences de la vie. Heureusement, le « code source » des organismes est extrêmement bien conservé dans une molécule bien connue, l'ADN. Décrypter les messages qu'elle porte est donc devenu un champ de recherche important et en constante expansion.

Les plus récentes avancées ont rendu accessible l'obtention du génome d'un organisme, c'est-à-dire l'ensemble de son matériel génétique. Ce fait autrefois extraordinaire (dix ans après la publication du premier génome, *Haemophilus influenzae*, 300 génomes bactériens avaient été séquencés) (Fleischmann *et al.*, 1995; Land *et al.*, 2015) est maintenant l'apanage de plus de 24 600 organismes vivants et virus, selon la classification de la base de données *Genomes* du *National Center for Biotechnology Information* (NCBI) (NCBI, 2014). Cette expansion de la quantité d'information génétique dans les dernières années est attribuable au développement et à la démocratisation de technologies permettant de générer des quantités massives de séquences nucléotidiques. Ces procédés techniques sont maintenant connus sous le nom de séquençage à haut débit.

Le terme « génomique », pour l'étude des éléments de génome(s), a été cerné par Thomas H. Roderick bien avant que ces technologies n'existent, soit en 1986 (Kuska, 1998). Toutefois, la génomique, tout comme les sciences « omiques » subséquentes, est devenue synonyme d'une vision intégrée d'un grand nombre d'éléments et des liens qui les rattachent (Kuska, 1998; Yadav, 2007). Les possibilités sont si nombreuses que la génomique contemporaine est elle-même divisée en plusieurs spécialisations qu'il serait impossible d'énumérer ici, mais qui font toutes appel aux données massives produites par le séquençage à haut débit. La génomique demande aussi l'utilisation d'une puissance computationnelle proportionnelle à la quantité de données traitée, et requiert l'utilisation ou la création de solutions bio-informatiques pour ordonner l'information et en tirer des conclusions (Land *et al.*, 2015; Vincent et Charette, 2015).

1.1.1 LA GÉNOMIQUE BACTÉRIENNE

Les bactéries ont des génomes relativement simples, courts et compacts comparativement à la plupart des eucaryotes (Ochman et Davalos, 2006). L'étude de leur génome est donc moins coûteuse en termes de temps et de ressources, et par conséquent plus accessible. Cela dit, la génomique bactérienne génère tout de même un flux important d'information et répond à des questions fort intéressantes. Par exemple, la génomique peut apporter une grande quantité d'information à l'avancement des connaissances sur la phylogénie, la relation gène/fonction et les mécanismes de l'évolution bactérienne (Bentley et Parkhill, 2015; Gray *et al.*, 2015; Klemm et Dougan, 2016). Cette discipline est aussi très utile à la lutte aux agents pathogènes. Par exemple, elle peut mener à l'identification de marqueurs pour l'épidémiologie ou l'évaluation de la virulence. La génomique est aussi une discipline tout indiquée pour étudier l'émergence et le transfert des gènes de résistance aux antibiotiques (Bayliss *et al.*, 2017). En effet, ces gènes sont fréquemment transférés horizontalement et retrouvés en association avec des éléments mobiles. La génomique permet de classifier les gènes impliqués dans les mécanismes de résistance et d'en déterminer le contexte (Beatson et Walker, 2014).

L'avancement de ces connaissances demande des ressources techniques, expérimentales et humaines pour le séquençage et l'analyse des génomes (Vincent *et al.*, 2017). Dans une première partie de cette introduction, les techniques de séquençage et d'analyse seront présentées. Une deuxième section de l'introduction traitera des éléments mobiles de l'ADN, puisqu'ils sont le sujet principal de cette thèse. Il sera aussi question du rôle de ce type d'élément génétique dans l'évolution bactérienne. Finalement, l'organisme modèle de l'étude, l'agent pathogène *Aeromonas salmonicida* sous-espèce (ssp.) *salmonicida*, sera présenté. Les éléments importants de son mode de vie pathogène et de son génome seront introduits.

1.2 LE SÉQUENÇAGE À HAUT DÉBIT

La génomique dépend des informations tirées du séquençage. Même si obtenir la séquence d'un brin d'ADN est possible depuis la fin des années 1970 (Sanger *et al.*, 1977), de

nombreuses améliorations ont été faites pour arriver au séquençage de génomes entiers. Les techniques les plus récentes, regroupées sous le terme de séquençage à haut débit, permettent de lire des millions de fragments d'ADN en parallèle, ce qui totalise jusqu'à des milliards de paires de bases lues par expérience. Ces techniques ont grandement révolutionné plusieurs domaines d'études en rendant accessible et abordable l'obtention d'une grande quantité d'information génétique. Cependant, elles requièrent des équipements spécialisés et font appel à des méthodes complexes qui peuvent introduire des biais et des erreurs (O'Rawe *et al.*, 2015).

Le séquençage à haut débit requiert des étapes à la fois expérimentales et bio-informatiques pour la planification, la préparation, le séquençage, l'assemblage et l'analyse des résultats (Vincent *et al.*, 2017). Brièvement, le séquençage d'un génome entier demande la préparation d'un échantillon de fragments d'ADN appelé librairie. La détermination de la séquence de chaque fragment se fait à l'aide d'un séquenceur et produit des lectures. Différents algorithmes sont ensuite utilisés pour assembler les lectures en des séquences plus longues, nommées contigs. Finalement, des étapes de finition comme l'annotation des séquences ou la jonction des contigs par des méthodes expérimentales sont effectuées (Vincent *et al.*, 2017).

Pour ce projet, le séquençage à haut débit a été utilisé afin d'étudier les gènes et l'architecture génomique de certains plasmides d'*A. salmonicida* ssp. *salmonicida*. Dans un tel cas, l'importance doit être accordée à la précision et à la longueur des lectures afin de produire un assemblage de qualité (Vincent *et al.*, 2017). Les séquences d'ADN répétées, qui peuvent être issues des éléments mobiles de l'ADN qui sont justement à l'étude, limitent l'achèvement de l'assemblage (Ricker *et al.*, 2012; Treangen et Salzberg, 2011). Les étapes qui influencent donc le résultat sont la préparation de l'ADN (la construction des librairies), la technique de séquençage, et les étapes d'assemblage et de finition. Les techniques pertinentes pour la suite de la thèse seront présentées ici.

1.2.1 COMPARAISON DU SÉQUENÇAGE À HAUT DÉBIT AVEC LA MÉTHODE SANGER

La première méthode de séquençage qui fut largement utilisée en génomique est le séquençage Sanger, nommé ainsi en raison de son créateur (Sanger *et al.*, 1977). Cette technique permet de déterminer la séquence d'un échantillon de molécules d'ADN identiques à l'aide d'une réaction de polymérisation modifiée. Les étapes d'assemblage et d'analyse sont comparables à son successeur, le séquençage à haut débit, mais la préparation des bibliothèques diffère (Land *et al.*, 2015).

Cette méthode a l'avantage d'être robuste, précise et de produire de plus longues lectures que certaines technologies à haut débit (1 kilobase [kb] avec les dernières technologies développées) (Heather et Chain, 2016). C'est elle qui a été utilisée pour les projets de séquençage de génomes jusqu'à l'avènement des techniques à haut débit. Toutefois, la méthode Sanger demande un temps de manipulation non négligeable lors de la création de la bibliothèque d'ADN, puisqu'elle n'accepte que des échantillons homogènes. De plus, elle produit substantiellement moins de nucléotides par réaction que les techniques plus récentes. Le séquençage de génome par méthode Sanger a donc été délaissé au profit des méthodes à haut débit (Land *et al.*, 2015; Metzker, 2010). Toutefois, la procédure reste utile pour déterminer de courtes séquences d'ADN purifié, comme des produits provenant de réactions de polymérisation en chaîne (PCR).

1.2.2 PRÉPARATION DE L'ADN

Toutes les techniques de séquençage présentées demandent des purifications d'ADN de grande qualité. Pour ce projet, puisque l'espèce bactérienne à séquencer était clairement définie et isolée, le matériel utilisé pour la purification est une culture pure de l'organisme en question. Cependant, la purification peut aussi être réalisée à partir de méthodes sans culture. L'ADN extrait est d'abord fragmenté en segments dont la taille dépend des méthodes qui seront utilisées (van Dijk *et al.*, 2014). Les fragments sont ensuite liés à des séquences d'ADN adaptatrices qui varient selon le séquenceur utilisé. Ces adaptateurs assurent la fixation des échantillons au support du séquenceur et la reconnaissance par la polymérase (Metzker, 2010). Finalement, il est recommandé, pour certaines techniques,

d'effectuer une amplification PCR des fragments avec adaptateurs afin d'augmenter la quantité de matériel génétique reconnue par le séquenceur (Vincent *et al.*, 2017). Bien que cette étape améliore les rendements, elle peut introduire des biais dans la représentativité des populations de fragments d'ADN de l'échantillon (O'Rawe *et al.*, 2015; van Dijk *et al.*, 2014).

Lors de la préparation de l'ADN, il est aussi possible de combiner fragments et adaptateurs afin de générer un type de lectures en particulier : les lectures à extrémités appariées (*paired-end reads*). Leur utilité est double. Certaines technologies, comme le séquençage Illumina, utilisent des lectures appariées sur de courts fragments d'ADN. Dans ce cas-ci, les lectures se chevauchent, ce qui augmente la qualité de la séquence (Vincent *et al.*, 2017). Il est aussi possible de créer des paires de lectures sur de plus longs fragments d'ADN à l'aide d'un adaptateur central. Dans ce cas, les lectures ne se chevauchent pas, mais elles sont physiquement liées, car elles proviennent d'un fragment d'ADN commun (Vincent *et al.*, 2017). Les informations de ce type d'appariement sont utilisées lors de l'assemblage pour associer des contigs lorsqu'ils partagent des lectures d'une même paire. Ceci permet d'éliminer les ambiguïtés causées par les éléments répétés (Wetzel *et al.*, 2011). Les suites de contigs jointes de cette façon se nomment des échafaudages (*scaffolds*).

1.2.3 LE PYROSÉQUENÇAGE

La première technique de séquençage haut débit qui a été utilisée dans les travaux de cette thèse est le pyroséquençage. C'est aussi la première technologie de ce type à avoir été développée (Margulies *et al.*, 2005). Pour cette technologie, le séquenceur suit des réactions de polymérisation en puits à l'aide de bioluminescence. Le signal lumineux est produit par cascade enzymatique à partir du pyrophosphate relâché lors de l'incorporation d'un nucléotide et est proportionnel au nombre de bases ajoutées. L'immobilisation et l'amplification de fragments d'ADN sur bille ainsi que les patrons de lavage de l'appareil permettent de générer un signal cohérent qui est ensuite converti en suites de nucléotides (Metzker, 2010).

Les réactions de pyroséquençage ont l'avantage de produire des lectures légèrement plus longues que d'autres technologies à haut débit (Tableau 1.1) (van Dijk *et al.*, 2014). Par

contre, les homopolymères (suites d'un même nucléotide) sont plus susceptibles aux erreurs en raison de la façon dont le signal lumineux est généré et interprété. Le pyroséquençage produit moins de nucléotides par expérience que les autres technologies à courtes lectures (Tableau 1.1) (van Dijk *et al.*, 2014). De plus, cette technologie n'est plus supportée par le fabricant des appareils (van Dijk *et al.*, 2014; Vincent *et al.*, 2014a). Son utilisation est donc en fort déclin.

1.2.4 SÉQUENÇAGE PAR SYNTHÈSE DE TYPE ILLUMINA

Le séquençage Illumina tient son nom de la compagnie qui commercialise cette technologie. Elle est devenue une des techniques de séquençage à haut débit les plus disponibles, les plus adaptables et les plus utilisées (Metzker, 2010; van Dijk *et al.*, 2014; Vincent *et al.*, 2017). Les différents séquenceurs Illumina fonctionnent sur le même principe réactionnel, mais produisent des longueurs de lecture et un nombre de bases par expérience différents. L'Illumina MiSeq offre une couverture suffisante pour le séquençage de génomes bactériens et ses lectures légèrement plus longues facilitent l'assemblage (Vincent *et al.*, 2017).

Comme le pyroséquençage, le séquençage Illumina convertit une source lumineuse en appel de bases lors d'une réaction de polymérisation. Cependant, la lumière y est générée par l'excitation de fluorophores de couleurs différentes liés aux nucléotides. Tous les nucléotides sont aussi modifiés par l'ajout d'un terminateur réversible. Ainsi, la polymérase incorpore une base par cycle de lavage même si les quatre nucléotides sont présents en même temps dans le mélange réactionnel. Les fragments sont immobilisés et amplifiés sur plaque afin de produire un signal suffisamment fort pour être détecté (Bentley *et al.*, 2008).

Le séquençage Illumina n'a pas le problème de résolution des homopolymères du pyroséquençage. Cependant, le temps d'expérience est plus long que pour d'autres séquenceurs (Tableau 1.1) (Buermans et den Dunnen, 2014). Il s'est cependant grandement amélioré, et la vitesse de réaction est compensée par la quantité et la qualité des données produites (van Dijk *et al.*, 2014).

1.2.5 SÉQUENÇAGE DE MOLÉCULE UNIQUE DE TYPE PACBIO

Le séquençage PacBio fonctionne de façon très différente des deux autres types présentés. Il se distingue par la longueur des lectures et par la méthode d'acquisition des bases. D'abord, les fragments d'ADN utilisés pour le PacBio peuvent être beaucoup plus longs que pour les autres méthodes, puisque les réactions peuvent séquencer jusqu'à 20 kb et même plus (van Dijk *et al.*, 2014). La méthode de purification employée en amont doit donc être assez douce pour éviter la fragmentation des brins d'ADN et doit produire plus de matériel. Aucune amplification PCR n'est requise lors de la préparation de la librairie ou de l'immobilisation du produit, ce qui élimine un type de biais causé par l'amplification réduite des séquences GC ou AT-riches (Metzker, 2010).

Le séquençage PacBio utilise des nucléotides modifiés par l'ajout d'un fluorophore comme substrat à une réaction de polymérisation. Cependant, le système d'excitation-détection du séquenceur permet de suivre l'activité d'une série de puits microscopiques en continu et en simultané. Puisque le signal d'une seule réaction de polymérisation est suffisant pour inférer l'ordre des bases, la longueur de la lecture n'est limitée que par l'efficacité de la polymérase, ce qui permet d'obtenir des séquences plus longues. Une réaction sur séquenceur PacBio est, de plus, très rapide (Tableau 1.1) (Metzker, 2010; van Dijk *et al.*, 2014).

Un des grands désavantages du séquençage PacBio est son taux d'erreur, qui varie entre 10 à 15% par base (van Dijk *et al.*, 2014). Une grande partie de ces erreurs peut être éliminée en faisant des librairies de fragments plus courts. Chaque fragment sera ensuite lu plusieurs fois par la même polymérase. L'algorithme d'appel des bases de PacBio fait le consensus des différents passages. Il est possible d'obtenir un génome assemblé avec un taux de confiance de 99,999% en utilisant cette méthode (Chin *et al.*, 2013). Une autre façon de compenser le taux d'erreur de cette technique est de combiner les longues lectures Pacbio avec des lectures plus courtes et produites par une autre technique (ex. : Illumina) qui donne un plus faible taux d'erreur (Koren *et al.*, 2013).

Tableau 1.1 Comparaison des techniques de séquençage utilisées au cours de la thèse

	Pyroséquençage ^a	Illumina ^{a,b}	PacBio ^{a,c}
Étapes de préparation			
Fragmentation	Mécanique, enzymatique ou tagmentation	Mécanique, enzymatique ou tagmentation	Mécanique Grands fragments nécessaires
Extrémités appariées	Possible (avec adaptateur au centre)	Possible (courts ou long)	Impossible
Amplification pré-fixation	Possible	Possible	Impossible
Fixation des échantillons	Billes dans puits	Fragments sur une plaque	Aucune, c'est la polymérase qui est fixe
Formation des polonies (grappes de séquence identiques)	PCR en émulsion	Amplification en pont sur la plaque	Aucune : signal attribuable à une seule molécule
Réaction de séquençage			
Modification des bases	Aucune	Fluorophore spécifique à la base Terminateur réversible en 3'	Fluorophore spécifique à la base
Distribution des réactifs	Lavage d'une base à la fois	Lavage de toutes les bases, puis clivage terminateur	Tous les éléments dans le mélange réactionnel
Signal détecté	Fluorescence des polonies (détection du pyrophosphate relâché)	Fluorescence des polonies (détection du fluorophore)	Fluorescence d'une seule molécule détectée par microscope confocal
Mode d'acquisition	Photo à chaque cycle	Photo à chaque cycle	Acquisition continue et puits-indépendante
Finalité			
Longueur de lectures	1000 bases	2 x 300 bases	Médiane de 20 kb
Bases par expérience	0,7 Gb	13,2 à 15 Gb	500 Mb à 1 Gb
Temps de l'expérience	10 heures	56 heures	4 heures

^a Informations tirées de Metzker (2010)

^b Performances suggérées pour un appareil Illumina MiSeq (utilisé au cours du projet) <http://www.illumina.com/systems/sequencing-platforms/miseq/specifications.html>

^c Performances suggérées pour un appareil PacBio RS II http://www.pacb.com/wp-content/uploads/2015/09/PacBio_RS_II_Brochure.pdf

1.2.6 ASSEMBLAGE ET ANALYSE DES SÉQUENCES

Les différents séquenceurs fournissent comme résultat des listes de lectures accompagnées d'un indice de qualité pour chaque nucléotide prédit. La qualité étant de mise pour

l'assemblage *de novo* de génomes, un premier algorithme est utilisé afin de filtrer les lectures trop courtes et les nucléotides de mauvaise qualité (Vincent *et al.*, 2017). Le résultat est ensuite pris en charge par un assembleur, un programme requérant un algorithme complexe (Vincent *et al.*, 2017). Ce dernier cherche des similitudes entre les lectures et les regroupe sous forme de contigs, qui sont des consensus de lectures. La validité de chaque base d'un contig est appuyée par un nombre n de lectures, ce qui représente la couverture à cet endroit donné.

Cependant, des ambiguïtés impossibles à résoudre surviennent autour des séquences répétées. En effet, il est impossible d'assembler avec certitude les fragments qui partagent des bases avec la répétition si la longueur des lectures est plus petite que celle-ci. L'assemblage du contig se termine donc à l'extrémité de la séquence répétée. Par conséquent, un génome sera divisé en un nombre de contigs qui dépend non seulement de sa taille, mais surtout de son nombre d'éléments répétés. Les ARN ribosomiaux, les ARN de transfert et certains éléments mobiles de l'ADN sont fréquemment présents en plusieurs copies dans un génome et fragmentent l'assemblage (Ricker *et al.*, 2012).

Plusieurs solutions permettent de contourner ce problème. Par exemple une librairie à extrémités appariées dans laquelle les lectures sont suffisamment éloignées permettra d'associer les séquences flanquant les différentes copies de la répétition (voir section 1.2.2). Il est aussi possible d'inférer l'ordre des contigs en réalisant un assemblage à partir d'un génome de référence, puis en vérifiant expérimentalement la présence des jonctions par PCR (Vincent *et al.*, 2017). Cette étape requiert toutefois un génome assemblé et fermé au préalable. Il faut aussi que cet assemblage soit fiable et que le génome assemblé soit un représentant typique pour l'espèce en question.

L'utilisation des données de séquençage et la finition acceptable d'un génome dépendent toujours de la question de recherche qui avait été posée au départ. Parfois, une comparaison entre contigs, ce qui met en évidence les éléments supplémentaires ou manquants entre les génomes, est suffisante. L'étude de l'architecture des génomes demande un assemblage et une finition manuelle complète ou partielle, là où le temps et les ressources ne doivent pas être sous-estimés (Vincent *et al.*, 2017). Finalement, les séquences obtenues doivent généralement être annotées afin d'y trouver une pertinence biologique. Plusieurs pipelines

automatisés et bases de données sont disponibles à ce sujet, mais doivent être utilisés consciencieusement par le génomicien en regard à la qualité de l'information présente.

1.3 LES ÉLÉMENTS MOBILES DE L'ADN

Contrairement à d'autres organismes, les bactéries ne bénéficient pas du brassage de gènes produit lors de la reproduction sexuée. Elles doivent donc compter sur d'autres moyens pour produire l'innovation génétique nécessaire à leur évolution (Treangen et Rocha, 2011). En effet, un modèle évolutif basé exclusivement sur l'apparition de mutations dans une population asexuée est voué à l'échec (Lynch, 2007).

Toutefois, les bactéries utilisent une série de tactiques regroupées sous le nom de transfert horizontal afin d'échanger du matériel génétique au sein d'une même espèce ou entre celles-ci. Elles sont au nombre de trois : la transformation, qui est l'intégration d'ADN nu extracellulaire, la conjugaison, qui est le transfert d'ADN intercellulaire par contact direct, et la transduction, qui est le transfert d'ADN par l'intermédiaire de bactériophages (Snyder et Champness, 2007). Le transfert horizontal contribue à l'évolution bactérienne par la propagation rapide de nouveaux phénotypes complexes (Hao, 2012; Jackson *et al.*, 2011). Les grandes populations de plusieurs bactéries et leur propension à partager des environnements avec d'autres espèces rendent cette stratégie viable (Shapiro, 2014). Un des champs d'application de la génomique est la détection et la caractérisation des événements de transfert horizontal.

Les éléments mobiles de l'ADN sont d'autres déterminants importants pour l'évolution des génomes bactériens (Jackson *et al.*, 2011). Ces séquences d'ADN sont capables de transfert intracellulaire ou intercellulaire, et existent dans tous les domaines de la vie (Frost *et al.*, 2005). Les éléments mobiles à propagation intercellulaire sont donc par définition impliqués dans le transfert horizontal de gènes. Les éléments mobiles de l'ADN à transfert intracellulaire sont eux aussi importants pour l'évolution bactérienne puisqu'ils se greffent à des éléments mobiles à propagation intercellulaire. Dans ce cas-ci, c'est leur capacité à transporter des gènes accessoires ou à favoriser des recombinaisons dans le génome qui leur permettent de générer de la diversité (Frost *et al.*, 2005; Toussaint et Merlin, 2002). La génomique permet de détecter et d'étudier l'impact des éléments mobiles de l'ADN sur la

plasticité du génome à condition d'utiliser des méthodes adéquates (Beatson et Walker, 2014; Ricker *et al.*, 2012). La présente section se concentrera sur les éléments mobiles bactériens qui ont été rencontrés au cours du projet.

1.3.1 LES TRANSPOSONS

Le transposon (Tn) est un élément mobile de l'ADN capable de transfert intracellulaire. C'est la transposase, protéine modificatrice de l'ADN dont le gène est encodé sur l'élément mobile, qui permet son déplacement. Afin d'aiguiller correctement leur transposase dans sa fonction, les transposons sont généralement bordés par des séquences inversées répétées qui sont reconnues par l'enzyme (Figure 1.1A). Certains de ces éléments mobiles portent aussi d'autres gènes codant pour la régulation de la transposition. Les transposons peuvent changer d'emplacement sur l'ADN par plusieurs mécanismes, comme la transposition conservative et la transposition répllicative (Snyder et Champness, 2007).

En plus de leurs gènes caractéristiques, certains transposons portent aussi des gènes accessoires, tels des gènes de résistance aux antibiotiques (Figure 1.1A). Au contraire, la forme la plus simple, mais autonome de transposon ne possède qu'une transposase ainsi que des répétitions inversées (Figure 1.1B) (Snyder et Champness, 2007). Elle est nommée séquence d'insertion (IS) et fera l'objet d'une section plus loin (section 1.3.5).

Une catégorie spéciale de transposon, le transposon composite, fait intervenir deux IS identiques en lieu d'extrémités (Figure 1.1C). Une des IS porte la transposase fonctionnelle, l'autre copie étant le plus souvent inactivée, et ce sont les extrémités inversées des deux IS qui sont reconnues lors de la transposition. Ce type de structure permet de transposer efficacement de grands fragments d'ADN (Snyder et Champness, 2007).

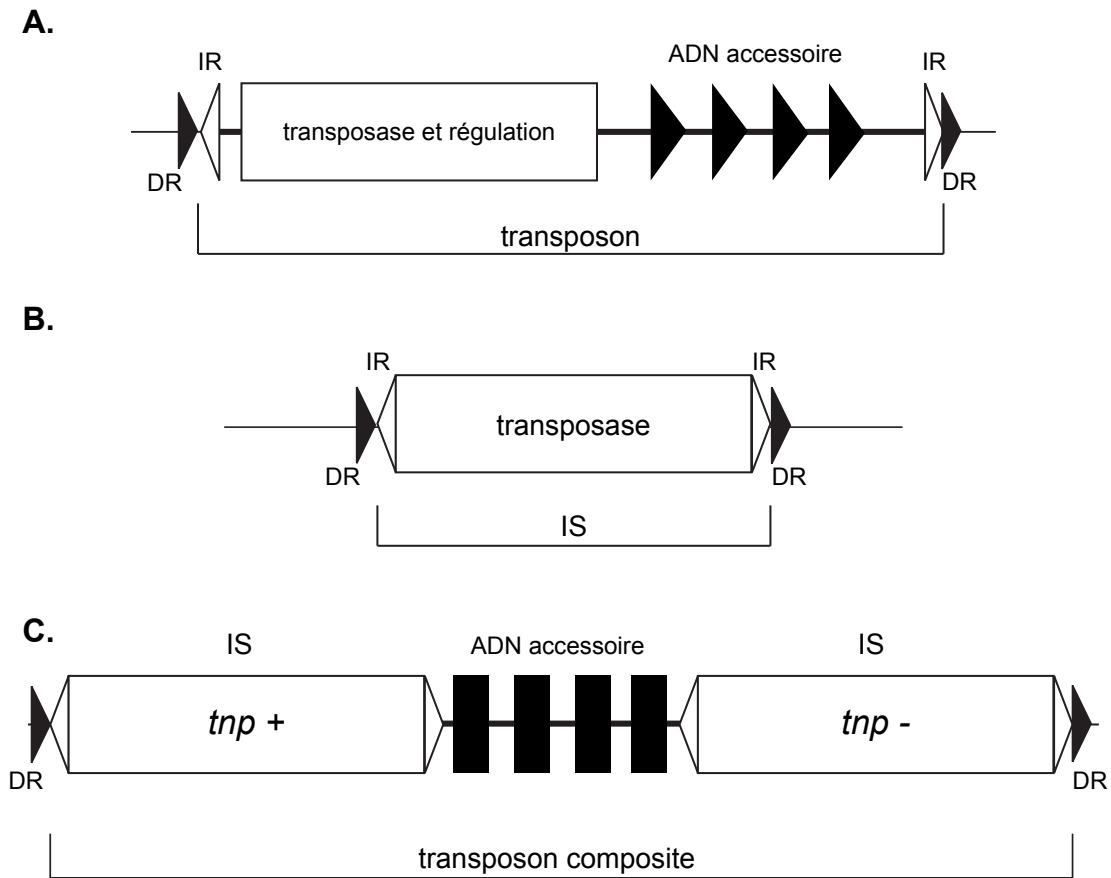


Figure 1.1 Différents types de transposons. **A.** Le transposon non composite est habituellement délimité par des répétitions inversées (IR, *inverted repeats*). Il contient les gènes nécessaires à sa transposition et à la régulation de celle-ci, en plus de gènes accessoires (l'orientation est arbitrairement représentée dans cette figure). Les répétitions directes (DR, *direct repeats*) sont le produit de la réparation de l'ADN après la transposition et ne font pas partie du transposon. **B.** La séquence d'insertion (IS, *insertion sequence*) est un court transposon ne portant pas de gène accessoire. Sa structure condensée comprend le gène de la transposase et des IR. **C.** Le transposon composite se forme lorsqu'une suite de gènes accessoires est encadrée par deux copies de la même IS. Habituellement, la transposase d'une des deux IS est active (*tnp +*), alors que l'autre est inactive (*tnp -*), ce qui constitue une forme de régulation de la transposition (Snyder et Champness, 2007).

1.3.2 LES INTÉGRONS

Les intégrons (In) sont des éléments spécialisés dans l'accumulation de gènes accessoires. Ils ne sont pas capables de transfert intra- ou intercellulaire, mais sont associés à d'autres

éléments mobiles (Domingues *et al.*, 2015). Leur bon fonctionnement dépend d'une intégrase, dont le gène est encodé dans la section 5' de l'élément, d'un site d'intégration nommé *attI* et d'un promoteur responsable de l'expression de tous les gènes de la section variable (Domingues *et al.*, 2012). Cette région est formée lorsque l'intégrase ajoute des cassettes à sa structure d'expression par recombinaison avec le site *attI*. Les cassettes sont de courts fragments d'ADN circulaire portant un gène sans promoteur et un site de recombinaison, nommé *attC*, compatible avec *attI* (Hall et Collis, 1995). Les cassettes intégrées les unes à la suite de l'autre sont séparées par des répétitions partielles d'*attC* qui se créent lors de la recombinaison d'*attI* avec *attC* (Hall et Collis, 1995). L'intégrase peut aussi, par une série de recombinaison, changer l'ordre des cassettes de l'intégron (Cambray *et al.*, 2010). Les intégrons de classe 1 sont d'une importance particulière puisqu'ils sont les plus fréquemment retrouvés dans les génomes d'agents pathogènes (Cambray *et al.*, 2010).

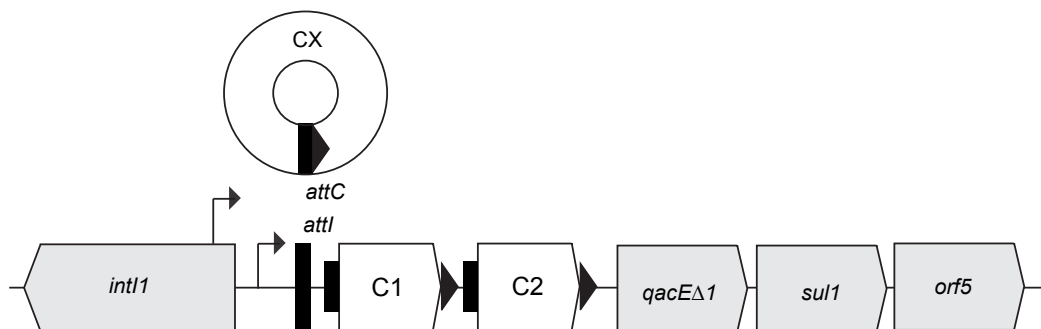


Figure 1.2 Schéma d'un intégron de classe 1. Les différents gènes d'un intégron de classe 1 sont représentés par les boîtes orientées. L'intégrase (*intI1*) assure une recombinaison homologue entre le site *attI* et le site *attC* d'une cassette circulaire. Elle sera intégrée en aval d'*attI* et en amont de C1, la cassette la plus récente. Les cassettes peuvent être transcrites à partir de deux promoteurs (flèches). Les trois éléments en gris en 3' de l'intégron (*qacEΔ1*, version tronquée d'un gène qui confère une résistance à certains détergents, *sul1*, qui confère une résistance aux sulfamides, et *orf5*, un cadre de lecture ouvert) ne font pas partie de la structure d'expression des cassettes, mais sont présents chez la majorité des intégrons de classe 1 (Domingues *et al.*, 2015).

1.3.2.1 ÉTUDE DE CAS : LE TRANSPOSON Tn21

Tn21 est un transposon non composite (tel qu'illustré à la Figure 1.1A) qui possède à la base ses gènes de transposition et un opéron codant pour la résistance au mercure. De plus, un intégron de classe 1, In2, est inséré dans Tn21, augmentant les combinaisons de gènes de résistance possibles pour ce transposon. L'intégron possède le gène de son intégrase, ainsi qu'une cassette *aadA1*, qui code pour une résistance à la streptomycine et à la spectinomycine. Deux autres éléments en aval de la structure d'expression, *qacEΔ1* et *sul1*, ne sont pas présents sous forme de cassettes. Ils confèrent des résistances aux composés d'ammoniums quaternaires et aux sulfamides, respectivement. Finalement, les gènes de transposition putatifs d'In2 sont non fonctionnels, rendant cet élément mobile dépendant de la transposition de Tn21 (Figure 1.2) (Liebert *et al.*, 1999).

Tn21 et son intégron ont été fréquemment étudiés et ils ont été retrouvés sur de nombreux plasmides (Liebert *et al.*, 1999; Partridge *et al.*, 2001). Ce transposon semble aussi être un site d'insertion d'autres éléments mobiles de l'ADN, comme des IS, ce qui augmente encore sa capacité à générer de la diversité. Tn21 est donc un agent important d'accumulation et de dissémination de résistances aux antimicrobiens (Gillings *et al.*, 2017; Liebert *et al.*, 1999).

1.3.3 LES ÎLOTS GÉNOMIQUES

Des preuves d'intégration d'ADN issu de transfert horizontal sont parfois trouvées dans les chromosomes bactériens. Ces éléments sont regroupés sous le terme d'îlots génomiques. La bio-informatique et la génomique ont permis d'enrichir le répertoire d'îlots génomiques au cours des dernières années grâce à la détection de signatures typiques à ces éléments d'ADN, comme un contenu en G/C différent, la présence de nombreux gènes accessoires regroupés sous forme de loci et des sites d'insertion spécifiques (Dobrindt *et al.*, 2004).

Les îlots génomiques sont le produit évolutif de deux voies d'intégration principales : les bactériophages et les plasmides (Bellanger *et al.*, 2011; Dobrindt *et al.*, 2004; Novick *et al.*, 2010). Dépendant de leur degré de dégradation, certains îlots conservent des fonctions

relatives à leur élément ancestral. Dans ce cas, un nom plus proche de leur réalité ou dépendant des fonctions qu'ils encodent peut leur être attribué (Carraro *et al.*, 2017).

1.3.4 LES PLASMIDES

Ces molécules autorépliquatives et généralement circulaires, qui ont été trouvées dans toutes les familles bactériennes, sont des éléments majeurs du génome accessoire (Snyder et Champness, 2007). Les plasmides sont des éléments mobiles de l'ADN capables de transfert intercellulaire, mais qui sont aussi porteurs d'éléments à transfert intracellulaire. Les plasmides sont donc des vecteurs de propagation des IS et des transposons, ainsi que de gènes accessoires biologiquement pertinents, comme des gènes de résistance aux antibiotiques (Jackson *et al.*, 2011).

Les plasmides sont pourvus de leur propre système de réplication et de transmission aux cellules-filles (par leur nombre de copies ou par leur système de ségrégation) (Snyder et Champness, 2007). Puisque la régulation des plasmides se fait par le contrôle de leur réplication et parfois de leur partage, deux plasmides ayant exactement l'une ou l'autre des machineries à l'identique seront incompatibles, ce qui signifie qu'ils ne pourront persister dans la même cellule. L'incompatibilité est une des façons de classifier les plasmides (Snyder et Champness, 2007).

L'apport de gènes accessoires par les plasmides est généralement bénéfique pour la bactérie, ce qui favorise leur maintien dans la cellule. Cependant, certains plasmides portent des mécanismes de dépendance, comme les systèmes toxine-antitoxine. Ces modules produisent deux éléments exprimés constitutivement: une toxine et une antitoxine qui neutralise l'action de la précédente. Puisque la toxine a une durée de vie plus longue que l'antitoxine, une perte du système de dépendance entraîne une accumulation de toxines non neutralisées qui cause la mort de la cellule (Darmon et Leach, 2014; Hayes, 2003). Ce mécanisme permet aux éléments génétiques qui portent des systèmes toxine-antitoxine, comme les plasmides, d'être conservés par la bactérie et d'être transmis correctement aux cellules-filles (Darmon et Leach, 2014; Hayes, 2003; Pandey et Gerdes, 2005). Ensemble, les systèmes toxines-antitoxines et les systèmes de stabilité permettent un transfert adéquat du plasmide de cellule-mère aux cellules-filles. Cependant, ils favorisent aussi le maintien

des plasmides malgré une pression de sélection qui tendrait à les supprimer de la cellule, puisqu'ils s'ajoutent au coût métabolique (Baltrus, 2013; Hayes, 2003).

Finalement, les plasmides sont très importants pour l'évolution des génomes bactériens, car ils sont capables de transfert intercellulaire. Pour se faire, les plasmides conjugatifs codent pour la machinerie nécessaire à la reconnaissance entre cellules, à leur arrimage (Mpf, *mating pair formation*) et au transfert de molécules d'ADN (Dtr, *DNA transfer and replication*). Chez les bactéries à Gram négatif, le Mpf s'apparente au système de sécrétion de type IV et est caractérisé par un pilus, la structure permettant la proximité entre cellule donatrice et réceptrice. Certains plasmides ne possèdent pas de système de Mpf, mais seulement un Dtr. Ils sont dits mobilisables (Snyder et Champness, 2007). Certains plasmides possèdent de larges spectres d'hôtes, indiquant que la conjugaison peut se faire entre bactéries distantes évolutivement (Musovic *et al.*, 2006; Petersen et Wagner-Dobler, 2017).

Le séquençage à haut débit a apporté une nouvelle mine d'information pour l'étude de la distribution, du contenu et de l'architecture des plasmides. Cependant, la séquence même de ces éléments ne peut apporter que des informations limitées sur leur classification et leur histoire évolutive. En effet, puisque les plasmides subissent fréquemment des recombinaisons, leur architecture est composée d'une mosaïque de plusieurs modules qui peuvent provenir d'origines variées (Hall *et al.*, 2015; Petersen, 2011).

1.3.5 LES SÉQUENCES D'INSERTION

Les IS se distinguent des autres transposons précédemment décrits par leur absence de gènes accessoires (Mahillon et Chandler, 1998). Leur taille se situe généralement entre 700 et 2 500 paires de bases (pb) (Siguier *et al.*, 2014; Siguier *et al.*, 2015). Leur structure est condensée ; un ou plusieurs cadres de lectures ouverts codent pour la transposase de l'IS et sont encadrés par des répétitions inversées imparfaites (IR, *inverted repeats*) (Figure 1.1B) (Mahillon et Chandler, 1998). Ces répétitions sont reconnues lors de l'attachement et du clivage de l'ADN par la transposase (Mahillon et Chandler, 1998). Tout comme chez le transposon, la transposition engendre généralement une courte répétition directe (de 2 à 14 pb) de part et d'autre de l'IS à la suite de la réparation de la séquence au site d'insertion

(Figure 1.1B) (Mahillon et Chandler, 1998). Ces répétitions directes ne font pas partie de l'élément mobile, mais sont caractéristiques de l'IS impliquée et peuvent donc aider à sa détection (Mahillon et Chandler, 1998).

Les IS peuvent être rassemblées en familles, puis en sous-groupes sur la base de leur transposase, de leur taille et de leurs répétitions (Siguier *et al.*, 2015). Le nombre d'IS documentées a grandement augmenté depuis leur description originale (Siguier *et al.*, 2015). Une base de données, ISFinder, consigne ces séquences depuis 1998, mais n'a accès qu'à une fraction des IS, puisqu'elles ne sont pas annotées automatiquement lors du séquençage à haut débit (Siguier *et al.*, 2015; Siguier *et al.*, 2006b).

1.3.5.1 LES TRANSPOSASES

Les transposases des IS peuvent être regroupées selon les acides aminés de leur site catalytique et la réaction qu'ils engendrent (Siguier *et al.*, 2015; Snyder et Champness, 2007). Deux types de transposition sont essentiellement produits : la transposition conservative (nombre de copies stable) et la transposition réplivative (augmentation du nombre de copies) (Siguier *et al.*, 2015; Snyder et Champness, 2007). Certains groupes, comme les transposases DDE (site catalytique composé de deux acides aspartiques [D] et un acide glutamique [E]) possèdent des membres qui peuvent catalyser l'un ou l'autre des modes de transposition (Siguier *et al.*, 2015).

La transposition doit être étroitement régulée, puisqu'elle peut mener à une accumulation de mutations chez l'organisme hôte et ultimement à son extinction (Doolittle *et al.*, 1984; Gueguen *et al.*, 2005). Cependant, les stratégies de régulation des IS sont condensées dans la structure compacte de cet élément mobile (Mahillon et Chandler, 1998). Par exemple, la position du promoteur de la transposase dans l'IR en amont de celle-ci pourrait permettre une autorégulation de la transcription du gène (Mahillon et Chandler, 1998). Aussi, les transposases de certaines familles d'IS (comme IS3) sont encodées sur deux cadres de lecture, ce qui demande un changement du cadre de lecture lors de la traduction et limite le nombre de protéines fonctionnelles produites (Mahillon et Chandler, 1998).

1.3.5.2 RÉPLICATION EN CERCLE ROULANT : LE CAS PARTICULIER D'IS91 ET LES ISCR

Les IS de la famille IS91 ont une structure et un mode de fonctionnement particuliers. Leur transposase utilise des tyrosines comme acides aminés catalytiques et produisent une transposition par réplication en cercle roulant (Bernales *et al.*, 1999). Leurs extrémités ne sont pas des répétitions inversées, mais sont tout de même le site de reconnaissance de la transposase (Siguier *et al.*, 2015). Leur mode de propagation peut amener la transposition accidentelle de séquences adjacentes par reconnaissance d'un site de terminaison erroné pour la réplication en cercle roulant (Tavakoli *et al.*, 2000). Ces IS ont été retrouvées préférentiellement aux abords de gènes de virulence (Garcillán-Barcia et de la Cruz, 2002). Certaines structures, les ISCR (*insertion sequence common region*) possèdent une transposase similaire à celle d'IS91, mais sont accompagnées de gène(s) de résistance aux antibiotiques. Elles sont trouvées en association avec des intégrons de classe 1 (Toleman *et al.*, 2006b).

1.3.5.3 IMPACT DES IS SUR LES GÈNES ET LES GÉNOMES

Même sans conférer de phénotype particulier, les IS peuvent avoir des effets profonds sur leurs hôtes. La perturbation la plus simple qui peut être causée par une IS est l'interruption d'un gène ou d'un opéron par insertion dans la séquence codante ou dans les séquences régulatrices (Figure 1.3B). Ce type de mutation est difficilement détectable, sauf lorsque l'interruption du gène produit une variation de phénotype évidente (Vandecraen *et al.*, 2017). Au contraire, certaines IS peuvent activer des gènes, car elles possèdent des promoteur qui initient une transcription vers l'ADN avoisinant (Mahillon et Chandler, 1998; Vandecraen *et al.*, 2017).

En plus d'avoir un impact ponctuel sur les gènes avoisinants, la présence de plusieurs copies d'une même IS dans un génome peut avoir des conséquences sur l'architecture. Elles peuvent être ciblées par la machinerie de recombinaison homologue bactérienne et engendrer des réarrangements (Siguier *et al.*, 2014; Vandecraen *et al.*, 2017). Des fusions seront produites si les IS sont situées sur des molécules d'ADN différentes, des délétions si les IS sont sur la même molécule et dans le même sens, et des inversions si les IS sont en sens contraire (Figure 1.3A) (Parkhill *et al.*, 2003b). L'intégration du plasmide F dans le

chromosome d'*Escherichia coli* par recombinaison des IS3 est un exemple classique de fusion (Deonier et Hadley, 1980). Les délétions de grands fragments d'ADN par recombinaison sont régulièrement observées dans le processus de réduction des génomes d'organismes pathogènes ou symbiotes obligatoires (Cole *et al.*, 2001; McCutcheon et Moran, 2012). Les inversions, quant à elles, ne changent pas le contenu en gène. Cependant, elles peuvent en perturber la régulation et même rendre cryptiques des gènes ou des portions d'opérons (Zinser *et al.*, 2003).

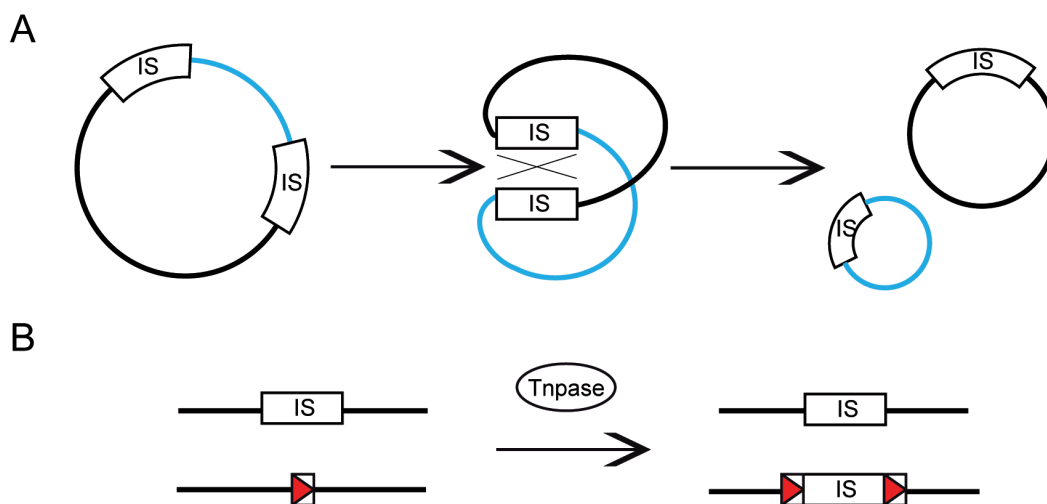


Figure 1.3 Deux exemples de perturbation impliquant les IS. **A.** Une recombinaison homologue entre deux copies de la même IS sur un élément d'ADN circulaire mène à un fractionnement de cette molécule si les IS sont dans le même sens (dans le cas contraire, une inversion aura lieu). La section incapable de répliquer sera perdue, menant à une délétion. **B.** La transposition d'une IS fait intervenir une protéine, la transposase (Tnpase), encodée sur cet élément. Cette enzyme reconnaît un site sur un fragment d'ADN et y déplace l'IS, dupliquant le site (ici, une transposition répliquative est schématisée). L'insertion d'une IS dans un gène ou une séquence régulatrice peut en perturber le fonctionnement. Cette figure est une reproduction de la Figure 2 de Tanaka *et al.* (2013), reproduite ici selon sa licence *Creative Commons Attribution-NonCommercial 3.0 Unported License* (<http://creativecommons.org/licenses/by-nc/3.0>). L'article original est disponible en ligne à l'adresse suivante : (<http://doi.org/10.4161/mge.23498>).

1.3.5.4 LA THÉORIE DE RÉDUCTION DES GÉNOMES

La prolifération d'IS et d'autres éléments mobiles est observée dans les génomes d'organismes pathogènes. Les petites populations et l'abondance des nutriments dans le microenvironnement créé par l'hôte limitent les effets de la sélection et amplifient les conséquences de la dérive génétique (l'accumulation de mutations aléatoires) (Ochman et Davalos, 2006). Plusieurs gènes codant pour des fonctions accessoires deviennent des pseudogènes et les répétitions d'éléments mobiles contribuent à l'instabilité du génome (Ochman et Davalos, 2006). Ultiment, les éléments mobiles restants et les pseudogènes sont érodés, et le génome résultant, minimal, rend l'organisme dépendant de son hôte (McCutcheon et Moran, 2012; Ochman et Davalos, 2006). Des exemples de bactéries se trouvant à un degré intermédiaire de réduction, et possédant donc encore des éléments mobiles et des pseudogènes, incluent *Yersinia pestis* (Chain *et al.*, 2004), *Lactobacillus helveticus* (Kaleta *et al.*, 2010), *Bordetella parapertussis* (Parkhill *et al.*, 2003a) et *Aliivibrio salmonicida* (Hjerde *et al.*, 2008).

1.3.5.5 PARADIGME DE LA COÉVOLUTION IS-BACTÉRIES

Les IS sont responsables de leur propre propagation au sein du génome et bénéficient de la présence d'autres éléments mobiles ubiquitaires, les plasmides, pour leur transfert intercellulaire. Leur contrôle échappe donc à l'organisme qui les héberge, bien qu'elles puissent avoir un effet délétère sur la cellule (par inactivation de gènes essentiels ou diminution de la valeur adaptative, par exemple). De plus, étant des séquences ni codantes, ni régulatrices en soi, leur présence en de multiples copies dans un génome représente un fardeau métabolique en terme de quantité d'ADN supplémentaire à répliquer (Doolittle et Sapienza, 1980). Ces constats nourrissent une vision des IS et autres transposons cryptiques comme éléments d'ADN égoïstes (*selfish DNA*), des parasites se propageant au détriment (toutefois léger) de l'organisme-hôte (Doolittle et Sapienza, 1980; Orgel et Crick, 1980). Dans cette vision, il se forme un équilibre entre la répllication des éléments d'ADN égoïstes et leur inactivation (Orgel et Crick, 1980).

Une contribution purement négative des IS au génome et à la valeur adaptative des bactéries explique toutefois mal les informations rassemblées sur ces éléments d'ADN à ce

jour : leur ubiquité dans les génomes bactériens et leur nombre parfois important. Plusieurs autres explications ont donc été formulées afin de raffiner ou d'infirmier ce modèle.

Certaines études utilisent les théories et les modèles mathématiques de la génétique des populations pour expliquer la présence des IS dans les génomes bactériens. La grande similarité entre copies de même IS au sein d'un génome pourrait indiquer une acquisition récente et que la sélection effacerait les copies d'IS avant qu'elles n'aient le temps de diverger. Dans ce modèle, des séries d'acquisitions/multiplications, suivies d'érosions/extinctions expliqueraient bien la présence d'IS complètes et partielles dans les génomes bactériens (Cerveau *et al.*, 2011; Touchon et Rocha, 2007; Wagner, 2006; Wu *et al.*, 2015). La mobilisation des IS par transfert horizontal de plasmides et d'autres éléments mobiles est nécessaire à leur survie dans ce modèle (Nelson *et al.*, 2011; Touchon et Rocha, 2007). Une conséquence observable de ce phénomène est une densité d'IS plus forte dans les plasmides (15 à 20% de la séquence des plasmides de plus de 20 kilopaires de bases [kpb]) que dans les chromosomes bactériens (3%) (Siguier *et al.*, 2006a; Siguier *et al.*, 2014; Siguier *et al.*, 2015).

Des études ont aussi montré que le coût métabolique associé aux nucléotides supplémentaires des IS était négligeable, mais que leur nombre était plutôt gouverné par le risque d'induire des mutations délétères (Touchon et Rocha, 2007). Cependant, dans la majorité des cas, les IS ont plutôt un apport neutre sur la valeur adaptative bactérienne. Dans ce cas, leur nombre est davantage modulé par un état d'équilibre entre duplications et délétions plutôt que par une sélection purificatrice (Iranzo *et al.*, 2014).

Une théorie étudiée est que les IS apportent une plasticité génomique qui est bénéfique. Les évènements produits par les IS auraient des conséquences favorables sur la valeur adaptative de la bactérie qui les portent. Des études montrant le rôle des IS sur l'augmentation de résistances aux antibiotiques, sur la modulation de la virulence, sur la production de biofilm et sur l'adaptation au système immunitaire de l'hôte en donnent des exemples (révisé dans (Vandecraen *et al.*, 2017)). En fait, peu importe les hypothèses sur la persistance des IS qui sont invoquées, il n'est pas possible de nier qu'elles apportent des variations génétiques bénéfiques dans certains cas.

1.4 L'ORGANISME À L'ÉTUDE : *AEROMONAS SALMONICIDA* SSP. *SALMONICIDA*

Les bactéries du genre *Aeromonas* appartiennent à la classe des Gammaprotéobactéries, à l'ordre des *Aeromonadales* et à la famille des *Aeromonadaceae*. Ce sont des bâtonnets ou des coccobacilles à Gram négatif anaérobies facultatives. Dépendant des espèces, elles peuvent être motiles et sont psychrophiles ou mésophiles (Cipriano et Austin, 2011; Janda et Abbott, 2010; Martin-Carnahan et Joseph, 2005). Les bactéries du genre *Aeromonas* sont ubiquitaires des environnements aquatiques (Cipriano et Austin, 2011; Janda et Abbott, 2010; Martin-Carnahan et Joseph, 2005). Elles ont été détectées dans des eaux douces, saumâtres et dans les cours d'eau pollués, ainsi que dans des systèmes d'eaux potables, d'eaux usées et sur des aliments destinés à la consommation (Martin-Carnahan et Joseph, 2005; Martínez-Murcia *et al.*, 2016; Nagar *et al.*, 2011; Pavan *et al.*, 2015; Piotrowska et Popowska, 2014). Selon la classification de 2005 du *Bergey's Manual of Systematic Bacteriology*, le genre est divisé en quatorze espèces (Martin-Carnahan et Joseph, 2005). Cependant, ce nombre a grandement augmenté puisque près de 40 espèces d'*Aeromonas* ont été proposées à ce jour (Beaz-Hidalgo *et al.*, 2015a; Beaz-Hidalgo *et al.*, 2015b; Marti et Balcázar, 2015; Martínez-Murcia *et al.*, 2016). La taxonomie de ce genre continuera de s'enrichir et de se clarifier au cours des prochaines années par l'obtention et l'analyse des données génomiques (Beaz-Hidalgo *et al.*, 2015a; Colston *et al.*, 2014; Janda et Abbott, 2010).

Plusieurs des espèces du genre *Aeromonas* sont des agents pathogènes opportunistes des animaux à sang chaud et à sang froid ainsi que des invertébrés (Janda et Abbott, 2010; Martin-Carnahan et Joseph, 2005). Des membres de certaines espèces, comme *Aeromonas hydrophila*, sont capables d'infecter une vaste gamme d'hôtes, dont des Poissons et les Humains (Janda et Abbott, 2010; Roger *et al.*, 2012). D'autres, comme *Aeromonas salmonicida* ssp. *salmonicida*, sont très spécifiques et n'affectent que les Salmonidés (Cipriano et Austin, 2011; Dallaire-Dufresne *et al.*, 2014b; Martin-Carnahan et Joseph, 2005). Puisque les *Aeromonas* ont le potentiel de côtoyer des microorganismes qui transitent par les systèmes aquatiques, elles sont des réservoirs et des vecteurs possibles de

transmission de gènes accessoires, comme des gènes de résistances aux antibiotiques (Piotrowska et Popowska, 2014).

Parmi le genre *Aeromonas*, l'espèce *salmonicida* regroupe des bactéries généralement psychrophiles et pathogènes chez le poisson. *Aeromonas salmonicida* est lui-même divisé en cinq sous-espèces, ssp. *salmonicida*, *masoucida*, *achromogenes*, *smithia* et *pectinolytica* (une sous-espèce mésophile), en plus d'isolats non classés (Dallaire-Dufresne *et al.*, 2014b; Martin-Carnahan et Joseph, 2005; Vincent *et al.*, 2016b). Historiquement, *A. salmonicida* était divisé en deux groupes selon les caractéristiques pathologiques des souches: les souches typiques, qui appartiennent à ssp. *salmonicida*, et les souches « atypiques », qui appartiennent aux autres sous-espèces (Hänninen et Hirvela-Koski, 1997). Les souches d'*A. salmonicida* ssp. *salmonicida* ont des caractéristiques phénotypiques homogènes : elles causent la furunculose chez les Salmonidés, sont psychrophiles, non motiles et produisent une pigmentation caractéristique brune (Bernoth, 1990; Donlon *et al.*, 1983; Martin-Carnahan et Joseph, 2005). Les souches atypiques causent quant à elles des infections moins typées chez une plus grande gamme d'hôtes (Cipriano et Austin, 2011). Plus récemment, des *A. salmonicida* mésophiles ont été isolées à l'extérieur du contexte vétérinaire qui caractérisait autrefois cette espèce (Pavan *et al.*, 2015; Vincent *et al.*, 2016b). Une récente étude suggère que les souches mésophiles et psychrophiles d'*A. salmonicida* ont des différences majeures dans leur génome. En effet, les souches psychrophiles possèdent un plus grand nombre d'IS. De plus, ces éléments mobiles appartiennent à des familles différentes que les IS des autres *A. salmonicida*. Puisqu'un grand nombre d'IS augmente le risque d'introduire des mutations délétères par transposition ou recombinaison, et que les IS d'*A. salmonicida* semblent plus impliquées dans ce type d'évènement à haute température, il a été suggéré que le nombre et le type d'IS présents chez les *A. salmonicida* psychrophiles les contraignaient à un mode de vie à basse température (Vincent *et al.*, 2016b).

Le modèle à l'étude pour cette thèse est la bactérie *Aeromonas salmonicida* ssp. *salmonicida*. Cette sous-espèce est psychrophile, avec un optimum de croissance entre 22°C et 25°C (Martin-Carnahan et Joseph, 2005). Cependant, puisque son intégrité génomique peut être compromise à cette plage de température, une culture en laboratoire à 18°C et moins est recommandée (Daher *et al.*, 2011; Moki *et al.*, 1995). *A. salmonicida* ssp.

salmonicida est un agent pathogène des Salmonidés chez lesquels il cause la furonculose (Cipriano et Austin, 2011; Dallaire-Dufresne *et al.*, 2014b; Martin-Carnahan et Joseph, 2005). La maladie a été répertoriée sur tous les continents où se pratique l'aquaculture, excepté l'Océanie (Cipriano et Austin, 2011). Par conséquent, les isolats de cette bactérie proviennent généralement de poissons malades (Cipriano et Austin, 2011) et plus rarement de l'environnement (Martin-Carnahan et Joseph, 2005).

1.4.1 LA PROBLÉMATIQUE DE LA FURONCULOSE

La culture des végétaux et l'élevage des animaux, en eaux douces comme en eaux salées, se nomme aquaculture. Depuis quelques années, ce secteur est en constante expansion; en 2014, 73,8 millions de tonnes de poissons, mollusques et crustacés ont été produits par aquaculture, pour une valeur estimée de 160,2 milliards de dollars américains (FAO, 2016). De plus, la production aquatique totale dépend de plus en plus de l'aquaculture. En 2014, 44,1% du tonnage mondial en animaux aquatiques était issu de cette industrie (FAO, 2016). Ce nombre est en constante augmentation, et, dans 35 pays, la production aquicole a dépassé celle des pêches (FAO, 2016). Cependant, ce type d'industrie engendre certaines problématiques. Les conditions d'élevage en aquaculture imposent un stress sur les poissons qui les rend plus susceptibles aux infections (Barton et Iwama, 1991; Cabello *et al.*). De plus, les antibiotiques doivent être utilisés de façon responsable afin de limiter la sélection et la dissémination des gènes de résistances dans les communautés bactériennes transitant par l'eau (Andersson et Hughes, 2014; Cabello *et al.*). Dans cette situation, le contrôle des agents pathogènes par de nouvelles approches répond à deux problématiques : la réduction des pertes et la diminution de l'utilisation des antibiotiques (Cabello *et al.*; FAO, 2016).

La furonculose est une infection locale ou systémique causée par *A. salmonicida* ssp. *salmonicida*. Elle peut toucher tous les Salmonidés (famille *Salmonidae*), mais elle affecte plus particulièrement les poissons d'élevage, puisqu'ils sont plus susceptibles aux infections (Barton et Iwama, 1991). La furonculose se décline en deux formes : l'infection aiguë et l'infection chronique. L'infection aiguë, de type septicémie, se déclare habituellement chez les poissons juvéniles et entraîne une forte mortalité en deux ou trois

jours (Cipriano et Austin, 2011; Dallaire-Dufresne *et al.*, 2014b). De son côté, la furonculose chronique se déclare surtout chez les poissons adultes qui ont survécu à un épisode de furonculose aigüe (Dallaire-Dufresne *et al.*, 2014b). Elle se caractérise par l'apparition de furoncles, des lésions nécrotiques sous-cutanées qui ont donné leur nom à la maladie (Dallaire-Dufresne *et al.*, 2014b). Puisque l'agent responsable de la furonculose est ubiquitaire de l'eau, le contrôle de cette maladie requiert de bonnes mesures de biosécurité pour la qualité de l'eau et l'importation de matériel vivant (Cipriano et Austin, 2011; Dallaire-Dufresne *et al.*, 2014b). La forme aigüe et la forme chronique sont responsables d'une fraction non négligeable des infections déclarées en aquaculture (entre 30 et 60% au Québec, annuellement) et engendrent de lourdes pertes pour l'industrie (Morin, 2010).

Deux types de traitements sont généralement utilisés de façon prophylactique ou thérapeutique contre la furonculose : la vaccination et l'antibiothérapie, respectivement. La vaccination est généralement faite par injection intrapéritonéale de bactéries atténuées ou tuées et d'adjuvants (Cipriano et Austin, 2011; Menanteau-Ledouble *et al.*, 2016). La vaccination est efficace, mais le choix d'adjuvant, l'âge des poissons au moment de la procédure et la température d'élevage peuvent mener à des effets secondaires (Cipriano et Austin, 2011; Villumsen *et al.*, 2015). Pour réduire ces inconvénients, des études ont été entreprises afin de formuler un vaccin à base d'antigènes (Lutwyche *et al.*, 1995; Marana *et al.*, 2017; Menanteau-Ledouble et El-Matbouli, 2016), ou administré par balnéation (Villumsen et Raida, 2013).

Au Québec, les antibiotiques homologués pour l'usage animal ne peuvent être administrés que sous l'ordonnance d'un médecin vétérinaire (Morin *et al.*, 2004). Quatre antibiotiques provenant de trois familles sont homologués pour l'aquaculture au Québec: une tétracycline (oxytétracycline), un chloramphénicol (florfénicol), et deux sulfamides ((sulfadiméthoxine/ormétoprime et sulfadiazine/triméthoprimine) (Morin *et al.*, 2004). Puisque les produits destinés à la consommation doivent être exempts de résidus médicamenteux, l'antibiothérapie demande des périodes de retrait variables avant la commercialisation du produit, ce qui pénalise le pisciculteur (Morin, 2010). Malheureusement, *A. salmonicida* ssp. *salmonicida* compte des souches résistantes et même multirésistantes aux antibiotiques. Des bactéries résistantes ont été isolées dès 1959 (Aoki *et al.*, 1971). En 2014, Vincent *et al.* (2014b) faisaient état de souches

d'*A. salmonicida* ssp. *salmonicida* résistantes à tous les antibiotiques homologués pour l'aquaculture dans leur région d'origine, la province de Québec.

En plus de la vaccination et de l'antibiothérapie, d'autres types de traitements ont déjà été expérimentés contre la furunculose. Entre autres, quelques essais *in vivo* ont été réalisés à partir de bactériophages (Imbeault *et al.*, 2006; Kim *et al.*, 2015; Verner-Jeffreys *et al.*, 2007). L'introduction de bactériophages en suspension dans l'aquarium ou par injection intrapéritonéale permettrait de ralentir l'apparition des symptômes de furunculose sans éliminer la maladie (Imbeault *et al.*, 2006; Kim *et al.*, 2015; Verner-Jeffreys *et al.*, 2007). De plus, des essais avec des produits naturels (extraits de plantes) ont été réalisés contre *A. salmonicida* (toutes sous-espèces confondues) *in vitro* (Starlíper *et al.*, 2015) et *in vivo* (Breyer *et al.*, 2015; Okmen *et al.*, 2012). Même si certains produits semblent efficaces contre cette bactérie, les concentrations utilisées peuvent aussi être toxiques chez le poisson (Okmen *et al.*, 2012). La diversification des traitements contre la furunculose est souhaitable et l'étude de traitements alternatifs mériterait d'être bonifiée.

1.4.2 LES FACTEURS DE VIRULENCE D'*A. SALMONICIDA* SSP. *SALMONICIDA*

La relation hôte-pathogène est une danse complexe où les deux parties passent par toute une série d'états. Un agent pathogène a donc besoin d'une variété de tactiques pour envahir son hôte, échapper à son système immunitaire, puis coloniser des environnements aux paramètres physico-chimiques différents (Brogden *et al.*, 2000). Les facteurs de virulence sont ces molécules et ces systèmes qui permettent de remplir les mandats précédents (Brogden *et al.*, 2000). Certains sont entièrement destinés à la sécrétion de toxines et entrent facilement dans la définition. D'autres éléments qui favorisent la survie de l'agent pathogène dans l'hôte, comme des sidérophores, une capsule ou des adhésines, sont aussi des facteurs de virulence (Brogden *et al.*, 2000). Dans l'ère pré-génomique, les études sur *A. salmonicida* ssp. *salmonicida* étaient davantage dirigées sur les caractéristiques de la bactérie et leur rôle au cours de l'infection: plusieurs facteurs de virulence ont donc été caractérisés. Dans la plupart des cas, les mécanismes ont été observés *in vitro* puis leur implication a été testée *in vivo* afin de mesurer leur impact sur la virulence totale de la bactérie (Adams *et al.*, 1988; Boyd *et al.*, 2008; Ellis *et al.*, 1997; Noonan et Trust, 1996).

1.4.2.1 TOXINES EXTRACELLULAIRES

La bactérie peut sécréter plusieurs protéines qui ont des activités protéases, hémolysines ou lipases (Arnesen et Eggset, 1999 ; Arnesen *et al.*, 1995; Eggset *et al.*, 1994; Fyfe *et al.*, 1987; Hirono et Aoki, 1993 ; Lee et Ellis, 1990; Price *et al.*, 1990; Shieh et Maclean, 1975). Deux enzymes ont été plus particulièrement étudiées, une sérine protéase de 70 kiloDaltons et une lipase, la glycérophospholipide-cholestérol acyltransférase (GCAT) (Eggset *et al.*, 1994; Fyfe *et al.*, 1987; Price *et al.*, 1990). La GCAT doit être activée par la sérine protéase et les deux enzymes causent des thromboses chez l'hôte (Eggset *et al.*, 1994; Fyfe *et al.*, 1987; Salte *et al.*, 1992). Cependant, les activités protéases, lipases et hémolysines ne semblent pas être des déterminants essentiels de la virulence (Adams *et al.*, 1988; Ellis *et al.*, 1988; Vipond *et al.*, 1998).

1.4.2.2 LA COUCHE DE SURFACE A ET LES LIPOPOLYSACCHARIDES

Les couches de surface (couches S ou *S-layers*) sont composées de protéines qui s'agencent de façon ordonnée à la surface des certaines bactéries à Gram positif et à Gram négatif ainsi que de plusieurs *Archaea* (Fagan et Fairweather, 2014). Cette structure a de nombreuses fonctions dépendantes de l'organisme qui la porte (Fagan et Fairweather, 2014). Chez *A. salmonicida ssp. salmonicida*, elle est nommée couche de surface A (*A-layer*) et sa sous-unité protéique est encodée sur le chromosome par le gène *vapA* (Chu *et al.*, 1991a; Kay *et al.*, 1981; Udey et Fryer, 1978). L'exportation de VapA à la membrane externe d'*A. salmonicida ssp. salmonicida* requiert un système de sécrétion de type II qui est encodé en amont de *vapA* (Belland et Trust, 1985, 1987; Noonan et Trust, 1996; Reith *et al.*, 2008). De plus, chez *A. salmonicida ssp. salmonicida*, l'ancrage de la couche de surface A à la membrane externe se fait par l'intermédiaire des lipopolysaccharides (LPS) (Belland et Trust, 1985; Dooley *et al.*, 1989).

La couche de surface A provoque l'agrégation (autoagglutination) des cellules (Kay *et al.*, 1981) et a une activité hémagglutinine, ce qui pourrait permettre d'échapper au système immunitaire (Adams *et al.*, 1988). Elle favorise l'adhésion, mais pas l'invasion, des cellules de l'hôte (Garduño *et al.*, 2000). La couche de surface A confère une virulence de 2 500 à 5 000 fois plus élevée d'une souche porteuse par rapport à son mutant atténué qui a été

généralisé spontanément par culture (Kay *et al.*, 1981). Cependant, cette protéine n'est pas le déterminant principal de la virulence chez *A. salmonicida* ssp. *salmonicida*, puisque certaines bactéries n'ayant pas de couche de surface A sont virulentes (Adams *et al.*, 1988; Ellis *et al.*, 1988).

Le LPS est une composante majeure de la membrane externe des bactéries à Gram négatif (Maldonado *et al.*, 2016). Il peut être divisé en trois sections : le lipide A, qui ancre la molécule au côté externe de la membrane, le polysaccharide central, et la chaîne latérale O (aussi appelée antigène O), qui est aussi un polysaccharide (Maldonado *et al.*, 2016). Les LPS, en particulier le lipide A et l'antigène O, sont importants dans la relation hôte-pathogène (Maldonado *et al.*, 2016). Les études biochimiques sur les sections du LPS d'*A. salmonicida* ssp. *salmonicida* ont montré que leur composition est conservée entre les souches (Chart *et al.*, 1984; Jimenez *et al.*, 2009; Merino *et al.*, 2015; Wang *et al.*, 2004; Wang *et al.*, 2006a; Wang *et al.*, 2006b; Wang *et al.*, 2007; Wang *et al.*, 2005). Les LPS et la couche de surface A d'*A. salmonicida* ssp. *salmonicida* sont liés fonctionnellement et génétiquement. La couche de surface A masque une grande partie du LPS, excepté certaines sections de la chaîne O (Chart *et al.*, 1984). Tous les deux protègent contre l'activité bactéricide du sérum (Munn *et al.*, 1982). De plus, les LPS sont importants pour l'attachement de la couche de surface A à la bactérie (Belland et Trust, 1985; Dooley *et al.*, 1989).

1.4.2.3 AUTRES ASPECTS EN LIEN AVEC LA MEMBRANE

A. salmonicida ssp. *salmonicida* possède des gènes codant pour deux types de flagelles (polaire et péritriche) (Reith *et al.*, 2008). Cependant, le flagelle polaire n'est observable qu'en de très rares occasions (moins de 1% des cellules) (Umelo et Trust, 1997), probablement parce que plusieurs des gènes nécessaires à sa synthèse sont inactivés (Reith *et al.*, 2008). La bactérie est décrite comme non motile (Martin-Carnahan et Joseph, 2005). Des gènes pour plusieurs types de pili (trois de type IV, un de type I) sont aussi présents, même si les structures ne sont pas observées (Boyd *et al.*, 2008; Reith *et al.*, 2008). Un de ces pili, le pilus Tap de type IV, aiderait à l'invasion de l'hôte sans toutefois être essentiel (Boyd *et al.*, 2008; Masada *et al.*, 2002).

A. salmonicida ssp. *salmonicida* est généralement décrite comme non encapsulée (Martin-Carnahan et Joseph, 2005). Cependant, une capsule et des exopolysaccharides ont été observés après culture de la bactérie dans des conditions très spécifiques (en phase stationnaire et en excès de glucose) (Bonet *et al.*, 1993; Garrote *et al.*, 1992; Wang *et al.*, 2004). La capsule semble promouvoir l'invasion de cellules de poisson *in vitro* (Merino *et al.*, 1996). *A. salmonicida* ssp. *salmonicida* peut aussi faire des nanovésicules qui pourraient être impliquées dans la pathogénicité (Lusta et Kozlovskii, 2011; Lusta et Kozlovsky, 2011).

1.4.2.4 MÉTABOLISME DANS L'HÔTE

Les souches typiques d'*A. salmonicida* ssp. *salmonicida* croissent *in vitro* en présence de chélateurs de Fe³⁺, des conditions qui répliquent la séquestration du fer présente chez l'hôte. Plusieurs souches produisent des sidérophores (Chart et Trust, 1983; Fernandez *et al.*, 1998; Hirst *et al.*, 1991). D'autres peuvent acquérir le fer par contact avec un récepteur membranaire dont l'identité est inconnue (Chart et Trust, 1983; Hirst *et al.*, 1994). *A. salmonicida* ssp. *salmonicida* produit aussi des superoxydes dismutases (SodA et SodB) ainsi qu'une catalase qui confèrent une protection contre des conditions oxydatives exogènes (Barnes *et al.*, 1999; Barnes *et al.*, 1996; Dacanay *et al.*, 2003). Ces enzymes protégeraient la bactérie des changements oxydatifs brusques produits en guise de réponse immunitaire par les macrophages de poisson (Barnes *et al.*, 1999; Barnes *et al.*, 1996; Dacanay *et al.*, 2003).

1.4.2.5 LE SYSTÈME DE SÉCRÉTION DE TYPE TROIS

Le système de sécrétion de type trois (SSTT) est un facteur de virulence retrouvé chez plusieurs bactéries à Gram négatif qui sont des agents pathogènes d'animaux et de plantes (Buttner, 2012; Gophna *et al.*, 2003). C'est un appareil transmembranaire complexe qui permet d'injecter des molécules toxiques, appelées effecteurs, dans les cellules de l'hôte (Buttner, 2012). Les gènes de SSTT sont généralement groupés chez un organisme et ont probablement été acquis par transfert horizontal (Buttner, 2012; Gophna *et al.*, 2003).

La composante structurale du SSTT est un assemblage complexe de plus de 20 protéines (Figure 1.4) (Buttner, 2012). Il est composé de protéines d'ancrage aux membranes, d'une pompe ATP-dépendante nécessaire à la translocation des effecteurs, d'une structure polymérique en forme d'aiguille qui fait le pont entre la bactérie et l'hôte et d'un complexe permettant de percer la membrane de l'hôte (Buttner, 2012). Cette structure permet la translocation dans les cellules de l'hôte de la seconde classe de protéines du SSTT : les effecteurs. Contrairement à d'autres toxines extracellulaires, les effecteurs du SSTT entraînent des perturbations « subtiles » chez l'hôte (Dean, 2011). En effet, les effecteurs ont des cibles cellulaires très variées, comme le contrôle de la phagocytose, le cytosquelette, l'apoptose, et plusieurs sentiers de signalisation (Dean, 2011). La dernière classe de protéines des SSTT regroupe des chaperonnes, des molécules liées aux effecteurs dans la bactérie jusqu'à leur translocation dans le système en forme d'aiguille. Les chaperonnes assurent le bon repliement de leur protéine associée et la bonne reconnaissance de l'effecteur par l'appareil de translocation (Buttner, 2012).

Chez *A. salmonicida* ssp. *salmonicida*, l'étude du SSTT a commencé plus tard que celle des autres facteurs de virulence (Burr *et al.*, 2002). Cependant, des études *in vitro* et *in vivo* ont montré que le SSTT était essentiel à la virulence de cette bactérie (Burr *et al.*, 2005; Burr *et al.*, 2003a; Burr *et al.*, 2003b). La perte du SSTT rend également la bactérie avirulente dans le modèle alternatif *Dictyostelium discoideum* (Froquet *et al.*, 2007). Les protéines du SSTT sont davantage exprimées à 28°C qu'à 17°C (Ebanks *et al.*, 2006). Le SSTT fonctionnel ou non interférerait avec la réponse immunitaire, ce qui augmenterait la gravité de l'infection, mais aussi préviendrait la construction d'une immunité adaptative solide post-vaccination (Bergh *et al.*, 2013; Origgi *et al.*, 2017).

Les protéines structurales du SSTT sont toutes encodées dans le même locus, lui-même situé sur le plasmide pAsa5 (Figure 1.4). Les effecteurs du SSTT d'*A. salmonicida* ssp. *salmonicida* sont au nombre de six : AopH, Ati2, AopO, AexT, AopN et AopP. Les gènes d'autres effecteurs potentiels ont subi une accumulation de mutations les transformant en pseudogènes. Des six effecteurs, quatre sont encodés sur le même plasmide que les gènes structuraux (AopH, Ati2, AopN et AopO), un est encodé sur le chromosome (AexT) et le dernier est encodé sur un autre plasmide (AopP, plasmide pAsa11). Le régulateur ExsE, la

chaperonne Ati1 et la protéine de fonction inconnue AscX sont aussi sécrétés par le SSTT d'*A. salmonicida* ssp. *salmonicida* (Vanden Bergh *et al.*, 2013).

La fonction des effecteurs AopP, AexT et Ati2 a été étudiée plus en détail. AexT est une activatrice de GTPase et une ADP ribosyltransferase, les deux fonctions permettent de causer la dépolymérisation de l'actine. AexT aurait donc un rôle à jouer dans la nécrose tissulaire (Burr *et al.*, 2003a). AopP est un inhibiteur de la voie des NF- κ B, et par conséquent, de l'apoptose (Fehr *et al.*, 2006).

Ati2 est un effecteur homologue à VPA0450 de *Vibrio parahaemolyticus* (Broberg *et al.*, 2010; Dallaire-Dufresne *et al.*, 2013). Il a une activité inositol polyphosphate 5-phosphatase (Dallaire-Dufresne *et al.*, 2013). Les inositol polyphosphate sont des glycolipides de la face interne de la membrane cytoplasmique des cellules eucaryotes qui agissent comme molécules de signalisation (Wilson et Saiardi, 2017). Les inositol polyphosphate phosphatase peuvent donc modifier des voies de signalisation moléculaire. Pour Ati2, l'activité suspectée est l'inhibition de la phagocytose (Dallaire-Dufresne *et al.*, 2013). De plus, la protéine Ati2 exprimée chez le modèle *Dictyostelium discoideum* est toxique (Dallaire-Dufresne *et al.*, 2013).

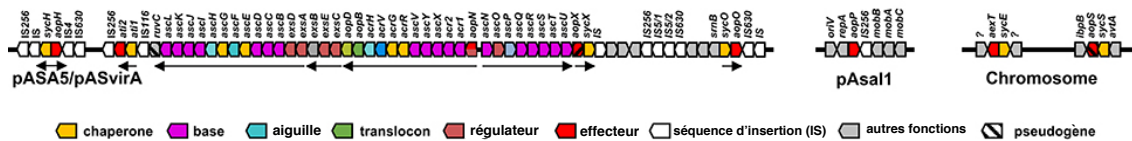
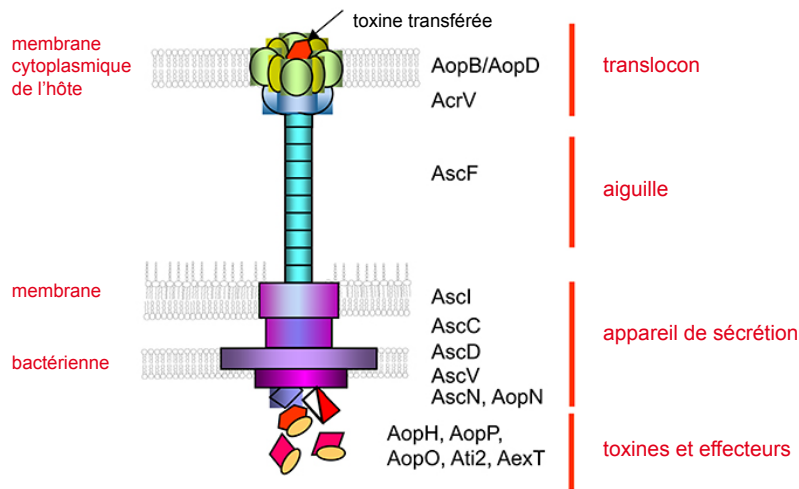


Figure 1.4 Le système de sécrétion (SSTT) d'*A. salmonicida ssp. salmonicida*. La structure d'un SSTT est schématisée, avec le nom des différentes protéines qui y sont associées. L'emplacement des gènes codant pour ces protéines est donné en bas : pAsa5/pASvirA et pAsal1 sont des plasmides. Les gènes non fonctionnels sont barrés. Cette figure est une traduction de la Figure 1 de Frey et Origgi (2016), reproduite ici selon sa licence *Creative Commons Attribution Licence* (CC BY) (<http://creativecommons.org/licenses/by/4.0/>). L'article original est disponible en ligne à l'adresse suivante : (<http://doi.org/10.3389/fmars.2016.00130>).

1.4.3 LE GÉNOME D'*A. SALMONICIDA SSP. SALMONICIDA* EST RICHE EN INFORMATION

Les premières études épidémiologiques sur *A. salmonicida ssp. salmonicida* laissaient présager que cette sous-espèce était très homogène génétiquement (Aoki et Holland, 1985; Hänninen *et al.*, 1995). Cependant, l'étude de cet agent pathogène a été profondément bouleversée par l'arrivée du séquençage à haut débit. Il est possible de constater à ce moment un changement de paradigme : de l'étude de caractéristiques visibles et mesurables (phénotypes, protéines produites), l'étude d'*A. salmonicida ssp. salmonicida* est plutôt passée à une caractérisation systématique de ses nombreux éléments génétiques (gènes, éléments mobiles, plasmides). Ce changement permet l'intégration des avantages propres

aux technologies à haut débit : grande quantité d'information générée, découvertes moins dirigées et plus générales, mais quantité d'analyse augmentée. Cette transition a probablement commencé bien avant l'accessibilité du séquençage haut débit : en effet, la caractérisation de nombreux plasmides (Boyd *et al.*, 2003; Fehr *et al.*, 2006; Stuber *et al.*, 2003), ainsi que la publication du génome de référence (séquencé selon la méthode Sanger) (Reith *et al.*, 2008) ont jeté les bases de l'étude génomique de cette bactérie. Afin de réconcilier ces deux façons d'étudier, une attention particulière doit être portée à l'impact des éléments génomiques obtenus sur les caractéristiques observables de la bactérie.

1.4.4 LE GÉNOME DE RÉFÉRENCE D'*A. SALMONICIDA* SSP. *SALMONICIDA*

Le premier génome obtenu pour *A. salmonicida* ssp. *salmonicida* est celui d'A449. Cette souche, isolée en 1975 d'une truite brune française, a régulièrement été étudiée avant de voir son génome séquencé (Belland et Trust, 1985; Belland et Trust, 1989; Chart *et al.*, 1984; Gustafson *et al.*, 1994; Gustafson *et al.*, 1992; Ishiguro *et al.*, 1981). Le projet du génome d'A449 comprenait le chromosome ainsi que deux grands plasmides de cette souche (pAsa4 et pAsa5). Ils ont été séquencés à l'aide de la technologie Sanger (Reith *et al.*, 2008). La séquence des autres petits plasmides de cette bactérie, pAsa1, pAsa2 et pAsa3, avait déjà fait l'objet d'une publication (Boyd *et al.*, 2003).

Le chromosome d'A449 fait 4,7 mégapaires de bases et code pour 4 388 gènes. Son pourcentage en G+C est de 58,5. La déviation en GC indique une inversion de séquence entre les bases 3 963 279 à 4 158 772, ce qui a été confirmé par PCR. Sur les 4 388 gènes prédits, 155 (3,5%) sont des pseudogènes. De ces inactivations, 20 sont causées par l'insertion d'IS dans les gènes, mais la plupart sont causées par de petites délétions ou des mutations ponctuelles. Le chromosome d'A449 est aussi le lieu d'insertion de deux prophages (prophage 1 et prophage 2) (Reith *et al.*, 2008).

Dans le même article, le chromosome d'A449 a été comparé à celui d'*A. hydrophila* ATCC 7966^T, un agent pathogène opportuniste de plusieurs animaux à sang froid et à sang chaud, dont l'humain. Excepté quelques inversions, la structure et le contenu des deux molécules sont semblables (Reith *et al.*, 2008). Cependant, *A. salmonicida* ssp. *salmonicida* a un nombre nettement plus élevé de pseudogènes qu'*A. hydrophila* (155 contre 7,

respectivement) (Reith *et al.*, 2008). Cette perte de gènes et de fonctions peut être un signe de spécialisation de la bactérie vers un hôte précis (Reith *et al.*, 2008). Cette caractéristique est aussi conséquence des plages d'hôtes de ces deux espèces.

1.4.5 LE PLASMIDOME D'*A. SALMONICIDA* SSP. *SALMONICIDA*

Dans les premières études sur *A. salmonicida* ssp. *salmonicida*, la description et la caractérisation des plasmides se faisaient par électrophorèse, et parfois par clonage ou conjugaison (Adams *et al.*, 1998; Bast *et al.*, 1988; Belland et Trust, 1989; Sørum *et al.*, 1993; Sørum *et al.*, 2003). Par la suite, l'arrivée du séquençage à haut débit a permis d'étendre les connaissances sur le plasmidome de cette bactérie et d'identifier sans équivoque les molécules étudiées.

Il est possible de séparer les plasmides d'*A. salmonicida* ssp. *salmonicida* en fonction de leur taille, de leurs gènes accessoires ou de leur occurrence dans les souches étudiées. Certains plasmides ont aussi des liens évolutifs importants qui ont été déduits par leur séquence. Il devient donc difficile de présenter une image intégrée de ce plasmidome, qui est d'ailleurs toujours en expansion. La présente section tâchera de décrire les plasmides d'*A. salmonicida* ssp. *salmonicida* qui ont fait l'objet de publications en se basant d'abord sur leurs caractéristiques fonctionnelles, puis sur leurs liens évolutifs putatifs (Tableau 1.2).

1.4.5.1 LES PETITS PLASMIDES À HAUT NOMBRE DE COPIES

Les plasmides pAsa1, pAsa2 et pAsa3 (Belland et Trust, 1989; Boyd *et al.*, 2003) ainsi que le petit plasmide pAsa11 (Fehr *et al.*, 2006) sont fréquemment trouvés dans les souches d'*A. salmonicida* ssp. *salmonicida* (Attéré *et al.*, 2015). Les trois premiers réplicons sont cryptiques, c'est-à-dire qu'ils ne codent pour aucune fonction connue sauf leur maintien et leur mobilisation. Ils possèdent toutefois des gènes codant pour des protéines hypothétiques (Boyd *et al.*, 2003). Le plasmide pAsa11 possède un gène codant pour un effecteur du SSTT (*aopP*) ainsi qu'une IS (*ISAS11*) (Fehr *et al.*, 2006). Les quatre plasmides sont généralement stables (Attéré *et al.*, 2015; Boyd *et al.*, 2003), mais une exposition à une température de 25°C pourrait favoriser la perte du plasmide pAsa11 (Tanaka *et al.*, 2012).

Trois de ces petits plasmides, pAsa1, pAsa3 et pAsa11, sont des réplicons de types ColE2 (Boyd *et al.*, 2003). Leur machinerie de réplication et de contrôle du nombre de copies est donc similaire, mais pas au point de les rendre incompatibles (Boyd *et al.*, 2003). pAsa1 et pAsa3 codent aussi pour des systèmes toxine-antitoxine homologues au système *relBE* (Boyd *et al.*, 2003). Le plasmide pAsa2 est un réplicon de type ColE1 (Boyd *et al.*, 2003).

Des similarités existent entre certains de ces petits plasmides. D'abord, pAsa3 et pAsa11 sont très similaires sur 3,8 kpb de leur séquence (Tanaka *et al.*, 2012). Cette caractéristique peut en compliquer l'assemblage, menant à l'écart d'un ou de l'autre des plasmides lors du séquençage à haut débit (Charette *et al.*, 2012). De plus, trois variants du plasmide pAsa11, nommés pAsa11B, pAsa11C et pAsa11D, ont été découverts et séquencés (Attéré *et al.*, 2015; Trudel *et al.*, 2013). Dans les trois cas, la différence vient d'une insertion d'IS (*ISAS5*) dans le plasmide, parfois accompagnée de duplications de séquences (Attéré *et al.*, 2015; Trudel *et al.*, 2013).

La structure de pAsa2 est très similaire (94,1%) à celle d'un plasmide nommé pAsa7, qui confère une résistance au chloramphénicol et qui n'a été retrouvé jusqu'à maintenant que dans une seule souche (Vincent *et al.*, 2016a). Pour pAsa7, une section codant pour deux protéines hypothétiques chez pAsa2 a été substituée par un autre fragment qui code pour une chloramphénicol acétyltransférase. Les deux plasmides sont des réplicons de type ColE1, mais sont tout de même compatibles à cause de mutations dans la structure qui contrôle la réplication (Vincent *et al.*, 2016a).

Les précédentes études montrent que les petits plasmides d'*A. salmonicida* ssp. *salmonicida* sont relativement stables et fréquents, mais peuvent toutefois faire l'objet de variations. La proximité entre pAsa3 et pAsa11 ainsi que la découverte des variants de pAsa11 suggèraient déjà que les plasmides cryptiques d'*A. salmonicida* ssp. *salmonicida* pouvaient servir de gabarit pour de nouvelles combinaisons de gènes (Attéré *et al.*, 2015; Trudel *et al.*, 2013). La découverte de pAsa7 vient renforcer cette hypothèse (Vincent *et al.*, 2016a).

1.4.5.2 LES PLASMIDES PORTANTS DES GÈNES DE RÉSISTANCE

C'est cette catégorie qui comporte le plus d'éléments (en plus de pAsa7, qui a été présenté dans la dernière section). Cependant, plusieurs de ces plasmides ont été nommés, identifiés,

mais pas séquencés. Par conséquent, il existe une certaine incertitude sur le nombre de plasmides distincts dans cette catégorie, puisque deux réplicons pourraient être connus sous le même nom.

1.4.5.2.1 LE PLASMIDE PAsA8

Ce plasmide de 110 kpb a été caractérisé à partir de deux souches québécoises présentant de multiples résistances aux antibiotiques (Trudel *et al.*, 2016). Ces gènes de résistance aux antibiotiques sont portés par un élément d'ADN mobile complexe, Tn1721. Ce transposon porte lui-même un intégron complexe de classe 1, In104, qui possède deux systèmes de cassettes. Cet élément a précédemment été observé sur l'îlot génomique SG11 de *Salmonella enterica* et de *Proteus mirabilis*, avec quelques variations. pAsa8 confère des résistances à tous les antibiotiques utilisés en aquaculture dans sa région géographique de détection, la province du Québec (voir section 1.4.1). Aussi, ce plasmide possède une bêta-lactamase *bla*_{PSE-1}, un agent préoccupant pour la santé publique. Le reste de la structure de pAsa8 possède peu d'homologie de séquence avec la base de données non redondante en nucléotides Genbank (nr/nt). Il est donc difficile d'inférer des fonctions à ses séquences codantes (Trudel *et al.*, 2016).

1.4.5.2.2 PRAS1, PAR-32 ET LES PASOT : DES PLASMIDES APPARENTÉS ?

Les plasmides pRAS1 (45 kpb) et pAr-32 sont des réplicons conjugatifs à large spectre (Aoki *et al.*, 1971; Bradley *et al.*, 1982; Sørum *et al.*, 2003). Ils ont été isolés en Europe (Écosse, Norvège) et en Asie (Japon), respectivement (Sandaa et Enger, 1994; Sørum *et al.*, 2003). Ils appartiennent au groupe d'incompatibilité IncU, qui a été défini à partir de pAr-32 et ils ont une structure commune (Bradley *et al.*, 1982; Sørum *et al.*, 2003). Cependant, puisqu'aucun des deux plasmides n'a été séquencé, leur structure reste nébuleuse, à l'exception des éléments mobiles qui portent les gènes de résistance aux antibiotiques. pAr-32 possède un intégron de type 1 apparenté à In6 qui possède une cassette *aadA2*. Un gène de résistance au chloramphénicol *catAIII* est aussi retrouvé hors de la structure d'intégration (Sørum *et al.*, 2003). pRAS1 possède un fragment du Tn1721 auquel est couplé un intégron de classe 1 apparenté à In4 (Sørum *et al.*, 2003).

Des plasmides conférant des résistances à l'oxytétracycline (OT) ont été caractérisés à partir de souches d'*A. salmonicida* d'Europe et d'Amérique et nommés pASOTX (Tableau 1.2). À la suite de l'obtention de profils de digestion, ces plasmides ont été classés en trois groupes : pASOT, pASOT2 et pASOT3. Même si leur séquence n'a pas été obtenue, les plasmides de type pASOT et pASOT2 auraient une structure apparentée à pRAS1, puisque leurs profils de digestion présentent des similitudes (Adams *et al.*, 1998).

1.4.5.2.3 AUTRES PLASMIDES DE LA SUITE PRAS

Les plasmides pRAS2 et pRAS4 ont été isolés en Norvège en même temps que pRAS1. Ce sont tous les deux des plasmides conjugatifs (L'Abée-Lund et Sørum, 2002). La structure et le contenu de pRAS4 restent largement inconnus (L'Abée-Lund et Sørum, 2002). Le plasmide pRAS2 confère une forte résistance à la streptomycine, en plus d'une résistance aux sulfamides et aux tétracyclines (L'Abée-Lund et Sørum, 2000). Il porte le transposon Tn5393c, qui contient les gènes de résistance *strA* et *strB* responsables de la résistance à la streptomycine. Fait intéressant, les gènes sont surexprimés en raison d'une IS, IS1133, insérée en amont de ceux-ci (L'Abée-Lund et Sørum, 2000).

Un troisième plasmide, pRAS3.1, a été isolé en même temps que pRAS1 chez *A. salmonicida* ssp. *salmonicida*. Ce plasmide est plus petit (11,8 kpb) que les autres pRAS et non conjugatif. Un variant de ce plasmide, pRAS3.2, a été isolé chez des *A. salmonicida* atypiques. Les pRAS3 confèrent une résistance aux tétracyclines (L'Abée-Lund et Sørum, 2002). Plus récemment, deux autres variants de pRAS3 (pRAS3.3 et pRAS3.4) ont été isolés au Nouveau-Brunswick (Tableau 1.2) (Vincent *et al.*, 2014b). Les quatre variants de pRAS3 diffèrent par leur nombre de répétitions dans deux régions nommées RegA et RegB (Vincent *et al.*, 2014b).

1.4.5.2.4 UNE ACCUMULATION DE RÉSISTANCES : PAB5S9B ET PSN254B

Le plasmide pAb5S9b a été isolé de deux souches néo-brunswickoises. Il confère des résistances à cinq antibiotiques (Tableau 1.2). Les gènes des quatre dernières résistances

sont situés dans une région similaire à la région SXT de *Vibrio cholerae*¹. pAb5S9b est un variant du plasmide pAb5S9 isolé en Europe d'une souche d'*Aeromonas bestiarum*. Les différences entre ces deux plasmides sont une transposition entre deux segments et une substitution des gènes codant pour la résistance à la tétracycline, les gènes *tet* (Vincent *et al.*, 2014b).

Le plasmide pSN254b est un autre variant d'un plasmide décrit chez une autre espèce, ici pSN254 de *S. enterica* (McIntosh *et al.*, 2008 ; Vincent *et al.*, 2014b). Les plasmides pSN254 font partie du groupe d'incompatibilité IncA/C (plus particulièrement du type 2 des IncA/C₂), qui regroupe de grands plasmides conjugatifs ayant une structure conservée, mais des régions de résistances variables (Fricke *et al.*, 2009; Harmer et Hall, 2015; Welch *et al.*, 2007). Le variant b est plus court que la version originale de 24 kpb à cause de l'absence d'un transposon composite putatif et de régions dans le Tn21 (Vincent *et al.*, 2014b). pSN254b confère toutefois des résistances à nombre d'antimicrobiens (Tableau 1.2) (Vincent *et al.*, 2014b). Les souches possédant pAB5S9b ou pSN254b sont résistantes à tous les antibiotiques autorisés pour l'aquaculture au Québec (Trudel *et al.*, 2016). pSN254b est aussi le plus prévalent des plasmides de résistance aux antibiotiques chez les souches québécoises des huit dernières années (Trudel *et al.*, 2016).

1.4.5.2.5 pAsa4, UN PLASMIDE SPÉCIFIQUE D'*A. SALMONICIDA* SSP. *SALMONICIDA*

La présence du plasmide pAsa4 dans la souche de référence A449 est connue depuis longtemps (Belland et Trust, 1989) et encore plus depuis la publication du génome de cette souche (Reith *et al.*, 2008). pAsa4 est un plasmide de 167 kpb qui confère des résistances aux tétracyclines, au mercure, à la streptomycine et à la spectinomycine, aux sulfamides, au chloramphénicol et aux composés d'ammonium quaternaire (Tableau 1.2) (Reith *et al.*, 2008). Tous les gènes codant pour ces résistances, excepté celui codant pour celle à la tétracycline, sont situés sur un transposon Tn21. pAsa4 possède aussi des gènes impliqués dans la répllication, la ségrégation et le transfert conjugatif du plasmide (Reith *et al.*, 2008). Son origine de répllication est compatible avec *E. coli*, mais le plasmide n'est pas

¹ Dans ce cas-ci, SXT réfère à une élément intégratif et conjugatif présent chez *V. cholerae*, et non à la combinaison d'antibiotiques sulfaméthoxazole-triméthoprime

transmissible par conjugaison à cette espèce (Belland et Trust, 1989). Sa structure ressemble à l'architecture partagée à haut pourcentage d'identité par les plasmides IncA/C. Par exemple, pAsa4 partage avec ces plasmides un système toxine/antitoxine *hipBA* (Fricke *et al.*, 2009). Cependant, pAsa4 ne fait pas partie de ce groupe puisque le pourcentage d'identité de sa structure par rapport au tronc commun des IncA/C n'est pas assez élevé (Fricke *et al.*, 2009). pAsa4 n'est pas très prévalent dans les souches génotypées ou séquencées à ce jour. Par contre quelques souches du Québec pourraient porter des variants (Trudel *et al.*, 2016; Vincent *et al.*, 2014b). Il s'agirait de la première détection de ce plasmide en Amérique de Nord et de la première identification de variants pour ce réplicon (Trudel *et al.*, 2016; Vincent *et al.*, 2014b).

1.4.5.2.6 UN AUTRE GRAND PLASMIDE D'A449, pAsa5

Un second plasmide a été séquencé lors de l'obtention du génome de référence : pAsa5 (ou pASvirA) (Reith *et al.*, 2008; Stuber *et al.*, 2003). Ce plasmide de 155 kpb contient tous les gènes structuraux du SSTT d'*A. salmonicida* ssp. *salmonicida*, répartis en cinq opérons polycistroniques rapprochés (Burr *et al.*, 2003a; Frey et Origgi, 2016). Trois des cinq effecteurs de ce SSTT sont aussi encodés sur ce plasmide (Frey et Origgi, 2016). Ce plasmide est sensible à la chaleur et la virulence de la bactérie est perdue lors de cultures en laboratoire à 25°C (Daher, 2010; Stuber *et al.*, 2003). Les modalités de perte du SSTT chez *A. salmonicida* ssp. *salmonicida* seront décrites dans la prochaine section.

Un dernier plasmide, pAsa6, a été isolé d'une souche d'*A. salmonicida* ssp. *salmonicida* infectant le turbot (Najimi *et al.*, 2009). pAsa6 est un petit plasmide (18 kpb) qui partage presque toute sa séquence avec des segments de pAsa5. Les séquences partagées incluent le gène de l'effecteur AopH, le gène de sa chaperonne et plusieurs IS. Au moment de la découverte du plasmide, les auteurs ne pouvaient inférer si pAsa6 était un dérivé de pAsa5 par délétion, ou si pAsa5 était un dérivé de pAsa6 par acquisition de gènes (Najimi *et al.*, 2009).

Tableau 1.2 Le plasmidome d'*A. salmonicida* ssp. *salmonicida*

Plasmide	Taille (pb)	Numéro d'accession	Groupe/ réplicon	AbiOR	Autres gènes accessoires	Éléments mobiles	Références
Petits plasmides à haut nombre de copies							
pAsa1	5 424	NC_004923	ColE2	-	-	-	(Belland et Trust, 1989; Boyd <i>et al.</i> , 2003)
pAsa3	5 616	NC_004924	ColE2	-	-	-	(Belland et Trust, 1989; Boyd <i>et al.</i> , 2003)
pAsa11	6 371	NC_004338	ColE2	-	Effecteur du SSTT <i>aoppP</i>	ISAS11	(Fehr <i>et al.</i> , 2006)
pAsa11B	8 989	NC_025018	ColE2	-	<i>id.</i>	ISAS5, ISAS11	(Trudel <i>et al.</i> , 2013)
pAsa11C	8 989	KT781680	ColE2	-	<i>id.</i>	ISAS5, ISAS11	(Attéré <i>et al.</i> , 2015)
pAsa11D	9 138	KT781681	ColE2	-	<i>id.</i>	ISAS5, ISAS11	(Attéré <i>et al.</i> , 2015)
pAsa2	5 247	NC_004925	ColE1	-	-	-	(Belland et Trust, 1989; Boyd <i>et al.</i> , 2003)
pAsa7*	5 276	NZ_KU499859	ColE1	CHL	-	-	(Vincent <i>et al.</i> , 2016a)
Plasmides de résistance aux antibiotiques							
pAsa8	110 577	KX364409	ND	AMP/CB, CHL, FLO, STR, TET, SUL		Tn1721, IS5	(Trudel <i>et al.</i> , 2016)
pRAS1	~ 45 000	-	IncU	SUL, TMP	<i>tetA</i> non fonctionnel?	Tn1721 tronqué, Intégron similaire à In4, IS6100	(Sørum <i>et al.</i> , 2003)
pAr-32	ND	-	IncU	CHL, STR, SUL	-	Intégron similaire à In6	(Aoki <i>et al.</i> , 1971)
pASOT1	< 47 000	-	ND	OT, STR	-	-	(Adams <i>et al.</i> , 1998)
pASOT3	< 47 000	-	ND	OT	-	-	(Adams <i>et al.</i> , 1998)
pASOT3	< 39 000	-	ND	OT, STR	-	-	(Adams <i>et al.</i> , 1998)
pRAS2	~ 48 000	-	ND	STR, SUL, TET	-	Tn5393c	(L'Abée-Lund et Sørum,

									2000)
pRAS4	> 45 000	-	ND	ND	-	-	-	-	(L'Abéc-Lund et Sørum, 2002)
pRAS3.1	11 851	AY043298	ND	TET	-	-	-	-	(L'Abéc-Lund et Sørum, 2002)
pRAS3.3	11 845	-	ND	TET	-	-	-	-	(Vincent <i>et al.</i> , 2014b)
pRAS3.4	ND	KJ909291	ND	TET	-	-	-	-	(Vincent <i>et al.</i> , 2014b)
pAb55S9b	25 540	KJ909292	ND	CHL, FLO, STR, SUL, TET	-	-	ISCR2 (complète et partielle)	-	(Vincent <i>et al.</i> , 2014b)
pSN254b	152 216	KJ909290	A/C ₂	BLA, CHL, FLO, STR, SUL, TET	Opéron <i>mer</i> , <i>qacEΔI</i>	Tn21, IS26, IS4321, IS6100, ISEc9	-	-	(McIntosh <i>et al.</i> , 2008 ; Vincent <i>et al.</i> , 2014b)
pAsa4	166 749	NC_009349	Apparenté à A/C	CHL, STR, SUL, TET	Opéron <i>mer</i> , <i>qacEΔI</i>	Tn21, ISAS1, ISAS2, ISAS5, ISAS9	-	-	(Belland et Trust, 1989; Reith <i>et al.</i> , 2008)
Plasmides reliés à pAsa5									
pAsa5	155 098	NC_009350	ND	-	SSTT, <i>aopH</i> , <i>ant2</i> , <i>aopO</i> (effecteurs)	ISAS3, ISAS4, ISAS5, ISAS10, ISAS11	-	-	(Reith <i>et al.</i> , 2008; Stubber <i>et al.</i> , 2003)
pAsa6**	18 536	NC_009352	ND	-	Effecteur <i>aopH</i>	ISAS2, ISAS11 partielle	-	-	(Najimi <i>et al.</i> , 2009)

* Plasmide apparenté à pAsa2

** Plasmide apparenté à pAsa5

ND : Non disponible ; *id.* : *idem* ; SSTT : système de sécrétion de type trois

pAsa1B, pAsa1C, pAsa1D, pAb55S9b et pSN254b sont tous des variants d'autres plasmides (voir dans le texte)

Abréviations associées aux antibiotiques : AMP/CB : ampicilline/carbénicilline ; BLA : bêta-lactamases ; CHL : chloramphénicol ;

FLO : floriténicol, OT : oxytétracycline, STR : streptomycine/spectinomycine, SUL : sulfamides, TET : tétracycline

1.4.6 LES SÉQUENCES D'INSERTION CHEZ *A. SALMONICIDA* SSP. *SALMONICIDA*

Lors de l'obtention du génome de référence, le nombre important d'IS (88 complètes et 13 partielles) a été remarqué par les auteurs (Reith *et al.*, 2008). Étant donné que le nombre moyen de transposases par génome bactérien qui en possède au moins une copie (86% des génomes analysés lors de l'étude) est d'environ 40, ce taux d'IS est inhabituel (Aziz *et al.*, 2010). Les IS d'A449 sont réparties en dix types (Tableau 1.3) (Reith *et al.*, 2008). Elles sont présentes sur le chromosome et sur les deux grands plasmides de cette souche, pAsa4 et pAsa5 (Reith *et al.*, 2008). On en retrouve aussi sur d'autres plasmides portés par *A. salmonicida* ssp. *salmonicida*, comme pAsa1 (Fehr *et al.*, 2006), pAsa6 (Najimi *et al.*, 2009), ou pSN254b (Vincent *et al.*, 2014b). Ce nombre important d'IS est aussi un problème lors de l'assemblage de ces génomes, puisqu'elles fragmentent les contigs (Charette *et al.*, 2012; Vincent *et al.*, 2014a; Vincent *et al.*, 2015).

Les multiples IS du génome d'*A. salmonicida* ssp. *salmonicida* sont d'autant de possibilités de transpositions et de réarrangements qui peuvent avoir des conséquences sur le contenu, l'architecture du génome et les fonctions qui en découlent (voir section 1.3.5.3 et Figure 1.3). Des études ont d'ailleurs mis en lumière différents exemples de variations génétiques et phénotypiques causées par les IS chez cette bactérie. En voici un résumé.

Tableau 1.3 Séquences d'insertion retrouvées dans le génome d'*A. salmonicida* ssp. *salmonicida* A449

Nom*	Taille (pb)	Famille	Nombre de copies, A449†	Activité	Implication	Autres noms	Source
IS <i>AsI</i>	1223	IS <i>AsI</i>	2	Oui	Inactivation, <i>A-layer</i> (Gustafson <i>et al.</i> , 1994)	IS <i>AsI</i> (Reith <i>et al.</i> , 2008)	(Gustafson <i>et al.</i> , 1994)
IS <i>As2</i>	1084	IS30	5	Oui	Inactivation, <i>A-layer</i> (Gustafson <i>et al.</i> , 1994)	IS <i>As2</i> (Reith <i>et al.</i> , 2008)	(Gustafson <i>et al.</i> , 1994)
IS <i>As3</i>	1233	IS3	12	-	-	IS <i>As5</i> (Reith <i>et al.</i> , 2008), IS <i>As7</i> NC_009348	(Merino <i>et al.</i> , 2003)
IS <i>As4</i>	1165	IS630	37	Suspectée	Transfert horizontal (Studer <i>et al.</i> , 2013)	IS <i>As7</i> (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
IS <i>As5</i>	2614	IS21	12	Oui	Plasmide p <i>Asal</i> 1B, p <i>Asal</i> 1C et p <i>Asal</i> 1D (Attéré <i>et al.</i> , 2015; Trudel <i>et al.</i> , 2013)	IS <i>As11</i> (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
IS <i>As6</i>	1240	IS3	7	-	-	IS <i>As6</i> (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
IS <i>As8</i>	754	IS1	3	-	-	IS <i>As8</i> (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
IS <i>As9</i>	1624	IS3	4	-	-	IS <i>As9</i> (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
IS <i>As10</i>	1229	IS30	2	-	-	IS <i>As10</i> (Reith <i>et al.</i> , 2008)	(Reith <i>et al.</i> , 2008)
IS <i>As11</i>	1326	IS256	4	Oui	Délétion, locus SSTT (Tanaka <i>et al.</i> , 2012)	IS <i>As3</i> (Fehr <i>et al.</i> , 2006) et NC_004338	(Fehr <i>et al.</i> , 2006)

* Nom tel qu'employé dans les fichiers GenBank de *A. salmonicida* subsp. *salmonicida* (chromosome A449, NC_009348; p*Asal*, NC_004923; p*Asa2*, NC_004925; p*Asa3*, NC_004924; p*Asal1*, NC_004338; p*Asa4*, NC_009349; p*Asa5*, NC_009350) du *National Center for Biotechnology Information*, sauf si mentionné dans autres noms.

† Informations tirées de (Reith *et al.*, 2008).

1.4.6.1 VARIATIONS CRYPTIQUES DANS LES PLASMIDES

Le plasmide pAsal1 se décline en quatre variants : la séquence originale pAsal1, pAsal1B, pAsal1C et pAsal1D (Attéré *et al.*, 2015; Fehr *et al.*, 2006; Trudel *et al.*, 2013). Les variants possèdent une ISAS5 en plus de l'ISAS11 déjà présente dans pAsal1. Chez pAsal1B, l'ISAS5 est insérée dans le gène de mobilisation *mobA*. Un autre fragment (121 pb) d'ISAS5 est aussi retrouvé dans une région non codante en amont du même gène (Trudel *et al.*, 2013). Chez pAsal1C et pAsal1D, l'ISAS5 est insérée au même site dans l'ISAS11. Cependant, pour pAsal1D, une duplication dans la séquence de l'ISAS11 permet d'obtenir une copie complète de cette IS (Attéré *et al.*, 2015). Dans chaque cas, le variant n'a été identifié que dans une seule souche, toutes sont géographiquement isolées les unes des autres (Attéré *et al.*, 2015; Trudel *et al.*, 2013). Ceci suggère que la transposition d'ISAS5 dans pAsal1 peut survenir de façon indépendante dans un petit nombre de souches (Attéré *et al.*, 2015). De plus, les transpositions d'ISAS5 pourraient converger vers pAsal1 puisqu'il est généralement présent en plus grand nombre de copies que tout autre réplicon et que l'insertion dans une IS pourrait minimiser l'impact délétère de la transposition (Attéré *et al.*, 2015).

1.4.6.2 TRANSPOSITION DES ISAS1 ET ISAS2 DANS LE GÈNE *VAPA* ET *ABCA*

Ishiguro *et al.* (1981) ont été les premiers à observer qu'*A. salmonicida* ssp. *salmonicida* perdait sa virulence lorsqu'elle était cultivée à des températures trop élevées. Dans leur étude, une fraction (moins de 10%) des bactéries cultivées à plus de 25°C perdait leur couche de surface A, ce qui était accompagné d'une augmentation de 100 à 1 000 fois de la dose létale médiane chez *Oncorhynchus kisutch* (Ishiguro *et al.*, 1981). Plus tard, des études ont montré que le gène qui code pour la protéine de surface A, *vapA*, pouvait être inactivé de deux façons. D'abord, *vapA* peut subir des délétions internes diverses, possiblement par recombinaison de séquences similaires (Chu *et al.*, 1991b; Gustafson *et al.*, 1992). Un autre type d'inactivation fait intervenir deux IS, ISAS1 ou ISAS2, dans la séquence de *vapA* et dans son promoteur (Gustafson *et al.*, 1994). Des insertions à plusieurs endroits ont été détectées sur des souches (A449 et A450) cultivées à 30°C (Gustafson *et al.*, 1994). De

plus, une insertion d'*ISASI* a aussi été retrouvée dans le gène directement en aval de *vapA*, *abcA* (Chu *et al.*, 1995). La protéine AbcA d'*A. salmonicida* ssp. *salmonicida* est responsable de l'assemblage et de l'exportation de l'antigène O des LPS (Chu *et al.*, 1995).

1.4.6.3 PERTE DU SSTT ET DE LA VIRULENCE

La perte des gènes du SSTT a été observée assez rapidement après la découverte du plasmide porteur, pAsa5 (pASvirA) (Stuber *et al.*, 2003). Tout comme l'inactivation de *vapA*, cette délétion se manifeste après une culture à une température trop élevée (Daher *et al.*, 2011; Stuber *et al.*, 2003). Une température de 25°C est suffisante pour supprimer le SSTT, et, par conséquent, rendre *A. salmonicida* ssp. *salmonicida* complètement avirulente (Daher *et al.*, 2011; Stuber *et al.*, 2003). Cependant, le temps de culture nécessaire pour produire la délétion semble variable. En effet, une perte du SSTT chez toute la population a été observée en seulement 6 heures à 25°C pour la souche JF2267 (Stuber *et al.*, 2003). Pour la souche A449, 30 heures à 28°C n'ont pas été suffisantes pour éliminer le SSTT (Ebanks *et al.*, 2006). Daher (2010) ont montré que les souches A449, 01-B526 et 01-B516 pouvaient perdre leur SSTT après une culture de deux semaines à 25°C. Cependant, cette perte n'était expérimentée que par une fraction de la population, nécessitant une culture en milieu gélosé pour isoler les mutants avirulents (Daher, 2010). Le taux de réarrangement observé au cours de cette étude était variable d'une souche à l'autre et même d'une expérience à l'autre (Daher, 2010).

La même étude a aussi montré que les délétions de pAsa5 suivaient un certain patron. Ceci a été détecté par géotypage PCR en utilisant 14 cibles réparties sur le plasmide, dont six dans la région du SSTT (Figure 1.5, sous la carte linéaire). Pour des souches cultivées de façon routinière à 25°C, puis acquises et analysées *a posteriori*, plusieurs cas de figure ont été obtenus, mais tous comprenaient une délétion des gènes structuraux du SSTT (Daher, 2010). Cependant, pour les souches SSTT négatives produites spécifiquement pour cette étude, deux patrons de délétion étaient observés. Les souches présentant une délétion de type 1 avaient perdu uniquement les gènes du locus du SSTT. Les souches présentant une délétion de type 2 avaient aussi perdu les gènes en amont du SSTT jusqu'à la première cible de géotypage utilisée sur la séquence du plasmide (Figure 1.5) (Tanaka *et al.*, 2012).

Une comparaison entre les profils de génotypage et la position des IS sur le plasmide pAsa5 a montré que les délétions étaient systématiquement encadrées par des IS de type IS*ASI1* ou IS*AS4* (Figure 1.5) (Tanaka *et al.*, 2012). De plus, les séquences d'insertion étaient dans la même orientation les unes par rapport aux autres, ce qui, pour une molécule circulaire, peut mener à des délétions par recombinaison homologue des séquences identiques (Figure 1.3A) (Tanaka *et al.*, 2012).

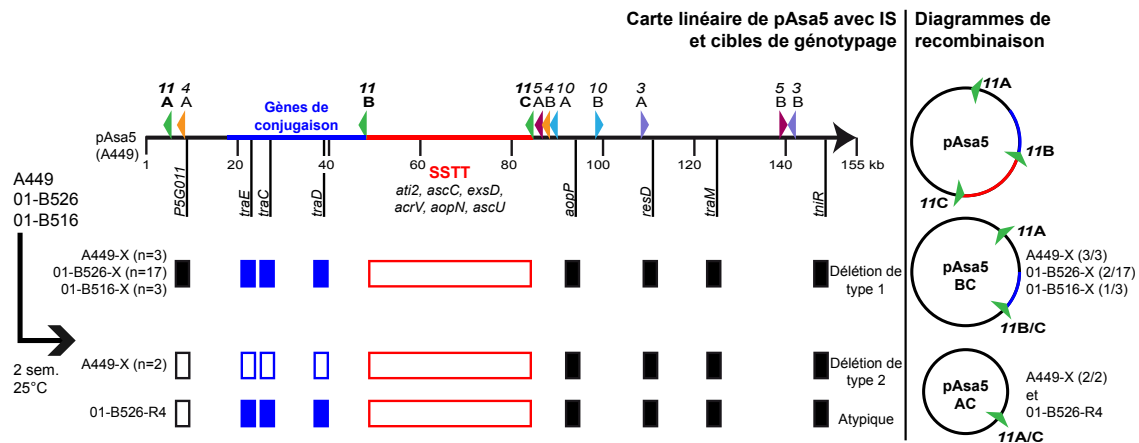


Figure 1.5 Résumé des études sur la délétion du locus du SSTT par recombinaison homologue. Cette figure résume des travaux effectués précédemment (Daher *et al.*, 2011; Tanaka *et al.*, 2012). 14 cibles de génotypage, dont 6 dans le locus du SSTT, ont été créées pour le plasmide pAsa5 (indications en italiques sous la carte linéaire). En parallèle, des mutants SSTT négatifs avirulents ont été produits par culture prolongée à 25°C de trois souches parentales (à gauche). Leurs profils de génotypage ont été classés en deux types plus un profil atypique (rectangles pleins, positifs; vides, négatifs). Dans tous les cas, la région du SSTT (rouge) était manquante. Pour les deux profils principaux, les délétions étaient encadrées par des IS (flèches colorées sur la carte linéaire) (Daher *et al.*, 2011). Dans une autre étude, des amorces directement en amont des IS*ASI1A* et B, et directement en aval de l'*ISASI1C* (non montrées) ont été utilisées pour tester une hypothèse de recombinaison homologue de ces éléments. Le test a montré qu'une fraction (exprimée à droite) des souches à délétion de type 1 possédait un pAsa5 BC et que toutes les souches à délétion de type 2 et la souche à délétion atypique (01-B526-R4) possédaient un pAsa5 AC (Tanaka *et al.*, 2012).

Une recombinaison homologue entre copies d'une même IS laisse un exemplaire intact au site de délétion. Pour montrer que les pAsa5 des souches dérivées d'A449, de 01-B526 et de 01-B516 répondaient bien à ce modèle, un test PCR utilisant des amorces directement en

amont ou en aval des copies concernées a été créé (Tanaka *et al.*, 2012). Ce test permettait d'amplifier un fragment d'ADN correspondant à l'IS, à la section en amont d'une des copies et à la section en aval de l'autre si recombinaison avait eu lieu. Pour certains des plasmides des souches réarrangées étudiées, la délétion était bel et bien causée par recombinaison de deux copies de séquences d'insertion (Figure 1.5) (Tanaka *et al.*, 2012). Seules les IS*AS11* étaient impliquées dans la recombinaison. Une fraction des plasmides à délétion de type 1 avait subi un réarrangement BC (recombinaison des IS*AS11B* et IS*AS11C*). Tous les plasmides à délétion de type 2 (beaucoup moins fréquents) avaient subi un réarrangement AC (recombinaison IS*AS11A*-IS*AS11C*) (Figure 1.5) (Tanaka *et al.*, 2012). La seule souche dérivée avec un profil de génotypage atypique, 01-B526-R4, avait subi un réarrangement de type AC (Figure 1.5) (Daher, 2010; Tanaka *et al.*, 2012). Par contre, plusieurs souches réarrangées n'ont pas vu leur délétion dans pAsa5 expliquée par recombinaison. Les auteurs avaient suggéré qu'elles avaient subi des séries d'évènements plus complexes et ne pouvant être analysées par PCR (Tanaka *et al.*, 2012).

1.5 ÉNONCÉ DE LA PROBLÉMATIQUE DE RECHERCHE

Au moment de commencer le projet, à l'été 2012, les données en génomique sur *A. salmonicida* ssp. *salmonicida* étaient encore minimales. Outre le génome de référence, obtenu par séquençage Sanger, une seconde tentative de séquençage avait été réalisée sur le génome de 01-B526 en pyroséquençage (Charette *et al.*, 2012). Cependant, l'utilisation d'une librairie de type *shotgun* à l'époque avait donné un assemblage fragmenté annonciateur des difficultés d'analyse de ce génome causées par les IS. Depuis, les données génomiques sur *A. salmonicida* ssp. *salmonicida* ont explosé, avec 23 souches dont le génome est disponible et 16 plasmides assemblés. Plusieurs études ont aussi mis en lumière la diversité des éléments mobiles de tous types (plasmides, transposons et îlots génomiques) pour cette bactérie. Certains de ces exemples, comme mentionnés précédemment, impliquent des IS.

Mon projet de doctorat s'imbriquait donc dans une suite plus large d'études visant à renforcer les connaissances sur *A. salmonicida* ssp. *salmonicida* pour mieux comprendre son épidémiologie, son interaction hôte-pathogène et son génome accessoire. Ultiment,

ces informations serviront à développer des mesures de suivi de cet agent pathogène et un plus large éventail de solutions prophylactiques et thérapeutiques contre la furonculose. Cette thèse a donc pu s'appuyer sur un tronc commun d'information obtenu par séquençage à haut débit.

Les IS du génome d'*A. salmonicida* ssp. *salmonicida* entraînent cette bactérie dans le paradigme de coévolution qui a été décrit précédemment (section 1.3.5.5). Un équilibre doit être établi entre la prolifération et l'inactivation des IS, sans quoi l'élément mobile ou la bactérie seront en danger d'extinction. Quoi qu'il en soit, le destin évolutif d'*A. salmonicida* ssp. *salmonicida* est déjà influencé par ses IS si elles limitent effectivement sa température de croissance (Vincent *et al.*, 2016b).

Le nombre élevé d'IS entraîne les délétions et les pertes de fonctions qui ont été observées. Cependant, les conséquences qu'elles peuvent avoir sur des fonctions cruciales de l'agent pathogène sont en inadéquation avec l'efficacité avec laquelle la bactérie cause la furonculose. Tous les exemples de mobilité des IS dans *A. salmonicida* ssp. *salmonicida* semblent négatifs (perte de virulence) ou neutre (variation de plasmides) du point de vue de la bactérie. Cependant, les pertes de virulence causées par des IS surviennent à des températures où l'infectivité est habituellement plus élevée (Groberg *et al.*, 1978).

Le nombre d'IS du génome d'*A. salmonicida* ssp. *salmonicida* est peut-être le résultat de plusieurs infections récentes de cette classe d'élément mobile, mais qu'elles sont destinées à être inactivées rapidement. Il est aussi possible que les IS de ce génome confèrent à la fois des avantages positifs et négatifs à cette bactérie. Puisque cette bactérie est en bonne position pour échanger du matériel génétique par transfert horizontal, cette contribution positive pourrait être dans l'intégration de gènes accessoires utiles à son mode de vie pathogène. L'apport positif des IS au génome d'*A. salmonicida* ssp. *salmonicida* pourrait aussi venir de l'augmentation de la plasticité génomique causée par les transpositions et les recombinaisons. La contribution à l'adaptation de cet agent pathogène à son milieu peut être positive même s'il y a délétion de gènes, comme ce fut observé pour plusieurs organismes en réduction de génome (Chain *et al.*, 2004; Kaleta *et al.*, 2010; Parkhill *et al.*, 2003a).

L'hypothèse de ce projet est donc que les IS peuvent apporter des modifications à valeur adaptative positive au génome d'*A. salmonicida* ssp. *salmonicida*. La génomique comparative a été utilisée pour obtenir les informations sur de nouveaux événements impliquant les IS. Une partie du projet a été consacrée à faire une adéquation entre les informations obtenues par la génomique et les phénotypes observés.

Les premiers résultats de séquençage ont mené à l'assemblage de variants du plasmide pAsa4. Les IS et d'autres éléments mobiles de l'ADN sont impliqués dans des variations qui affectent la capacité de conjugaison des souches porteuses du plasmide et leurs résistances aux antibiotiques. Ces travaux ont été publiés dans *PeerJ* (Tanaka *et al.*, 2016) et sont présentés, sous forme d'article, au Chapitre 2.

Une seconde série de résultats de séquençage a mené à l'assemblage d'un nouveau plasmide pour *A. salmonicida* ssp. *salmonicida* 01-B526, pAsa9, ainsi qu'à l'assemblage du plasmide pAsa5 de la même souche. Or, ces nouvelles connaissances sur les plasmides de 01-B526 ont permis de réviser le mécanisme de délétion du locus du SSTT par recombinaison homologue des IS, qui avait été proposé avant le début de mon doctorat (Tanaka *et al.*, 2012). Le Chapitre 3 présente donc une suite de cette étude, qui a été publiée dans *BMC Genomics* (Tanaka *et al.*, 2017).

Le Chapitre 4 revient sur les résultats des deux articles précédents et leur contribution à l'hypothèse de départ. L'impact de chaque événement impliquant des IS sera discuté plus en détail. Finalement, la contribution des IS au passé, présent et futur d'*A. salmonicida* ssp. *salmonicida* sera abordée dans un chapitre de conclusion (Chapitre 5).

Chapitre 2 LES CONSÉQUENCES SUR LA RÉSISTANCE AUX ANTIBIOTIQUES DE L'ARCHITECTURE EN MOSAÏQUE DES PLASMIDES DE LA FAMILLE DE PASA4.

RÉSUMÉ

Certains isolats d'*Aeromonas salmonicida* ssp. *salmonicida*, l'agent étiologique de la furunculose, sont résistants à plusieurs antibiotiques. Le plasmide pAsa4 de cette bactérie porte plusieurs de ces gènes de résistance. Des résultats récents suggèrent que des formes alternatives de ce réplicon existent. Dans cette étude, nous avons identifié deux variants de pAsa4, qui ont été nommés pAsa4b et pAsa4c. Les trois plasmides possèdent des différences d'architecture dont certaines ont une influence directe sur le contenu en gènes de résistance. La plasticité génomique au sein de la famille des pAsa4 est partiellement expliquée par la présence et l'activité des éléments mobiles de l'ADN, dont les séquences d'insertion. Aussi, seuls pAsa4b et pAsa4c sont transférables vers *Escherichia coli*, ce qui reflète leur composition en gènes de conjugaison. Les plasmides de la famille de pAsa4 pourraient être un élément facilitant la propagation d'*A. salmonicida* ssp. *salmonicida* multirésistantes et l'acquisition de nouveaux gènes.

**THE MOSAIC ARCHITECTURE OF *AEROMONAS SALMONICIDA*
SUBSP. *SALMONICIDA* P_{ASA4} PLASMID AND ITS CONSEQUENCES ON
ANTIBIOTIC RESISTANCE**

Katherine H. Tanaka^{1,2,3}, Antony T. Vincent^{1,2,3}, Mélanie V. Trudel^{1,2,3}, Valérie E. Paquet^{1,2,3}, Michel Frenette^{2,4}, and Steve J. Charette^{1,2,3}

¹Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Quebec City, Quebec, Canada

²Département de biochimie, de microbiologie et de bio-informatique, Faculté des sciences et de génie, Université Laval, Quebec City, Quebec, Canada

³Centre de recherche de l'Institut universitaire de cardiologie et de pneumologie de Québec, Quebec City, Quebec, Canada

⁴Groupe de Recherche en Écologie Buccale (GREB), Faculté de médecine dentaire, Université Laval, Quebec City, Quebec, Canada

ABSTRACT

Aeromonas salmonicida subsp. *salmonicida*, the causative agent of furunculosis in salmonids, is an issue especially because many isolates of this bacterium display antibiotic resistances, which limit treatments against the disease. Recent results suggested the possible existence of alternative forms of pAsa4, a large plasmid found in *A. salmonicida* subsp. *salmonicida* and bearing multiple antibiotic resistance genes. The present study reveals the existence of two newly detected pAsa4 variants, pAsa4b and pAsa4c. We present the extensive characterization of the genomic architecture, the mobile genetic elements and the antimicrobial resistance genes of these plasmids in addition to the reference pAsa4 from the strain A449. The analysis showed differences between the three architectures with consequences on the content of resistance genes. The genomic plasticity of the three pAsa4 variants could be partially explained by the action of mobile genetic elements like insertion sequences. Eight additional isolates from Canada and Europe that bore similar antibiotic resistance patterns as pAsa4-bearing strains were genotyped and specific pAsa4 variants could be attributed to phenotypic profiles. pAsa4 and pAsa4c were found in Europe, while pAsa4b was found in Canada. In accordance with their content in conjugative transfer genes, only pAsa4b and pAsa4c can be transferred by conjugation in *Escherichia coli*. The plasticity of pAsa4 variants related to the acquisition of antibiotic resistance indicates that these plasmids may pose a threat in terms of the dissemination of antimicrobial-resistant *A. salmonicida* subsp. *salmonicida* bacteria.

INTRODUCTION

The ubiquitous waterborne Gram-negative bacterium *Aeromonas salmonicida* subsp. *salmonicida* is the causative agent of furunculosis, a disease that affects aquaculture operations worldwide (Derome *et al.*, 2016). The main treatments for this disease are vaccination and antibiotics. Vaccination was shown to be efficient but is expensive and may cause major side effects (Dallaire-Dufresne *et al.*, 2014b). Antibiotics are becoming increasingly less effective against *A. salmonicida* subsp. *salmonicida* due to the spread of antibiotic resistance genes. For example, more and more antibiotic-resistant *A. salmonicida* subsp. *salmonicida* strains are being isolated and characterized, many of them bearing resistance genes on plasmids (McIntosh *et al.*, 2008; Piotrowska and Popowska, 2015; Sørum *et al.*, 2003; Vincent *et al.*, 2016a; Vincent *et al.*, 2014b).

In *A. salmonicida* subsp. *salmonicida*, insertion sequences (ISs) are responsible for several genomic modifications (Vincent *et al.*, 2016b). ISs are made of a transposase gene and inverted repeats. Some ISs are involved in virulence loss when *A. salmonicida* subsp. *salmonicida* is under stressful conditions (ISAS1, ISAS2 and ISAS11) (Gustafson *et al.*, 1994; Tanaka *et al.*, 2012). Furthermore, many plasmid variants display transpositions or IS-mediated recombinations when compared to their reference (ISAS5 in many plasmids, ISEc9 in pSN254b) (Attéré *et al.*, 2015; Najimi *et al.*, 2009; Trudel *et al.*, 2013; Vincent *et al.*, 2014b). Given the high number of ISs in the genome of this bacterium (Studer *et al.*, 2013; Vincent *et al.*, 2016b), we hypothesize that ISs play a role in plasmid reshaping (Tanaka *et al.*, 2013).

The large plasmid pAsa4 from *A. salmonicida* subsp. *salmonicida* carries genes that provide resistance against chloramphenicol, spectinomycin, streptomycin, sulfonamides, tetracycline, mercury, and quaternary ammonium compounds (Reith *et al.*, 2008). Except for tetracycline resistance, these genes are located in Tn21, a non-composite transposon. Tn21 is a widespread replicative transposon that also carries another mobile element, the integron In2 (Liebert *et al.*, 1999). The complete sequence of pAsa4 was first described in reference strain A449, which originated from France (Reith *et al.*, 2008). Genotyping done in a previous study has shown that some *A. salmonicida* subsp. *salmonicida* isolates likely

bear pAsa4 but do not display the expected antibiotic resistance profile (Vincent *et al.*, 2014b). This suggests that pAsa4 variants may have evolved from a common replicon backbone, but do not share the same antibiotic resistance genes.

We used next-generation sequencing (NGS) on two isolates, one from the province of Quebec (Canada) and one from Switzerland, suspected of carrying pAsa4 variants based on preliminary genotyping and antibiotic resistance profiles, to obtain the complete sequences of the two plasmids. Both plasmids exhibited marked differences from the original pAsa4 plasmid from the reference strain A449 and from each other. A detailed analysis of these pAsa4 variants is presented.

MATERIAL AND METHODS

BACTERIAL ISOLATES, GROWTH CONDITIONS, ANTIBIOTIC RESISTANCE PROFILES, AND CONJUGATION ASSAYS

The 129 *A. salmonicida* subsp. *salmonicida* strains listed in Table S1 were included in this study. All strains were grown on furunculosis agar (10 g of Bacto-Tryptone, 5 g of yeast extract, 1 g of L-tyrosine, 2.5 g of NaCl, and 15 g of agar per liter of distilled water) or tryptic soy agar (TSA) for two or three days at 18°C (Hänninen and Hirvela-Koski, 1997). *Escherichia coli* DH5 α was grown on lysogeny broth (LB) agar for one day at 37°C. Disk diffusion assay using chloramphenicol (30 μ g), florfenicol (30 μ g), sulfamethoxazole/trimethoprim (SXT) (23.75 μ g/1.25 μ g), and tetracycline (5 μ g) disks (Becton Dickinson, USA) were performed for strains listed in Table S1 as done previously (Vincent *et al.*, 2014b).

Bacterial conjugation assays have been done as previously described (Boyd *et al.*, 2008). *A. salmonicida* A449, 01-B522 and JF2267 (donor strains) were pre-cultivated in 2 ml of tryptic soy broth (TSB) at 18°C overnight. *E. coli* DH5 α (recipient strain) was pre-cultivated in 2 ml of LB at 37°C for the same period of time. For each conjugation experiment, cultures of donor and recipient cells (1 ml each) were harvested by centrifugation at 17,200 x g for 1 minute, suspended in 20 μ l TSB, mixed together, and spotted on TSA without selection for 24 hours at 18°C. Afterwards, the culture was

suspended in TSB, diluted and plated on TSA with either 5 µg/ml tetracycline (pAsa4 and pAsa4b) or 5 µg/ml chloramphenicol (pAsa4 and pAsa4c). Plates were incubated overnight at 37°C to select against *A. salmonicida* which is psychrophilic (Vincent *et al.*, 2016b). Large colonies were picked and suspended in TSB with appropriate selection. The presence of pAsa4 variants in transformants was confirmed by PCR using primer pairs *traG*, 2, 3, 9 and 10 (Table S2), *A. salmonicida* absence was confirmed by *tapA*. Conjugation assays were performed twice for every pAsa4 variant.

DNA EXTRACTION AND SEQUENCING

The total genomic DNA of two isolates (01-B522 and JF2267) was extracted using DNeasy Blood and Tissue kits (Qiagen, Canada) and was sequenced at the Plateforme d'Analyse Génomique of the Institut de Biologie Intégrative et des Systèmes (IBIS, Université Laval). For JF2267, a 650-bp shotgun library was sequenced using 454 GS-FLX+ technology. Isolate 01-B522 was sequenced as previously described (454 GS-FLX+ technology, mate-pair library with 5 kbp fragment size and 1500 bp library size) (Vincent *et al.*, 2014a). The reads were assembled *de novo* using Newbler version 2.5.3 with default parameters (Margulies *et al.*, 2005).

SEQUENCE ANALYSIS

Contigs resulting from the assembly of 01-B522 and JF2267 were initially mapped locally on the sequence of the pAsa4 from A449 (GenBank accession number NC_009349.1) using CONTIGuator version 2.7.4 (Galardini *et al.*, 2011). All contig junctions were manually verified by PCR and Sanger sequencing and links were joined using Consed version 27 (Gordon and Green, 2013).

The assembled plasmids were annotated as follows. Open reading frames (ORFs) were predicted by getorf (available as a part of EMBOSS 6.6.0.0) (Rice *et al.*, 2000). All the detected ORFs were then compared to pAsa4 coding sequences using fasta36 (Pearson and Lipman, 1988). Lastly, the remaining ORFs were annotated using Blastn and Blastp (Altschul *et al.*, 1990) against the NCBI non-redundant (nr/nt) database and, if necessary,

against the whole genome shotgun database (wgs, Gammaproteobacteria (taxid:1236)). Annotations were manually verified using the Artemis version 16.0.0 visualization tool, and alignments between the assembled pAsa4 were visualized using EasyFig 2.1 and ACT 13.0.0 (Rutherford *et al.*, 2000; Sullivan *et al.*, 2011). Insertion sequence nomenclature follows the one of *A. salmonicida* A449 original Genbank annotation (“ISAS” names, which differs from “ISAs” nomenclature) (Reith *et al.*, 2008). Antibiotic resistance genes were validated with The Comprehensive Antibiotic Resistance Database (CARD) (McArthur *et al.*, 2013). The annotated sequences of pAsa4b and pAsa4c were deposited in GenBank under accession numbers KT033469 and KT033470, respectively.

The average copy number per cell for pAsa4b in 01-B522 and pAsa4c in JF2267 were estimated by mapping the sequencing reads using TAPyR v1.3-beta4 (Fernandes *et al.*, 2011) and by calculating the average coverage using Qualimap 2.0 (García-Alcalde *et al.*, 2012). The copy numbers were standardized against the average coverage of the *gyrB* housekeeping gene (single copy per chromosome).

Contigs from two other *A. salmonicida* strains, RS 534 (NCBI wgs JYFF00000000) (Vincent *et al.*, 2016b) and JF3517 that had been sequenced previously (Attéré *et al.*, 2015) were mapped against pAsa4, pAsa4b, and pAsa4c using CONTIGuator version 2.7.4.

A global alignment of pAsa4b and pAsa4c was performed using *stretcher* (available as a part of EMBOSS 6.6.0.0) (Rice *et al.*, 2000), and a custom R script (R Core Team, 2015) was used to visualize the number of substitutions by 1000-bp sliding windows (Data S1) (Zeileis and Grothendieck, 2005). For the heatmap representations, all the ORFs from pAsa4b were compared to the NCBI nucleotide collection (nr/nt) using tBlastn (Altschul *et al.*, 1990). The data was ordered and visualized using a custom R script (Wickham, 2009). k-means clustering was used to group target sequence identifiers based on the matrix results in as many clusters that could create reproducible grouping (Data S2) (Hartigan and Wong, 1979).

PCR ANALYSES

The DNA templates, PCR mixtures, and program cycles were performed as previously described (Trudel *et al.*, 2013), with the exception of the elongation time, which was 1 min

per kbp of amplicon. The PCR assays were performed at least twice, and appropriate positive and negative controls were included with each assay. The PCR primers are listed in Table S2. Genotyping primers were designed using PrimerBlast (Ye *et al.*, 2012) at plasmid insertion/deletion sites (junction between segments, Fig. 2.1B).

RESULTS

COMPLETE SEQUENCES OF THE NEW pAsa4 VARIANTS

A. salmonicida subsp. *salmonicida* isolate 01-B522 harbored a potential pAsa4 variant based on the genotyping results and antibiotic-resistance profile (Vincent *et al.*, 2014b) (Table 2.1). Isolate JF2267 displayed genotyping results similar to those of 01-B522, but had a different antibiotic resistance profile (Table 2.1). To determine the complete sequences of these potential pAsa4 variants, pyrosequencing, PCR and Sanger sequencing were used to assemble the complete plasmid sequences. The reference plasmid pAsa4 is composed of 166,749 bp and 173 ORFs and has a G+C content of 52.8% (Reith *et al.*, 2008) compared to 181,933 bp, 175 ORFs, and a G+C content of 52.48% for 01-B522 pAsa4 variant (pAsa4b), and 163,022 bp, 156 ORFs, and a G+C content of 53.42% for JF2267 variant (pAsa4c). Based on the relative coverage of the sequenced reads compared to *gyrB* coverage, the estimated number of copies of the pAsa4 variants in 01-B522 and JF2267 was 1 in both cases. The contigs alignments of previously sequenced strains RS 534 and JF3517 indicated that they had the same content as pAsa4 and pAsa4c, respectively.

pAsa4 and its variants bear many ORFs coding for hypothetical proteins (Fig. 2.1A, cyan arrows). The plasmids also carry ORFs for their replication and partition and for proteins with other functions (all shown in Fig. 2.1A as dark grey arrows). Furthermore, two regions bear resistance antibiotic genes (Fig. 2.1A, orange arrows). *Tn21*, a transposon whose presence was already acknowledged in pAsa4, carries most of the resistance genes via its built-in integron, *In2* (Liebert *et al.*, 1999; Reith *et al.*, 2008). A tetracycline resistance gene and its repressor are located elsewhere on the plasmid (Fig. 2.1A). Finally, the conjugation-related genes (Fig. 2.1A, blue arrows) are separated in two loci.

Table 2.1 *Aeromonas salmonicida* subsp. *salmonicida* strains bearing pAsa4 variants

Strain	Source (host)*	Origin*	Antibiotic resistance determined by antibiogram†	pAsa4 variant determined by genotyping	Reference
A449	Brown trout	France	TET, CHL	pAsa4	(Reith <i>et al.</i> , 2008)
RS 534 (A450)	INA	France	TET, CHL	pAsa4	(Kay <i>et al.</i> , 1981)
01-B522	Brook trout	Quebec (Canada)	SXT, TET	pAsa4b	(Daher <i>et al.</i> , 2011)
RS 1458	Rainbow trout	Ontario (Canada)	TET	pAsa4b	(Attéré <i>et al.</i> , 2015)
SHY13-2627	Brook trout	Quebec (Canada)	TET	pAsa4b	(Attéré <i>et al.</i> , 2015)
SHY13-3799	Brook trout	Quebec (Canada)	TET	pAsa4b	(Attéré <i>et al.</i> , 2015)
HER1107	INA	INA	TET	pAsa4b	(Daher <i>et al.</i> , 2011)
JF2267	Arctic char	Switzerland	CHL	pAsa4c	(Braun <i>et al.</i> , 2002)
JF3517	Turbot	Norway	CHL	pAsa4c	(Burr and Frey, 2007)
JF3518	Turbot	Norway	CHL	pAsa4c	(Burr and Frey, 2007)
JF2869	Arctic char	INA	CHL	pAsa4c	(Studer <i>et al.</i> , 2013)

Notes:

*INA: Information not available or not traceable

†SXT = sulfamethoxazole/trimethoprim, TET = tetracycline, CHL = chloramphenicol

Large insertions or deletions, as well as an inversion, have occurred between the pAsa4 variants, as shown in the sequences alignment (Fig. 2.1A). These events have mainly occurred between each plasmid's first conjugative loci, *Tn21s*, and *tetA(E)* flanking sequences. Insertion sequences have caused alignment gaps as well. Otherwise, the three pAsa4 variants displayed a high level of sequence identity (from 94 to 99%) for syntenic regions, with pAsa4b being more similar to pAsa4 than pAsa4c. Base substitution count by 1-kbp window between pAsa4b and pAsa4c showed that some regions are more prone to mutations (Fig. S1). In fact, more than 50 substitutions per kilobase occurred upstream from the first transfer genes (Fig. 2.1B, start of orange segment and Fig. S1 at 45-50 kbp) and in a long ORF only predicted in pAsa4c's first conjugative block (Fig. 2.1A, longest hypothetical protein in this region and Fig. S1 at 72 kbp) and in a single long ORF found in all plasmids (Fig. S1 at 112 kbp). On the other hand, almost no mismatches were found in

the 60-kbp region that comprised *Tn21* and the region downstream from it (Fig. S1, between 125 and 165 kbp).

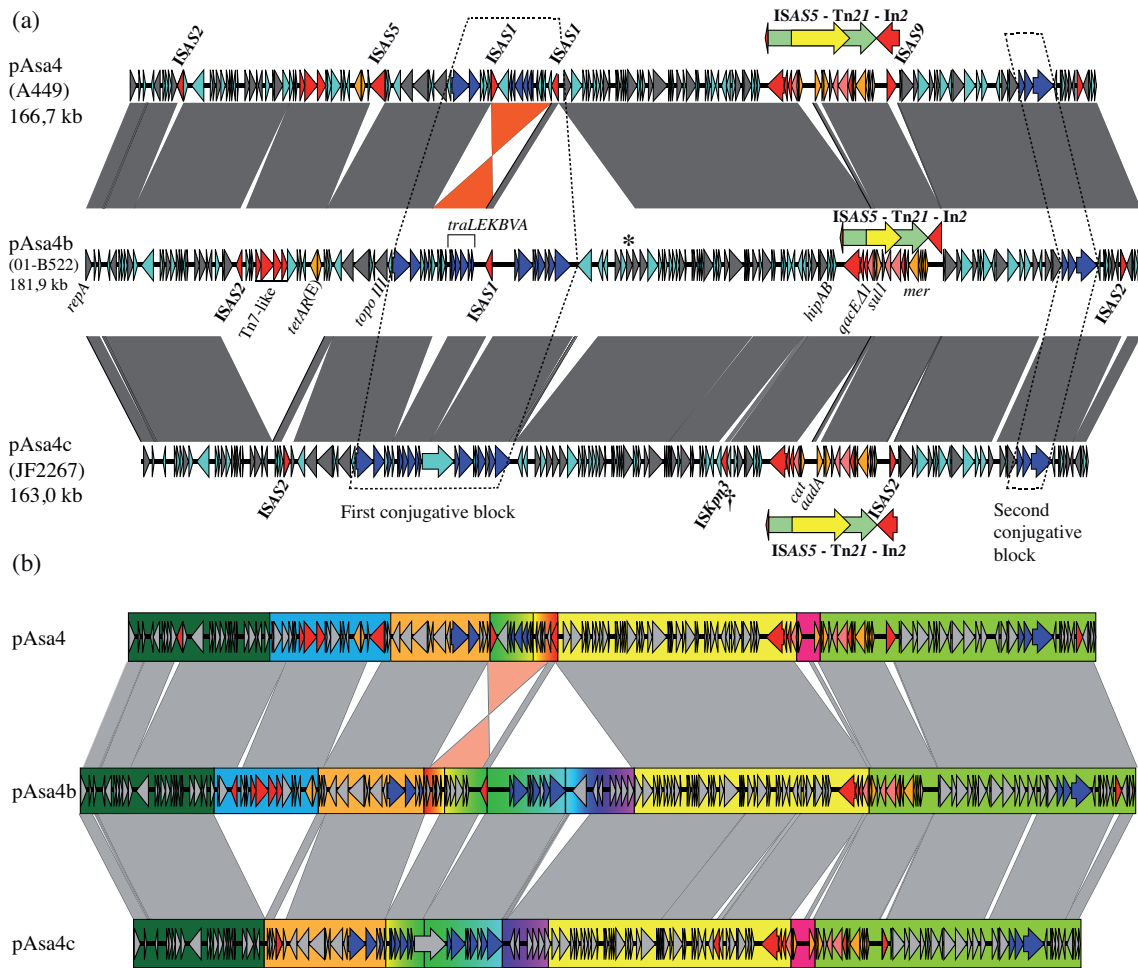


Figure 2.1 Nucleotide alignment of three plasmid variants: pAsa4, pAsa4b, and pAsa4c. (a) Plasmid alignments and ORF representations were done using EasyFig (Sullivan *et al.*, 2011). The dark grey bands denote regions of identity. Overall, the identity was more than 94%. The bands of non-contiguous repeat regions were removed for clarity. An inversion between pAsa4 and pAsa4b/c is marked in orange. Open reading frames (ORFs) are indicated by colored arrows that indicate their deduced function: Cyan: hypothetical protein; Dark grey: maintenance and replication; Blue: type IV secretion system-like conjugative system, contained in two conjugative blocks (dashed lines); Red: insertion sequence, transposition-associated genes; Orange: antimicrobial resistance. The following features have also been annotated: Asterisk: phage endonuclease, similar to pAsa4 pseudogene; Dagger: IS1595-family ISKpn3. A transposon/integron structure (Tn21/In2) that was common to all pAsa4 plasmids and that is integrated into an ISAS5 is indicated over each plasmid by nested red, green, and yellow arrows. Besides the transposon, but inside ISAS5, an ISAS9 and an ISAS2 insertion could be seen for pAsa4 and pAsa4c, respectively. (b) Segments of large insertion/deletion or recombination sequences are highlighted in color. Segments of particular significance are: Blue: an insertion/deletion in pAsa4 comprising tetracycline resistance genes *tetAR(E)*; Rainbow progression: multiple insertions/deletions and an inversion encompassing a conjugative gene region; Pink: two

events: an ISCR insertion comprising the chloramphenicol resistance gene, *cat* and an *aadA1* cassette.

INSERTION SEQUENCES

All pAsa4 variants carry insertions sequences (ISs) (Fig. 2.1A, named red arrows). pAsa4b and pAsa4c retained the same IS types that were described in pAsa4, namely ISAS1, ISAS2, ISAS5 and ISAS9 (Reith *et al.*, 2008) (see also GenBank accession number NC_009349.1). However, no IS shared the same location among all variants, except for the disrupted ISAS5 nesting the Tn21 copy. In pAsa4 and pAsa4c, two different ISs (ISAS9 and ISAS2, respectively) are inserted in this disrupted IS (Fig. 2.1A, downstream of the transposon).

A comparison of transposase sequences using Blast and of inverted repeats using the IS Finder database indicated that there was a member of the IS1595-family (Siguier *et al.*, 2006b) in pAsa4c (Fig. 2.1A, dagger). This IS, ISKpn3, has been originally identified in *Klebsiella pneumoniae* plasmid pRDDHA (Verdet *et al.*, 2006). To our knowledge, this was the first identification of this IS in *A. salmonicida*. Based on the Blast search results against the NCBI nr/nt and wgs databases, ISKpn3 is present in the *Aeromonas* genus, namely in *Aeromonas media* WS plasmid (accession number CP007567.1) and in *Aeromonas dhakensis* SSU strain (accession number JDWD00000000.1).

DETAILED PLASMID ARCHITECTURE

We compared all three pAsa4's architecture to assess their impact on the plasmid function, including antibiotic resistance (Fig. 2.1). To facilitate the analysis and the following genotyping, syntenic regions among the variants were grouped together as empirical segments (Fig. 2.1B, colored rectangles). We investigated the features in each segment as well as their boundaries to infer the causes of these large-scale rearrangements.

A first segment (Fig. 2.1B, blue rectangle) contained an ISAS2, Tn7-like transposition protein genes (ABCD), and tetracycline resistance genes (*tetAR*(E)). It was absent from pAsa4c compared to pAsa4 and pAsa4b, which explains why JF2267 was not resistant to tetracycline (Table 2.1). An imperfect 36-nucleotide inverted repeat flanking this segment

in pAsa4 and pAsa4b was not found in pAsa4c at the deletion site, suggesting that it could have been involved in the recombination-deletion process.

Tn21 and its In2 spanned over three segments based on this partition (Fig. 2.1B, yellow, pink and light green rectangles). Two contiguous variations in In2 are comprised in one segment (Fig. 2.1B, pink rectangle) that differentiated pAsa4b from pAsa4c and pAsa4, the latter two carrying identical integrons. pAsa4b In2 bears the integrase, a fused cassette *qacEΔ1 sulI*, a putative acetyltransferase and *tniABΔ3* (Fig. 2.2). In2 from pAsa4 and pAsa4c bears an additional *aadA* gene (synonym: *aadA1*) that codes for an aminoglycoside nucleotidyltransferase (ANT(3'')) (Ramirez and Tolmasky, 2010). Also, in pAsa4 and pAsa4c, the *cat* gene (synonym *catA1*, encoding a class A-1 chloramphenicol acetyltransferase) is not inserted as a cassette in In2. Instead, it is located between a hypothetical protein ORF and a partial ISCR that includes a partial transposase, a 3' ISCR and *oriIS*, but lacks the other components. This insertion is located between partial *intI* duplication. In pAsa4b, neither the insertion nor the duplication was found (Fig. 2.2).

The regions in the three plasmids harboring most of the conjugative transfer genes also displayed the most differences (Fig. 2.1B, rainbow colored section). An inversion of the *traLEKBVA* locus and the surrounding region seemed to have occurred in pAsa4. This inversion is flanked by two inverted ISASIs, which are in the appropriate position to have mediated the rearrangement. Only one of the two ISASI was found at this position in pAsa4b, while both are absent in pAsa4c. However, several genes of unknown function upstream from the transfer locus were deleted from this position in pAsa4c. The two new variants also have an insertion contiguous to the *traLEKBVA* locus that is not present in pAsa4 (Fig. 2.1B, rainbow colored section in pAsa4b, green to purple). This region is slightly longer in pAsa4b and harbors other transfer genes and coding sequences. Interestingly, the ultimate downstream gene in this segment is a putative phage-type endonuclease that shares identity with a pAsa4 pseudogene that, given the inversion and deletion in this region, is at the same location with respect to the other coding sequences in pAsa4 (Fig. 2.1A, asterisk). Conjugative transfer of all pAsa4 variants in *E. coli* was attempted. JF2267 (pAsa4c) and 01-B522 (pAsa4b) were able to produce transconjugants, but A449 (pAsa4) did not.

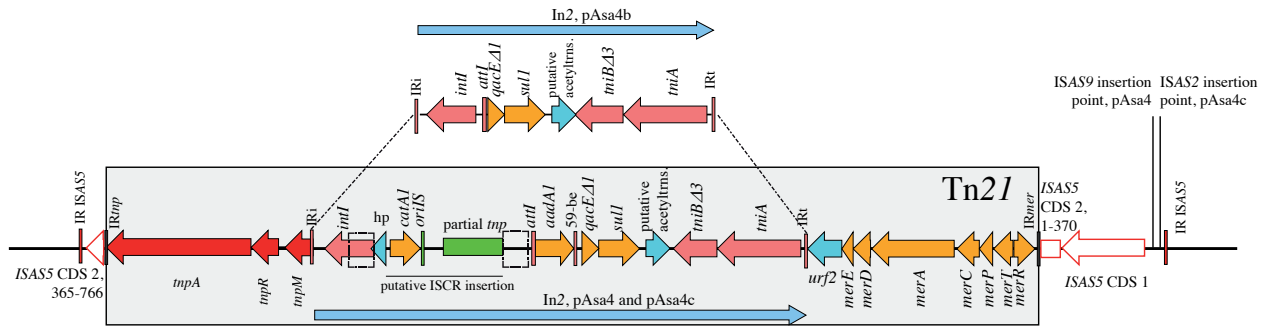


Figure 2.2 In2 comparison between pAsa4b and the other variants. Tn21 differences between the variants are all located in In2. Within Tn21 (shaded rectangle), features are colored using the following: Red: Tn21 transposition genes, Pink: In2 integrase, transposition features and insertion sites, Cyan: hypothetical proteins, Orange: Antibiotic and mercury resistance genes, Green: ISCR-related features. Two dotted rectangles represent a repeat region in pAsa4/pAsa4c In2, likely caused by *cat*-ISCR insertion. This IS is partial, lacking a *terIS* and part of its transposase, suggesting a complex insertion event. The figure also shows where Tn21 is inserted in ISAS5. Red outline: ISAS5 CDS, position called in nucleotide.

COMPARATIVE ANALYSIS OF THE pAsa4 ARCHITECTURE

In order to find similarities between empirically drawn regions shown in Fig. 2.1 and co-transferred block of genes, a tBlastn search of pAsa4b coding sequences (excluding IS transposases) against the NCBI non-redundant database was achieved to collect 516 uniquely identified sequences that were hit more than three times. By *k-means* clustering, those sequences were reproducibly clustered into four groups, one of which had two sub-groups (Fig. 2.3). Overall, identity percentage for the hits was between 20 and 80%, except for Group 2, where the identity was near 100%. Group 1 was divided into sub-groups a and b, which would always be differentiated by the *k-means* analysis. Group 1 (Fig. 2.3, red and orange) had hits for coding sequences scattered along pAsa4b against the A/C₂ family conserved backbone (Fricke *et al.*, 2009; Harmer and Hall, 2015). pRA1, a A/C₁ plasmid, also fell in this category (Harmer and Hall, 2015). The hits covered the majority of the plasmid, including the first and second conjugative block (Fig. 2.1A), but not the *tet* region (Fig. 2.1B, blue segment) nor the region directly downstream of ISAS5-

Tn21. Group 2 (Fig. 2.3, green) had hits targeted at Tn21/In2. However, its sequence identifiers were more disparate. Group 3 (Fig. 2.3, purple) had hits against integrative conjugative elements (ICE) and the *Vibrio* SXT-pathogenesis island for some of the coding sequences that provided hits in Group 1. Group 4 (Fig. 2.2, blue) had more heterogeneous identifiers and had hits for more specific coding sequences, including sequences for the Tn7-like transposition proteins and the *hipAB* toxin-antitoxin genes.

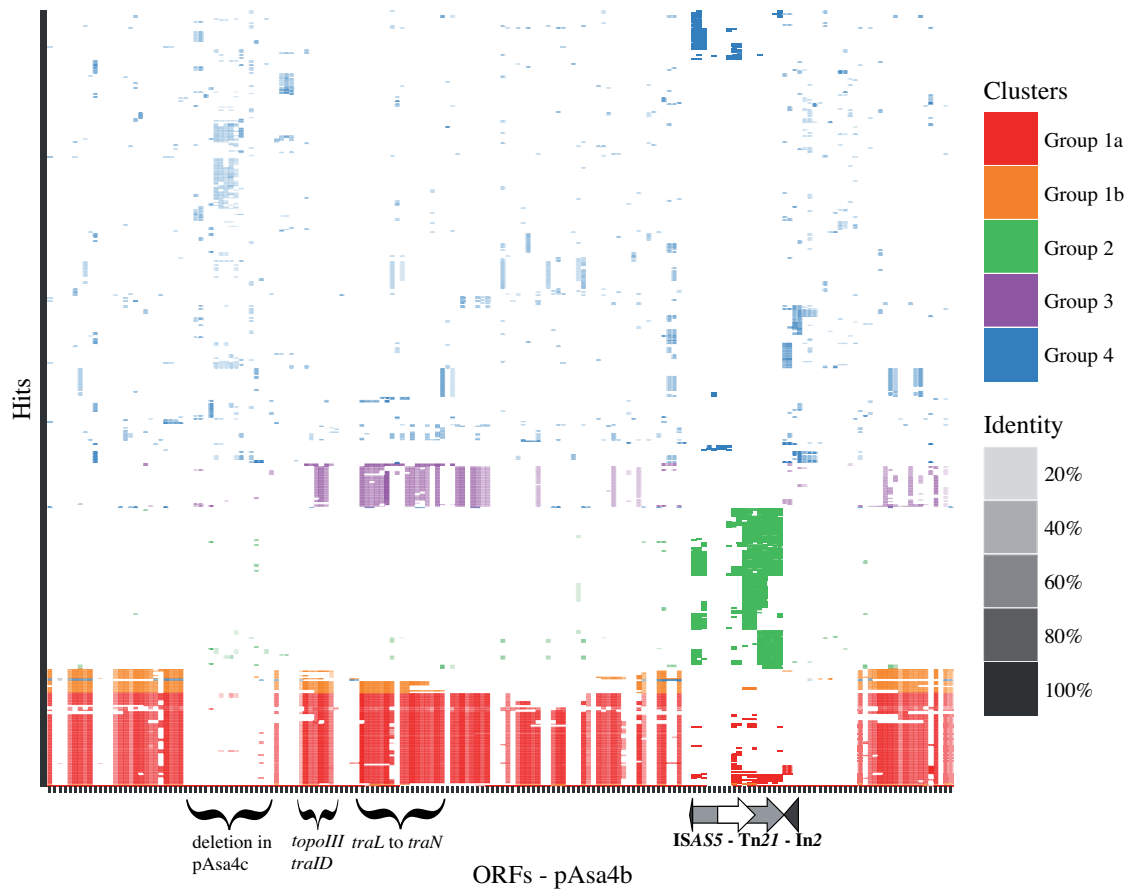


Figure 2.3 Clustering tBlastn results for pAsa4b. The shading denotes the maximum identity between the ORF query and the target. *k-means* clustered the molecules into four stable groups: Group 1 a and b is representative of incompatibility group IncA/C plasmids; Groups 2 and 4 do not encompass a specific type of sequence identifiers. However, Group 2 shares significant identity with Tn21 targets; Group 3 is representative of integrative and conjugative elements (ICEs). Some regions are less covered by tBlastn hits, such as Fig. 2.1B blue segment, and a region downstream from Tn21.

A final alignment was performed between pAsa4b and another *A. salmonicida* subsp. *salmonicida* plasmid, pSN254b (Fig. 2.4). pSN254b is a large IncA/C₂ plasmid that is also found in Canadian isolates (Vincent *et al.*, 2014b). The identity between continuous segments was between 59% and 81%, and the synteny between genes was well conserved, a feature that could not be analyzed by the heatmap. However, due to the dissimilarity of the backbone, pAsa4s cannot be considered as an IncA/C plasmid compared to what has previously been described (Fricke *et al.*, 2009). Again, the *tet*-containing segment (Fig. 2.4, blue rectangle) and a region directly downstream from Tn21 were not covered by the alignment (Fig. 2.4).

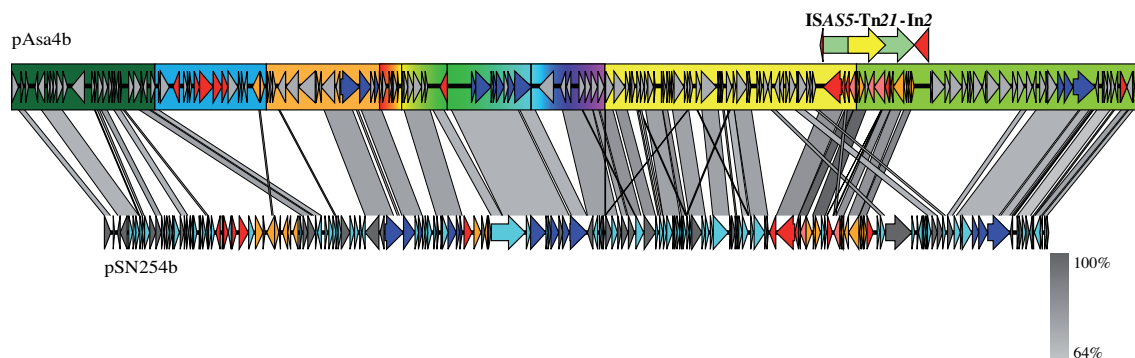


Figure 2.4 Nucleotide alignment between pAsa4b and pSN254b. The linear representations were aligned and visualized using EasyFig (Sullivan *et al.*, 2011). The color codes and segments are the same as in Fig 2.1.

VARIANT GENOTYPING AND ANTIBIOTIC RESISTANCES

Some insertions and deletions between pAsa4 variants changed their antibiotic resistance gene content. Consequently, A449, 01-B522 and JF2267 displayed different antibiotic resistances (Table 2.1). All resistance patterns but SXT, which is an antibiotic combination used in aquaculture (Morin, 2010), are directly explained by their respective pAsa4 architectures. JF2267 lack of tetracycline resistance is related to a segment deletion in pAsa4 that carries *tetA(E)* (Fig. 2.1B, blue segment). 01-B522 lack of chloramphenicol resistance is explained by its In2 structure, which does not bear the ISCR-*cat* insertion (Fig. 2). Since pAsa4 carries those two regions, A449 is resistant to both antibiotics.

Among a collection of 129 *A. salmonicida* subsp. *salmonicida* isolates from Canada and Europe, 11 (A449 included) were detected with pAsa4-positive PCR genotyping results

using a single pair of primers (Table 2.1, Table S2). These isolates had different resistance profiles for tetracycline, chloramphenicol and sulfamethoxazole-trimetoprim (SXT) resistance, and none were resistant to florfenicol, another aquaculture-relevant antibiotic whose resistance is conferred by pSN254b (Table 2.1, Table S1). 01-B522 was the only SXT-resistant strain, and since only the sulphonamide resistance is encoded on pAsa4, the 01-B522 genome has presumably another element to provide trimethoprim resistance. Otherwise, strains bearing pAsa4 variants could be clustered into three groups: tetracycline resistant, chloramphenicol resistant and resistant to both antibiotics.

We refined our genotyping of the pAsa4 variants by designing primers covering each segment junction (Fig. 2.1B, Table S2). All 11 pAsa4-positive isolates were associated with one variant version: pAsa4, pAsa4b, or pAsa4c (Table 2.1). All but one of the isolates (RS 1458) displayed the junction pattern (Fig. 2.1B) of their variant type and had a concordant antibiotic resistance profile (Table 2.1). The irregular strain RS 1458 had a pAsa4b pattern, except for *Tn21* (Fig. 2.1B, yellow to green junction). With exception of the two strains of unknown geographic origin, pAsa4b was found in Canada, while pAsa4 and pAsa4c were found in Europe (Table 2.1).

DISCUSSION

pAsa4 is a large antimicrobial resistance-encoding plasmid that was sequenced with *A. salmonicida* subsp. *salmonicida* reference genome (Reith *et al.*, 2008). In this study, we characterized two pAsa4 variants, pAsa4b and pAsa4c. The analysis of these variants highlighted the importance of mobile genetic elements in shaping the genomic landscape of this bacterium, in particular its antibiotic resistance and its ability to propagate by conjugation. Moreover, comparative genomics between variants and other plasmids as well as base substitution analysis were used to infer pAsa4 modular architecture.

The variable position of the ISs in the pAsa4 variants indicated that they were active and capable of transposition (Fig. 2.1A). This is an additional example of IS activity responsible for plasmid variations in *A. salmonicida* subsp. *salmonicida* (Attéré *et al.*, 2015; Najimi *et al.*, 2009; Trudel *et al.*, 2013; Vincent *et al.*, 2014b). Moreover, pAsa4c bears IS*Kpn3*, originally described on the *Klebsiella pneumoniae* plasmid pRDDHA. Based on its transposase annotation, up to twelve copies of this IS could be found in the *Aeromonas media* WS chromosome (accession number CP007567) and 1 to 3 copies could be found in *Aeromonas dhakensis* SSU (accession number JDWD00000000.1). However, *A. media* WS may be prone to “infection” by ISs since it bears 324 transposase-associated annotations (7.4% of the coding sequence), compared to an average of 38.42 transposase genes per bacterial genomes (Aziz *et al.*, 2010). IS and transposase proportion varies within *Aeromonas* sp. (Chai *et al.*, 2012; Vincent *et al.*, 2016b). All these ISs can disrupt genes and functions by subsequent transposition, or can be targeted by the recombination machinery to produce larger structural variations, and thus bring a genetic modification potential. In pAsa4, two IS*ASIs* were likely the mediators of the large inversion (Fig. 2.1A). This is a reason why pAsa4 is not conjugative compared to pAsa4b and pAsa4c (Fig. 2.1B). Similar IS-dependent recombinations have been observed in pAsa5 variants in *A. salmonicida* subsp. *salmonicida* and have been reproduced *in vitro* by growing the bacteria under stressful conditions (Daher *et al.*, 2011; Emond-Rheault *et al.*, 2015b; Tanaka *et al.*, 2012; Vincent *et al.*, 2016b).

In2 is both an active mobile element in pAsa4 and a site for complex IS integration. The cassette integration system is potentially active given the presence of *aadA* cassette in pAsa4 and pAsa4c. This region also contains the *cat* gene, encoding a chloramphenicol acetyltransferase, which is not integrated as a cassette, but rather as an ISCR-like insertion sequence (Fig. 2.2). This provides another example of phenotypic diversity driven by ISs. The ISCR elements, which are known to transpose *cis* resistance genes between class 1 integrons in non-standard transpositions, may also create integron fragment duplications during those events (Toleman *et al.*, 2006a). In pAsa4, the duplication of the integrase between the ISCR transposase fragment and the cassette structure is an indication of a complex transposition (Fig. 2.2).

Empirical segments representing insertion/deletion and inversion were created to facilitate plasmid visualization and genotyping. PCR across the segments junctions paired with an antimicrobial disk assay assigned plasmid variant types to strains that bore pAsa4-like plasmids (Table 2.1). Some inserted/deleted blocks between variants could be related to original genes series or metabolic functions that came together by horizontal gene transfer, and provide insights for this process in *A. salmonicida*. Thus, base substitution counts (Fig. S1) and tBlastn comparisons (Fig. 2.3) were used to conduct further our analysis. The heatmap and the resulting clustering revealed previously observed similarities between pAsa4 and the IncA/C incompatibility group, although pAsa4s fail the requirements to belong in either IncA/C₁ or IncA/C₂ (Fricke *et al.*, 2009; Harmer and Hall, 2015). However, two pAsa4b regions were poorly covered by hits in this analysis. One was the region deleted in pAsa4c, corresponding to blue segment in Fig. 1B. Since very few hits were found against these ORFs, their origin, although not IncA/C-related, could not be inferred. However, this module could bring specific accessory functions to pAsa4-bearing strains. The other region poorly covered by hits was a region immediately downstream from Tn21. Interestingly, the base substitution analysis also showed that this region was not prone to mutation (Fig. S1). This region contains many genes that code for hypothetical proteins, but their implication in pAsa4 maintenance or functions is unknown. However, given their presence in all variants and the region's low substitution rate, it could contain genes essential for pAsa4 maintenance and is another region unique to pAsa4, compared to the

A/C group. The pAsa4b to pSN254b alignment further highlight those two unique regions, as well as the synteny between the common ones (Fig. 2.4).

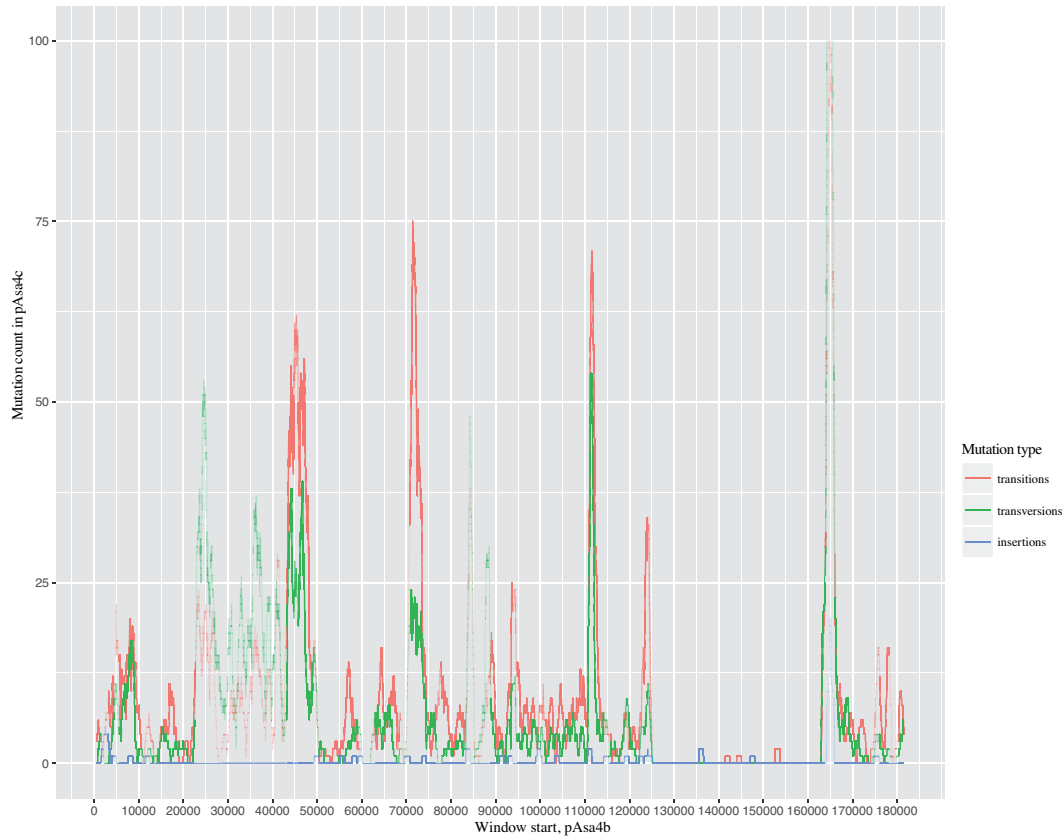
CONCLUSION AND PERSPECTIVES

Our results showed that pAsa4 variant architecture impacts resistance antibiotic genes, and identified active ISs and integration hotspots that could promote novel resistance combinations. Because of its ubiquitous nature, *A. salmonicida* subsp. *salmonicida* interacts with many other waterborne microbes. Therefore, it may serve as a reservoir for disseminating new plasmid-based combinations of antimicrobial resistance. Even if pAsa4 was not as prevalent as pSN254b in geographic regions included in the present study, it should be regarded as a potential threat to the propagation and shuffling of antibiotic resistance due to its modular and recombinant structure. The transmission of pAsa4 should thus be monitored, especially given the propagation of *A. salmonicida* subsp. *salmonicida* infections in fish farms.

ACKNOWLEDGMENTS

We thank Mathilde Goldschmitt for technical support. We also thank the Laboratoire de bactériologie clinique, Faculté de médecine vétérinaire, Université de Montréal (Montreal, QC, Canada), the Félix d'Hérelle Reference Center (Département de biochimie, de microbiologie et de bio-informatique, Université Laval, Quebec City, QC, Canada) as well as M. Reith (National Research Council laboratories, Halifax, Nova Scotia), J. Frey (University of Bern), and R.M.W. Stevenson (University of Guelph) for the *A. salmonicida* subsp. *salmonicida* isolates.

RELEVANT SUPPORTING INFORMATION²



Additional File 2.1 Substitution rate between pAsa4c (JF2267) and pAsa4b (01-B522) using 1-kbp windows. A global alignment of pAsa4b and pAsa4c was performed using *stretcher* (available as a part of EMBOSS 6.6.0.0) (Rice *et al.*, 2000), and a custom R script (R Core Team, 2015) was used to visualize the number of substitutions by 1000-bp sliding windows (Zeileis and Grothendieck, 2005). Transparent lines represent deletions in pAsa4c compared to pAsa4b and should not be considered.

² L'ensemble des fichiers supplémentaires associés à ce manuscrit peuvent être consultés à l'adresse suivante : <http://doi.org/10.7717/peerj.2595>

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- Att  r  , S.A., Vincent, A.T., Trudel, M.V., Chanut, R., Charette, S.J., 2015. Diversity and Homogeneity among Small Plasmids of *Aeromonas salmonicida* subsp. *salmonicida* Linked with Geographical Origin. *Frontiers in microbiology* 6, 1274.
- Aziz, R.K., Breitbart, M., Edwards, R.A., 2010. Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res* 38, 4207-4217.
- Boyd, J.M., Dacanay, A., Knickle, L.C., Touhami, A., Brown, L.L., Jericho, M.H., Johnson, S.C., Reith, M., 2008. Contribution of type IV pili to the virulence of *Aeromonas salmonicida* subsp. *salmonicida* in Atlantic salmon (*Salmo salar* L.). *Infect. Immun.* 76, 1445-1455.
- Chai, B., Wang, H., Chen, X., 2012. Draft genome sequence of high-melanin-yielding *Aeromonas media* strain WS. *J. Bacteriol.* 194, 6693-6694.
- Daher, R.K., Fillion, G., Tan, S.G., Dallaire-Dufresne, S., Paquet, V.E., Charette, S.J., 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.
- Dallaire-Dufresne, S., Tanaka, K.H., Trudel, M.V., Lafaille, A., Charette, S.J., 2014. Virulence, genomic features, and plasticity of *Aeromonas salmonicida* subsp. *salmonicida*, the causative agent of fish furunculosis. *Vet. Microbiol.* 169, 1-7.
- Derome, N., Boutin, S., Llewellyn, M., Gauthier, J. 2016. Opportunistic pathogens of fish, In: Hurst, C.J. (Ed.) *The Rasputin effect: when commensals and symbionts become parasitic*. Springer, New York, NY, 649.
- Emond-Rheault, J.G., Vincent, A.T., Trudel, M.V., Frey, J., Frenette, M., Charette, S.J., 2015. *AsaGEI2b*: a new variant of a genomic island identified in the *Aeromonas salmonicida* subsp. *salmonicida* JF3224 strain isolated from a wild fish in Switzerland. *FEMS Microbiol. Lett.* 362.
- Fernandes, F., da Fonseca, P.G., Russo, L.M., Oliveira, A.L., Freitas, A.T., 2011. Efficient alignment of pyrosequencing reads for re-sequencing applications. *BMC bioinformatics* 12, 163.
- Fricke, W.F., Welch, T.J., McDermott, P.F., Mammel, M.K., LeClerc, J.E., White, D.G., Cebula, T.A., Ravel, J., 2009. Comparative genomics of the IncA/C multidrug resistance plasmid family. *J. Bacteriol.* 191, 4750-4757.
- Galardini, M., Biondi, E.G., Bazzicalupo, M., Mengoni, A., 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* 6, 11.
- Garcia-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Gotz, S., Tarazona, S., Dopazo, J., Meyer, T.F., Conesa, A., 2012. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28, 2678-2679.

- Gordon, D., Green, P., 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29, 2936-2937.
- Gustafson, C.E., Chu, S.J., Trust, T.J., 1994. Mutagenesis of the paracrystalline surface protein array of *Aeromonas salmonicida* by endogenous insertion elements. *J. Mol. Biol.* 237, 452-463.
- Hanninen, M.L., Hirvela-Koski, V., 1997. Molecular and phenotypic methods for the characterization of atypical *Aeromonas salmonicida*. *Vet. Microbiol.* 56, 147-158.
- Harmer, C.J., Hall, R.M., 2015. The A to Z of A/C plasmids. *Plasmid* 80, 63-82.
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 100-108.
- Liebert, C.A., Hall, R.M., Summers, A.O., 1999. Transposon Tn21, flagship of the floating genome. *Microbiol. Mol. Biol. Rev.* 63, 507-522.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.
- McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M.A., Baylay, A.J., Bhullar, K., Canova, M.J., De Pascale, G., Ejim, L., Kalan, L., King, A.M., Koteva, K., Morar, M., Mulvey, M.R., O'Brien, J.S., Pawlowski, A.C., Piddock, L.J., Spanogiannopoulos, P., Sutherland, A.D., Tang, I., Taylor, P.L., Thaker, M., Wang, W., Yan, M., Yu, T., Wright, G.D., 2013. The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* 57, 3348-3357.
- McIntosh, D., Cunningham, M., Ji, B., Fekete, F.A., Parry, E.M., Clark, S.E., Zalinger, Z.B., Gilg, I.C., Danner, G.R., Johnson, K.A., Beattie, M., Ritchie, R., 2008. Transferable, multiple antibiotic and mercury resistance in Atlantic Canadian isolates of *Aeromonas salmonicida* subsp. *salmonicida* is associated with carriage of an IncA/C plasmid similar to the *Salmonella enterica* plasmid pSN254. *J. Antimicrob. Chemother.* 61, 1221-1228.
- Morin, R., 2010. L'utilisation des antibiotiques pour combattre la furonculose chez l'omble de fontaine génère de l'antibiorésistance chez *Aeromonas salmonicida*. *L'Aquicole - Le Bulletin de l'Association des aquaculteurs du Québec* 15.
- Najimi, M., Balado, M., Lemos, M.L., Osorio, C.R., 2009. Genetic characterization of pAsa6, a new plasmid from *Aeromonas salmonicida* subsp. *salmonicida* that encodes a type III effector protein AopH homolog. *Plasmid* 61, 176-181.
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U. S. A.* 85, 2444-2448.

- Piotrowska, M., Popowska, M., 2015. Insight into the mobilome of *Aeromonas* strains. *Frontiers in microbiology* 6, 494.
- R Core Team 2015. R: A Language and Environment for Statistical Computing (Vienna, Austria, R Foundation for Statistical Computing).
- Ramirez, M.S., Tolmasky, M.E., 2010. Aminoglycoside modifying enzymes. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy* 13, 151-171.
- Reith, M.E., Singh, R.K., Curtis, B., Boyd, J.M., Bouevitch, A., Kimball, J., Munholland, J., Murphy, C., Sarty, D., Williams, J., Nash, J.H., Johnson, S.C., Brown, L.L., 2008. The genome of *Aeromonas salmonicida* subsp. *salmonicida* A449: insights into the evolution of a fish pathogen. *BMC genomics* 9, 427.
- Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276-277.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944-945.
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., Chandler, M., 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34, D32-36.
- Sørum, H., L'Abée-Lund, T., Solberg, A., Wold, A., 2003. Integron-containing IncU R plasmids pRAS1 and pAr-32 from the fish pathogen *Aeromonas salmonicida*. *Antimicrob. Agents Chemother.* 47, 1285-1290.
- Studer, N., Frey, J., Vanden Bergh, P., 2013. Clustering subspecies of *Aeromonas salmonicida* using IS630 typing. *BMC microbiology* 13, 36.
- Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009-1010.
- Tanaka, K.H., Dallaire-Dufresne, S., Daher, R.K., Frenette, M., Charette, S.J., 2012. An insertion sequence-dependent plasmid rearrangement in *Aeromonas salmonicida* causes the loss of the type three secretion system. *PLoS one* 7, e33725.
- Tanaka, K.H., Frenette, M., Charette, S.J., 2013. IS-mediated loss of virulence by *Aeromonas salmonicida*: A tangible piece of an evolutionary puzzle. *Mobile genetic elements* 3, e23498.
- Toleman, M.A., Bennett, P.M., Walsh, T.R., 2006. Common regions e.g. *orf513* and antibiotic resistance: IS91-like elements evolving complex class 1 integrons. *J. Antimicrob. Chemother.* 58, 1-6.
- Trudel, M.V., Tanaka, K.H., Filion, G., Daher, R.K., Frenette, M., Charette, S.J., 2013. Insertion sequence *AS5* (ISAS5) is involved in the genomic plasticity of *Aeromonas salmonicida*. *Mobile genetic elements* 3, e25640.
- Verdet, C., Benzerara, Y., Gautier, V., Adam, O., Ould-Hocine, Z., Arlet, G., 2006. Emergence of DHA-1-producing *Klebsiella spp.* in the Parisian region: genetic organization of the *ampC* and *ampR* genes originating from *Morganella morganii*. *Antimicrob. Agents Chemother.* 50, 607-617.

- Vincent, A.T., Boyle, B., Derome, N., Charette, S.J., 2014a. Improvement in the DNA sequencing of genomes bearing long repeated elements. *J. Microbiol. Methods* 107, 186-188.
- Vincent, A.T., Emond-Rheault, J.G., Barbeau, X., Attéré, S.A., Frenette, M., Lagüe, P., Charette, S.J., 2016a. Antibiotic resistance due to an unusual ColE1-type replicon plasmid in *Aeromonas salmonicida*. *Microbiology* 162, 942-953.
- Vincent, A.T., Trudel, M.V., Freschi, L., Nagar, V., Gagné-Thivierge, C., Levesque, R.C., Charette, S.J., 2016b. Increasing genomic diversity and evidence of constrained lifestyle evolution due to insertion sequences in *Aeromonas salmonicida*. *BMC genomics* 17, 44.
- Vincent, A.T., Trudel, M.V., Paquet, V.E., Boyle, B., Tanaka, K.H., Dallaire-Dufresne, S., Daher, R.K., Frenette, M., Derome, N., Charette, S.J., 2014b. Detection of variants of the pRAS3, pAB5S9, and pSN254 plasmids in *Aeromonas salmonicida* subsp. *salmonicida*: multidrug resistance, interspecies exchanges, and plasmid reshaping. *Antimicrob. Agents Chemother.* 58, 7367-7374.
- Wickham, H., 2009. *ggplot2: elegant graphics for data analysis*. Springer New York.
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T.L., 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics* 13, 134.
- Zeileis, A., Grothendieck, G., 2005. zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software* 14, 1-27.

Chapitre 3 UN NOUVEAU PATRON DE RECOMBINAISON DES IS DU PLASMIDE PAsa5 CHEZ *AEROMONAS SALMONICIDA* SSP. *SALMONICIDA* 01-B526.

RÉSUMÉ

Le plasmide pAsa5 d'*Aeromonas salmonicida* ssp. *salmonicida* porte plusieurs gènes du système de sécrétion de type trois (SSTT), un facteur de virulence essentiel pour cette bactérie, en une seule région. Celle-ci peut être perdue lorsque la bactérie est cultivée à 25°C. Dans une étude précédente, certains événements de perte du SSTT avaient été expliqués par une recombinaison homologue entre séquences d'insertion (IS). Dans cette étude, le séquençage du génome de 01-B526 a permis de découvrir un nouveau plasmide similaire à pAsa5 sur 40 kpb de sa séquence, pAsa9, ainsi qu'une nouvelle IS sur pAsa5. Cette IS permet un nouveau patron de recombinaison homologue présent chez des souches dont la perte du SSTT n'avait pas été expliquée. Le plasmide pAsa9 interférait avec la détection de ce type de délétion. Ces résultats permettent de regrouper les pertes du SSTT étudiées sous un seul mécanisme : la recombinaison homologue entre copies identiques d'IS.

**PLASMID COMPOSITION IN *AEROMONAS*
SALMONICIDA SUBSP. *SALMONICIDA* 01-B526
UNRAVELS UNSUSPECTED TYPE THREE SECRETION
SYSTEM LOSS PATTERNS**

Katherine H. Tanaka^{a,b,c}, Antony T. Vincent^{a,b,c}, Jean-Guillaume Emond-Rheault^{a,b,c},
Marcin Adamczuk^d, Michel Frenette^{b,e} and Steve J. Charette^{a,b,c#}

^aInstitut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, 1030, avenue de la Médecine, Quebec City, Quebec, Canada, G1V 0A6

^bDépartement de biochimie, de microbiologie et de bio-informatique, Faculté des sciences et de génie, Université Laval, 1045, avenue de la Médecine, Quebec City, Quebec, Canada, G1V 0A6

^cCentre de recherche de l'Institut universitaire de cardiologie et de pneumologie de Québec, 2725, chemin Sainte-Foy, Quebec City, Quebec, Canada, G1V 4G5

^dDepartment of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Ilii Miecznikowa 1, 02-096 Warsaw, Poland.

^eGroupe de Recherche en Écologie Buccale (GREB), Faculté de médecine dentaire, Université Laval, 2420, rue de la Terrasse, Quebec City, Quebec, Canada, G1V 0A6

ABSTRACT

Background

Aeromonas salmonicida subsp. *salmonicida* is a ubiquitous psychrophilic waterborne bacterium and a fish pathogen. The numerous mobile elements, especially insertion sequences (IS), in its genome promote rearrangements that impact its phenotype. One of the main virulence factors of this bacterium, its type three secretion system (TTSS), is affected by these rearrangements. In *Aeromonas salmonicida* subsp. *salmonicida* most of the TTSS genes are encoded in a single locus on a large plasmid called pAsa5, and may be lost when the bacterium is cultivated at a higher temperature (25°C), producing non-virulent mutants. In a previous study, pAsa5-rearranged strains that lacked the TTSS locus on pAsa5 were produced using parental strains, including 01-B526. Some of the generated deletions were explained by homologous recombination between ISs found on pAsa5, whereas the others remained unresolved. To investigate those rearrangements, short- and long-read high-throughput sequencing technologies were used on the *A. salmonicida* subsp. *salmonicida* 01-B526 whole genome.

Results

Whole genome sequencing of the 01-B526 strain revealed that its pAsa5 has an additional IS copy, an ISAS5, compared to the reference strain (A449) sequence, which allowed for a previously unknown rearrangement to occur. It also appeared that 01-B526 bears a second large plasmid, named pAsa9, which shares 40 kbp of highly similar sequences with pAsa5. Following these discoveries, previously unexplained deletions were elucidated by genotyping. Furthermore, in one of the derived strains a fusion of pAsa5 and pAsa9, involving the newly discovered ISAS5 copy, was observed.

Conclusion

The loss of TTSS and hence virulence is explained by one consistent mechanism: IS-driven homologous recombination. The similarities between pAsa9 and pAsa5 also provide another example of genetic diversity driven by ISs.

Keywords: Insertion sequence, Plasmid, Homologous recombination, Virulence, *Aeromonas salmonicida*

BACKGROUND

The bacterium *Aeromonas salmonicida* subsp. *salmonicida* is the causative agent of furunculosis, a disease that affects salmonids worldwide. It has a significant economic impact on the fish farming industry (Derome *et al.*, 2016). Vaccination and antibiotherapy are the available treatments for furunculosis. However, resistant *A. salmonicida* subsp. *salmonicida* strains have emerged (McIntosh *et al.*, 2008; Piotrowska et Popowska, 2015; Sørum *et al.*, 2003; Trudel *et al.*, 2016; Vincent *et al.*, 2016a; Vincent *et al.*, 2014b) and vaccination is not always effective (Dallaire-Dufresne *et al.*, 2014b). Alternative treatment options against this pathogen would thus be beneficial (Dallaire-Dufresne *et al.*, 2014b; Imbeault *et al.*, 2006; Tanaka *et al.*, 2013) but developing new treatments requires a better understanding of its underlying mechanisms, such as pathogenicity (Frey and Origgi, 2016; Piotrowska and Popowska, 2015).

This ubiquitous waterborne bacterium shows evidence of lateral gene transfer. *A. salmonicida* subsp. *salmonicida* strain genomes have been shown to bear a wide array of mobile elements (Piotrowska and Popowska, 2015; Reith *et al.*, 2008), genomic islands (Emond-Rheault *et al.*, 2015a; Emond-Rheault *et al.*, 2015b; Long *et al.*, 2016), and transferable plasmids (McIntosh *et al.*, 2008; Sørum *et al.*, 2003; Tanaka *et al.*, 2016; Vincent *et al.*, 2014b). The close relationship between the mobile elements from this species and those of *Salmonella enterica* have raised concerns about its ability to act as a reservoir of antibiotic resistance genes, which could transfer to human pathogens (Trudel *et al.*, 2016; Vincent *et al.*, 2014b). *Aeromonas salmonicida* subsp. *salmonicida* is also known for its genomic variability. Different kinds of recombination have formed plasmid and genomic island variants with altered gene contents that result in phenotypic changes (Attéré *et al.*, 2015; Long *et al.*, 2016; Tanaka *et al.*, 2012; Tanaka *et al.*, 2016; Trudel *et al.*, 2013; Vincent *et al.*, 2016a; Vincent *et al.*, 2014b). This bacterium, beyond mediating gene transfer, could thus provide a new combination of those genes.

In many cases, insertion sequences (ISs) and transposons are responsible for these variations. They were found to participate in plasmid rearrangements, virulence factor inactivation, and the exchange of antibiotic resistance genes (Attéré *et al.*, 2015; Gustafson *et al.*, 1994; Tanaka *et al.*, 2012; Tanaka *et al.*, 2016; Trudel *et al.*, 2013). It was even

suggested that the abundance of ISs (88 copies divided into 10 different types) in the reference genome A449 could maintain the bacterium in a psychrophilic lifestyle to preserve its genomic integrity (Vincent *et al.*, 2016b). Unfortunately, transposable elements also hinder genome sequencing and assembly, which are used to study structural variations (Beatson and Walker, 2014; Charette *et al.*, 2012; Vincent *et al.*, 2014a; Vincent *et al.*, 2015).

One structural variation mediated by ISs that has been studied in *A. salmonicida* subsp. *salmonicida* is the loss of its type three secretion system (TTSS). This needle-like apparatus injects proteins named effectors into the host cells (Buttner, 2012). An interesting point for furunculosis management is the fact that the TTSS is essential for the pathogenicity of this bacterium (Burr *et al.*, 2005; Frey and Origgi, 2016). In *A. salmonicida* subsp. *salmonicida*, the TTSS structural components and many effectors are encoded on a plasmid called pAsa5 (also known as pASvirA) (Reith *et al.*, 2008; Stuber *et al.*, 2003), which has been found in many strains (Burr *et al.*, 2005; Reith *et al.*, 2008; Stuber *et al.*, 2003). This 155-kilobase-pair (kbp) replicon harbours all TTSS structural genes in a single locus, four TTSS effector genes (*aopH*, *ati2*, *aopN*, and *aopO*), conjugative transfer genes, many uncharacterized open reading frames, and ISs (Frey and Origgi, 2016; Reith *et al.*, 2008). Interestingly, pAsa5 is a thermolabile plasmid that is known to lose segments, including the TTSS locus, when the cells are exposed to 25°C and above (Daher *et al.*, 2011; Stuber *et al.*, 2003). The resulting mutants are non-virulent in fish and in model hosts, making the TTSS loss mechanism an interesting target for avirulence-producing treatment development (Burr *et al.*, 2005; Burr *et al.*, 2002; Burr *et al.*, 2003b; Dacanay *et al.*, 2010; Dacanay *et al.*, 2006; Daher *et al.*, 2011; Froquet *et al.*, 2007).

In a previous study, a collection of TTSS-negative mutants were produced by prolonged cultivation at 25°C of three *A. salmonicida* subsp. *salmonicida* parental strains: A449, 01-B526, and 01-B516 (Daher *et al.*, 2011). These plasmid-rearranged strains were shown to have two different types of truncated pAsa5.

Plasmids displaying a type 1 deletion profile had lost their TTSS locus, which includes the essential structural proteins, effectors *ati2* and *aopN*, and their respective chaperones (Fig. 3.1a). Plasmids with a type 2 deletion profile lost their TTSS locus as well as 40 kbp

upstream of it. Both deletion types led to complete virulence loss against an alternative host (Daher *et al.*, 2011).

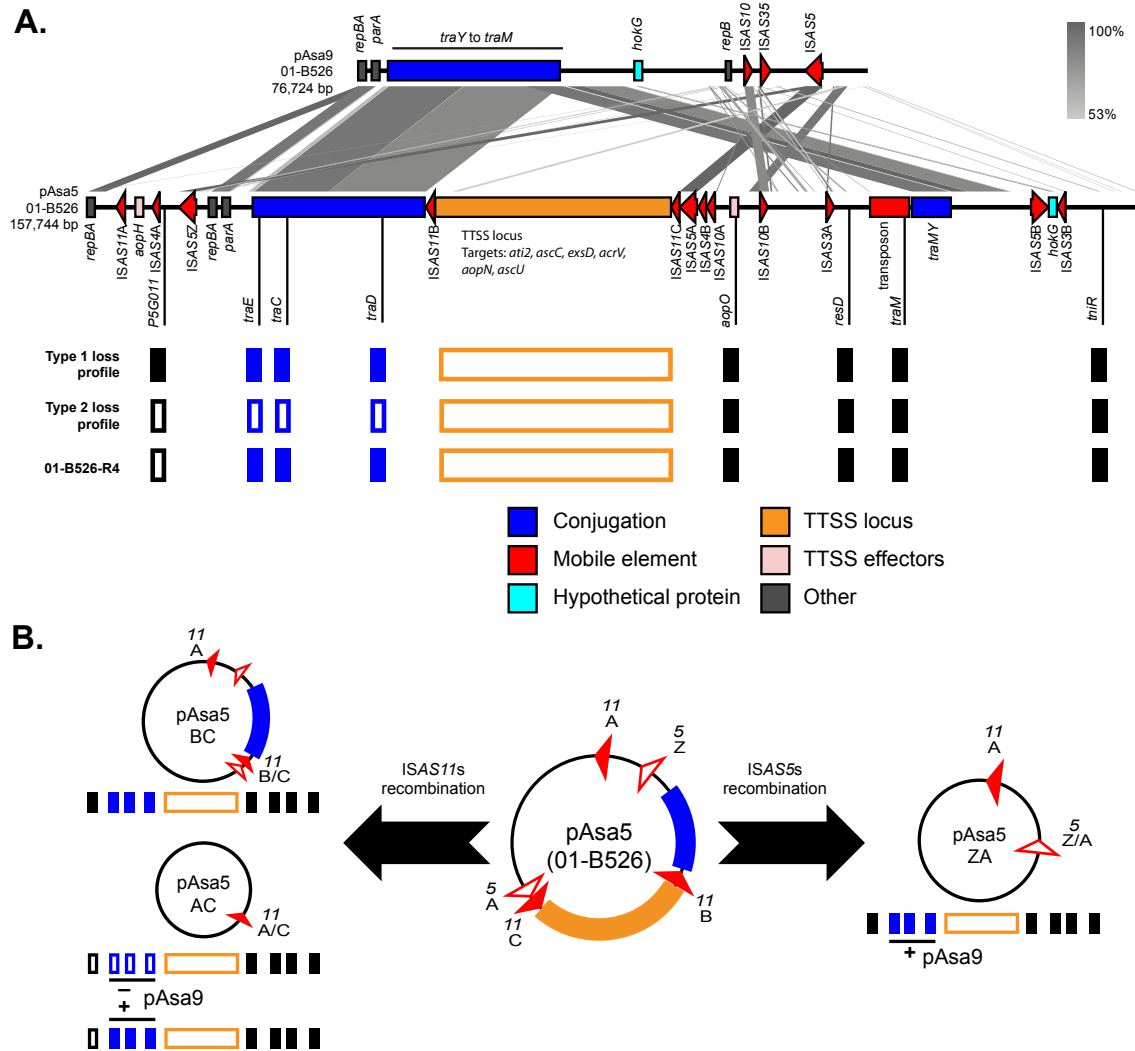


Figure 3.1 Alignment of 01-B526 pAsa5 and pAsa9, and proposed recombination patterns for 01-B526 pAsa5. **A.** The plasmid sequences were aligned and visualized using EasyFig (Sullivan *et al.*, 2011). Featured ORFs and regions were coloured using the following code: Red: transposases or transposase fragments, Blue: conjugative transfer protein, Black: ORFs with detailed functions or domains, Cyan: hypothetical protein, Orange: TTSS structural genes; Light pink: TTS effectors and chaperones. This alignment shows that shared region between pAsa5 and pAsa9 are also syntenic. Moreover, pAsa9 has its conjugative transfer genes in a single locus, compared to two loci with a *traI* repetition for pAsa5. Below the alignment, PCR targets locations used by Daher *et al.* to assess pAsa5 integrity at stressful growth temperature, and the resulting loss profile, are shown (Daher *et al.*, 2011; Tanaka *et al.*, 2012). **B.** IS locations on 01-B526 pAsa5 allow for two types of recombination patterns: ISAS11-targetted, which was described previously (Tanaka *et al.*,

2012), and ISAS5-targetted. ISAS11B-ISAS11C recombination leads to pAsa5 BC, which has lost its TTSS locus (top left circle). ISAS11A-ISAS11C recombination leads to pAsa5 AC, which has lost its TTSS and conjugative transfer loci (bottom left circle). However, in 01-B526-R4, the conjugative locus is detected by PCR since a similar copy is carried by pAsa9 (bottom left rectangles). ISAS5Z-ISAS5A recombination leads to pAsa5 ZA, which has also lost its TTSS and conjugative transfer loci (right circle). Again, in this case the pAsa9 conjugative locus is detected by PCR, effectively masking the deletion.

Further work revealed that all type 2 and some type 1 deletions could be explained by the same mechanism based on IS-mediated homologous recombination driven by the copies of ISAS11 found on pAsa5 (Tanaka *et al.*, 2012). However, many strains that show a type 1 deletion profile could not be explained by the known IS recombination (Tanaka *et al.*, 2012). It was thus hypothesized that their plasmids had undergone incomplete or more complex recombinations.

As mentioned before, a better understanding of the genomic content and rearrangements in *A. salmonicida* subsp. *salmonicida* would be beneficial for virulence-targeted treatment development and for assessing the contributions of this species to gene flow by horizontal transfers. Different TTSS-loss mechanisms seem to occur in strains 01-B526 and A449. The genome of A449 has been already sequenced and fully assembled (Reith *et al.*, 2008). In this study, we combined Illumina and single molecule real-time (SMRT) sequencing for the 01-B526 genome to elucidate its plasmid composition. This helped us to identify another TTSS-loss scenario involved in unexplained pAsa5-rearranged strains.

RESULTS

CHARACTERIZATION OF PAsa9

A new plasmid sequence was obtained from the 01-B526 strain by combining short and long reads in a hybrid *de novo* assembly. The new plasmid, named pAsa9, is 76,724 base pairs with a G+C content of 52.76%, and 90 annotated open reading frames (ORF). It can be divided into sections according to its gene content and nucleotide similarities (Fig. 3.2). The first segment has replication-associated *repBA* genes (Fig. 3.2, 0 to 5 kbp) and has high nucleotide identity with two pAsa5 regions at 98.13% and 93.12%, respectively (Fig. 3.1A). The second region has many conjugation genes (*tra*, blue rectangles) and again

was over 83% similar to pAsa5 (Fig. 3.1A). The third region, directly following *tra* genes, had many ORFs coding for hypothetical proteins (cyan, Fig. 3.2) or with less described functions (black, Fig. 3.2), but was still 84% similar (Fig. 3.1A) to pAsa5. Finally, pAsa9, the last region, was mostly filled with fragmented and complete transposase genes (Fig. 3.2, red rectangles).

Since ISs are a prominent component of *A. salmonicida* subsp. *salmonicida* genome (Vincent *et al.*, 2016b), we investigated those of pAsa9. ISs and IS-related sequences were found solely in the last quarter of the sequence (Fig. 3.2, 57 kbp to end). More partial than complete ISs could be identified, suggesting ancestral transposition events (Siguier *et al.*, 2014). Of the three complete ISs found on pAsa9, two, ISAS10 and ISAS5, have already been identified and described in *A. salmonicida* subsp. *salmonicida* notably on the reference pAsa5 (Reith *et al.*, 2008). ISAS35, an IS that had not been described previously, was identified based on its transposase gene and inverted repeats homology. It belongs to the ISBst12 group, of which members have been found in *Klebsiella pneumoniae* (ISKpn15, best score 248) and *A. salmonicida* subsp. *salmonicida* (ISAs21, best score 167) based on the ISFinder database (Siguier *et al.*, 2006b). A blast analysis against deposited *Aeromonas* genomes (taxid:642, October 2016) showed ISAS35 is only found in three *A. salmonicida* subsp. *salmonicida* genomes: 01-B526 (accession number AGVO01000000), J231 (LSGY01000000), and J227 (LSGX01000000).

COMPARISON BETWEEN PAsA9 AND PAsA5

pAsa9 was sequenced from the strain 01-B526, which is already known to contain other plasmids: pAsa1, pAsa2, pAsa3, pAsa11 and pAsa5 (Charette *et al.*, 2012). Interestingly, pAsa9 shared long (up to 24 kbp), similar regions with pAsa5 (Fig. 3.1A). Shared similarities made it impossible to detect, assemble, and characterize pAsa9 in the previous studies focusing on 01-B526 genomics that were based on short-read sequencing only (Charette *et al.*, 2012; Vincent *et al.*, 2014a). High-throughput sequencing was sufficient to cover its specific regions since some contigs of *A. salmonicida* subsp. *salmonicida* J231 and J227 strains found in GenBank matched against pAsa9. However, the repeated regions between pAsa5 and pAsa9 did not allow for a complete assembly with previous 01-B526

datasets (Charette *et al.*, 2012; Vincent *et al.*, 2014a). SMRT and Illumina sequencing resulted in the assembly of both pAsa5 and pAsa9 as single contigs, which allowed us to distinguish them for the first time.

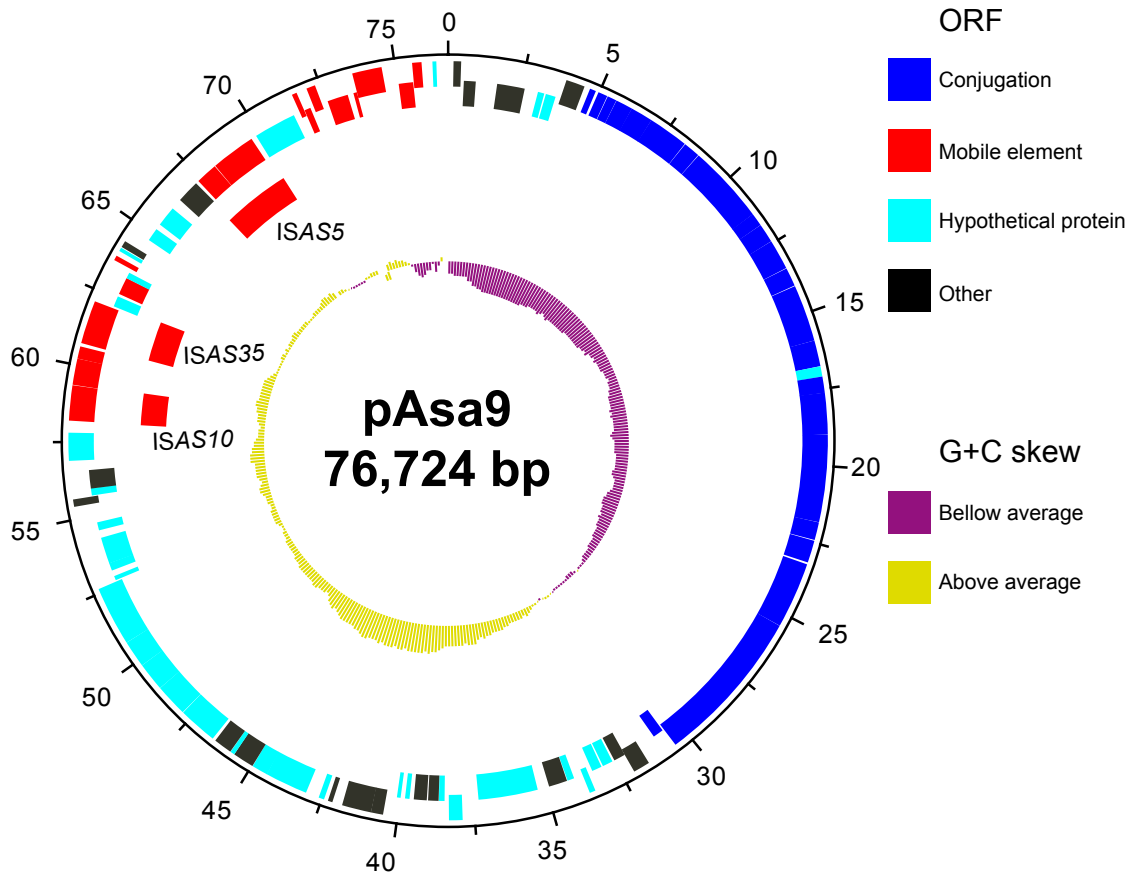


Figure 3.2 pAsa9 map. The pAsa9 sequence and features were visualized with Artemis and DNAPlotter (Carver *et al.*, 2009; Rutherford *et al.*, 2000). Going inwards from the scale (expressed in kbp), the first two circles mark ORFs (forward, then reverse strand) using the colour code introduced in Fig. 3.1. The third ring marks insertion sequences. The fourth marks GC skew, using the following colours: purple: below average; yellow: above average.

Apart from ISs, both plasmids share regions that encode conjugative transfer and plasmid replication genes. pAsa5 and pAsa9 have similar and contiguous set of *tra* genes (between 83% and 91%) and similar *repBA* genes (between 93% and 98%). Some regions, such as the *rep* surroundings and the *traI* gene, are repeated thrice between the plasmids: once on pAsa9 and twice on pAsa5. The distinct regions on pAsa9 compared to pAsa5 contained mostly hypothetical protein-coding sequences. On the other hand, pAsa5 unique regions

carried hypothetical protein-coding sequences, but also ISs and, more importantly, TTSS-related genes. pAsa9 ISs and IS-associated genes are concentrated in the same region, while pAsa5 ISs are scattered throughout its length, and some are proximal to TTSS genes. For instance, the virulence-essential TTSS locus is framed by two IS*AS11*, as described earlier (Tanaka *et al.*, 2012). This region also disrupts the G+C skew in pAsa5 compared to pAsa9, suggesting a recent acquisition (Additional File 1). A 10 kbp region comprising another TTSS effector gene, *aopO*, is also framed by ISs, two IS*AS10s*. The third TTSS effector gene, *aopH*, is framed by non-identical ISs (IS*AS11* and IS*AS4*).

pAsa5 also shares similarities with pAsa6 (18.5 kbp), another *A. salmonicida* subsp. *salmonicida* plasmid that also has the *aopH* gene (Najimi *et al.*, 2009). Sequence alignment showed that the genetic content and architecture of pAsa6 are closer to pAsa5 than pAsa9, since it carries the *aopH* section (Additional File 2). Furthermore, its similarities are concentrated around pAsa5's first *rep* region and insertion sequences. The significant similarities between pAsa5 and pAsa9 suggest that pAsa6 has probably derived from pAsa5 through successive deletions instead of pAsa5 evolved from the fusion of pAsa6 with other plasmid(s).

GENOTYPING UNEXPLAINED TYPE 1 LOSS PROFILE AND PAsa9 OCCURRENCE

01-B526 pAsa5 was aligned against the reference plasmid from the strain A449. The two plasmids had a high nucleotide identity (more than 99.5%) across their length. The only structural variation is an additional IS*AS5* inserted between the first IS*AS4* and the second *rep* region on 01-B526 pAsa5 (Fig. 3.1A). Since this IS copy is upstream from IS*AS5A* from the previous annotation (Tanaka *et al.*, 2012), it was designated as IS*AS5Z* for the purpose of this paper. IS*AS5Z* (upstream from *tra* genes) and IS*AS5A* (downstream from the TTSS locus) are in the same orientation, hence they form a valid pair for IS-mediated loss by recombination.

This deletion would encompass both conjugation and TTSS loci, and should result in another loss profile for pAsa5 genotyping, a feature that had not been observed previously (Fig. 3.1B) (Daher *et al.*, 2011; Tanaka *et al.*, 2012). However, due to its similarity with pAsa9, primers hybridizing in the conjugation locus (*traE*, *traC*, and *traD*) can also

hybridize on pAsa9, leading to false positives for this region (Fig. 3.1). ISAS5Z-ISAS5A recombination thus leads to a type 1 deletion profile in pAsa9-bearing strains when the previous genotyping primer set is used (Fig. 3.1B). The unexplained type 1 deletion profiles found in pAsa5-rearranged strains could thus be explained by ISAS5Z-A recombination, but only if the strains bore pAsa9 and an ISAS5Z.

All parental and derived strains were directly tested with primers flanking ISAS5Z 3' and ISAS5A 5' to assess the new ZA recombination. All parental strains and daughter strains for which a recombination pattern had already been found were negative to that new primer pair.

In the unexplained derived strains, all but one (01-B526-R19) were positive for the new ZA amplification, which gave a 2888 bp amplification product (Table 3.1). Sanger sequencing for that amplicon showed that it encloses an ISAS5, ISAS5Z flanking 3', and ISAS5A flanking 5', confirming the recombination. Those strains had thus experienced a homologous recombination between ISAS5Z and ISAS5A, leading to the loss of both TTSS and conjugation loci. All derived strains were then tested for the presence of pAsa9 using a set of genotyping primers that targeted its unique region (Additional File 3). A449-derived strains were negative for pAsa9, which was expected, since A449 does not carry this plasmid. All the other pAsa5-rearranged strains were positive for pAsa9. Since all newly assigned ZA pAsa5-bearing strains also carried pAsa9, it explained why the conjugation locus loss had not been detected in those strains previously (Tanaka *et al.*, 2012).

Table 3.1 *A. salmonicida* subsp. *salmonicida* isolates used in this study

Isolates	Source	Previous A-C rearrangement genotyping	Previous B-C rearrangement genotyping	ISAS5 Z-A rearrangement genotyping	pAsa9 genotyping
Parental strains					
A449	(Dacanay <i>et al.</i> , 2006)	-	-	-	-
01-B526	(Dautremepuits <i>et al.</i> , 2006)	-	-	-	+
01-B516	(Daher <i>et al.</i> , 2011)	-	-	-	-
Displaying Type 1 Loss profile					
A449-R1	(Daher <i>et al.</i> , 2011)	-	+	-	-
A449-R3	(Daher <i>et al.</i> , 2011)	-	+	-	-
A449-R4	(Daher <i>et al.</i> , 2011)	-	+	-	-
01-B526-R2	(Daher <i>et al.</i> , 2011)	-	+	-	+
01-B526-R3	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R5	(Daher <i>et al.</i> , 2011)	-	+	-	+
01-B526-R6	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R7	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R8	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R9	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R10	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R11	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R12	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R13	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R14	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R15	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R16	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R17	(Daher <i>et al.</i> , 2011)	-	-	+	+

01-B526-R18	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-R19	(Daher <i>et al.</i> , 2011)	-	-	-	+
01-B526-3 (mistaken for 01-B516-3)	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-11 (mistaken for 01-B516-11)	(Daher <i>et al.</i> , 2011)	-	-	+	+
01-B526-30 (mistaken for 01-B516-30)	(Daher <i>et al.</i> , 2011)	-	+	-	+
Displaying Type 2 loss profile					
A449-R2	(Daher <i>et al.</i> , 2011)	+	-	-	-
A449-R5	(Daher <i>et al.</i> , 2011)	+	-	-	-
01-B526-R4	(Daher <i>et al.</i> , 2011)	+	-	-	+

01-B516 AND ITS ASSIGNED DAUGHTER STRAINS

The parental strains, A449, 01-B526 and 01-B516 were then tested for the presence of pAsa9. It was found that neither A449 (as expected) nor 01-B516 had the plasmid. However, all three 01-B516-derived strains (01-B516-3, 01-B516-11, and 01-B516-30) had pAsa9, creating a disparity between parental and offspring strains. It was thus hypothesized that during mutant production, some 01-B526-derived strains had been mistakenly labelled as 01-B516.

To verify that 01-B516 did not have any form of pAsa9, plasmid profiles of all parental strains, one BC pAsa5-bearing strain (01-B526-R2), and one possibly mislabelled derived strain (01-B516-11), were visualized in agarose gel electrophoresis (Fig. 3.3A). As expected, A449 showed one band corresponding to pAsa5 and pAsa4 (155 kbp and 167 kbp, respectively). 01-B526 showed two bands, one for pAsa5 and one for pAsa9. 01-B516 showed a single large band whose weight matched pAsa5 (Fig. 3.3A). 01-B526-R2 showed the pAsa9 band as well as a BC-rearranged pAsa5 band. However, 01-B516-11, which was found to bear a ZA-rearranged pAsa5 according to our previous genotyping, gave two bands: one for the ZA pAsa5 and one whose size matched pAsa9. This result suggested this strain has a 01-B526-derived background and was thus mislabelled.

01-B516 genomic DNA was then sequenced using short-read, high-throughput sequencing (Illumina). When the resulting contigs were aligned on pAsa5 and pAsa9, it showed 01-B516 did not have sequences unique to pAsa9. Moreover, 01-B516 did not contain *ISAS5Z*, making an *ISAS5Z-ISAS5A* recombination impossible in this strain, nor in any daughter strain.

Derived strains 01-B516-3, 01-B516-11 and 01-B516-30 were also genotyped for *AsaGEI1a* (accession number KJ626178) presence and were found positive for that genomic island, whereas 01-B516 carries *AsaGEI2a* (Emond-Rheault *et al.*, 2015a). *AsaGEIs* (for *Aeromonas salmonicida* GENomic Island) are found in some *A. salmonicida* subsp. *salmonicida* genomes. Each carrying strain only has one copy of one type of *AsaGEIs*. They have variable insertion sites (indicated by the number in their name) and gene content (indicated by their letter). They contain some phage-related genes and many

ORFs coding for hypothetical protein. Consequently, their role remains to be determined (Emond-Rheault *et al.*, 2015a; Emond-Rheault *et al.*, 2015b; Long *et al.*, 2016).

Moreover, *AsaGEI* types are linked with their geographic origin and strains bearing *AsaGEI1a* are found in North America (Emond-Rheault *et al.*, 2015a). Since all 01-B516 ‘derived’ strains genotyped positive for pAsa9 and *AsaGEI1a*, and that two of them have a ZA-rearranged pAsa5, we concluded these three strains were mislabelled when they were obtained, and are instead 01-B526-derived strains.

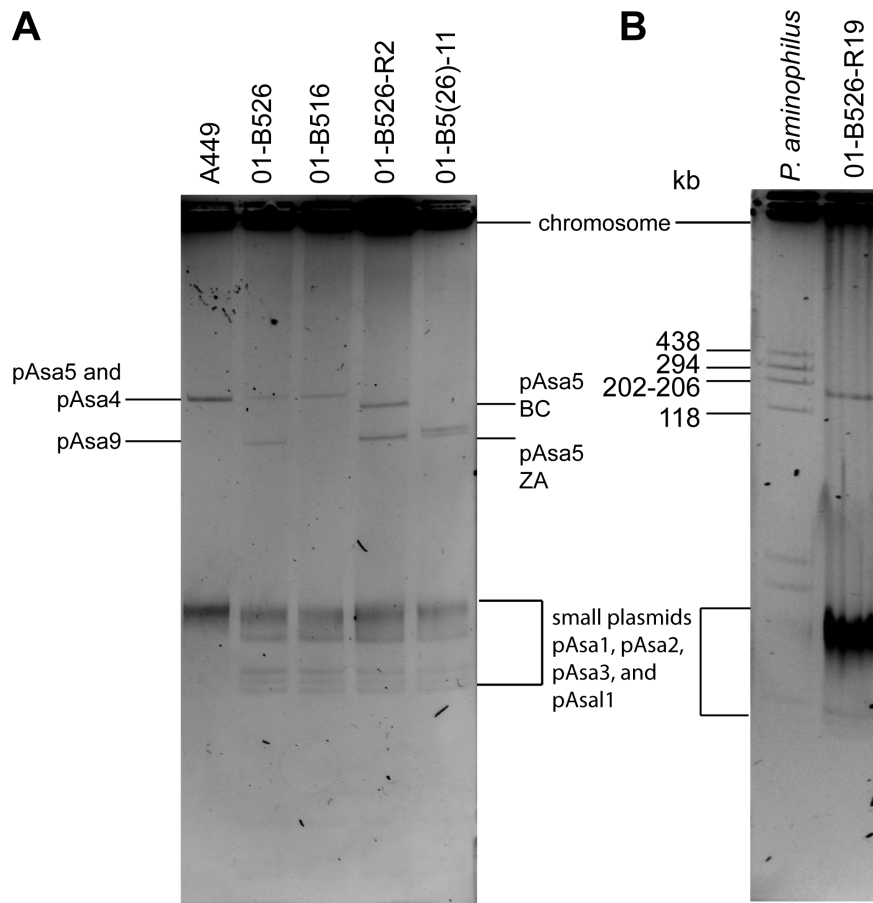


Figure 3.3 Large plasmids from parental and derived strains. A protocol adapted from Wheatcroft (Wheatcroft *et al.*, 1990) was used to visualize large plasmids by agarose gel electrophoresis. **A.** Some daughter strain profiles were compared to the parental strains. A449 displays a single band, corresponding to plasmids pAsa5 and pAsa4; 01-B526 displays pAsa5 and pAsa9 bands; and 01-B516 displays a single band, pAsa5. The 01-B526-R2 pAsa5 band is lower than 01-B526 one, since it underwent an ISAS1/B-C recombination. 01-B516-11 displays a ZA-rearranged pAsa5 band as well as a pAsa9 one, whereas its parent does not have the pAsa9 plasmid. It was confirmed that all 01-B516-

derived strains were mislabelled in the previous studies, and are in fact 01-B526-derived. **B.** 01-B526-R19, the only derived strain whose deletion profile on pAsa5 cannot be explained by IS recombination, was also put on a gel beside *Paracoccus aminophilus* JCM 7686 (Wheatcroft *et al.*, 1990). The 01-B526-R19 single band has a molecular weight corresponding to its ZA-rearranged pAsa5 and pAsa9 fused together.

PASA9 PREVALENCE

The prevalence of pAsa9 in a collection of *A. salmonicida* subsp. *salmonicida* isolates was then determined by genotyping with primers covering its specific region (Additional File 3). Among the 154 genotyped strains, 35 had the plasmid and two generated amplicons using some primers only, suggesting another plasmid (Additional File 4). Furthermore, in our strain set, pAsa9 was always found with another element, the genomic island *AsaGEIIa*.

PLASMID FUSION IN 01-B526-R19

One type 1 deletion profile strain, 01-B526-R19, could not be explained by any IS recombination, even with the ZA recombination discovery. 01-B526-R19 genomic DNA was thus sequenced by Illumina technology to detect potential new genomic content. Contig alignment to 01-B526 pAsa5 and pAsa9 showed it lacked the *tra* and TTSS regions on pAsa5, suggesting an IS*AS5Z*-A recombination, a fact which had been suspected but not proven by PCR. It also had a complete pAsa9, as shown by contig alignment.

The 01-B526-R19 recombination was deemed to be complex, so its plasmids were visualized on gel (Fig. 3.3B). 01-B526-R19 displayed a single large plasmid band at approximately 150 kbp, higher than a ZA-rearranged pAsa5 band, and no pAsa9 band was detected, even if its presence had been assessed by PCR and sequencing (Fig. 3.3B).

However, since the weight of the 01-B526-R19 band matched the sum of ZA-rearranged pAsa5 and pAsa9, and since both plasmids had one IS*AS5* copy, we hypothesized that both large plasmids had merged together by IS*AS5* recombination (Fig. 3.4A). Primers flanking IS*AS5ZA* in pAsa5 and pAsa9 IS*AS5* were used in a long amplification PCR (Additional File 3). Both sides of the merge were successfully amplified and confirmed by Sanger sequencing, indicating 01-B526-R19's rearranged pAsa5 and its pAsa9 had indeed

recombined into a single plasmid, and shedding light on the last unexplained strain recombination (Fig. 3.4B).

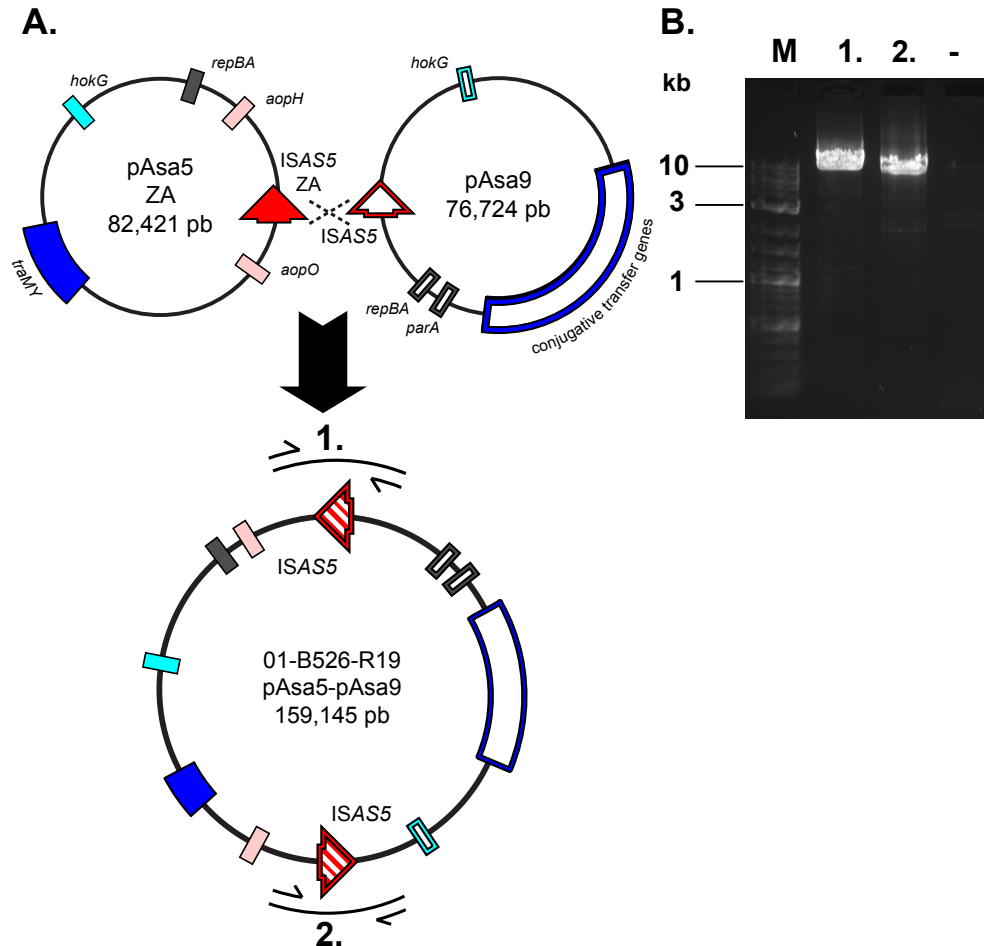


Figure 3.4 Fusion of the 01-B526-R19 pAsa5 plasmid with pAsa9 through their ISAS5.
A. The pAsa5 and pAsa9 fusion involved two ISAS5 (full and hollow red arrows): ISAS5ZA, which is the product of a previous recombination, and pAsa9 sole ISAS5. Features from both plasmids that are found in Fig. 3.1 are included in the diagram for comprehension purposes, and ISs other than the involved ISAS5s were omitted for clarity. The figure is not to scale. The pAsa5 ZA-pAsa9 plasmid presence was visualized by gel electrophoresis (see Fig. 3.3B) and the ISAS5 recombination and surrounding sequences were amplified by PCR (1 and 2). **B.** Two long PCRs show the junctions between pAsa5 ZA and pAsa9. The numbers match the amplicon shown in the previous diagram. Product 1 is 9,348 pb and Product 2 is 8,429 pb. A negative control (-) targeting a TTSS region (Additional File 3) was also included.

DISCUSSION

In this study, we used long-read sequencing (PacBio SMRT technology) to uncover a new *A. salmonicida* subsp. *salmonicida* plasmid, pAsa9. It would have been impossible to assemble pAsa9 with short-read sequencing information alone, due to its similarity with another well-studied, virulence-essential plasmid called pAsa5 (Burr *et al.*, 2002; Daher *et al.*, 2011; Reith *et al.*, 2008). Even with mate-pair sequencing, which allowed us, with a certain configuration, to assemble the *A. salmonicida* subsp. *salmonicida* chromosome in one scaffold (Vincent *et al.*, 2014a), the pAsa9 assembly was complicated by its 24 kbp-long repeated sequence with pAsa5.

SMRT sequencing can be used to reduce contig numbers and manual gap filling for most microbial genomes, especially those that have many long repeats; this feature applies to *A. salmonicida* subsp. *salmonicida* (Koren *et al.*, 2013). Single-read sequencing was previously used to identify and assemble plasmids that shared a common backbone in the same strain (Mehdizadeh Gohari *et al.*, 2016). PacBio or other long-read sequencing technologies should be used in future studies to facilitate the discovery of new plasmids in *A. salmonicida* subsp. *salmonicida*, especially if they are presumed to share repeats with the rest of the genome. Also, some species in the *Aeromonas* genus, such as *Aeromonas media* (Chai *et al.*, 2012), have complex genomes and would thus benefit from that technology (Vincent *et al.*, 2016b).

The newly discovered plasmid, pAsa9, was found in *A. salmonicida* subsp. *salmonicida* strains that also bear the genomic island *AsaGEIIa*. Thus, its presence is limited to North America so far based on the strains analysed. Interestingly, pAsa9 is found concurrently with pAsa5, even though their replication genes (*repBA*) are similar. Although the discovered plasmid, pAsa9, shares highly similar sequences with pAsa5, its unique content and architecture show that it is not a mere copy.

In fact, we propose that pAsa9 provides information on pAsa5 evolution and gene acquisition. The two plasmids are similar and have syntenic sequences, which suggests they share a common ancestor whose backbone should be closer to pAsa9 than pAsa5. Apart from its last quarter (Fig. 3.2), which encloses all mobile element-related sequences, pAsa9 is free from IS or sequence duplication. In contrast, pAsa5 has sequence duplications in

core genes (replication and transfer) (Garcillán-Barcia *et al.*, 2011), and many ISs. Furthermore, some IS localizations follow TTSS genes.

We thus suggest that ISs were key elements in TTSS-related gene acquisition in pAsa5, which arose on a plasmid backbone close to pAsa9. The discovery of pAsa9 also places pAsa6, another pAsa5-related plasmid, as a pAsa5-like descendant rather than an ancestor, since its structure is closer to the latter (Additional file 2). Together, those results further prove that insertion sequences have a role in *A. salmonicida* subsp. *salmonicida* genetic diversity, as proposed earlier (Vincent *et al.*, 2016b).

PacBio sequencing also provided new architectural details that allowed us to revisit previous experimental results on pAsa5 recombination and the associated loss of TTSS and virulence. With this information, all pAsa5 deletions that were generated in the study of Daher *et al.* can be regrouped under a consistent and simple mechanism, which is IS-mediated homologous recombination (Daher *et al.*, 2011). These results also allowed us to explain previous observations of pAsa5-rearranged strains that had lost their TTSS locus (Tanaka *et al.*, 2012).

For instance, while attempting to produce TTSS-negative mutants, strain 01-B526 was more prone to producing plasmid-rearranged strains than strain A449 (Daher *et al.*, 2011). This can be explained by the fact that in 01-B526 pAsa5, another IS (ISAS5Z) can be a target for recombination.

With 2.6 kbp, ISAS5 offers a large span of homologous sequences for recombination to occur, which could explain why ISAS5 rearrangements were observed more frequently in 01-B526-derived strains than ISAS11 rearrangements. ISAS5 was already known as an active insertion sequence in *A. salmonicida* subsp. *salmonicida*. It was found to have transposed at multiple locations on pAsa11 (Attéré *et al.*, 2015; Trudel *et al.*, 2013). The results thus suggest that ISAS5 is both an insertion sequence capable of transposition, and a target for host-mediated recombination. 01-B526 also gave only one ISAS11A-C recombinant with an atypical genotyping profile, 01-B526-R4, which can now be explained by the presence of pAsa9 and the false positive *tra* targets it creates (Fig. 3.1). Thus, PacBio sequencing both provided an explanation for shaded areas left by the previous study, and simplified some cases.

One exception is 01-B526-R19, which has a fusion between its ZA-rearranged pAsa5 and its pAsa9. This merged plasmid is a rarity since it resulted from two sequential recombinations, one between ISAS5Z and 5A, and one with ISAS5ZA and pAsa9 ISAS5. Such double recombinations between ISAS1Is were tested in 2012 with no success (Tanaka *et al.*, 2012). This may be due to the fact, as mentioned earlier, ISAS1I is half the size of ISAS5.

The reassignment of derived strains 01-B516-3, 01-B516-11 and 01-B516-30 to 01-B526 leaves 01-B516 without pAsa5-rearranged descendants that would prove the strain is capable of IS-targeted recombination. Difficulties in producing 01-B516 pAsa5-rearranged strains were observed in the original experiment (Daher *et al.*, 2011). This problem could be caused either by the structure of 01-B516 pAsa5 or by the lack of suitable recombinase. The Illumina sequencing showed that ISAS1IB and ISAS1IC are in the correct position relative to each other to mediate the deletion. BC recombination and TTSS loss should thus have been observed. Consequently, the lack of rearrangement in 01-B516 may be due to a missing or inactive recombinase.

Furunculosis management needs new alternative treatments to antibiotherapy and vaccination. Avirulent treatment suppresses essential pathogenicity mechanisms without killing the bacteria, and could be an option against *A. salmonicida* subsp. *salmonicida*. However, full understanding of the TTSS loss and the underlying IS-mediated recombination is a prerequisite to finding new ways to attenuate furunculosis-causing strains. Now, factors promoting these events can be investigated in the hopes of finding a specific, aquaculture-compatible condition that would trigger recombination of pAsa5.

CONCLUSIONS

Our results showed that PacBio sequencing could provide crucial information for plasmid architecture studies by allowing for proper separation of long repeated regions. In this case, it led to the discovery of a new plasmid, pAsa9, which in turn allowed refinement and completion of the recombination model described for pAsa5 in 2012 (Tanaka *et al.*, 2012).

Regrouping all TTSS-loss scenarios under a simple rule was an important step in investigating the mechanism for avirulence-producing treatments. The relationship between the pAsa9 and pAsa5 architectures demonstrates the importance of ISs in producing biologically-relevant genomic diversity in *A. salmonicida* subsp. *salmonicida*.

METHODS

BACTERIAL ISOLATES AND GROWTH CONDITIONS

The *A. salmonicida* subsp. *salmonicida* strains that were used in this study and are listed in Table 3.1 (strains listed in Tanaka *et al.*) and Additional File 4 (all other strains, used for genotyping). In Table 3.1, pAsa5-rearranged strains (that display type 1 or type 2 loss profiles) were obtained by culturing virulent parental strains (A449, 01-B516 and 01-B526) for two weeks at 25°C (Daher *et al.*, 2011). All *A. salmonicida* strains were grown on furunculosis agar (10 g of tryptone, 5 g of yeast extract, 1 g of L-tyrosine, 2.5 g of NaCl, and 15 g of agar per litre of distilled water) for 2 or 3 days at 18°C (Hänninen and Hirvela-Koski, 1997). *Paracoccus aminophilus* JCM 7686 was used as a DNA size marker in plasmid visualization (Dziewit *et al.*, 2014). It was grown on tryptic soy agar (EMD Millipore, Ontario, Canada) for 3 days at 30°C.

DNA EXTRACTION AND SEQUENCING

Total genomic DNA was extracted for strains 01-B526 and 01-B516, and derived strain 01-B526-R19 (Table 3.1) using DNeasy Blood and Tissue kits (Qiagen, Canada). Libraries were prepared with a KAPA Hyper Prep Kit and sequencing was done at the Plateforme d'Analyse Génomique of the Institut de Biologie Intégrative et des Systèmes (IBIS, Université Laval, Quebec City, Canada) using the Illumina MiSeq system. Strain 01-B526 total genomic DNA was also extracted by phenol/chloroform by following the protocol *Extracting DNA Using Phenol-Chloroform* provided by Pacific Biosciences (<http://www.pacb.com>) and was sequenced at the Génome Québec Innovation Centre (McGill University, Montreal, Canada) using a Pacific BioScience RS II system. A hybrid *de novo* assembly was used for 01-B526 reads using Spades version 3.6.0 (kmer lengths of

21, 33, 55, 77, 99, 127) (Bankevich *et al.*, 2012). 01-B516 and 01-B526-R19 reads were assembled using A5-miseq version 20150522 (Coil *et al.*, 2015).

SEQUENCE ANALYSIS

The 01-B526 assembly allowed us to obtain the large plasmids as complete closed sequences. pAsa5 and pAsa9 were annotated using the webserver RAST (February 2016) (Overbeek *et al.*, 2014). Manual verification was done using Artemis version 16.0.0 using pAsa5 reference sequence (NC_009350) for comparison and blastp (non-redundant protein sequences database, May 2016) (Altschul *et al.*, 1990; Carver *et al.*, 2012). Manual verification of the mobile genetic elements was done with ISFinder assistance (June 2016) (Siguier *et al.*, 2006b). For 01-B516 and 01-B526-R19, contigs were aligned on 01-B526 pAsa5 and pAsa9 sequences using CONTIGuator version 2.7.4 (Galardini *et al.*, 2011). Sequence alignment visualization was done with ACT version 13.0.0, DNA Plotter version 10.2 and EasyFig version 2.1 (Carver *et al.*, 2009; Carver *et al.*, 2005; Sullivan *et al.*, 2011).

PCR ANALYSES

For PCR, amplifications up to 3 kbp were performed as previously described (Trudel *et al.*, 2013), with the exception of the elongation time, which was 1 min per 1 kbp of amplicon. For amplicons longer than 3 kbp, each reaction was built as follows: 3.75 μ L of 2 mM dNTPs, 1 μ L of each 100 ng/ μ L forward and reverse primers, 1 μ L of 100 ng/ μ L template, 12.25 μ L of H₂O, 5 μ L of 5X LongAmp Taq buffer (New England BioLabs, Massachusetts, US), and 1 μ L LongAmp Taq (New England BioLabs). For those amplicons, the elongation time was 50 s per 1 kb of amplicon. New PCR primers were designed manually and verified with Oligoanalyser 3.1 from Integrated DNA Technologies (IDT, <http://www.idtdna.com/calc/analyzer>). The PCR assays were performed at least twice, and appropriate positive and negative controls were included with each assay. The PCR primers are listed in Additional File 3.

PLASMID VISUALIZATION BY AGAROSE GEL ELECTROPHORESIS

A. salmonicida subsp. *salmonicida* plasmids were visualized by agarose gel electrophoresis using a protocol adapted from Wheatcroft *et al.* (Wheatcroft *et al.*, 1990). 160 μ L of overnight *P. aminophilus* JCM 7686 in tryptic soy broth (EMD Millipore) or *A. salmonicida* subsp. *salmonicida* in lysogeny broth (EMD Millipore) were mixed with 1 mL ice-cold 2% N-lauroyl sarcosine and centrifuged at 17200 x g for 10 min at 4°C. Supernatants were discarded and pellets were suspended in 40 μ L 10 mM Tris, 10 mM EDTA, 4 mg/mL RNase, 1 mg/mL lysozyme, 1 mg/mL bromophenol blue and 15.2% Ficoll solution. Samples were incubated on ice for 45 min. A 0.75% agarose gel in Tris-borate-EDTA buffer (TBE 1X) was loaded with 40 μ L of SDS 10%, 0.5% xylene cyanol solution, which was migrated backwards at 100 V for 15 min. DNA samples were then centrifuged at 376 x g for 20 s, then 20 μ L were loaded on the gel. The polarity was inverted back and samples were migrated 30 min at 40V, and then 5–7 h at 100 V. Gels were stained in 1 μ g/mL ethidium bromide for 40 min, rinsed in water for 1 h to remove the excess of ethidium bromide, and photographed using short-length UV transilluminator.

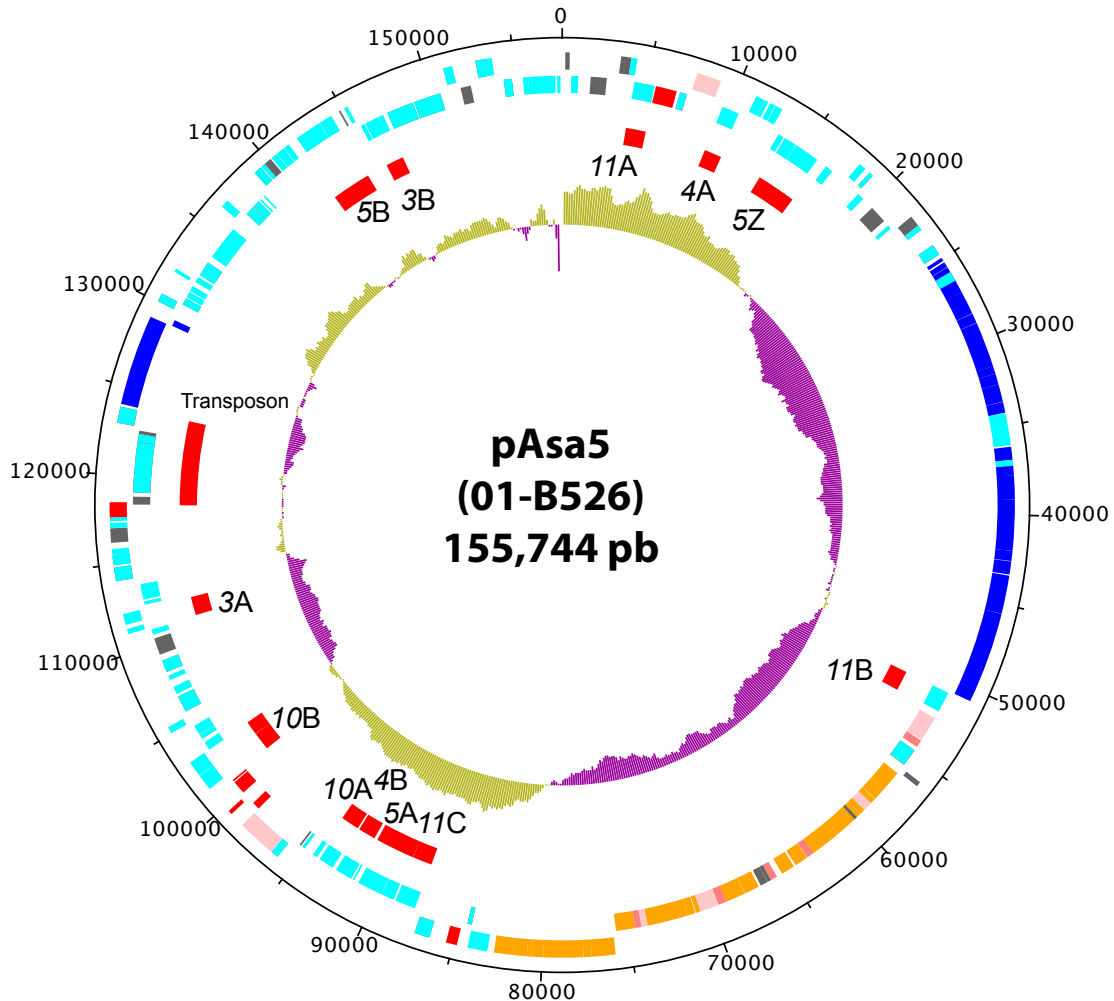
ABBREVIATIONS

IS : insertion sequence ; **ORF** : open reading frame ; **SMRT** : single molecule real-time ; **TTSS** : type three secretion system

ACKNOWLEDGMENTS

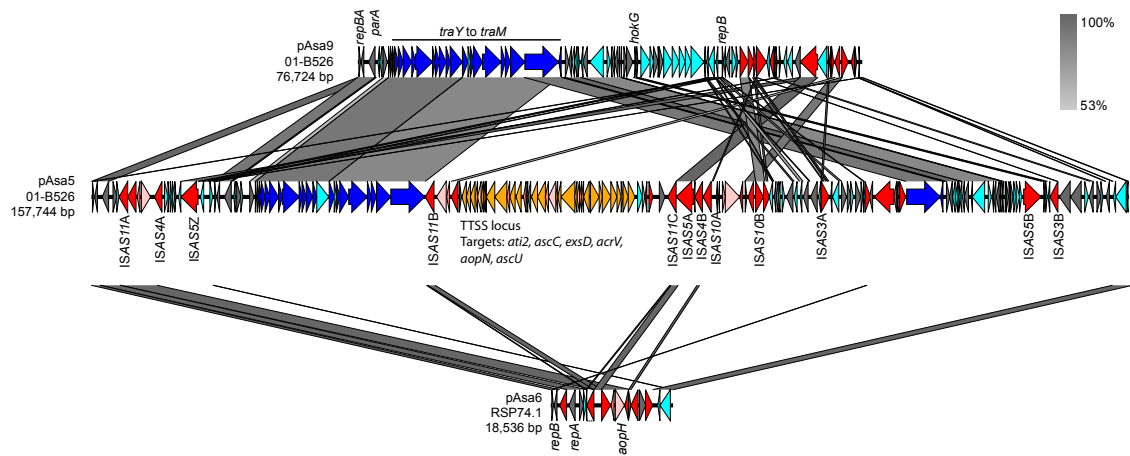
We thank Mégane Paccaud, Mathilde Goldschmitt, and Romain Chanut for technical support. We also thank the Laboratoire de bactériologie clinique, Faculté de médecine vétérinaire, Université de Montréal (Montreal, QC, Canada), the Félix d'Hérelle Reference Center (Département de biochimie, de microbiologie et de bio-informatique, Université Laval, Quebec City, QC, Canada), M. Reith (National Research Council laboratories, Halifax, Nova Scotia), J. Frey (University of Bern), and R.M.W. Stevenson (University of Guelph) for the *A. salmonicida* subsp. *salmonicida* isolates.

RELEVANT SUPPORTING INFORMATION³



Additional File 3.1 Circular map of pAsa5 of 01-B526 strain with G+C skew. pAsa5 sequence and features were visualized with Artemis and DNAPlotter (Carver *et al.*, 2009). From the outermost ring moving inwards, the first two circles shows open reading frames (forward, then reverse strand) in the colours described in Fig. 3.1. The third circle shows mobile elements. The fourth circle shows the G+C skew, using the following colours: purple: below average; yellow: above average.

³ L'ensemble des fichiers supplémentaires associés à ce manuscrit peuvent être consultés à l'adresse suivante : <http://doi.org/10.1186/s12864-017-3921-1>



Additional File 3.2 Alignment between pAsa5 and pAsa9 from 01-B526 with pAsa6. Methods and colour shown are the same as Fig. 3.1.

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- Att  r  , S.A., Vincent, A.T., Trudel, M.V., Chanut, R., Charette, S.J., 2015. Diversity and Homogeneity among Small Plasmids of *Aeromonas salmonicida* subsp. *salmonicida* Linked with Geographical Origin. *Frontiers in microbiology* 6, 1274.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455-477.
- Beatson, S.A., Walker, M.J., 2014. Tracking antibiotic resistance. *Science* 345, 1454-1455.
- Burr, S.E., Pugovkin, D., Wahli, T., Segner, H., Frey, J., 2005. Attenuated virulence of an *Aeromonas salmonicida* subsp. *salmonicida* type III secretion mutant in a rainbow trout model. *Microbiology* 151, 2111-2118.
- Burr, S.E., Stuber, K., Wahli, T., Frey, J., 2002. Evidence for a type III secretion system in *Aeromonas salmonicida* subsp. *salmonicida*. *J. Bacteriol.* 184, 5966-5970.
- Burr, S.E., Wahli, T., Segner, H., Pugovkin, D., Frey, J., 2003. Association of Type III secretion genes with virulence of *Aeromonas salmonicida* subsp. *salmonicida*. *Dis. Aquat. Organ.* 57, 167-171.
- Buttner, D., 2012. Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 76, 262-310.
- Carver, T., Harris, S.R., 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, T., Thomson, N., Bleasby, A., Berriman, M., Parkhill, J., 2009. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 25, 119-120.
- Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., Parkhill, J., 2005. ACT: the Artemis Comparison Tool. *Bioinformatics* 21, 3422-3423.
- Chai, B., Wang, H., Chen, X., 2012. Draft genome sequence of high-melanin-yielding *Aeromonas media* strain WS. *J. Bacteriol.* 194, 6693-6694.
- Charette, S.J., Brochu, F., Boyle, B., Fillion, G., Tanaka, K.H., Derome, N., 2012. Draft genome sequence of the virulent strain 01-B526 of the fish pathogen *Aeromonas salmonicida*. *J. Bacteriol.* 194, 722-723.
- Coil, D., Jospin, G., Darling, A.E., 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* 31, 587-589.
- Dacanay, A., Boyd, J.M., Fast, M.D., Knickle, L.C., Reith, M.E., 2010. *Aeromonas salmonicida* Type I pilus system contributes to host colonization but not invasion. *Dis. Aquat. Organ.* 88, 199-206.

- Dacanay, A., Knickle, L., Solanky, K.S., Boyd, J.M., Walter, J.A., Brown, L.L., Johnson, S.C., Reith, M., 2006. Contribution of the type III secretion system (TTSS) to virulence of *Aeromonas salmonicida* subsp. *salmonicida*. *Microbiology* 152, 1847-1856.
- Daher, R.K., Fillion, G., Tan, S.G., Dallaire-Dufresne, S., Paquet, V.E., Charette, S.J., 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.
- Dallaire-Dufresne, S., Tanaka, K.H., Trudel, M.V., Lafaille, A., Charette, S.J., 2014. Virulence, genomic features, and plasticity of *Aeromonas salmonicida* subsp. *salmonicida*, the causative agent of fish furunculosis. *Vet. Microbiol.* 169, 1-7.
- Dautremepuits, C., Fortier, M., Croisetièrre, 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.
- Derome, N., Boutin, S., Llewellyn, M., Gauthier, J. 2016. Opportunistic pathogens of fish, In: Hurst, C.J. (Ed.) *The Rasputin effect: when commensals and symbionts become parasitic*. Springer, New York, NY, 649.
- Dziewit, L., Czarnecki, J., Wibberg, D., Radlinska, M., Mrozek, P., Szymczak, M., Schluter, A., Puhler, A., Bartosik, D., 2014. Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. *BMC genomics* 15.
- Emond-Rheault, J.G., Vincent, A.T., Trudel, M.V., Brochu, F., Boyle, B., Tanaka, K.H., Attéré, S.A., Jubinville, E., Loch, T.P., Winters, A.D., Faisal, M., Frenette, M., Derome, N., Charette, S.J., 2015a. Variants of a genomic island in *Aeromonas salmonicida* subsp. *salmonicida* link isolates with their geographical origins. *Vet. Microbiol.* 175, 68-76.
- Emond-Rheault, J.G., Vincent, A.T., Trudel, M.V., Frey, J., Frenette, M., Charette, S.J., 2015b. *AsaGEI2b*: a new variant of a genomic island identified in the *Aeromonas salmonicida* subsp. *salmonicida* JF3224 strain isolated from a wild fish in Switzerland. *FEMS Microbiol. Lett.* 362.
- Frey, J., Origgi, F.C., 2016. Type III Secretion System of *Aeromonas salmonicida* Undermining the Host's Immune Response. *Frontiers in Marine Science* 3.
- Froquet, R., Cherix, N., Burr, S.E., Frey, J., Vilches, S., Tomas, J.M., Cosson, P., 2007. Alternative host model to evaluate *Aeromonas* virulence. *Appl. Environ. Microbiol.* 73, 5657-5659.
- Galardini, M., Biondi, E.G., Bazzicalupo, M., Mengoni, A., 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* 6, 11.
- Garcillan-Barcia, M.P., Alvarado, A., de la Cruz, F., 2011. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol. Rev.* 35, 936-956.
- Gustafson, C.E., Chu, S.J., Trust, T.J., 1994. Mutagenesis of the paracrystalline surface protein array of *Aeromonas salmonicida* by endogenous insertion elements. *J. Mol. Biol.* 237, 452-463.

- Hanninen, M.L., Hirvela-Koski, V., 1997. Molecular and phenotypic methods for the characterization of atypical *Aeromonas salmonicida*. *Vet. Microbiol.* 56, 147-158.
- Imbeault, S., Parent, S., Lagacé, M., Uhland, C.F., Blais, J.-F., 2006. Using Bacteriophages to prevent furunculosis caused by *Aeromonas salmonicida* in farmed brook trout. *Journal of Aquatic Animal Health* 18, 203-214.
- Koren, S., Harhay, G.P., Smith, T.P., Bono, J.L., Harhay, D.M., McVey, S.D., Radune, D., Bergman, N.H., Phillippy, A.M., 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol* 14, R101.
- Long, M., Nielsen, T.K., Leisner, J.J., Hansen, L.H., Shen, Z.X., Zhang, Q.Q., Li, A., 2016. *Aeromonas salmonicida* subsp. *salmonicida* strains isolated from Chinese freshwater fish contain a novel genomic island and possible regional-specific mobile genetic elements profiles. *FEMS Microbiol. Lett.* 363.
- McIntosh, D., Cunningham, M., Ji, B., Fekete, F.A., Parry, E.M., Clark, S.E., Zalinger, Z.B., Gilg, I.C., Danner, G.R., Johnson, K.A., Beattie, M., Ritchie, R., 2008. Transferable, multiple antibiotic and mercury resistance in Atlantic Canadian isolates of *Aeromonas salmonicida* subsp. *salmonicida* is associated with carriage of an IncA/C plasmid similar to the *Salmonella enterica* plasmid pSN254. *J. Antimicrob. Chemother.* 61, 1221-1228.
- Mehdizadeh Gohari, I., Kropinski, A.M., Weese, S.J., Parreira, V.R., Whitehead, A.E., Boerlin, P., Prescott, J.F., 2016. Plasmid Characterization and Chromosome Analysis of Two *netF*⁺ *Clostridium perfringens* Isolates Associated with Foal and Canine Necrotizing Enteritis. *PloS one* 11, e0148344.
- Najimi, M., Balado, M., Lemos, M.L., Osorio, C.R., 2009. Genetic characterization of pAsa6, a new plasmid from *Aeromonas salmonicida* subsp. *salmonicida* that encodes a type III effector protein AopH homolog. *Plasmid* 61, 176-181.
- Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens, R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42, D206-214.
- Piotrowska, M., Popowska, M., 2015. Insight into the mobilome of *Aeromonas* strains. *Frontiers in microbiology* 6, 494.
- Reith, M.E., Singh, R.K., Curtis, B., Boyd, J.M., Bouevitch, A., Kimball, J., Munholland, J., Murphy, C., Sarty, D., Williams, J., Nash, J.H., Johnson, S.C., Brown, L.L., 2008. The genome of *Aeromonas salmonicida* subsp. *salmonicida* A449: insights into the evolution of a fish pathogen. *BMC genomics* 9, 427.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944-945.
- Siguier, P., Gourbeyre, E., Chandler, M., 2014. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* 38, 865-891.
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., Chandler, M., 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34, D32-36.

- Sørum, H., L'Abée-Lund, T., Solberg, A., Wold, A., 2003. Integron-containing IncU R plasmids pRAS1 and pAr-32 from the fish pathogen *Aeromonas salmonicida*. *Antimicrob. Agents Chemother.* 47, 1285-1290.
- Stuber, K., Burr, S.E., Braun, M., Wahli, T., Frey, J., 2003. Type III secretion genes in *Aeromonas salmonicida* subsp. *salmonicida* are located on a large thermolabile virulence plasmid. *J. Clin. Microbiol.* 41, 3854-3856.
- Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009-1010.
- Tanaka, K.H., Dallaire-Dufresne, S., Daher, R.K., Frenette, M., Charette, S.J., 2012. An insertion sequence-dependent plasmid rearrangement in *Aeromonas salmonicida* causes the loss of the type three secretion system. *PLoS one* 7, e33725.
- Tanaka, K.H., Frenette, M., Charette, S.J., 2013. IS-mediated loss of virulence by *Aeromonas salmonicida*: A tangible piece of an evolutionary puzzle. *Mobile genetic elements* 3, e23498.
- Tanaka, K.H., Vincent, A.T., Trudel, M.V., Paquet, V.E., Frenette, M., Charette, S.J., 2016. The mosaic architecture of *Aeromonas salmonicida* subsp. *salmonicida* pAsa4 plasmid and its consequences on antibiotic resistance. *PeerJ* 4, e2595.
- Trudel, M.V., Tanaka, K.H., Fillion, G., Daher, R.K., Frenette, M., Charette, S.J., 2013. Insertion sequence *AS5* (*ISAS5*) is involved in the genomic plasticity of *Aeromonas salmonicida*. *Mobile genetic elements* 3, e25640.
- Trudel, M.V., Vincent, A.T., Attéré, S.A., Labbé, M., Derome, N., Culley, A.I., Charette, S.J., 2016. Diversity of antibiotic-resistance genes in Canadian isolates of *Aeromonas salmonicida* subsp. *salmonicida*: dominance of pSN254b and discovery of pAsa8. *Scientific reports* 6, 35617.
- Vincent, A.T., Boyle, B., Derome, N., Charette, S.J., 2014a. Improvement in the DNA sequencing of genomes bearing long repeated elements. *J. Microbiol. Methods* 107, 186-188.
- Vincent, A.T., Emond-Rheault, J.G., Barbeau, X., Attéré, S.A., Frenette, M., Lagüe, P., Charette, S.J., 2016a. Antibiotic resistance due to an unusual ColE1-type replicon plasmid in *Aeromonas salmonicida*. *Microbiology* 162, 942-953.
- Vincent, A.T., Tanaka, K.H., Trudel, M.V., Frenette, M., Derome, N., Charette, S.J., 2015. Draft genome sequences of two *Aeromonas salmonicida* subsp. *salmonicida* isolates harboring plasmids conferring antibiotic resistance. *FEMS Microbiol. Lett.* 362.
- Vincent, A.T., Trudel, M.V., Freschi, L., Nagar, V., Gagné-Thivierge, C., Levesque, R.C., Charette, S.J., 2016b. Increasing genomic diversity and evidence of constrained lifestyle evolution due to insertion sequences in *Aeromonas salmonicida*. *BMC genomics* 17, 44.
- Vincent, A.T., Trudel, M.V., Paquet, V.E., Boyle, B., Tanaka, K.H., Dallaire-Dufresne, S., Daher, R.K., Frenette, M., Derome, N., Charette, S.J., 2014b. Detection of variants of the pRAS3, pAB5S9, and pSN254 plasmids in *Aeromonas salmonicida* subsp. *salmonicida*: multidrug resistance, interspecies exchanges, and plasmid reshaping. *Antimicrob. Agents Chemother.* 58, 7367-7374.

Wheatcroft, R., McRae, D.G., Miller, R.W., 1990. Changes in the *Rhizobium meliloti* Genome and the Ability to Detect Supercoiled Plasmids During Bacteroid Development. *Mol. Plant. Microbe Interact.* 3, 9-17.

Chapitre 4 DISCUSSION

4.1 RAPPEL DE L'HYPOTHÈSE DE DÉPART ET DES RÉSULTATS

Au début de ce travail, l'hypothèse formulée était que certaines variations structurales causées par les IS pouvaient apporter des avantages adaptatifs à *A. salmonicida* ssp. *salmonicida*. Cette hypothèse avait pour but d'expliquer la présence de ce grand nombre d'éléments mobiles dans le génome. Elle s'appuie sur une position plus nuancée de l'impact des IS sur l'intégrité et la diversité génomique. Dans ce modèle, les événements délétères et bénéfiques causés par les IS leur permettent de rester en équilibre entre transpositions et inactivations sans que ni la population bactérienne ni la population d'IS ne s'écroulent.

Les résultats de cette thèse s'appuient sur un travail de séquençage à haut débit, de génomique comparative et de validations expérimentales. Ils ont été présentés en deux sections : la découverte de variants du plasmide de résistance pAsa4 (Chapitre 2), et la découverte du plasmide pAsa9 et de nouveaux modèles de recombinaison sur le plasmide pAsa5 (Chapitre 3). Dans les deux cas, plusieurs classes d'éléments mobiles de l'ADN (IS, plasmides, transposons, intégrons) sont impliquées. Pour cette discussion générale, les IS et les événements qui ont été identifiés dans les chapitres précédents seront mis en évidence et approfondis par rapport à l'hypothèse de départ. Finalement, puisqu'il est impossible de parler des IS sans parler de leurs vecteurs, l'architecture et l'évolution des plasmides qui ont été étudiées au cours de la thèse (pAsa4, pAsa5 et pAsa9) seront discutées.

4.2 LES VARIANTS DU PLASMIDE PAsa4 CONFÈRENT DIFFÉRENTES RÉSISTANCES AUX ANTIBIOTIQUES

Dans le Chapitre 2, nous avons confirmé qu'il existait bien des variants du plasmide pAsa4, tels que suspecté par Trudel *et al.* (2016) (Tanaka *et al.*, 2016). Aussi, les variations dans l'architecture du plasmide étaient bien responsables des différences en résistance aux antibiotiques qui avaient été observées dans l'étude précédente (Tanaka *et al.*, 2016; Trudel *et al.*, 2016). La famille des plasmides pAsa4 a été séparée en trois selon l'architecture des

variants : pAsa4 et pAsa4c, d'origine européenne, et pAsa4b, d'origine nord-américaine. L'analyse comparative de la séquence de ces trois plasmides a permis d'identifier des variations structurales qui influençaient le phénotype de résistance des souches porteuses. Ces variations présentées à la Figure 4.1 sont : [1] la délétion des gènes de résistance *tet(E)* chez pAsa4c, [2] l'insertion d'une cassette *aadA1* chez pAsa4 et pAsa4c et [3] la présence d'un fragment ISCR couplé à un gène *catA1* chez pAsa4 et pAsa4c (Figure 4.1). De plus, l'inversion ainsi que plusieurs délétions dans le locus de conjugaison de pAsa4 influencent négativement la capacité de conjugaison de ce plasmide [4]. Deux autres modifications apportées par les IS, mais associées à aucun phénotype particulier, ont aussi été notées dans cet article. Par exemple, l'ISAS5 interrompue par Tn21 est aussi le site d'insertion de deux IS, ISAS9 dans pAsa4 et ISAS2 dans pAsa4c [5]. Finalement, pAsa4c porte aussi une ISKpn3 [6] (Tanaka *et al.*, 2016).

4.2.1 RETOUR SUR LES VARIATIONS STRUCTURALES DE PAsa4 CAUSÉES PAR LES IS

De ces six événements, trois impliquent directement des IS et sont donc très pertinents pour l'hypothèse de départ. La recombinaison homologue créant l'inversion dans les gènes de conjugaison, l'insertion d'IS dans ISAS5 et la région ISCR seront donc examinées plus en détail (Figure 4.1, [3-4-5]).

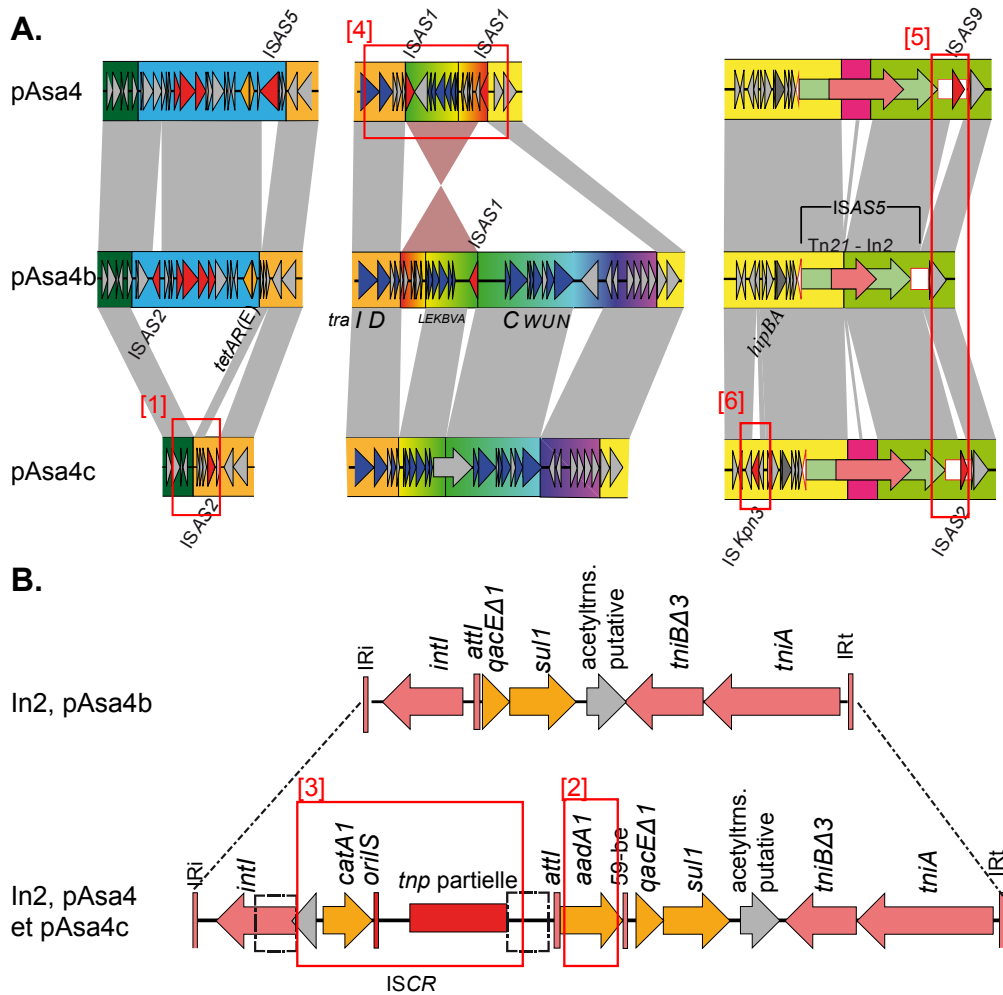


Figure 4.1 Mise en évidence des variations structurales entre les plasmides pAsa4. A. Trois sections d'un alignement entre les variants de pAsa4. Les blocs de couleur ont été définis empiriquement selon les lieux d'insertion et de délétion. Les flèches représentent des éléments sur les plasmides, selon les couleurs suivantes : **Rouge** : séquences d'insertion et transposases, **Orange** : gènes de résistance aux antibiotiques, **Bleu** : gènes impliqués dans la conjugaison, **Noir** : gènes *hipBA*, **Blanc** : *ISAS5* interrompue par *Tn21*, **Rose** : *In2*, **Gris** : codant pour des protéines hypothétiques et autres. **B.** Grossissement des *In2* de pAsa4b (haut) et pAsa4/pAsa4c (bas). Les encadrés pointillés montrent la répétition dans l'*In2* de pAsa4/pAsa4c. Les couleurs utilisées sont les mêmes qu'au Panneau A. **Tous les panneaux.** Les encadrés rouges marquent les variations structurales qui sont discutées dans le texte : [1] : Délétion du segment contenant *tetAR(E)* dans pAsa4c, [2] : Insertion de la cassette *aadA1*, [3] : insertion *ISCR*, [4] : Inversion et délétion dans les gènes de conjugaison de pAsa4, [5] : l'*ISAS5* interrompue par *Tn21* est aussi le site d'insertion de deux IS, [6] : pAsa4c possède une *ISKpn3*. Les éléments de la figure proviennent des Figures 1A et 2B de Tanaka *et al.* (2016) reproduits et modifiés selon sa licence CC BY (<http://creativecommons.org/licenses/by/4.0/>) (Figure 2.1 et 2.2 du Chapitre 2).

4.2.1.1 L'INVERSION DANS pAsa4

L'inversion d'une portion des gènes dans pAsa4 par rapport aux deux autres variants est probablement due à la présence de copies d'*ISASI* inversées l'un par rapport à l'autre et flanquant cette variation structurale. Dans ce cas-ci, c'est une recombinaison homologue entre les deux IS, causée par la machinerie de recombinaison de la bactérie, qui permet d'arriver à ce résultat. En ce sens, l'inversion dans pAsa4 entre dans la même catégorie d'évènements que les délétions du SSTT observées sur pAsa5 (Figure 1.5) (Tanaka *et al.*, 2012). Cependant, cette inversion à elle seule ne peut expliquer pourquoi pAsa4 est non conjugatif alors que pAsa4b et pAsa4c le sont (Tanaka *et al.*, 2016). En effet, en plus de l'inversion, plusieurs gènes environnants ont été supprimés sur pAsa4 (Figure 4.1, encadré 4). Même si certains de ces gènes semblent facultatifs pour la conjugaison (le plasmide pAsa4c est conjugatif malgré plusieurs petites délétions (Tanaka *et al.*, 2016)), il manque à pAsa4 plusieurs gènes étant identifiés comme essentiels à cette fonction.

Une analyse des séquences codantes des trois pAsa4 avec conjScan et une comparaison avec les gènes de conjugaison des plasmides A/C montrent que le système de sécrétion de type IV des pAsa4 est apparenté à celui du plasmide F (Abby *et al.*, 2016; Guglielmini *et al.*, 2014; Harmer et Hall, 2015). Sur le plasmide pAsa4, certains gènes essentiels à l'assemblage du pilus, dont *traU*, qui code pour une ATP synthase, sont absents (Guglielmini *et al.*, 2014). D'un autre côté, les différences entre les loci de conjugaison chez pAsa4b et pAsa4c concernent des gènes codants pour des protéines hypothétiques. En fait, conjScan identifie davantage de gènes de système de sécrétion de type IV à partir de pAsa4c qu'à partir de pAsa4b (Abby *et al.*, 2016). Cela peut donc expliquer pourquoi les deux plasmides sont conjugatifs.

Il est donc possible de diviser la perte de la capacité de transfert de pAsa4 en deux types d'évènements : une ou plusieurs délétions dans les gènes en 3', ainsi que l'inversion causée par la recombinaison homologue des *ISASI*. Le rôle des IS dans cette perte de fonction serait plus facilement estimable si la séquence des évènements était connue. En effet, si l'inversion a eu lieu en premier, elle a probablement causé la perte de fonction principale, le reste ayant eu lieu par érosion. Par contre, si les délétions ont eu lieu en premier, plus

aucune pression de sélection n'empêcherait l'inversion, et celle-ci serait non contributive à la perte de la conjugaison du plasmide pAsa4.

Dans les deux autres plasmides, cette région n'est pas encadrée par des *ISASI*, ce qui suggère des transpositions d'arrivée ou de départ pour en venir à cette conformation. Puisque *ISASI* est connue pour transposer en présence de stress thermique (c'est-à-dire, à partir d'une température de culture de 25°C) (Gustafson *et al.*, 1994; Ishiguro *et al.*, 1981), cette séquence d'évènements a peut-être été déclenchée par un stress. pAsa4 n'a été trouvé que dans deux souches (A449 et A450), qui sont possiblement clonales, puisqu'elles ont été isolées dans la même localité (Michel et Dubois-Darnaudpeys, 1980). Les transpositions d'*ISASI* à proximité des gènes de conjugaison, l'inversion observée dans pAsa4 et la délétion des gènes avoisinants sont peut-être des évènements peu fréquents qui ont été caractérisés par hasard.

La perte de la conjugaison de ce plasmide n'est pas vraiment un évènement favorable pour la distribution du plasmide pAsa4. En effet, la conjugaison favoriserait la persistance des grands plasmides en absence de sélection. Il s'agit d'un dernier mécanisme permettant la réinfection des bactéries ayant perdu le plasmide par défaut des systèmes de partition et de toxine/antitoxine (Bahl *et al.*, 2007; De Gelder *et al.*, 2007). Cependant, la conjugaison seule n'atteindrait pas des fréquences assez élevées pour permettre le maintien d'un plasmide dans une population (Bergstrom *et al.*, 2000; Harrison et Brockhurst, 2012). D'un autre côté, les plasmides non conjugatifs ou avec un faible taux de conjugaison sont moins coûteux métaboliquement (Harrison et Brockhurst, 2012). Donc, la perte de la conjugaison pourrait avoir réduit le coût métabolique de pAsa4 par rapport aux variants conjugatifs, ce qui est une coadaptation.

Pour tester cette hypothèse, une souche naïve d'*A. salmonicida* ssp. *salmonicida* pourrait être transformée avec chacun des variants de la famille de pAsa4. Une comparaison des courbes de croissance entre les transformants permettrait de mesurer une différence dans le fardeau métabolique. Des cultures mixtes de la même souche portant différents variants de pAsa4 pourraient aussi être faites. Puisque les plasmides portent différents gènes de résistance aux antibiotiques, un repiquage sur milieux sélectifs permettrait de calculer les fractions de la culture portant l'un ou l'autre des plasmides. Cette valeur permettrait

d'estimer la compétitivité de chaque transformant. Malheureusement, les outils moléculaires actuels pour cette espèce ne permettent pas d'introduire de l'ADN exogène de grande taille, et donc d'obtenir une souche portant l'un ou l'autre des variants de pAsa4.

Finalement, il serait possible de contourner le manque d'outils moléculaires par une expérience dérivée. Les IS*AS1* sont connues pour transposer à une température supérieure à 25°C (Gustafson *et al.*, 1994; Ishiguro *et al.*, 1981), et une telle condition favorise aussi les recombinaisons homologues entre IS (Tanaka *et al.*, 2012). Cependant, l'effet d'une température de culture élevée sur l'architecture de pAsa4 n'a pas encore été testé. Il serait envisageable et prometteur de cultiver une souche possédant un plasmide de la famille de pAsa4 (probablement 01-B522, puisque pAsa4b est le plus long des trois) de façon répétée à 25°C, à la manière de ce qui a été fait avec pAsa5 (Daher *et al.*, 2011). Cette expérience permettrait peut-être de reproduire la séquence des événements observée qui a mené à pAsa4. Si tel est le cas, c'est la souche 01-B522 porteuse d'un nouveau pAsa4 altéré par la chaleur qui pourrait ensuite être utilisée pour les tests de croissance et de culture mixte proposés au paragraphe précédent.

4.2.1.2 INSERTION DE DIFFÉRENTES FAMILLES D'IS DANS IS*AS5*

Une des copies d'IS*AS5* de la famille des pAsa4 est déjà interrompue par le transposon Tn21 (Reith *et al.*, 2008; Tanaka *et al.*, 2016). De plus, pour pAsa4 et pAsa4c, une autre IS (IS*AS2* pour pAsa4, IS*AS9* pour pAsa4c) est insérée dans cette même IS*AS5* (Reith *et al.*, 2008; Tanaka *et al.*, 2016). Il est donc possible que les IS*AS5* contiennent un site préférentiel d'insertion. En effet, les transposases de certaines IS, dont celles des familles IS30 et IS630, reconnaissent préférentiellement certaines suites ou combinaisons de nucléotides et ne transposent donc pas aléatoirement (Kiss *et al.*, 2007; Olsz *et al.*, 1998; Siguier *et al.*, 2014; Tenzen *et al.*, 1990).

Il est aussi possible que les transpositions d'IS*AS5* aient un impact assez négatif sur *A. salmonicida ssp. salmonicida* pour que ses copies soient fréquemment inactivées. IS*AS5* est déjà connue pour avoir transposé sur le plasmide pAsa1 (Attéré *et al.*, 2015; Trudel *et al.*, 2013) et sur le pAsa5 de 01-B526, ce qui permet d'ailleurs d'autres scénarios de recombinaison homologue (Tanaka *et al.*, 2017). L'inactivation des IS par transposition

d'une autre IS serait une façon de mitiger les conséquences de cette famille particulière dans le génome (Gueguen *et al.*, 2005; Nelson *et al.*, 2011).

Les données de génomique accumulées à ce jour ont été utilisées pour vérifier si, dans le génome, *ISAS5* est fréquemment interrompue par d'autres IS. Les résultats pour A449 montrent que cette IS est interrompue dans les 75 premiers nucléotides à trois reprises dans le chromosome, toutes les fois par une *ISAS9* (Annexe 1; Tableau 4.1). Sur le chromosome de 01-B526, ces trois IS sont continues. Cependant, 01-B526 ne possède pas de copies complètes d'*ISAS9* ou d'*ISAS2* (Tableau 4.1). La seule autre IS complète qui est interrompue par une autre IS dans le génome d'*A. salmonicida ssp. salmonicida* est l'*ISAS11C* sur le plasmide pAsa5 (ironiquement, elle est interrompue par l'*ISAS5A*). L'*ISAS5*, et particulièrement ses premiers nucléotides, semble donc être un lieu préférentiel d'insertion pour *ISAS9* et possiblement pour *ISAS2*.

Cet exercice a amené à une comparaison systématique des IS entre A449 et 01-B526, puis dans une moindre mesure avec JF2267 et 01-B522. Ces résultats sont présentés à la suite de la description des événements spécifiques à pAsa4, c'est-à-dire à la section 4.2.2.

4.2.1.3 L'ISCR ET L'INSERTION DE *CAT1*

Les intégrons de pAsa4 et de pAsa4c possèdent un gène partiel codant pour une transposase d'*ISCR* (Figure 4.1) (Tanaka *et al.*, 2016). Contrairement aux IS classiques, les *ISCR* ne possèdent pas de répétitions inversées. Leur transposase reconnaît deux sites, *oriIS* et *terIS* (dont la séquence est moins bien définie qu'*oriIS*), et catalyse une transposition en cercle roulant (Toleman *et al.*, 2006b). Les intégrons de pAsa4 et de pAsa4c possèdent une *oriIS*, mais pas de *terIS*. De plus, le segment *ISCR* est adjacent au gène *catA1*, qui code pour une chloramphénicol acétyltransférase (Figure 4.1B) (Tanaka *et al.*, 2016).

Des erreurs de reconnaissance du *terIS* peuvent mener à la mobilisation de séquences adjacentes par des *ISCR* (Toleman *et al.*, 2006a). *ISCR1*, en particulier, est connue pour mobiliser des gènes de résistance aux antibiotiques (*cat* et *qnr*, qui confère une résistance aux quinolones) et pour créer des intégrons de classe 1 complexes (Toleman *et al.*, 2006b). Cette mobilisation est rendue possible par une série de transpositions fautives et de recombinaison homologue entre l'intermédiaire *ISCR* circulaire et un intégron. La

recombinaison de séquences crée des duplications, comme ce qui est observé autour de la section ISCR-*catA1* des pAsa4 (Figure 4.1) (Naas *et al.*, 2008; Tanaka *et al.*, 2016; Toleman *et al.*, 2006b). Le fragment ISCR de pAsa4 et de pAsa4c, contenant déjà le gène *catA1*, est possiblement arrivé dans l'intégron des plasmides par une série d'évènements similaires. Cependant, contrairement aux cas précédents, le gène de la transposase dans ISCR-pAsa4 est partiel (Tanaka *et al.*, 2016). Il est donc possible que ce segment ait subi des délétions après son arrivée dans pAsa4.

Le segment ISCR-*catA1* de pAsa4 et pAsa4c confère un avantage sélectif aux bactéries porteuses du plasmide en présence de chloramphénicol. Puisque le gène de résistance a vraisemblablement été mobilisé par l'IS, il s'agit d'un cas où, conformément à l'hypothèse de départ de la thèse, l'IS vient apporter une diversité génétique avantageuse. Toutefois, certaines informations suggèrent de l'insertion de *catA1* est relativement ancienne puisque ce segment a subi de l'érosion, visible par la délétion d'une partie de la transposase. De plus, *catA1*-pAsa4 est fortement transcrit et confère la plus forte résistance au chloramphénicol observée à ce jour pour *A. salmonicida* ssp. *salmonicida* (Vincent *et al.*, 2016a). Le promoteur de *catA1*-pAsa4 doit donc être fortement adapté à la machinerie de transcription de la bactérie. Une façon de le vérifier serait de comparer les niveaux d'expression de *catA1* et les concentrations minimales inhibitrices de chloramphénicol pour une souche d'*A. salmonicida* ssp. *salmonicida* pAsa4 positive, et pour un transconjugant *E. coli* qui a été produit dans le cadre de l'étude sur ce plasmide (Tanaka *et al.*, 2016).

4.2.2 DISTRIBUTION DES IS ENTRE LES PAsa4, LES AUTRES PLASMIDES ET LE CHROMOSOME

Le séquençage des variants de pAsa4 a permis de répertorier une IS jusqu'ici non retrouvée chez *A. salmonicida* ssp. *salmonicida*, IS*Kpn3* (Tanaka *et al.*, 2016). Cette IS n'est retrouvée qu'en une seule copie dans le génome de JF2267, sur pAsa4c. Elle n'est présente dans aucun des génomes pleinement assemblés (A449 et 01-B526), ni dans 01-B522, la souche porteuse de pAsa4b (Tableau 4.1; Annexe 1).

Les plasmides sont une voie de propagation importante des IS entre les génomes (Nelson *et al.*, 2011; Touchon et Rocha, 2007). Certaines IS d'*A. salmonicida* ssp. *salmonicida* ne sont

pas uniformément distribuées entre les différents réplicons (*ISAS1*, par exemple, n'est présent dans le génome d'A449 que sur pAsa4) (Reith *et al.*, 2008). La découverte d'*ISKpn3* sur pAsa4c, celle d'*ISAS35* sur pAsa9 (Tanaka *et al.*, 2017) et les différences en site d'insertion d'*ISAS9* discutées au point 4.2.1.2 incitaient à développer un meilleur portrait de la distribution des IS entre les plasmides et les chromosomes de différentes souches d'*A. salmonicida* ssp. *salmonicida*.

Un compte du nombre d'IS pour les deux génomes complètement assemblés d'*A. salmonicida* ssp. *salmonicida*, A449 et 01-B526, a donc été fait en séparant les IS des plasmides de celles du chromosome (Annexe 1 ; Tableau 4.1). De plus, la présence, mais non le nombre, d'IS a été vérifiée sur les contigs de JF2267 et de 01-B522, deux souches porteuses de pAsa4 (Annexe 1; Tableau 4.1). Seules les IS complètes à plus de 95% ont été considérées, et les *ISAS5* interrompues par *ISAS9* chez A449 ont été écartées. Le nombre d'IS détectée varie grandement des résultats de Reith *et al.* (2008), possiblement puisque les IS partielles ont été exclues.

Tableau 4.1 Nombre ou présence d’IS dans les génomes dans une sélection de génomes d’*A. salmonicida* ssp. *salmonicida*

	A449		01-B526		JF2267*	01-B522*
	plasmides	chromosome	plasmides	chromosome		
ISAS1	2	0	0	0	0	+
ISAS2	2	3	0	0	+	+
ISAS3	2	12	2	13	+	+
ISAS4	2	12	2	13	+	+
ISAS5	3	5	4	8	+	+
ISAS6	0	7	0	7	+	+
ISAS8	0	3	0	2	+	+
ISAS9	1	3	1	0	0	1
ISAS10	2	0	3	0	+	+
ISAS11	2	1	3	1	+	+
ISAS19	0	1	0	1	+	+
ISAs21	0	0	1	0	0	1
ISAs34	0	1	0	1	1	1
ISAS35	0	0	1	0	0	1
ISKpn3	0	0	0	0	1	0
Total	16	48	17	46	ND	ND
Grand total		64		63	ND	ND

* Puisque les génomes de JF2267 et de 01-B522 sont sous forme de contigs, et que les IS représentent une cause importante de bris dans l’assemblage, il n’est pas possible de déterminer le nombre d’IS exact pour ces génomes. Cependant, lorsque les contigs sont alignés sur une base de données d’IS, il est possible de déterminer si un génome contient aucune (0), une copie (1) ou plusieurs copies (+). La méthode employée est expliquée à l’Annexe 1.

Selon les résultats obtenus, il est possible de classer les IS d’*A. salmonicida* ssp. *salmonicida* en trois groupes. [1] Certaines IS sont présentes dans tous les génomes étudiés dans un nombre de copies relativement constant (Tableau 4.1, orange). Certaines ont colonisé plusieurs réplicons. C’est le cas d’ISAS3, ISAS4, ISAS5, ISAS6, ISAS8, ISAS11, ISAS19 et ISAs34. Ces IS sont probablement arrivées dans le génome d’*A. salmonicida* ssp. *salmonicida* par le transfert d’un autre élément mobile, comme fréquemment suggéré (Iranzo *et al.*, 2014; Nelson *et al.*, 2011; Touchon et Rocha, 2007). Cependant, puisque ces IS sont probablement des résidentes de longue date du génome d’*A. salmonicida* ssp. *salmonicida*, d’autres transpositions entre plasmides et chromosome, puis inversement, ont

dû avoir eu lieu (surtout pour ISAS5, qui, comme mentionné précédemment, semble encore active). [2] Un autre groupe comprend les IS associées aux plasmides et présentes seulement dans certaines souches (Tableau 4.1, bleu). ISAS1 (pAsa4), ISAS10 (pAsa5), ISAs21 (pAsa9), ISAS35 (pAsa9) et ISKpn3 (pAsa4c) font partie de ce groupe. Ces IS ont probablement été acquises récemment par transfert horizontal du plasmide porteur, mais ne se sont pas particulièrement multipliées depuis. Ces IS gravitent donc fortement autour du génome accessoire et pourraient être considérées comme « transitoires ». [3] Finalement, deux classes d'IS, ISAS2 et ISAS9, méritent une attention particulière. En effet, elles ne sont retrouvées que dans certaines souches, et sont présentes sur des plasmides (pAsa4 et pAsa9). Contrairement aux IS du dernier groupe, ISAS2 et ISAS9 ont transposé en plusieurs copies dans un chromosome, celui d'A449 (Tableau 4.1, rouge). Ces IS sont-elles aussi transitoires, mais potentiellement plus actives que celles du dernier groupe ?

4.2.3 IMPLICATION DES AUTRES ÉLÉMENTS MOBILES SUR PAsA4

Comme mentionné dans le Chapitre 2, l'intégron In2, lui-même inclus dans Tn21, est lui aussi responsable de variations génétiques et phénotypiques observées entre les variants de pAsa4. En effet, les variants pAsa4 et pAsa4c portent une cassette supplémentaire par rapport à pAsa4b, *aadA1*, qui confère une résistance à la streptomycine et à la spectinomycine (Figure 4.1B) (Reith *et al.*, 2008; Tanaka *et al.*, 2016). La comparaison entre les gènes de cet In2 et la structure de référence (Liebert *et al.*, 1999), et la présence d'une cassette supplémentaire chez pAsa4 et pAsa4c suggèrent que cette structure est toujours capable de capturer des cassettes, et donc d'augmenter le bassin de résistance aux antibiotiques des souches d'*A. salmonicida* ssp. *salmonicida* porteuses de pAsa4.

Des essais avec des cassettes construites et circularisées *in vitro*, puis électroporées dans des souches porteuses de pAsa4 pourraient être utilisés pour montrer l'activité de l'intégrase (Gestal *et al.*, 2011). Compte tenu de la petite taille des cassettes, une méthode comme le protocole d'électroporation pour *A. salmonicida* ssp. *salmonicida* pourrait être adaptée à cette expérience (Dallaire-Dufresne *et al.*, 2014a). Le même essai pourrait être fait chez les souches d'*E. coli* porteuses de pAsa4c, le cas échéant (Tanaka *et al.*, 2016).

4.2.4 L'ÉVOLUTION DU PLASMIDE PAsa4 ET LES LIENS ÉVOLUTIFS AVEC D'AUTRES PLASMIDES

Tous les éléments mobiles concernés dans les variations structurales précédemment décrites sont dépendants du porteur, un plasmide de la famille des pAsa4, pour leur transfert intercellulaire et leur persistance dans le génome. Cette dépendance est autant plus importante qu'il semble exister une partition entre les IS des plasmides et celles du chromosome. L'étude présentée au Chapitre 2 nous a permis d'assembler des informations supplémentaires sur l'évolution, la prévalence, et la distribution des plasmides de la famille de pAsa4. Avec ces caractéristiques, le risque que posent les souches porteuses de pAsa4 pour l'aquaculture et la transmission de résistances en général peut être inféré.

4.2.4.1 LIENS ÉVOLUTIFS ENTRE LES PLASMIDES DE LA FAMILLE DE PAsa4 ET CEUX DU GROUPE D'INCOMPATIBILITÉ INCA/C

Les plasmides pAsa4 et les plasmides du groupe d'incompatibilité IncA/C (plus particulièrement le groupe A/C₂) partagent un ensemble de gènes dont l'ordre est conservé (Fricke *et al.*, 2009; Harmer et Hall, 2015; Tanaka *et al.*, 2016). Ces caractéristiques, qui sont mises en évidence dans les Figures 2.3 et 2.4, suggèrent que les pAsa4 et les plasmides IncA/C dérivent d'un ancêtre commun.

L'analyse par carte de densité (Figure 2.3) donne toutefois peu d'indications sur des plasmides qui, comme pAsa4, auraient divergé du groupe devenu IncA/C dans des temps évolutifs semblables. En fait, un retour sur l'analyse et une comparaison avec des listes de plasmides du groupe d'incompatibilité IncA/C montrent que des plasmides ayant divergé du groupe il y a plus longtemps, comme pRA1 ou pXNC1_p ont aussi été inclus dans l'analyse (données accessibles dans le matériel supplémentaire de l'article du Chapitre 2) (Harmer et Hall, 2015; Meinersmann *et al.*, 2013). Tous ont groupé avec les plasmides plus représentatifs du groupe A/C, ce qui suggère que l'analyse de départ était peu restrictive. Il est possible que de tels plasmides divergents de la branche A/C se trouvent dans des environnements qui sont sous-représentés dans la base de donnée utilisée (NCBI nr/nt).

D'ailleurs, si on exclut les réarrangements de grande taille, la famille des pAsa4 doit être divisée en deux sous-groupes : pAsa4 et pAsa4b d'un côté, pAsa4c, qui partage moins

d'identité, de l'autre. Cette divergence est particulière compte tenu des pourcentages d'identité habituellement très élevés entre les génomes d'*A. salmonicida* ssp. *salmonicida*. Cependant, cette différence n'affecte que certaines parties de pAsa4. De plus, les autres plasmides de JF2267 ne sont pas affectés par cette divergence, puisque leurs séquences sont très similaires (plus de 98%) aux plasmides équivalents dans d'autres souches d'*A. salmonicida* ssp. *salmonicida* (résultats non montrés obtenus par CONTIGUator 2.7.4 entre les contigs de JF2267 et pAsa5, pAsa5 01-B526, pAsa1, pAsa2 et pAsa11, selon les numéros d'accèsion donnés à l'Annexe 1). Les variations d'identité nucléotidique entre pAsa4c et pAsa4/pAsa4b sont possiblement le produit d'une pression de sélection différente selon les régions du plasmide.

4.2.5 LES PLASMIDES DE LA FAMILLE DE PAsa4, UN EXEMPLE DE COÉVOLUTION ENTRE PLASMIDE ET HÔTE ?

Contrairement à pSN254b, les variants de pAsa4 ont été trouvés sur plusieurs continents et sont différents les uns des autres (Reith *et al.*, 2008; Tanaka *et al.*, 2016; Vincent *et al.*, 2014b). Ils n'ont aussi pas été retrouvés à ce jour dans d'autres espèces qu'*A. salmonicida* ssp. *salmonicida*. D'autres plasmides de cette espèce ont aussi été répertoriés sur différents continents, notamment les petits plasmides à haut nombre de copies (pAsa1, pAsa2, pAsa3 et pAsa11) et pAsa5 (Attéré *et al.*, 2015; Boyd *et al.*, 2003; Charette *et al.*, 2012; Fehr *et al.*, 2006; Reith *et al.*, 2008). Comparativement aux plasmides pAsa4, ces plasmides sont habituellement présents dans toutes les souches analysées (Attéré *et al.*, 2015). Le plasmide pRAS3 ou ses variants, d'un autre côté, ont été répertoriés sur trois continents (L'Abée-Lund et Sørum, 2002; Vincent *et al.*, 2014b).

Le cantonnement des plasmides pAsa4 à *A. salmonicida* ssp. *salmonicida*, l'échange de certaines IS entre ce plasmide et le chromosome et sa distribution sur plusieurs continents sont peut-être des signes d'une cohabitation ancienne du plasmide et de la bactérie. Cela pourrait aussi signifier qu'une coadaptation des plasmides pAsa4 et d'*A. salmonicida* ssp. *salmonicida* s'est faite, ce qui pourrait être mesuré (Harrison et Brockhurst, 2012). La puissance d'expression de *catA1* de pAsa4 par rapport à la chloramphénicol acétyltransférase de pAsa7 en est un indice (Vincent *et al.*, 2016a). L'analyse normalisée de

l'usage de codons préférentiels pour certains gènes du plasmide par rapport au chromosome pourrait être un autre indice obtenu *in silico* (Suzuki *et al.*, 2005). L'analyse de la conjugabilité, de la persistance et du fardeau métabolique associé de pAsa4b et c chez *E. coli* par rapport à d'autres plasmides (pSN254b, si possible) permettrait aussi de répondre à cette question.

4.2.6 APPORT DE PAsA4 À LA PLASTICITÉ GÉNOMIQUE D'*A. SALMONICIDA* SSP. *SALMONICIDA*

L'analyse de la séquence des pAsa4 a apporté plusieurs exemples de variations structurales dont certaines ont un impact direct sur le phénotype exprimé par la souche porteuse (Tanaka *et al.*, 2016). Même si cette famille de plasmides n'est pas la plus prévalente chez *A. salmonicida* ssp. *salmonicida*, elle devrait être considérée lors d'études plus poussées sur les conséquences des éléments mobiles de l'ADN sur ce génome. La mobilisation des IS du plasmide vers le chromosome et la capacité de capture de l'intégron pourraient être mesurées expérimentalement dans une suite du projet. Une telle étude permettrait de valider les variations observées entre les plasmides de la famille de pAsa4 et d'approfondir la répartition des IS entre les différents réplicons.

4.3 NOUVELLES CIBLES DE RECOMBINAISON ET NOUVEAU PLASMIDE POUR LES *A. SALMONICIDA* SSP. *SALMONICIDA* QUÉBÉCOISES

Dans le Chapitre 3, deux éléments ont été introduits: la recombinaison homologue d'ISAS5 qui mène à une autre forme de perte du SSTT, et la présence d'un nouveau plasmide nommé pAsa9 (Tanaka *et al.*, 2017). Cet article est étroitement lié aux résultats décrits dans Tanaka *et al.* (2012), puisque les phénomènes décrits sont similaires (même mécanisme, cibles différentes), et que les résultats les plus récents permettent d'expliquer les disparités des résultats les plus anciens. Il est possible de refaire avec le Chapitre 3 la même chose qu'à la section précédente, c'est-à-dire de discuter si la perte du SSTT contribue positivement à la valeur adaptative de la bactérie, conformément à ce qui a été formulé dans l'hypothèse.

4.3.1 RETOUR SUR LES MÉCANISMES DE DÉLÉTION DU SSTT

La perte de la virulence par délétion du SSTT est un phénomène qui a été observé au même moment que la découverte de ce facteur de virulence chez cette bactérie (Burr *et al.*, 2003a; Stuber *et al.*, 2003). Cependant, le phénomène de perte de virulence en général, qui pourrait être associé soit à la perte du SSTT, soit à la perte de la couche de surface A, avait déjà été observé accidentellement dans d'autres études (Adams *et al.*, 1988; Michel et Dubois-Darnaudpeys, 1980).

Une série d'articles se concluant avec le Chapitre 3 rapporte que la perte de virulence causée par la délétion du locus du SSTT se produit de façon similaire pour les cas observés (Daher, 2010; Emond-Rheault *et al.*, 2015b; Stuber *et al.*, 2003; Tanaka *et al.*, 2012; Tanaka *et al.*, 2017). Des régions entre les IS*AS11* de pAsa5, dont les gènes du SSTT, sont supprimées lorsque la bactérie est cultivée à 25°C ou plus (Daher, 2010; Tanaka *et al.*, 2012). Une recombinaison homologue a été envisagée pour expliquer ces pertes de séquence, puisqu'une copie de l'IS reste au site de délétion (Tanaka *et al.*, 2012). Toutes les recombinaisons observées en 2012 ciblaient les IS*AS11* (Tanaka *et al.*, 2012). L'étude présentée au Chapitre 3 introduit une nouvelle cible de recombinaison, IS*AS5*, et permet d'expliquer les délétions pour toutes les souches n'ayant recombiné à partir de leurs IS*AS11* (Tanaka *et al.*, 2017).

Les réarrangements entre les IS ne se font pas tous à la même fréquence. En effet, pour les souches réarrangées produites par Daher (2010) puis analysées dans les études subséquentes, 16 recombinaisons IS*AS5* ont été établies contre 9 recombinaisons IS*AS11* (Tanaka *et al.*, 2012; Tanaka *et al.*, 2017). Puisque l'IS*AS5* est deux fois plus longue que l'IS*AS11*, il a été suggéré qu'elle avait deux fois plus de chances d'être ciblée par la recombinaison (Tanaka *et al.*, 2017). Cependant, la seule reconnaissance aléatoire de séquences répétées du plasmide n'explique pas la spécificité des patrons observés. En effet, pAsa5 possède d'autres régions répétées qui pourraient faire l'objet de recombinaisons.

Le plasmide pAsa5 porte plusieurs copies de cinq types d'IS différents qui sont autant de cibles de recombinaisons potentielles (Figure 4.2). Bien sûr, les copies en sens inverse les unes des autres (toutes les IS*AS10*, toutes les IS*AS3*, IS*AS5A* et IS*AS5B*) produiraient des inversions, et non des délétions, ce qui n'a pas été testé lors de la production de souches

réarrangées (Daher, 2010). Cependant, des délétions par recombinaison *ISAS4A-ISAS4B* ou *ISAS11A-ISAS11B* sont théoriquement possibles et produiraient des délétions détectables par notre technique de génotypage (Figure 4.2). La recombinaison *ISAS4* n'a pas été observée à ce jour, et la recombinaison *ISAS11A-ISAS11B* a potentiellement été observée chez deux souches d'*A. salmonicida* ssp. *salmonicida*, BG et YK (Long *et al.*, 2016).

De plus, une autre région répétée sur le plasmide pAsa5 permettrait une délétion du locus du SSTT par recombinaison. Il s'agit de la partie 3' du gène *traI* ainsi que du gène *traM* (Tanaka *et al.*, 2017; Vincent *et al.*, 2015). Cette délétion serait détectable par notre technique de génotypage, puisqu'elle entraînerait la perte du locus du SSTT ainsi que des sites d'hybridation pour les amorces *aopO*, *resD* et *traM* (Figure 4.2). Fait intéressant, un tel patron de délétion peut être observé pour le pAsa5 de la souche HER1108 (Daher, 2010).

Ces informations suggèrent que les recombinaisons sur le plasmide pAsa5 ciblent préférentiellement les IS par rapport à la région répétée *traIM*. Cette région fait 5 kpb avec 97% d'identité entre les répétitions (pour 01-B526). *ISAS11C* fait 1304 pb (à cause de l'insertion d'*ISAS5A*) et est 100% identique aux fragments correspondants d'*ISAS11A* et *ISAS11B*. *ISAS5Z* et *ISAS5A* font 2613 pb et sont à 99% identiques. La différence d'identité pourrait être suffisante pour favoriser les recombinaisons entre IS par rapport à *traIM*, puisque l'efficacité de recombinaison est dépendante de l'homologie de séquence (Shen et Huang, 1986).

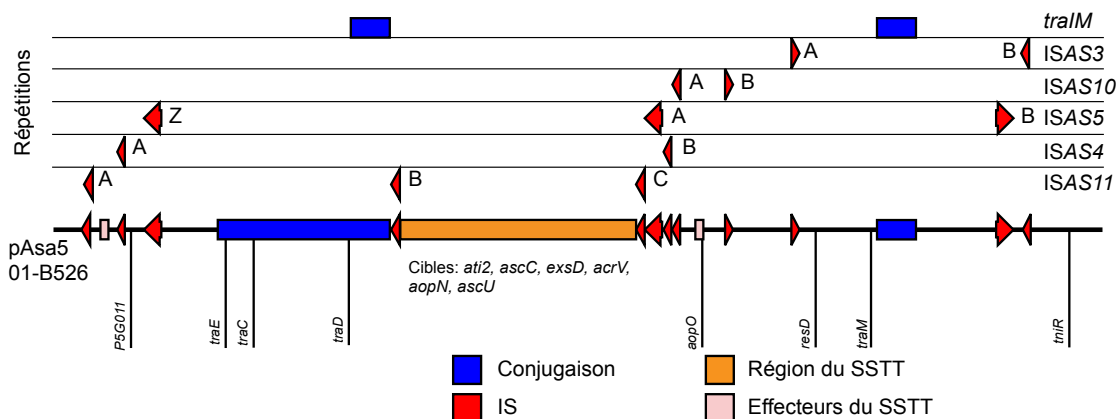


Figure 4.2 Les séquences répétées du pAsa5 de 01-B526. Cette carte linéaire montre quelques éléments importants du plasmide pAsa5 d'*A. salmonicida* ssp. *salmonicida* 01-

B526. La position des cibles PCR utilisées précédemment est indiquée sous la carte (Daher, 2010). Au-dessus de la carte, les répétitions de pAsa5 ont été isolées dans différentes pistes. Cette figure est une adaptation de la Figure 1 de Tanaka *et al.* (2017) (Figure 3.1 de la thèse), reproduite ici selon sa licence CC BY (<http://creativecommons.org/licenses/by/4.0/>).

Plusieurs études font état de délétions, d'inversions ou de formation de cointégrats causées par recombinaison homologue d'IS chez différentes espèces bactériennes (Casetta *et al.*, 1998 ; Clark *et al.*, 2014 ; Futagami *et al.*, 2006; He *et al.*, 2016 ; Selle *et al.*, 2015). Même les répétitions inversées des IS peuvent servir de gabarit à une recombinaison homologue (Ling et Cordaux, 2010). Cependant, beaucoup moins d'études ont comparé la fréquence de recombinaison entre IS par rapport à d'autres types de mutations. Dans une étude chez *Streptococcus thermophilus*, la recombinaison homologue entre IS était privilégiée à la création de mutations ponctuelles pour échapper au CRISPR-Cas ciblant un îlot génomique (Selle *et al.*, 2015). Dans une autre étude d'accumulation de mutations chez *E. coli*, le taux de recombinaison impliquant des IS a été estimé à $4,5 \times 10^{-5}$ événements par génome par génération (Lee *et al.*, 2016). Cette fréquence ne variait pas dans les lignées déficientes en système de réparation de l'ADN. Les délétions impliquant des recombinaisons d'IS étaient aussi deux fois plus fréquentes que des délétions n'impliquant aucune IS (Lee *et al.*, 2016).

Un autre plasmide, pINV de *Shigella* spp., subit des patrons de recombinaison très similaires à ceux observés chez pAsa5. Dans les deux cas, le locus du SSTT peut être supprimé du plasmide par des recombinaisons homologues entre IS adjacentes lorsque les bactéries sont exposées à des températures de croissance élevées (37°C pour *Shigella*, 25°C pour *A. salmonicida* ssp. *salmonicida*) (Pilla *et al.*, 2017; Tanaka *et al.*, 2012). Une plus faible perte du SSTT de pINV est aussi observable à des températures de croissance plus basses, ce qui pourrait aussi être envisageable pour pAsa5 (Pilla *et al.*, 2017). Finalement, deux autres événements retrouvés chez pINV attirent l'attention. Premièrement, ce plasmide peut s'intégrer dans le chromosome de *Shigella*. Dans un tel cas les gènes du SSTT sont sous-exprimés (Pilla *et al.*, 2017). À ce jour, un seul événement de fusion a été détecté pour pAsa5, c'est-à-dire la fusion des plasmides pAsa5 et pAsa9 dans 01-B526-R19 (Tanaka *et al.*, 2017). Cependant, la présence de telles fusions pourrait être explorée pour pAsa5, dans la mesure où un test rapide à effectuer sur de nombreux échantillons permettrait de les détecter. Deuxièmement, les patrons de recombinaison possibles chez

pINV sont dépendants de la position sur le plasmide de trois systèmes toxines-antitoxines, ce qui restreint le nombre d'évènements possibles (Pilla *et al.*, 2017). C'est peut-être une telle situation qui limite aussi le nombre de patrons de recombinaison chez pAsa5, et ce, malgré la présence de plusieurs copies d'IS identiques sur ce plasmide (Figure 4.2).

4.3.1.1 TROUVER LA RECOMBINASE RESPONSABLE DE LA DÉLÉTION

La fréquence de recombinaison entre les IS de pAsa5 est très variable entre les souches (problème partiellement résolu avec la découverte d'ISAS5Z) et entre différentes conditions de culture (Daher, 2010; Stuber *et al.*, 2003). Conséquemment, la production de nouvelles souches réarrangées s'en trouve compliquée par le manque de contrôle sur l'ensemble des conditions qui favorisent les réarrangements (V. Paquet et G. Byatt, comm. personnelle). Une meilleure connaissance de l'enzyme catalysant la recombinaison homologue des IS chez *A. salmonicida* ssp. *salmonicida* permettrait d'approfondir les conditions propices à la délétion du SSTT.

Des protéines responsables de la réparation de l'ADN comme RecA et RecBCD permettent aussi des recombinaisons homologues chez les bactéries (Redfield, 2001). Chez *A. salmonicida* ssp. *salmonicida*, le gène *recA* a déjà été cloné et caractérisé (Umelo *et al.*, 1996). Il code pour une protéine qui complète le mutant *recA* chez *E. coli* (Umelo *et al.*, 1996). Le gène *recA* d'*A. salmonicida* ssp. *salmonicida* possède une séquence régulatrice putative de type SOS (Umelo *et al.*, 1996).

Le gène *recA*, ainsi que les autres gènes de réparation de l'ADN du génome d'*A. salmonicida* ssp. *salmonicida* mériteraient une annotation *in silico* détaillée pour les génomes séquencés. Une analyse comparative permettrait de mettre en évidence des différences en gènes ou en séquences régulatrices qui pourraient ensuite être recoupées avec la faculté à produire des réarrangements dans différentes souches.

4.3.1.2 LE STRESS, DÉCLENCHÉUR DE LA PERTE DU SSTT ?

Le stress, sous forme de température de croissance élevée, est un dénominateur commun pour deux catégories d'évènements impliquant des IS : les inactivations par transposition

ISAS1/ISAS2 dans les gènes *vapA* et *abcA*, et les délétions du SSTT par réarrangement (Daher, 2010; Gustafson *et al.*, 1994). Le stress peut à la fois augmenter la fréquence de recombinaison et la transposition/mobilisation de certains éléments mobiles de l'ADN (Foster, 2007). Une des suites du projet sur la perte du SSTT pourrait donc être de tester l'impact de différents stress sur les fréquences de recombinaison des IS.

Quelques essais ont déjà été tentés sur le sujet. Ainsi, certains composés d'huiles essentielles, dont la cinnamaldéhyde et le carvacrol, se sont montrés inefficaces pour produire des réarrangements de pAsa5 (D. Drouin et M. Trudel, comm. personnelle). Les huiles essentielles sont des solutions de métabolites secondaires extraites de plantes. Liposolubles, elles peuvent traverser les membranes, perturber leur intégrité et leur fonction et amener des déséquilibres physico-chimiques à l'intérieur des cellules (Bakkali *et al.*, 2008; Radulovic *et al.*, 2013).

Une avenue intéressante serait de profiter du stress oxydatif et de la pauvreté en nutriments apportée par des conditions de croissance stationnaire, particulièrement en milieu liquide (G. Byatt et V. Paquet, comm. personnelle) (Foster, 2007). Ces conditions entraînent aussi la lyse des cellules chez *A. salmonicida* ssp. *salmonicida*. C'est aussi à ce moment que s'observe l'excision de l'îlot génomique *AsaGEIIa* (résultats non publiés).

4.3.2 CONSÉQUENCE D'UNE PERTE DU SSTT POUR *A. SALMONICIDA* SSP. *SALMONICIDA*

Le SSTT est un élément essentiel à la virulence d'*A. salmonicida* ssp. *salmonicida* (Burr *et al.*, 2005; Burr *et al.*, 2003a; Burr *et al.*, 2003b). Sa perte devrait donc placer la bactérie en désavantage lors du processus infectieux. Cependant, la position du locus du SSTT sur pAsa5 le rend vulnérable à plusieurs scénarios de délétion par recombinaison (Tanaka *et al.*, 2012; Tanaka *et al.*, 2016). De plus, la délétion du locus du SSTT a été retrouvée dans la majorité des cas de pAsa5 tronqués produits ou étudiés à ce jour (cependant, certaines études étaient biaisées vers la recherche de perte du SSTT) (Daher, 2010; Emond-Rheault *et al.*, 2015b; Long *et al.*, 2016; Stuber *et al.*, 2003).

Les loci du SSTT sont fréquemment retrouvés dans des régions instables des génomes bactériens, sur des plasmides et des îlots génomiques (Gophna *et al.*, 2003). De plus, la

production d'un système de sécrétion représente un fardeau métabolique significatif (Sturm *et al.*, 2011). Sa délétion pourrait donc être favorisée dans certaines conditions (Gophna *et al.*, 2003).

Des résultats obtenus initialement pour l'article de Tanaka *et al.* (2012), mais non publiés, montrent que les souches SSTT-négatives ont un avantage de croissance *in vitro* (S. Dallaire-Dufresne, comm. personnelle). De plus, les souches portant des pAsa5 recombinés ISAS5Z-A, dont la délétion est plus grande, ont aussi une croissance plus rapide qu'une des souches à recombinaison ISAS11B-C (Figure 4.3). Cette région représente un fardeau métabolique inutile aux conditions de croissance *in vitro*. Les souches d'*A. salmonicida ssp. salmonicida* cultivées en laboratoire sur de longues périodes, même en l'absence de stress, tendraient donc à perdre leur SSTT.

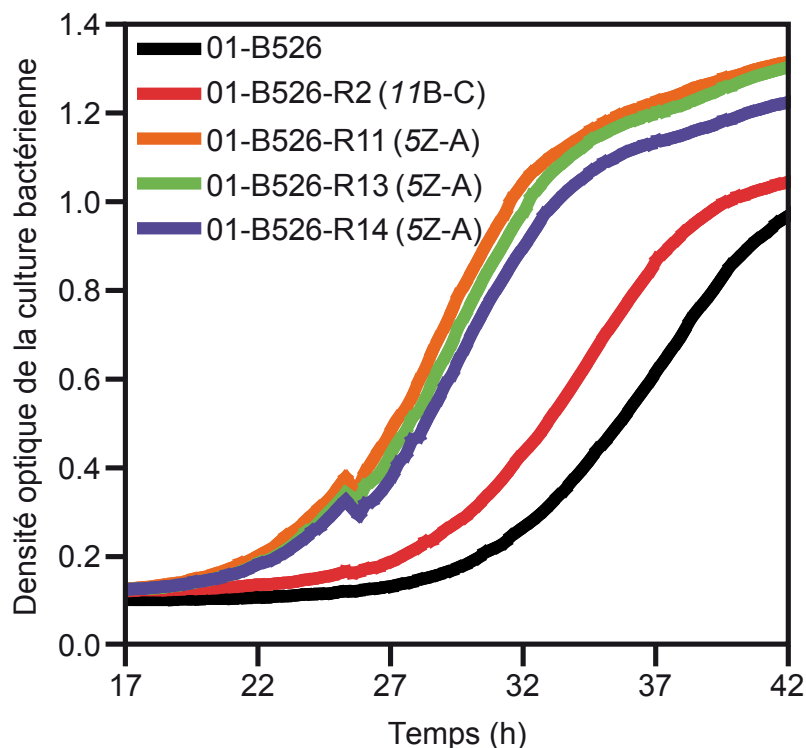


Figure 4.3 Courbes de croissance d'*A. salmonicida ssp. salmonicida* 01-B526 et de ses réarrangés. Les souches ont été inoculées à une densité optique à 595 nm de 0,1 en milieu *Lysogeny broth* (LB) et ont été incubées à 20°C avec une agitation de 200 rpm. La densité optique à 595 nm a été obtenue toutes les 15 minutes à l'aide de l'appareil Infinite 200 Pro (TECAN) (S. Dallaire-Dufresne, comm. personnelle). Les souches 01-B526-R2, -R11, -R13 et -R14 sont SSTT-négatives et ont été obtenues à partir d'une culture à 25°C de 01-B526 (Daher, 2010; Tanaka *et al.*, 2012). Le pAsa5 de 01-B526-R2 a subi une

recombinaison ISAS11B-ISAS11C, ce qui mène à la délétion du locus du SSTT (Tanaka *et al.*, 2012). Les pAsa5 des autres souches réarrangées ont subi une recombinaison ISAS5Z-ISAS5A, ce qui mène à la délétion du locus du SSTT ainsi qu'à celle de certains gènes en amont (Tanaka *et al.*, 2017).

Cependant, les conditions dans lesquelles la perte du SSTT a été observée, *in vitro*, sont différentes de la réalité *in vivo*. Conséquemment, il est difficile de prédire si la perte du SSTT est fréquente dans les populations naturelles, ou si elle apporte une contribution lors du cycle infectieux de la bactérie. Même si nous savons que le SSTT joue un rôle essentiel dans la furonculose, les études en ont surtout suivi la finalité (Burr *et al.*, 2005; Burr *et al.*, 2003a; Burr *et al.*, 2003b). Le SSTT pourrait servir à pénétrer la barrière épithéliale, puis à l'évasion ou à l'effondrement de la réponse immunitaire (Bergh *et al.*, 2013; Dallaire-Dufresne *et al.*, 2013; Origi *et al.*, 2017). Dans un tel modèle, les stages subséquents de l'infection seraient pris en charge par les autres facteurs de virulence d'*A. salmonicida* ssp. *salmonicida* (Vanden Bergh et Frey, 2014).

L'hypothèse de la population mixte pourrait concilier l'importance du SSTT pour *A. salmonicida* ssp. *salmonicida* et le mécanisme de perte observé dans les études précédentes. Dans un tel modèle, des sous-populations SSTT-positives et SSTT-négatives coexisteraient lors de l'infection. Une telle population hétérogène a déjà été observée chez le phytopathogène *Pseudomonas syringae* pv. *phaseolicola* (Arnold *et al.*, 2007; Lovell *et al.*, 2011; Pitman *et al.*, 2005). Dans ce cas précis, l'injection d'un effecteur de son SSTT entraîne une réponse hypersensible de la plante, qui relâche des espèces réactives de l'oxygène. Ce stress induit l'excision de l'îlot génomique codant pour le SSTT dans une partie de la population de *P. syringae*, ce qui mène à une délétion permanente du SSTT. Les deux sous-populations croissent à une vitesse réduite sous l'effet du stress oxydatif, mais lorsque la population SSTT-négative atteint des cellules qui n'ont pas déclenché la réponse hypersensible, elle prolifère (Arnold *et al.*, 2007). Chez *A. salmonicida* ssp. *salmonicida*, la présence d'une sous-population SSTT-négatives dans les conditions de culture normale (à 18°C) de souches SSTT-positives a déjà été indirectement observée par PCR (Tanaka *et al.*, 2012).

Il est aussi possible qu'*A. salmonicida* ssp. *salmonicida* perde son SSTT lors de l'établissement de l'infection chronique. La perte de facteurs de virulence sur de courtes

périodes évolutives survient, entre autres, dans les infections respiratoires chroniques à *Pseudomonas aeruginosa* chez les patients atteints de fibrose kystique (Smith *et al.*, 2006). Cependant, la présence du locus du SSTT dans presque toutes les souches d'*A. salmonicida* ssp. *salmonicida* issues de poissons malades et disponibles dans notre collection vient infirmer cette hypothèse. Toutefois, notre collection est biaisée pour des souches québécoises, infectieuses, et possiblement pour des cas de furonculose aigus. Il nous manque des métadonnées sur les conditions de collecte des isolats de notre collection pour inférer toutes ces hypothèses.

4.3.3 LIENS ÉVOLUTIFS ENTRE pAsa5, pAsa9, ET L'ORIGINE DU LOCUS DU SSTT

Dans l'article présenté au Chapitre 3, nous émettons l'hypothèse que pAsa9 a une architecture similaire au plasmide ancestral qui a mené à pAsa5 et pAsa9 puisque sa structure est plus condensée (Tanaka *et al.*, 2017). Dans le modèle proposé, le locus du SSTT d'*A. salmonicida* ssp. *salmonicida* ainsi que d'autres segments apparentés (les régions des effecteurs *aopH* et *aopO*, par exemple) auraient été acquis par transfert horizontal et se seraient greffés à l'architecture du plasmide par l'intervention des IS proximales à ces régions (Tanaka *et al.*, 2017). Cependant, cette hypothèse est en contradiction avec la distribution de pAsa5 puisque bien que ce plasmide a été retrouvé dans plusieurs régions géographique (Daher, 2010; Long *et al.*, 2016; Reith *et al.*, 2008; Stuber *et al.*, 2003), alors que pAsa9 est pour le moment exclusif au Québec et aux souches possédant un îlot génomique *AsaGEIIa* (Tanaka *et al.*, 2017). Ces deux éléments sont peut-être interdépendants. En effet, il est possible pour un plasmide et un îlot génomique de partager des fonctions relatives à leur mobilisation (Carraro *et al.*, 2017; Osorio *et al.*, 2008). Sinon, les souches pAsa9 positives sont peut-être des descendantes rapprochées d'un même ancêtre qui possédait aussi cette version d'*AsaGEI*.

Chapitre 5 CONCLUSION

5.1 LE PASSÉ, PRÉSENT ET FUTUR D'*A. SALMONICIDA* SSP. *SALMONICIDA* ET DE SES IS

Cette thèse a présenté quelques exemples de variations génotypiques et phénotypiques impliquant des IS, des transposons et des plasmides. Au moins une de ces variations, l'intégration du gène *catA1* dans le plasmide pAsa4, est directement responsable d'une augmentation de la valeur adaptative dans une condition donnée. Dans la discussion, nous avons aussi inféré sur les implications des variations architecturales des plasmides pAsa4 et pAsa5. Il en ressort que même si l'effet des éléments mobiles de l'ADN sur *A. salmonicida* ssp. *salmonicida*, particulièrement *in vivo*, sont méconnus, leur impact sur l'architecture génomique est certain et se précise à chaque étude.

La découverte de plusieurs variants de plasmides impliquant un nombre limité d'IS et les résultats préliminaires présentés à la discussion suggèrent que toutes les IS du génome d'*A. salmonicida* ssp. *salmonicida* ne sont pas prompte à la transposition de la même façon. Curieusement, les IS présentes en haut nombre de copies, comme ISAS4, ne sont pas nécessairement les plus actives. D'un autre côté, certaines familles plus « transitoires » et résidant principalement sur les plasmides pourraient avoir un plus haut taux de transposition. Il est possible que les IS « résidentes » du génome *A. salmonicida* ssp. *salmonicida* descendent d'une cohabitation plus ancienne et que leur activité réduite reflète un état d'équilibre entre la bactérie et l'élément mobile.

Comment cette interaction IS-bactérie peut-elle évoluer? *A. salmonicida* ssp. *salmonicida* répond à certains critères qui la place au début d'un processus de réduction de génome (Moran et Plague, 2004; Siguier *et al.*, 2014). C'est un agent pathogène spécialisé par rapport à d'autres bactéries de son genre, et son génome possède beaucoup d'éléments mobiles et beaucoup de pseudogènes (Reith *et al.*, 2008). On suppose aussi que ses IS ont déjà commencé à contraindre son mode de vie en la restreignant à la psychrophilie, contrairement aux *A. salmonicida* mésophiles et à d'autres espèces du genre *Aeromonas* (Vincent *et al.*, 2016b). Le futur de la cohabitation IS-*A. salmonicida* ssp. *salmonicida* implique peut-être plus de grands réarrangements médiés par les IS afin d'éliminer des

gènes inactivés ou accessoires. La recombinaison des IS sur pAsa5 et la perte des gènes du SSTT pourraient être une de ces étapes et devenir fixée dans une certaine partie de la population. Si *A. salmonicida* ssp. *salmonicida* est bien dans les premières étapes d'une réduction de génome, son association avec les Salmonidés pourrait devenir encore plus étroite. En effet, l'implication des IS dans la réduction de génomes a été observée pour quelques agents pathogènes spécialisés, comme *Bordetella pertussis*, *Yersinia pestis* et *Mycobacterium ulcerans*, ainsi que des endosymbiontes, comme *Wolbachia* sp. (Cerveau *et al.*, 2011; Chain *et al.*, 2004; Parkhill *et al.*, 2003a; Röltgen *et al.*, 2012).

BIBLIOGRAPHIE

- Abby, S.S., Cury, J., Guglielmini, J., Neron, B., Touchon, M., Rocha, E.P., 2016. Identification of protein secretion systems in bacterial genomes. *Scientific reports* 6, 23080.
- Adams, A., Bundy, A., Thompson, K., Horne, M.T., 1988. The association between virulence and cell-surface characteristics of *Aeromonas salmonicida*. *Aquaculture* 69, 1-14.
- Adams, C.A., Austin, B., Meaden, P.G., McIntosh, D., 1998. Molecular characterization of plasmid-mediated oxytetracycline resistance in *Aeromonas salmonicida*. *Appl. Environ. Microbiol.* 64, 4194-4201.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- Andersson, D.I., Hughes, D., 2014. Microbiological effects of sublethal levels of antibiotics. *Nature reviews. Microbiology* 12, 465-478.
- Aoki, T., Egusa, S., Kimura, T., Watanabe, T., 1971. Detection of R factors in naturally occurring *Aeromonas salmonicida* strains. *Appl. Microbiol.* 22, 716-717.
- Aoki, T., Holland, B.I., 1985. The outer-membrane proteins of the fish pathogens *Aeromonas hydrophila*, *Aeromonas salmonicida* and *Edwardsiella tarda*. *FEMS Microbiol. Lett.* 27, 299-305.
- Arnesen, J.A., Eggset, G., 1999. Isolation and characterisation of two extracellular metalloproteases from *Aeromonas salmonicida* ssp. *salmonicida*. *J. Fish Dis.* 22, 35-43.
- Arnesen, J.A., Eggset, G., Jorgensen, T.O., 1995. Partial-purification and characterization of extracellular metalloproteases from *Aeromonas salmonicida* ssp. *salmonicida*. *J. Fish Dis.* 18, 283-295.
- Arnold, D.L., Jackson, R.W., Waterfield, N.R., Mansfield, J.W., 2007. Evolution of microbial virulence: the benefits of stress. *Trends Genet.* 23, 293-300.
- Att  r  , S.A., Vincent, A.T., Trudel, M.V., Chanut, R., Charette, S.J., 2015. Diversity and Homogeneity among Small Plasmids of *Aeromonas salmonicida* subsp. *salmonicida* Linked with Geographical Origin. *Frontiers in microbiology* 6, 1274.
- Aziz, R.K., Breitbart, M., Edwards, R.A., 2010. Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res* 38, 4207-4217.
- Bahl, M.I., Hansen, L.H., S  rensen, S.J., 2007. Impact of conjugal transfer on the stability of IncP-1 plasmid pKJK5 in bacterial populations. *FEMS Microbiol. Lett.* 266, 250-256.
- Bakkali, F., Averbeck, S., Averbeck, D., Waomar, M., 2008. Biological effects of essential oils - A review. *Food Chem. Toxicol.* 46, 446-475.
- Baltrus, D.A., 2013. Exploring the costs of horizontal gene transfer. *Trends in ecology & evolution* 28, 489-495.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455-477.

- Barnes, A.C., Bowden, T.J., Horne, M.T., Ellis, A.E., 1999. Peroxide-inducible catalase in *Aeromonas salmonicida* subsp. *salmonicida* protects against exogenous hydrogen peroxide and killing by activated rainbow trout, *Oncorhynchus mykiss* L., macrophages. *Microb. Pathog.* 26, 149-158.
- Barnes, A.C., Horne, M.T., Ellis, A.E., 1996. Effect of iron on expression of superoxide dismutase by *Aeromonas salmonicida* and associated resistance to superoxide anion. *FEMS Microbiol. Lett.* 142, 19-26.
- Barton, B.A., Iwama, G.K., 1991. Physiological changes in fish from stress in aquaculture with emphasis on the response and effects of corticosteroids. *Annual Review of Fish Disease* 1, 3-26.
- Bast, L., Daly, J.G., Degrandis, S.A., Stevenson, R.M.W., 1988. Evaluation of profiles of *Aeromonas salmonicida* as epidemiological markers of furunculosis infections in fish. *J. Fish Dis.* 11, 133-145.
- Bayliss, S.C., Verner-Jeffreys, D.W., Bartie, K.L., Aanensen, D.M., Sheppard, S.K., Adams, A., Feil, E.J., 2017. The Promise of Whole Genome Pathogen Sequencing for the Molecular Epidemiology of Emerging Aquaculture Pathogens. *Frontiers in microbiology* 8, 121.
- Beatson, S.A., Walker, M.J., 2014. Tracking antibiotic resistance. *Science* 345, 1454-1455.
- Beaz-Hidalgo, R., Hossain, M.J., Liles, M.R., Figueras, M.J., 2015a. Strategies to avoid wrongly labelled genomes using as example the detected wrong taxonomic affiliation for *Aeromonas* genomes in the GenBank database. *PloS one* 10, e0115813.
- Beaz-Hidalgo, R., Latif-Eugenín, F., Hossain, M.J., Berg, K., Niemi, R.M., Rapala, J., Lyra, C., Liles, M.R., Figueras, M.J., 2015b. *Aeromonas aquatica* sp. nov., *Aeromonas finlandiensis* sp. nov. and *Aeromonas lacus* sp. nov. isolated from Finnish waters associated with cyanobacterial blooms. *Syst. Appl. Microbiol.* 38, 161-168.
- Belland, R.J., Trust, T.J., 1985. Synthesis, export, and assembly of *Aeromonas salmonicida* A-layer analyzed by transposon mutagenesis. *J. Bacteriol.* 163, 877-881.
- Belland, R.J., Trust, T.J., 1987. Cloning of the gene for the surface array protein of *Aeromonas salmonicida* and evidence linking loss of expression with genetic deletion. *J. Bacteriol.* 169, 4086-4091.
- Belland, R.J., Trust, T.J., 1989. *Aeromonas salmonicida* plasmids: plasmid-directed synthesis of proteins in vitro and in *Escherichia coli* minicells. *J. Gen. Microbiol.* 135, 513-524.
- Bellanger, X., Morel, C., Gonot, F., Puymege, A., Decaris, B., Guedon, G., 2011. Site-specific accretion of an integrative conjugative element together with a related genomic island leads to cis mobilization and gene capture. *Mol. Microbiol.* 81, 912-925.
- Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J.M., Bryant, J., Carter, R.J., Keira Cheetham, R., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Humphray, S.J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Liu, X., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Parkinson, M.L., Pratt, M.R., Rasolonjatovo, I.M., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S.V., Scally, A., Schroth, G.P., Smith, M.E., Smith, V.P., Spiridou, A., Torrance, P.E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Alam,

- M.D., Anastasi, C., Aniebo, I.C., Bailey, D.M., Bancarz, I.R., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Benson, K.F., Bevis, C., Black, P.J., Boodhun, A., Brennan, J.S., Bridgham, J.A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Burrows, J.C., Carter, N.P., Castillo, N., Chiara, E.C.M., Chang, S., Neil Cooley, R., Crake, N.R., Dada, O.O., Diakoumakos, K.D., Dominguez-Fernandez, B., Earnshaw, D.J., Egbujor, U.C., Elmore, D.W., Etchin, S.S., Ewan, M.R., Fedurco, M., Fraser, L.J., Fuentes Fajardo, K.V., Scott Furey, W., George, D., Gietzen, K.J., Goddard, C.P., Golda, G.S., Granieri, P.A., Green, D.E., Gustafson, D.L., Hansen, N.F., Harnish, K., Haudenschild, C.D., Heyer, N.I., Hims, M.M., Ho, J.T., Horgan, A.M., Hoschler, K., Hurwitz, S., Ivanov, D.V., Johnson, M.Q., James, T., Huw Jones, T.A., Kang, G.D., Kerelska, T.H., Kersey, A.D., Khrebtukova, I., Kindwall, A.P., Kingsbury, Z., Kokko-Gonzales, P.I., Kumar, A., Laurent, M.A., Lawley, C.T., Lee, S.E., Lee, X., Liao, A.K., Loch, J.A., Lok, M., Luo, S., Mammen, R.M., Martin, J.W., McCauley, P.G., McNitt, P., Mehta, P., Moon, K.W., Mullens, J.W., Newington, T., Ning, Z., Ling Ng, B., Novo, S.M., O'Neill, M.J., Osborne, M.A., Osnowski, A., Ostadan, O., Paraschos, L.L., Pickering, L., Pike, A.C., Pike, A.C., Chris Pinkard, D., Pliskin, D.P., Podhasky, J., Quijano, V.J., Raczy, C., Rae, V.H., Rawlings, S.R., Chiva Rodriguez, A., Roe, P.M., Rogers, J., Rogert Bacigalupo, M.C., Romanov, N., Romieu, A., Roth, R.K., Rourke, N.J., Ruediger, S.T., Rusman, E., Sanches-Kuiper, R.M., Schenker, M.R., Seoane, J.M., Shaw, R.J., Shiver, M.K., Short, S.W., Sizto, N.L., Sluis, J.P., Smith, M.A., Ernest Sohna Sohna, J., Spence, E.J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C.L., Turcatti, G., Vandevondele, S., Verhovsky, Y., Virk, S.M., Wakelin, S., Walcott, G.C., Wang, J., Worsley, G.J., Yan, J., Yau, L., Zuerlein, M., Rogers, J., Mullikin, J.C., Hurler, M.E., McCooke, N.J., West, J.S., Oaks, F.L., Lundberg, P.L., Klenerman, D., Durbin, R., Smith, A.J., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53-59.
- Bentley, S.D., Parkhill, J., 2015. Genomic perspectives on the evolution and spread of bacterial pathogens. *Proceedings. Biological sciences* 282, 20150488.
- Bergh, P.V., Burr, S.E., Benedicenti, O., von Siebenthal, B., Frey, J., Wahli, T., 2013. Antigens of the type-three secretion system of *Aeromonas salmonicida* subsp. *salmonicida* prevent protective immunity in rainbow trout. *Vaccine* 31, 5256-5261.
- Bergstrom, C.T., Lipsitch, M., Levin, B.R., 2000. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics* 155, 1505-1519.
- Bernales, I., Mendiola, M.V., de la Cruz, F., 1999. Intramolecular transposition of insertion sequence IS91 results in second-site simple insertions. *Mol. Microbiol.* 33, 223-234.
- Bernoth, E.M., 1990. Autoagglutination, growth on tryptone-soy-coomassie-agar, outer-membrane protein-patterns and virulence of *Aeromonas salmonicida* strains. *J. Fish Dis.* 13, 145-155.
- Bonet, R., Simonpujol, M.D., Congregado, F., 1993. Effects of nutrients on exopolysaccharide production and surface-properties of *Aeromonas salmonicida*. *Appl. Environ. Microbiol.* 59, 2437-2441.
- 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.

- Boyd, J.M., Dacanay, A., Knickle, L.C., Touhami, A., Brown, L.L., Jericho, M.H., Johnson, S.C., Reith, M., 2008. Contribution of type IV pili to the virulence of *Aeromonas salmonicida* subsp. *salmonicida* in Atlantic salmon (*Salmo salar* L.). *Infect. Immun.* 76, 1445-1455.
- Bradley, D.E., Aoki, T., Kitao, T., Arai, T., Tschäpe, H., 1982. Specification of characteristics for the classification of plasmids in incompatibility group U. *Plasmid* 8, 89-93.
- Braun, M., Stuber, K., Schlatter, Y., Wahli, T., Kuhnert, P., Frey, J., 2002. Characterization of an ADP-ribosyltransferase toxin (AexT) from *Aeromonas salmonicida* subsp. *salmonicida*. *J. Bacteriol.* 184, 1851-1858.
- Breyer, K.E., Getchell, R.G., Cornwell, E.R., Wooster, G.A., Ketola, H.G., Bowser, P.R., 2015. Efficacy of an Extract from Garlic, *Allium sativum*, Against Infection with the Furunculosis Bacterium, *Aeromonas salmonicida*, in Rainbow Trout, *Oncorhynchus mykiss*. *Journal of the World Aquaculture Society* 46, 273-282.
- Broberg, C.A., Zhang, L., Gonzalez, H., Laskowski-Arce, M.A., Orth, K., 2010. A *Vibrio* effector protein is an inositol phosphatase and disrupts host cell membrane integrity. *Science* 329, 1660-1662.
- Brogden, K.A., Roth, J.A., Stanton, T.B., Bolin, C.A., Minion, C.F., Wannemuehler, M.J., 2000. Virulence mechanisms of bacterial pathogens.
- Buermans, H.P., den Dunnen, J.T., 2014. Next generation sequencing technology: Advances and applications. *Biochim. Biophys. Acta* 1842, 1932-1941.
- Burr, S.E., Frey, J., 2007. Analysis of type III effector genes in typical and atypical *Aeromonas salmonicida*. *J. Fish Dis.* 30, 711-714.
- Burr, S.E., Pugovkin, D., Wahli, T., Segner, H., Frey, J., 2005. Attenuated virulence of an *Aeromonas salmonicida* subsp. *salmonicida* type III secretion mutant in a rainbow trout model. *Microbiology* 151, 2111-2118.
- Burr, S.E., Stuber, K., Frey, J., 2003a. The ADP-ribosylating toxin, AexT, from *Aeromonas salmonicida* subsp. *salmonicida* is translocated via a type III secretion pathway. *J. Bacteriol.* 185, 6583-6591.
- Burr, S.E., Stuber, K., Wahli, T., Frey, J., 2002. Evidence for a type III secretion system in *Aeromonas salmonicida* subsp. *salmonicida*. *J. Bacteriol.* 184, 5966-5970.
- Burr, S.E., Wahli, T., Segner, H., Pugovkin, D., Frey, J., 2003b. Association of Type III secretion genes with virulence of *Aeromonas salmonicida* subsp. *salmonicida*. *Dis. Aquat. Organ.* 57, 167-171.
- Buttner, D., 2012. Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 76, 262-310.
- Cabello, F.C., Godfrey, H.P., Buschmann, A.H., Dölz, H.J., Aquaculture as yet another environmental gateway to the development and globalisation of antimicrobial resistance. *The Lancet Infectious Diseases* 16, e127-e133.
- Cambray, G., Guerout, A.M., Mazel, D., 2010. Integrins. *Annu. Rev. Genet.* 44, 141-166.
- Carraro, N., Rivard, N., Burrus, V., Ceccarelli, D., 2017. Mobilizable genomic islands, different strategies for the dissemination of multidrug resistance and other adaptive traits. *Mobile genetic elements* 7, 1-6.
- Carver, T., Harris, S.R., 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, T., Thomson, N., Bleasby, A., Berriman, M., Parkhill, J., 2009. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 25, 119-120.
- Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., Parkhill, J., 2005. ACT: the Artemis Comparison Tool. *Bioinformatics* 21, 3422-3423.
- Casetta, A., Hoi, A.B., de Cespedes, G., Horaud, T., 1998. Diversity of structures carrying the high-level gentamicin resistance gene (*aac6-aph2*) in *Enterococcus faecalis* strains isolated in France. *Antimicrob. Agents Chemother.* 42, 2889-2892.
- Cerveau, N., Leclercq, S., Leroy, E., Bouchon, D., Cordaux, R., 2011. Short- and long-term evolutionary dynamics of bacterial insertion sequences: insights from *Wolbachia* endosymbionts. *Genome biology and evolution* 3, 1175-1186.
- Chai, B., Wang, H., Chen, X., 2012. Draft genome sequence of high-melanin-yielding *Aeromonas media* strain WS. *J. Bacteriol.* 194, 6693-6694.
- Chain, P.S., Carniel, E., Larimer, F.W., Lamerdin, J., Stoutland, P.O., Regala, W.M., Georgescu, A.M., Vergez, L.M., Land, M.L., Motin, V.L., Brubaker, R.R., Fowler, J., Hinnebusch, J., Marceau, M., Medigue, C., Simonet, M., Chenal-Francisque, V., Souza, B., Dacheux, D., Elliott, J.M., Derbise, A., Hauser, L.J., Garcia, E., 2004. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13826-13831.
- Charette, S.J., Brochu, F., Boyle, B., Filion, G., Tanaka, K.H., Derome, N., 2012. Draft genome sequence of the virulent strain 01-B526 of the fish pathogen *Aeromonas salmonicida*. *J. Bacteriol.* 194, 722-723.
- Chart, H., Shaw, D.H., Ishiguro, E.E., Trust, T.J., 1984. Structural and immunochemical homogeneity of *Aeromonas salmonicida* lipopolysaccharide. *J. Bacteriol.* 158, 16-22.
- Chart, H., Trust, T.J., 1983. Acquisition of iron by *Aeromonas salmonicida*. *J. Bacteriol.* 156, 758-764.
- Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., Turner, S.W., Korlach, J., 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature methods* 10, 563-569.
- Chu, S., Cavaignac, S., Feutrier, J., Phipps, B.M., Kostrzynska, M., Kay, W.W., Trust, T.J., 1991a. Structure of the tetragonal surface virulence array protein and gene of *Aeromonas salmonicida*. *J. Biol. Chem.* 266, 15258-15265.
- Chu, S., Cavaignac, S., Feutrier, J., Phipps, B.M., Kostrzynska, M., Kay, W.W., Trust, T.J., 1991b. Structure of the tetragonal surface virulence array protein and gene of *Aeromonas salmonicida*. *J. Biol. Chem.* 266, 15258-15265.
- Chu, S.J., Noonan, B., Cavaignac, S., Trust, T.J., 1995. Endogenous mutagenesis by an insertion-sequence element identifies *Aeromonas salmonicida* ABCA as an ATP-binding cassette transport protein required for biogenesis of smooth lipopolysaccharide. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5754-5758.
- Cipriano, R.C., Austin, B., 2011. Furunculosis and Other Aeromonad Diseases. *Fish Diseases and Disorders, Vol 3: Viral, Bacterial and Fungal Infections, 2nd Edition*, 424-483.
- Clark, I.C., Melnyk, R.A., Iavarone, A.T., Novichkov, P.S., Coates, J.D., 2014. Chlorate reduction in *Shewanella algae* ACDC is a recently acquired metabolism

- characterized by gene loss, suboptimal regulation and oxidative stress. *Mol. Microbiol.* 94, 107-125.
- Coil, D., Jospin, G., Darling, A.E., 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* 31, 587-589.
- Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R.M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.A., Rutherford, K.M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitehead, S., Woodward, J.R., Barrell, B.G., 2001. Massive gene decay in the leprosy bacillus. *Nature* 409, 1007-1011.
- Colston, S.M., Fullmer, M.S., Beka, L., Lamy, B., Gogarten, J.P., Graf, J., 2014. Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. *mBio* 5, e02136.
- Dacanay, A., Boyd, J.M., Fast, M.D., Knickle, L.C., Reith, M.E., 2010. *Aeromonas salmonicida* Type I pilus system contributes to host colonization but not invasion. *Dis. Aquat. Organ.* 88, 199-206.
- Dacanay, A., Johnson, S.C., Bjornsdottir, R., Ebanks, R.O., Ross, N.W., Reith, M., Singh, R.K., Hiu, J., Brown, L.L., 2003. Molecular characterization and quantitative analysis of superoxide dismutases in virulent and avirulent strains of *Aeromonas salmonicida* subsp. *salmonicida*. *J. Bacteriol.* 185, 4336-4344.
- Dacanay, A., Knickle, L., Solanky, K.S., Boyd, J.M., Walter, J.A., Brown, L.L., Johnson, S.C., Reith, M., 2006. Contribution of the type III secretion system (TTSS) to virulence of *Aeromonas salmonicida* subsp. *salmonicida*. *Microbiology* 152, 1847-1856.
- Daher, R., 2010. Heat-induced rearrangement of the pAsa5 plasmid of *Aeromonas salmonicida* is associated with loss of virulence. *Mémoire (M.Sc.)*. Université Laval, Québec.
- Daher, R.K., Fillion, G., Tan, S.G., Dallaire-Dufresne, S., Paquet, V.E., Charette, S.J., 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.
- Dallaire-Dufresne, S., Barbeau, X., Sarty, D., Tanaka, K.H., Denoncourt, A.M., Lagüe, P., Reith, M.E., Charette, S.J., 2013. *Aeromonas salmonicida* Ati2 is an effector protein of the type three secretion system. *Microbiology* 159, 1937-1945.
- Dallaire-Dufresne, S., Emond-Rheault, J.G., Attéré, S.A., Tanaka, K.H., Trudel, M.V., Frenette, M., Charette, S.J., 2014a. Optimization of a plasmid electroporation protocol for *Aeromonas salmonicida* subsp. *salmonicida*. *J. Microbiol. Methods* 98, 44-49.
- Dallaire-Dufresne, S., Tanaka, K.H., Trudel, M.V., Lafaille, A., Charette, S.J., 2014b. Virulence, genomic features, and plasticity of *Aeromonas salmonicida* subsp. *salmonicida*, the causative agent of fish furunculosis. *Vet. Microbiol.* 169, 1-7.
- Darmon, E., Leach, D.R., 2014. Bacterial genome instability. *Microbiol. Mol. Biol. Rev.* 78, 1-39.

- Dautremepuits, C., Fortier, M., Croisetièrre, 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.
- De Gelder, L., Ponciano, J.M., Joyce, P., Top, E.M., 2007. Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology* 153, 452-463.
- Dean, P., 2011. Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiol. Rev.* 35, 1100-1125.
- Deonier, R.C., Hadley, R.G., 1980. IS2-IS2 and IS3-IS3 Relative Recombination Frequencies in F-Integration. *Plasmid* 3, 48-64.
- Derome, N., Boutin, S., Llewellyn, M., Gauthier, J. 2016. Opportunistic pathogens of fish, In: Hurst, C.J. (Ed.) *The Rasputin effect: when commensals and symbionts become parasitic*. Springer, New York, NY, 649.
- Dobrindt, U., Hochhut, B., Hentschel, U., Hacker, J., 2004. Genomic islands in pathogenic and environmental microorganisms. *Nature reviews. Microbiology* 2, 414-424.
- Domingues, S., da Silva, G.J., Nielsen, K.M., 2012. Integrons: Vehicles and pathways for horizontal dissemination in bacteria. *Mobile genetic elements* 2, 211-223.
- Domingues, S., da Silva, G.J., Nielsen, K.M., 2015. Global dissemination patterns of common gene cassette arrays in class 1 integrons. *Microbiology* 161, 1313-1337.
- Donlon, J., McGettigan, S., O'Brien, P., Carra, P.O., 1983. Reappraisal of the Nature of the Pigment Produced by *Aeromonas salmonicida*. *FEMS Microbiol. Lett.* 19, 285-290.
- Dooley, J.S.G., Engelhardt, H., Baumeister, W., Kay, W.W., Trust, T.J., 1989. 3-dimensional structure of an open form of the surface-layer from the fish pathogen *Aeromonas salmonicida*. *J. Bacteriol.* 171, 190-197.
- Doolittle, W.F., Kirkwood, T.B., Dempster, M.A., 1984. Selfish DNAs with self-restraint. *Nature* 307, 501-502.
- Doolittle, W.F., Sapienza, C., 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284, 601-603.
- Dziewit, L., Czarnecki, J., Wibberg, D., Radlinska, M., Mrozek, P., Szymczak, M., Schluter, A., Puhler, A., Bartosik, D., 2014. Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. *BMC genomics* 15.
- Ebanks, R.O., Knickle, L.C., Goguen, M., Boyd, J.M., Pinto, D.M., Reith, M., Ross, N.W., 2006. Expression of and secretion through the *Aeromonas salmonicida* type III secretion system. *Microbiology* 152, 1275-1286.
- Eggset, G., Bjornsdottir, R., Leifson, R.M., Arnesen, J.A., Coucheron, D.H., Jorgensen, T.O., 1994. Extracellular glycerophospholipid-cholesterol acyltransferase from *Aeromonas salmonicida* - activation by serine-protease. *J. Fish Dis.* 17, 17-29.
- Ellis, A.E., Burrows, A.S., Stapleton, K.J., 1988. Lack of relationship between virulence of *Aeromonas salmonicida* and the putative virulence factors: A-layer, extracellular proteases and extracellular hemolysins. *J. Fish Dis.* 11, 309-323.
- Ellis, A.E., doVale, A., Bowden, T.J., Thompson, K., Hastings, T.S., 1997. In vivo production of A-protein, lipopolysaccharide, iron-regulated outer membrane proteins and 70-kDa serine protease by *Aeromonas salmonicida* subsp. *salmonicida*. *FEMS Microbiol. Lett.* 149, 157-163.
- Emond-Rheault, J.G., Vincent, A.T., Trudel, M.V., Brochu, F., Boyle, B., Tanaka, K.H., Att  r  , S.A., Jubinville, E., Loch, T.P., Winters, A.D., Faisal, M., Frenette, M.,

- Derome, N., Charette, S.J., 2015a. Variants of a genomic island in *Aeromonas salmonicida* subsp. *salmonicida* link isolates with their geographical origins. *Vet. Microbiol.* 175, 68-76.
- Emond-Rheault, J.G., Vincent, A.T., Trudel, M.V., Frey, J., Frenette, M., Charette, S.J., 2015b. *AsaGEI2b*: a new variant of a genomic island identified in the *Aeromonas salmonicida* subsp. *salmonicida* JF3224 strain isolated from a wild fish in Switzerland. *FEMS Microbiol. Lett.* 362.
- Fagan, R.P., Fairweather, N.F., 2014. Biogenesis and functions of bacterial S-layers. *Nature reviews. Microbiology* 12, 211-222.
- FAO 2016. The State of World Fisheries and Aquaculture 2016. Contributing to food security and nutrition for all. (Rome), 200.
- Fehr, D., Casanova, C., Liverman, A., Blazkova, H., Orth, K., Dobbelaere, D., Frey, J., Burr, S.E., 2006. AopP, a type III effector protein of *Aeromonas salmonicida*, inhibits the NF-kappaB signalling pathway. *Microbiology* 152, 2809-2818.
- Fernandes, F., da Fonseca, P.G., Russo, L.M., Oliveira, A.L., Freitas, A.T., 2011. Efficient alignment of pyrosequencing reads for re-sequencing applications. *BMC bioinformatics* 12, 163.
- Fernandez, A.I.G., Fernández, A.F., Pérez, M.J., Nieto, T.P., Ellis, A.E., 1998. Siderophore production by *Aeromonas salmonicida* subsp. *salmonicida*. Lack of strain specificity. *Dis. Aquat. Organ.* 33, 87-92.
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocyne, J.D., Scott, J., Shirley, R., Liu, L., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O., Venter, J.C., 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496-512.
- Foster, P.L., 2007. Stress-induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.* 42, 373-397.
- Frey, J., Origgi, F.C., 2016. Type III Secretion System of *Aeromonas salmonicida* Undermining the Host's Immune Response. *Frontiers in Marine Science* 3.
- Fricke, W.F., Welch, T.J., McDermott, P.F., Mammel, M.K., LeClerc, J.E., White, D.G., Cebula, T.A., Ravel, J., 2009. Comparative genomics of the IncA/C multidrug resistance plasmid family. *J. Bacteriol.* 191, 4750-4757.
- Froquet, R., Cherix, N., Burr, S.E., Frey, J., Vilches, S., Tomas, J.M., Cosson, P., 2007. Alternative host model to evaluate *Aeromonas* virulence. *Appl. Environ. Microbiol.* 73, 5657-5659.
- Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: the agents of open source evolution. *Nature reviews. Microbiology* 3, 722-732.
- Futagami, T., Tsuboi, Y., Suyama, A., Goto, M., Furukawa, K., 2006. Emergence of two types of nondechlorinating variants in the tetrachloroethene-halo-respiring *Desulfitobacterium* sp. strain Y51. *Appl. Microbiol. Biotechnol.* 70, 720-728.
- Fyfe, L., Coleman, G., Munro, A.L.S., 1987. Identification of major common extracellular proteins secreted by *Aeromonas salmonicida* strains isolated from diseased fish. *Appl. Environ. Microbiol.* 53, 722-726.

- Galardini, M., Biondi, E.G., Bazzicalupo, M., Mengoni, A., 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* 6, 11.
- García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Gotz, S., Tarazona, S., Dopazo, J., Meyer, T.F., Conesa, A., 2012. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28, 2678-2679.
- Garcillán-Barcia, M.P., Alvarado, A., de la Cruz, F., 2011. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol. Rev.* 35, 936-956.
- Garcillán-Barcia, M.P., de la Cruz, F., 2002. Distribution of IS91 family insertion sequences in bacterial genomes: evolutionary implications. *FEMS microbiology ecology* 42, 303-313.
- Garduño, R., Moore, A., Olivier, G., Lizama, A.L., Garduño, E., Kay, W.W., 2000. Host cell invasion and intracellular residence by *Aeromonas salmonicida*: Role of the S-layer. *Can. J. Microbiol.* 46, 660-668.
- Garrote, A., Bonet, R., Merino, S., Simonpujol, M.D., Congregado, F., 1992. Occurrence of a capsule in *Aeromonas salmonicida*. *FEMS Microbiol. Lett.* 95, 127-132.
- Gestal, A.M., Liew, E.F., Coleman, N.V., 2011. Natural transformation with synthetic gene cassettes: new tools for integron research and biotechnology. *Microbiology* 157, 3349-3360.
- Gillings, M.R., Paulsen, I.T., Tetu, S.G., 2017. Genomics and the evolution of antibiotic resistance. *Ann. N. Y. Acad. Sci.* 1388, 92-107.
- Gophna, U., Ron, E.Z., Graur, D., 2003. Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. *Gene* 312, 151-163.
- Gordon, D., Green, P., 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29, 2936-2937.
- Gray, A.N., Koo, B.M., Shiver, A.L., Peters, J.M., Osadnik, H., Gross, C.A., 2015. High-throughput bacterial functional genomics in the sequencing era. *Curr. Opin. Microbiol.* 27, 86-95.
- Groberg, W.J., McCoy, R.H., Pilcher, K.S., Fryer, J.L., 1978. Relation of Water Temperature to Infections of Coho Salmon (*Oncorhynchus kisutch*), Chinook Salmon (*O. tshawytscha*), and Steelhead Trout (*Salmo gairdneri*) With *Aeromonas salmonicida* and *Aeromonas hydrophila*. *Journal of the Fisheries Research Board of Canada* 35, 1-7.
- Gueguen, E., Rousseau, P., Duval-Valentin, G., Chandler, M., 2005. The transpososome: control of transposition at the level of catalysis. *Trends Microbiol.* 13, 543-549.
- Guglielmini, J., Neron, B., Abby, S.S., Garcillan-Barcia, M.P., de la Cruz, F., Rocha, E.P., 2014. Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res* 42, 5715-5727.
- Gustafson, C.E., Chu, S.J., Trust, T.J., 1994. Mutagenesis of the paracrystalline surface protein array of *Aeromonas salmonicida* by endogenous insertion elements. *J. Mol. Biol.* 237, 452-463.
- Gustafson, C.E., Thomas, C.J., Trust, T.J., 1992. Detection of *Aeromonas salmonicida* from fish by using polymerase chain reaction amplification of the virulence surface array protein gene. *Appl. Environ. Microbiol.* 58, 3816-3825.
- Hall, J.P., Harrison, E., Lilley, A.K., Paterson, S., Spiers, A.J., Brockhurst, M.A., 2015. Environmentally co-occurring mercury resistance plasmids are genetically and

- phenotypically diverse and confer variable context-dependent fitness effects. *Environmental microbiology* 17, 5008-5022.
- Hall, R.M., Collis, C.M., 1995. Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. *Mol. Microbiol.* 15, 593-600.
- Hänninen, M.L., Hirvelä-Koski, V., 1997. Molecular and phenotypic methods for the characterization of atypical *Aeromonas salmonicida*. *Vet. Microbiol.* 56, 147-158.
- Hänninen, M.L., 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.
- Hao, W. 2012. Fast rates of evolution in bacteria due to horizontal gene transfer, In: Singh, R.S., Xu, J., Kulathinal, R.J. (Eds.) *Rapidly Evolving Genes and Genetic Systems*. Oxford University Press, Oxford, United Kingdom, 288.
- Harmer, C.J., Hall, R.M., 2015. The A to Z of A/C plasmids. *Plasmid* 80, 63-82.
- Harrison, E., Brockhurst, M.A., 2012. Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol.* 20, 262-267.
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 100-108.
- Hayes, F., 2003. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science* 301, 1496-1499.
- He, S., Chandler, M., Varani, A.M., Hickman, A.B., Dekker, J.P., Dyda, F., 2016. Mechanisms of Evolution in High-Consequence Drug Resistance Plasmids. *mBio* 7.
- Heather, J.M., Chain, B., 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics* 107, 1-8.
- Hirono, I., Aoki, T., 1993. Cloning and characterization of 3 hemolysin genes from *Aeromonas salmonicida*. *Microb. Pathog.* 15, 269-282.
- Hirst, I.D., Hastings, T.S., Ellis, A.E., 1991. Siderophore production by *Aeromonas salmonicida*. *J. Gen. Microbiol.* 137, 1185-1192.
- Hirst, I.D., Hastings, T.S., Ellis, A.E., 1994. Utilization of heme compounds by *Aeromonas salmonicida*. *J. Fish Dis.* 17, 365-373.
- Hjerde, E., Lorentzen, M.S., Holden, M.T., Seeger, K., Paulsen, S., Bason, N., Churcher, C., Harris, D., Norbertczak, H., Quail, M.A., Sanders, S., Thurston, S., Parkhill, J., Willassen, N.P., Thomson, N.R., 2008. The genome sequence of the fish pathogen *Aliivibrio salmonicida* strain LFI1238 shows extensive evidence of gene decay. *BMC genomics* 9, 616.
- Imbeault, S., Parent, S., Lagacé, M., Uhland, C.F., Blais, J.-F., 2006. Using Bacteriophages to prevent furunculosis caused by *Aeromonas salmonicida* in farmed brook trout. *Journal of Aquatic Animal Health* 18, 203-214.
- Iranzo, J., Gómez, M.J., López de Saro, F.J., Manrubia, S., 2014. Large-scale genomic analysis suggests a neutral punctuated dynamics of transposable elements in bacterial genomes. *PLoS computational biology* 10, e1003680.
- Ishiguro, E.E., Kay, W.W., Ainsworth, T., Chamberlain, J.B., Austen, R.A., Buckley, J.T., Trust, T.J., 1981. Loss of virulence during culture of *Aeromonas salmonicida* at high temperature. *J. Bacteriol.* 148, 333-340.
- Jackson, R.W., Vinatzer, B., Arnold, D.L., Dorus, S., Murillo, J., 2011. The influence of the accessory genome on bacterial pathogen evolution. *Mobile genetic elements* 1, 55-65.

- Janda, J.M., Abbott, S.L., 2010. The genus *Aeromonas*: taxonomy, pathogenicity, and infection. *Clin. Microbiol. Rev.* 23, 35-73.
- Jimenez, N., Lacasta, A., Vilches, S., Reyes, M., Vazquez, J., Aquillini, E., Merino, S., Regué, M., Tomás, J.M., 2009. Genetics and proteomics of *Aeromonas salmonicida* lipopolysaccharide core biosynthesis. *J. Bacteriol.* 191, 2228-2236.
- Kaleta, P., O'Callaghan, J., Fitzgerald, G.F., Beresford, T.P., Ross, R.P., 2010. Crucial role for insertion sequence elements in *Lactobacillus helveticus* evolution as revealed by interstrain genomic comparison. *Appl. Environ. Microbiol.* 76, 212-220.
- Kay, W.W., Buckley, J.T., Ishiguro, E.E., Phipps, B.M., Monette, J.P.L., Trust, T.J., 1981. Purification and disposition of a surface protein associated with virulence of *Aeromonas salmonicida*. *J. Bacteriol.* 147, 1077-1084.
- Kim, J.H., Choresca, C.H., Shin, S.P., Han, J.E., Jun, J.W., Park, S.C., 2015. Biological control of *Aeromonas salmonicida* subsp. *salmonicida* infection in rainbow trout (*Oncorhynchus mykiss*) using *Aeromonas* phage PAS-1. *Transboundary and emerging diseases* 62, 81-86.
- Kiss, J., Nagy, Z., Tóth, G., Kiss, G.B., Jakab, J., Chandler, M., Olasz, F., 2007. Transposition and target specificity of the typical IS30 family element IS1655 from *Neisseria meningitidis*. *Mol. Microbiol.* 63, 1731-1747.
- Klemm, E., Dougan, G., 2016. Advances in Understanding Bacterial Pathogenesis Gained from Whole-Genome Sequencing and Phylogenetics. *Cell host & microbe* 19, 599-610.
- Koren, S., Harhay, G.P., Smith, T.P., Bono, J.L., Harhay, D.M., McVey, S.D., Radune, D., Bergman, N.H., Phillippy, A.M., 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol* 14, R101.
- Kuska, B., 1998. Beer, Bethesda, and biology: how "genomics" came into being. *J. Natl. Cancer Inst.* 90, 93.
- L'Abée-Lund, T.M., Sørum, H., 2000. Functional Tn5393-like transposon in the R plasmid pRAS2 from the fish pathogen *Aeromonas salmonicida* subspecies *salmonicida* isolated in Norway. *Appl. Environ. Microbiol.* 66, 5533-5535.
- L'Abée-Lund, T.M., Sørum, H., 2002. A global non-conjugative Tet C plasmid, pRAS3, from *Aeromonas salmonicida*. *Plasmid* 47, 172-181.
- Land, M., Hauser, L., Jun, S.R., Nookaew, I., Leuze, M.R., Ahn, T.H., Karpinets, T., Lund, O., Kora, G., Wassenaar, T., Poudel, S., Ussery, D.W., 2015. Insights from 20 years of bacterial genome sequencing. *Functional & integrative genomics* 15, 141-161.
- Lee, H., Doak, T.G., Popodi, E., Foster, P.L., Tang, H., 2016. Insertion sequence-caused large-scale rearrangements in the genome of *Escherichia coli*. *Nucleic Acids Res* 44, 7109-7119.
- Lee, K.K., Ellis, A.E., 1990. Glycerophospholipid:cholesterol acyltransferase complexed with lipopolysaccharide (LPS) is a major lethal exotoxin and cytolysin of *Aeromonas salmonicida*: LPS stabilizes and enhances toxicity of the enzyme. *J. Bacteriol.* 172, 5382-5393.
- Liebert, C.A., Hall, R.M., Summers, A.O., 1999. Transposon Tn21, flagship of the floating genome. *Microbiol. Mol. Biol. Rev.* 63, 507-522.
- Ling, A., Cordaux, R., 2010. Insertion sequence inversions mediated by ectopic recombination between terminal inverted repeats. *PloS one* 5, e15654.
- Long, M., Nielsen, T.K., Leisner, J.J., Hansen, L.H., Shen, Z.X., Zhang, Q.Q., Li, A., 2016. *Aeromonas salmonicida* subsp. *salmonicida* strains isolated from Chinese

- freshwater fish contain a novel genomic island and possible regional-specific mobile genetic elements profiles. *FEMS Microbiol. Lett.* 363.
- Lovell, H.C., Jackson, R.W., Mansfield, J.W., Godfrey, S.A., Hancock, J.T., Desikan, R., Arnold, D.L., 2011. In planta conditions induce genomic changes in *Pseudomonas syringae* pv. *phaseolicola*. *Molecular plant pathology* 12, 167-176.
- Lusta, K.A., Kozlovskii, Y.E., 2011. Outer membrane nanovesicles of gram-negative bacteria *Aeromonas hydrophila* and *Aeromonas salmonicida*. *Microbiology* 80, 519-524.
- Lusta, K.A., Kozlovsky, Y.E., 2011. Excretion of extracellular membrane nanovesicles by *Aeromonas*. *Zhurnal Mikrobiologii Epidemiologii i Immunobiologii*, 77-79.
- Lutwyche, P., Exner, M.M., Hancock, R.E.W., Trust, T.J., 1995. A conserved *Aeromonas salmonicida* porin provides protective immunity to rainbow trout. *Infect. Immun.* 63, 3137-3142.
- Lynch, M., 2007. The origins of genome architecture. Sinauer Associates, Sunderland, Mass, xvi, 494 p pp.
- Mahillon, J., Chandler, M., 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62, 725-774.
- Maldonado, R.F., Sa-Correia, I., Valvano, M.A., 2016. Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiol. Rev.* 40, 480-493.
- Marana, M.H., Jørgensen, L.V., Skov, J., Chettri, J.K., Holm Mattsson, A., Dalsgaard, I., Kania, P.W., Buchmann, K., 2017. Subunit vaccine candidates against *Aeromonas salmonicida* in rainbow trout *Oncorhynchus mykiss*. *PloS one* 12, e0171944.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.
- Marti, E., Balcázar, J.L., 2015. *Aeromonas rivipollensis* sp. nov., a novel species isolated from aquatic samples. *J. Basic Microbiol.* 55, 1435-1439.
- Martin-Carnahan, A., Joseph, S.W. 2005. *Aeromonas*, In: Brenner, D.J., Krieg, N.R., Staley, J.T., Garrity, G.M. (Eds.) *Bergey's Manual of Systematic Bacteriology*. William and Wilkins, New York.
- Martínez-Murcia, A., Beaz-Hidalgo, R., Navarro, A., Carvalho, M.J., Aravena-Román, M., Correia, A., Figueras, M.J., Saavedra, M.J., 2016. *Aeromonas lusitana* sp. nov., Isolated from Untreated Water and Vegetables. *Curr. Microbiol.* 72, 795-803.
- Masada, C.L., LaPatra, S.E., Morton, A.W., Strom, M.S., 2002. An *Aeromonas salmonicida* type IV pilin is required for virulence in rainbow trout *Oncorhynchus mykiss*. *Dis. Aquat. Organ.* 51, 13-25.
- McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M.A., Baylay, A.J., Bhullar, K., Canova, M.J., De Pascale, G., Ejim, L., Kalan, L., King, A.M., Koteva, K., Morar, M., Mulvey, M.R., O'Brien, J.S., Pawlowski, A.C., Piddock, L.J.,

- Spanogiannopoulos, P., Sutherland, A.D., Tang, I., Taylor, P.L., Thaker, M., Wang, W., Yan, M., Yu, T., Wright, G.D., 2013. The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* 57, 3348-3357.
- McCutcheon, J.P., Moran, N.A., 2012. Extreme genome reduction in symbiotic bacteria. *Nature reviews. Microbiology* 10, 13-26.
- McIntosh, D., Cunningham, M., Ji, B., Fekete, F.A., Parry, E.M., Clark, S.E., Zalinger, Z.B., Gilg, I.C., Danner, G.R., Johnson, K.A., Beattie, M., Ritchie, R., 2008. Transferable, multiple antibiotic and mercury resistance in Atlantic Canadian isolates of *Aeromonas salmonicida* subsp. *salmonicida* is associated with carriage of an IncA/C plasmid similar to the *Salmonella enterica* plasmid pSN254. *J. Antimicrob. Chemother.* 61, 1221-1228.
- Mehdizadeh Gohari, I., Kropinski, A.M., Weese, S.J., Parreira, V.R., Whitehead, A.E., Boerlin, P., Prescott, J.F., 2016. Plasmid Characterization and Chromosome Analysis of Two *netF*⁺ *Clostridium perfringens* Isolates Associated with Foal and Canine Necrotizing Enteritis. *PloS one* 11, e0148344.
- Meinersmann, R.J., Lindsey, R.L., Bono, J.L., Smith, T.P., Oakley, B.B., 2013. Proposed model for the high rate of rearrangement and rapid migration observed in some IncA/C plasmid lineages. *Appl. Environ. Microbiol.* 79, 4806-4814.
- Menanteau-Ledouble, S., El-Matbouli, M., 2016. Antigens of *Aeromonas salmonicida* subsp. *salmonicida* specifically induced in vivo in *Oncorhynchus mykiss*. *J Fish Dis* 39, 1015-1019.
- Menanteau-Ledouble, S., Kumar, G., Saleh, M., El-Matbouli, M., 2016. *Aeromonas salmonicida*: updates on an old acquaintance. *Dis. Aquat. Organ.* 120, 49-68.
- Merino, S., Aguilar, A., Rubires, X., SimonPujol, D., Congregado, F., Tomás, J.M., 1996. The role of the capsular polysaccharide of *Aeromonas salmonicida* in the adherence and invasion of fish cell lines. *FEMS Microbiol. Lett.* 142, 185-189.
- Merino, S., de Mendoza, E., Canals, R., Tomás, J.M., 2015. Functional Genomics of the *Aeromonas salmonicida* Lipopolysaccharide O-Antigen and A-Layer from Typical and Atypical Strains. *Marine drugs* 13, 3791-3808.
- Merino, S., Gavin, R., Vilches, S., Shaw, J.G., Tomas, J.M., 2003. A colonization factor (Production of lateral flagella) of mesophilic *Aeromonas* spp. is inactive in *Aeromonas salmonicida* strains. *Appl. Environ. Microbiol.* 69, 663-667.
- Metzker, M.L., 2010. Sequencing technologies - the next generation. *Nature reviews. Genetics* 11, 31-46.
- Michel, C., Dubois-Darnaudpeys, A., 1980. Persistence of the virulence of *Aeromonas salmonicida* strains kept in river sediments. *Ann. Rech. Vet.* 11, 375-380.
- Moki, S.T., Nomura, T., Yoshimizu, M., 1995. Effect of Incubation Temperature for Isolation on Auto-Agglutination of *Aeromonas salmonicida*. *Fish Pathol.* 30, 67-68.
- Moran, N.A., Plague, G.R., 2004. Genomic changes following host restriction in bacteria. *Curr. Opin. Genet. Dev.* 14, 627-633.
- Morin, R., 2010. L'utilisation des antibiotiques pour combattre la furunculose chez l'omble de fontaine génère de l'antibiorésistance chez *Aeromonas salmonicida*. *L'Aquicole - Le Bulletin de l'Association des aquaculteurs du Québec* 15.
- Morin, R., Uhland, C.F., Lévesque, G., 2004. L'utilisation des antibiotiques en pisciculture au Québec. *L'Aquicole* 9.

- Munn, C.B., Ishiguro, E.E., Kay, W.W., Trust, T.J., 1982. Role of surface components in serum resistance of virulent *Aeromonas salmonicida*. *Infect. Immun.* 36, 1069-1075.
- Musovic, S., Oregaard, G., Kroer, N., Sørensen, S.J., 2006. Cultivation-independent examination of horizontal transfer and host range of an IncP-1 plasmid among gram-positive and gram-negative bacteria indigenous to the barley rhizosphere. *Appl. Environ. Microbiol.* 72, 6687-6692.
- Naas, T., Namdari, F., Bogaerts, P., Huang, T.D., Glupczynski, Y., Nordmann, P., 2008. Genetic structure associated with *bla*_{OXA-18}, encoding a clavulanic acid-inhibited extended-spectrum oxacillinase. *Antimicrob. Agents Chemother.* 52, 3898-3904.
- Nagar, V., Shashidhar, R., Bandekar, J.R., 2011. Prevalence, Characterization, and Antimicrobial Resistance of *Aeromonas* Strains from Various Retail Food Products in Mumbai, India. *J. Food Sci.* 76, M486-M492.
- Najimi, M., Balado, M., Lemos, M.L., Osorio, C.R., 2009. Genetic characterization of pAsa6, a new plasmid from *Aeromonas salmonicida* subsp. *salmonicida* that encodes a type III effector protein AopH homolog. *Plasmid* 61, 176-181.
- NCBI 2014. National Center for Biotechnology Information Genome Browser.
- Nelson, W.C., Wollerman, L., Bhaya, D., Heidelberg, J.F., 2011. Analysis of insertion sequences in thermophilic cyanobacteria: exploring the mechanisms of establishing, maintaining, and withstanding high insertion sequence abundance. *Appl. Environ. Microbiol.* 77, 5458-5466.
- Noonan, B., Trust, T.J., 1996. An *Aeromonas salmonicida* gene required for the establishment of infection in rainbow trout (*Oncorhynchus mykiss*). *Gene* 175, 127-131.
- Novick, R.P., Christie, G.E., Penades, J.R., 2010. The phage-related chromosomal islands of Gram-positive bacteria. *Nature reviews. Microbiology* 8, 541-551.
- O'Rawe, J.A., Ferson, S., Lyon, G.J., 2015. Accounting for uncertainty in DNA sequencing data. *Trends Genet.* 31, 61-66.
- Ochman, H., Davalos, L.M., 2006. The nature and dynamics of bacterial genomes. *Science* 311, 1730-1733.
- Okmen, G., Ugur, A., Sarac, N., Arslan, T., 2012. *In vivo* and *In vitro* Antibacterial Activities of Some Essential Oils of Lamiaceae Species on *Aeromonas salmonicida* Isolates from Cultured Rainbow Trout, *Oncorhynchus mykiss*. *J. Anim. Vet. Adv.* 11, 2762-2768.
- Olasz, F., Kiss, J., König, P., Buzás, Z., Stalder, R., Arber, W., 1998. Target specificity of insertion element IS30. *Mol. Microbiol.* 28, 691-704.
- Orgel, L.E., Crick, F.H., 1980. Selfish DNA: the ultimate parasite. *Nature* 284, 604-607.
- Origi, F.C., Benedicenti, O., Segner, H., Sattler, U., Wahli, T., Frey, J., 2017. *Aeromonas salmonicida* type III secretion system-effectors-mediated immune suppression in rainbow trout (*Oncorhynchus mykiss*). *Fish & shellfish immunology* 60, 334-345.
- Osorio, C.R., Marrero, J., Wozniak, R.A., Lemos, M.L., Burrus, V., Waldor, M.K., 2008. Genomic and functional analysis of ICEPdaSpa1, a fish-pathogen-derived SXT-related integrating conjugative element that can mobilize a virulence plasmid. *J. Bacteriol.* 190, 3353-3361.
- Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens,

- R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42, D206-214.
- Pandey, D.P., Gerdes, K., 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* 33, 966-976.
- Parkhill, J., Sebaihia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T., Churcher, C.M., Bentley, S.D., Mungall, K.L., Cerdeno-Tarraga, A.M., Temple, L., James, K., Harris, B., Quail, M.A., Achtman, M., Atkin, R., Baker, S., Basham, D., Bason, N., Cherevach, I., Chillingworth, T., Collins, M., Cronin, A., Davis, P., Doggett, J., Feltwell, T., Goble, A., Hamlin, N., Hauser, H., Holroyd, S., Jagels, K., Leather, S., Moule, S., Norberczak, H., O'Neil, S., Ormond, D., Price, C., Rabinowitsch, E., Rutter, S., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Unwin, L., Whitehead, S., Barrell, B.G., Maskell, D.J., 2003a. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat. Genet.* 35, 32-40.
- Parkhill, J., Sebaihia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T., Churcher, C.M., Bentley, S.D., Mungall, K.L., Cerdeño-Tarraga, A.M., Temple, L., James, K., Harris, B., Quail, M.A., Achtman, M., Atkin, R., Baker, S., Basham, D., Bason, N., Cherevach, I., Chillingworth, T., Collins, M., Cronin, A., Davis, P., Doggett, J., Feltwell, T., Goble, A., Hamlin, N., Hauser, H., Holroyd, S., Jagels, K., Leather, S., Moule, S., Norberczak, H., O'Neil, S., Ormond, D., Price, C., Rabinowitsch, E., Rutter, S., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Unwin, L., Whitehead, S., Barrell, B.G., Maskell, D.J., 2003b. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat. Genet.* 35, 32-40.
- Partridge, S.R., Brown, H.J., Stokes, H.W., Hall, R.M., 2001. Transposons Tn1696 and Tn21 and their integrons In4 and In2 have independent origins. *Antimicrob. Agents Chemother.* 45, 1263-1270.
- Pavan, M.E., Pavan, E.E., Lopez, N.I., Levin, L., Pettinari, M.J., 2015. Living in an Extremely Polluted Environment: Clues from the Genome of Melanin-Producing *Aeromonas salmonicida* subsp. *pectinolytica* 34meIT. *Appl. Environ. Microbiol.* 81, 5235-5248.
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U. S. A.* 85, 2444-2448.
- Petersen, J., 2011. Phylogeny and compatibility: plasmid classification in the genomics era. *Arch. Microbiol.* 193, 313-321.
- Petersen, J., Wagner-Dobler, I., 2017. Plasmid Transfer in the Ocean - A Case Study from the Roseobacter Group. *Frontiers in microbiology* 8, 1350.
- Pilla, G., McVicker, G., Tang, C.M., 2017. Genetic plasticity of the *Shigella* virulence plasmid is mediated by intra- and inter-molecular events between insertion sequences. *PLoS genetics* 13, e1007014.
- Piotrowska, M., Popowska, M., 2014. The prevalence of antibiotic resistance genes among *Aeromonas* species in aquatic environments. *Ann. Microbiol.* 64, 921-934.
- Piotrowska, M., Popowska, M., 2015. Insight into the mobilome of *Aeromonas* strains. *Frontiers in microbiology* 6, 494.

- Pitman, A.R., Jackson, R.W., Mansfield, J.W., Kaitell, V., Thwaites, R., Arnold, D.L., 2005. Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. *Curr. Biol.* 15, 2230-2235.
- Price, N.C., Banks, R.M., Campbell, C., Duncan, D., Stevens, L., 1990. The specificity of the major (70 kDa) protease secreted by *Aeromonas salmonicida*. *J. Fish Dis.* 13, 49-58.
- R Core Team 2015. R: A Language and Environment for Statistical Computing (Vienna, Austria, R Foundation for Statistical Computing).
- Radulovic, N.S., Blagojevic, P.D., Stojanovic-Radic, Z.Z., Stojanovic, N.M., 2013. Antimicrobial Plant Metabolites: Structural Diversity and Mechanism of Action. *Curr. Med. Chem.* 20, 932-952.
- Ramirez, M.S., Tolmasky, M.E., 2010. Aminoglycoside modifying enzymes. Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy 13, 151-171.
- Redfield, R.J., 2001. Do bacteria have sex? *Nature reviews. Genetics* 2, 634-639.
- Reith, M.E., Singh, R.K., Curtis, B., Boyd, J.M., Bouevitch, A., Kimball, J., Munholland, J., Murphy, C., Sarty, D., Williams, J., Nash, J.H., Johnson, S.C., Brown, L.L., 2008. The genome of *Aeromonas salmonicida* subsp. *salmonicida* A449: insights into the evolution of a fish pathogen. *BMC genomics* 9, 427.
- Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276-277.
- Ricker, N., Qian, H., Fulthorpe, R.R., 2012. The limitations of draft assemblies for understanding prokaryotic adaptation and evolution. *Genomics* 100, 167-175.
- Roger, F., Marchandin, H., Jumas-Bilak, E., Kodjo, A., Lamy, B., 2012. Multilocus genetics to reconstruct aeromonad evolution. *BMC microbiology* 12, 62.
- Röltgen, K., Stinear, T.P., Pluschke, G., 2012. The genome, evolution and diversity of *Mycobacterium ulcerans*. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 12, 522-529.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944-945.
- Salte, R., Norberg, K., Arnesen, J.A., Ødegaard, O.R., Eggset, G., 1992. Serine protease and glycerophospholipid: cholesterol acyltransferase of *Aeromonas salmonicida* work in concert in thrombus formation; *in vitro* the process is counteracted by plasma antithrombin and α 2-macroglobulin. *J. Fish Dis.* 15, 215-227.
- Sandaa, R.A., Enger, O., 1994. Transfer in Marine Sediments of the Naturally Occurring Plasmid pRAS1 Encoding Multiple Antibiotic Resistance. *Appl. Environ. Microbiol.* 60, 4234-4238.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* 74, 5463-5467.
- Selle, K., Klaenhammer, T.R., Barrangou, R., 2015. CRISPR-based screening of genomic island excision events in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8076-8081.
- Shapiro, B.J. 2014. Signatures of Natural Selection and Ecological Differentiation in Microbial Genomes, In: Landry, C.R., Aubin-Horth, N. (Eds.) *Ecological Genomics: Ecology and the Evolution of Genes and Genomes*. Springer Netherlands, Dordrecht, 339-359.
- Shen, P., Huang, H.V., 1986. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* 112, 441-457.

- Shieh, H.S., Maclean, J.R., 1975. Purification and Properties of an Extracellular Protease of *Aeromonas salmonicida*, Causative Agent of Furunculosis. *Int. J. Biochem.* 6, 653-656.
- Siguier, P., Filee, J., Chandler, M., 2006a. Insertion sequences in prokaryotic genomes. *Curr. Opin. Microbiol.* 9, 526-531.
- Siguier, P., Gourbeyre, E., Chandler, M., 2014. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* 38, 865-891.
- Siguier, P., Gourbeyre, E., Varani, A., Ton-Hoang, B., Chandler, M., 2015. Everyman's Guide to Bacterial Insertion Sequences. *Microbiology spectrum* 3, MDNA3-0030-2014.
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., Chandler, M., 2006b. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34, D32-36.
- Smith, E.E., Buckley, D.G., Wu, Z., Saenphimmachak, C., Hoffman, L.R., D'Argenio, D.A., Miller, S.I., Ramsey, B.W., Speert, D.P., Moskowitz, S.M., Burns, J.L., Kaul, R., Olson, M.V., 2006. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8487-8492.
- Snyder, L., Champness, W., 2007. Molecular genetics of bacteria, Third Edition Edition. ASM Press, Washington, 735 p.
- Sørum, H., Kvello, J., Hastein, T., 1993. Occurrence and stability of plasmids in *Aeromonas salmonicida* ss *salmonicida* isolated from salmonids with furunculosis. *Dis. Aquat. Organ.* 16, 199-206.
- Sørum, H., L'Abée-Lund, T., Solberg, A., Wold, A., 2003. Integron-containing IncU R plasmids pRAS1 and pAr-32 from the fish pathogen *Aeromonas salmonicida*. *Antimicrob. Agents Chemother.* 47, 1285-1290.
- Starliper, C.E., Ketola, H.G., Noyes, A.D., Schill, W.B., Henson, F.G., Chalupnicki, M.A., Dittman, D.E., 2015. An investigation of the bactericidal activity of selected essential oils to *Aeromonas spp.* *Journal of advanced research* 6, 89-97.
- Stuber, K., Burr, S.E., Braun, M., Wahli, T., Frey, J., 2003. Type III secretion genes in *Aeromonas salmonicida* subsp. *salmonicida* are located on a large thermolabile virulence plasmid. *J. Clin. Microbiol.* 41, 3854-3856.
- Studer, N., Frey, J., Vanden Bergh, P., 2013. Clustering subspecies of *Aeromonas salmonicida* using IS630 typing. *BMC microbiology* 13, 36.
- Sturm, A., Heinemann, M., Arnoldini, M., Benecke, A., Ackermann, M., Benz, M., Dormann, J., Hardt, W.D., 2011. The cost of virulence: retarded growth of *Salmonella* Typhimurium cells expressing type III secretion system 1. *PLoS pathogens* 7, e1002143.
- Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009-1010.
- Suzuki, H., Saito, R., Tomita, M., 2005. A problem in multivariate analysis of codon usage data and a possible solution. *FEBS Lett.* 579, 6499-6504.
- Tanaka, K.H., Dallaire-Dufresne, S., Daher, R.K., Frenette, M., Charette, S.J., 2012. An insertion sequence-dependent plasmid rearrangement in *Aeromonas salmonicida* causes the loss of the type three secretion system. *PloS one* 7, e33725.
- Tanaka, K.H., Frenette, M., Charette, S.J., 2013. IS-mediated loss of virulence by *Aeromonas salmonicida*: A tangible piece of an evolutionary puzzle. *Mobile genetic elements* 3, e23498.

- Tanaka, K.H., Vincent, A.T., Emond-Rheault, J.-G., Adamczuk, M., Frenette, M., Charette, S.J., 2017. Plasmid composition in *Aeromonas salmonicida* subsp. *salmonicida* 01-B526 unravels unsuspected type three secretion system loss patterns. *BMC genomics* 18, 528.
- Tanaka, K.H., Vincent, A.T., Trudel, M.V., Paquet, V.E., Frenette, M., Charette, S.J., 2016. The mosaic architecture of *Aeromonas salmonicida* subsp. *salmonicida* pAsa4 plasmid and its consequences on antibiotic resistance. *PeerJ* 4, e2595.
- Tavakoli, N., Comanducci, A., Dodd, H.M., Lett, M.C., Albiger, B., Bennett, P., 2000. IS1294, a DNA element that transposes by RC transposition. *Plasmid* 44, 66-84.
- Tenzen, T., Matsutani, S., Ohtsubo, E., 1990. Site-Specific Transposition of Insertion Sequence IS630. *J. Bacteriol.* 172, 3830-3836.
- Toleman, M.A., Bennett, P.M., Walsh, T.R., 2006a. Common regions e.g. orf513 and antibiotic resistance: IS91-like elements evolving complex class 1 integrons. *J. Antimicrob. Chemother.* 58, 1-6.
- Toleman, M.A., Bennett, P.M., Walsh, T.R., 2006b. ISCR elements: novel gene-capturing systems of the 21st century? *Microbiol. Mol. Biol. Rev.* 70, 296-316.
- Touchon, M., Rocha, E.P., 2007. Causes of insertion sequences abundance in prokaryotic genomes. *Mol. Biol. Evol.* 24, 969-981.
- Toussaint, A., Merlin, C., 2002. Mobile elements as a combination of functional modules. *Plasmid* 47, 26-35.
- Treangen, T.J., Rocha, E.P., 2011. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS genetics* 7, e1001284.
- Treangen, T.J., Salzberg, S.L., 2011. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature reviews. Genetics* 13, 36-46.
- Trudel, M.V., Tanaka, K.H., Fillion, G., Daher, R.K., Frenette, M., Charette, S.J., 2013. Insertion sequence AS5 (ISAS5) is involved in the genomic plasticity of *Aeromonas salmonicida*. *Mobile genetic elements* 3, e25640.
- Trudel, M.V., Vincent, A.T., Att  r  , S.A., Labb  , M., Derome, N., Culley, A.I., Charette, S.J., 2016. Diversity of antibiotic-resistance genes in Canadian isolates of *Aeromonas salmonicida* subsp. *salmonicida*: dominance of pSN254b and discovery of pAsa8. *Scientific reports* 6, 35617.
- Udey, L.R., Fryer, J.L., 1978. Immunization of fish with bacterins of *Aeromonas salmonicida*. *Marine Fisheries Review* 40, 12-17.
- Umelo, E., Noonan, B., Trust, T.J., 1996. Cloning, characterization and expression of the *recA* gene of *Aeromonas salmonicida*. *Gene* 175, 133-136.
- Umelo, E., Trust, T.J., 1997. Identification and molecular characterization of two tandemly located flagellin genes from *Aeromonas salmonicida* A449. *J. Bacteriol.* 179, 5292-5299.
- van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C., 2014. Ten years of next-generation sequencing technology. *Trends Genet.* 30, 418-426.
- Vandecraen, J., Chandler, M., Aertsen, A., Van Houdt, R., 2017. The impact of insertion sequences on bacterial genome plasticity and adaptability. *Crit. Rev. Microbiol.*, 1-22.
- Vanden Bergh, P., Frey, J., 2014. *Aeromonas salmonicida* subsp. *salmonicida* in the light of its type-three secretion system. *Microbial biotechnology* 7, 381-400.
- Vanden Bergh, P., Heller, M., Braga-Lagache, S., Frey, J., 2013. The *Aeromonas salmonicida* subsp. *salmonicida* exoproteome: determination of the complete

- repertoire of Type-Three Secretion System effectors and identification of other virulence factors. *Proteome science* 11, 42.
- Verdet, C., Benzerara, Y., Gautier, V., Adam, O., Ould-Hocine, Z., Arlet, G., 2006. Emergence of DHA-1-producing *Klebsiella spp.* in the Parisian region: genetic organization of the *ampC* and *ampR* genes originating from *Morganella morganii*. *Antimicrob. Agents Chemother.* 50, 607-617.
- Verner-Jeffreys, D.W., Algoet, M., Pond, M.J., Virdee, H.K., Bagwell, N.J., Roberts, E.G., 2007. Furunculosis in Atlantic salmon (*Salmo salar* L.) is not readily controllable by bacteriophage therapy. *Aquaculture* 270, 475-484.
- Villumsen, K.R., Koppang, E.O., Raida, M.K., 2015. Adverse and long-term protective effects following oil-adjuvanted vaccination against *Aeromonas salmonicida* in rainbow trout. *Fish & shellfish immunology* 42, 193-203.
- Villumsen, K.R., Raida, M.K., 2013. Long-lasting protection induced by bath vaccination against *Aeromonas salmonicida* subsp. *salmonicida* in rainbow trout. *Fish & shellfish immunology* 35, 1649-1653.
- Vincent, A.T., Boyle, B., Derome, N., Charette, S.J., 2014a. Improvement in the DNA sequencing of genomes bearing long repeated elements. *J. Microbiol. Methods* 107, 186-188.
- Vincent, A.T., Charette, S.J., 2015. Who qualifies to be a bioinformatician? *Frontiers in genetics* 6, 164.
- Vincent, A.T., Derome, N., Boyle, B., Culley, A.I., Charette, S.J., 2017. Next-generation sequencing (NGS) in the microbial world: How to make the most of your money. *J. Microbiol. Methods* 138, 60-71.
- Vincent, A.T., Emond-Rheault, J.G., Barbeau, X., Attéré, S.A., Frenette, M., Lagüe, P., Charette, S.J., 2016a. Antibiotic resistance due to an unusual ColE1-type replicon plasmid in *Aeromonas salmonicida*. *Microbiology* 162, 942-953.
- Vincent, A.T., Tanaka, K.H., Trudel, M.V., Frenette, M., Derome, N., Charette, S.J., 2015. Draft genome sequences of two *Aeromonas salmonicida* subsp. *salmonicida* isolates harboring plasmids conferring antibiotic resistance. *FEMS Microbiol. Lett.* 362.
- Vincent, A.T., Trudel, M.V., Freschi, L., Nagar, V., Gagné-Thivierge, C., Levesque, R.C., Charette, S.J., 2016b. Increasing genomic diversity and evidence of constrained lifestyle evolution due to insertion sequences in *Aeromonas salmonicida*. *BMC genomics* 17, 44.
- Vincent, A.T., Trudel, M.V., Paquet, V.E., Boyle, B., Tanaka, K.H., Dallaire-Dufresne, S., Daher, R.K., Frenette, M., Derome, N., Charette, S.J., 2014b. Detection of variants of the pRAS3, pAB5S9, and pSN254 plasmids in *Aeromonas salmonicida* subsp. *salmonicida*: multidrug resistance, interspecies exchanges, and plasmid reshaping. *Antimicrob. Agents Chemother.* 58, 7367-7374.
- Vipond, R., Bricknell, I.R., Durant, E., Bowden, T.J., Ellis, A.E., Smith, M., MacIntyre, S., 1998. Defined deletion mutants demonstrate that the major secreted toxins are not essential for the virulence of *Aeromonas salmonicida*. *Infect. Immun.* 66, 1990-1998.
- Wagner, A., 2006. Periodic extinctions of transposable elements in bacterial lineages: evidence from intragenomic variation in multiple genomes. *Mol. Biol. Evol.* 23, 723-733.
- Wang, Z., Larocque, S., Vinogradov, E., Brisson, J.R., Dacanay, A., Greenwell, M., Brown, L.L., Li, J.J., Altman, E., 2004. Structural studies of the capsular

- polysaccharide and lipopolysaccharide O-antigen of *Aeromonas salmonicida* strain 80204-1 produced under in vitro and in vivo growth conditions. Eur. J. Biochem. 271, 4507-4516.
- Wang, Z., Li, J., Vinogradov, E., Altman, E., 2006a. Structural studies of the core region of *Aeromonas salmonicida* subsp. *salmonicida* lipopolysaccharide. Carbohydr. Res. 341, 109-117.
- Wang, Z., Li, J.J., Altman, E., 2006b. Structural characterization of the lipid A region of *Aeromonas salmonicida* subsp. *salmonicida* lipopolysaccharide. Carbohydr. Res. 341, 2816-2825.
- Wang, Z., Liu, X., Dacanay, A., Harrison, B.A., Fast, M., Colquhoun, D.J., Lund, V., Brown, L.L., Li, J.J., Altman, E., 2007. Carbohydrate analysis and serological classification of typical and atypical isolates of *Aeromonas salmonicida*: A rationale for the lipopolysaccharide-based classification of *A. salmonicida*. Fish & shellfish immunology 23, 1095-1106.
- Wang, Z., Vinogradov, E., Larocque, S., Harrison, B.A., Li, J.J., Altman, E., 2005. Structural and serological characterization of the O-chain polysaccharide of *Aeromonas salmonicida* strains A449, 80204 and 80204-1. Carbohydr. Res. 340, 693-700.
- Welch, T.J., Fricke, W.F., McDermott, P.F., White, D.G., Rosso, M.L., Rasko, D.A., Mammel, M.K., Eppinger, M., Rosovitz, M.J., Wagner, D., Rahalison, L., Leclerc, J.E., Hinshaw, J.M., Lindler, L.E., Cebula, T.A., Carniel, E., Ravel, J., 2007. Multiple antimicrobial resistance in plague: an emerging public health risk. PloS one 2, e309.
- Wetzel, J., Kingsford, C., Pop, M., 2011. Assessing the benefits of using mate-pairs to resolve repeats in de novo short-read prokaryotic assemblies. BMC bioinformatics 12, 95.
- Wheatcroft, R., McRae, D.G., Miller, R.W., 1990. Changes in the *Rhizobium meliloti* Genome and the Ability to Detect Supercoiled Plasmids During Bacteroid Development. Mol. Plant. Microbe Interact. 3, 9-17.
- Wickham, H., 2009. ggplot2: elegant graphics for data analysis. Springer New York.
- Wilson, M.S.C., Saiardi, A., 2017. Importance of Radioactive Labelling to Elucidate Inositol Polyphosphate Signalling. Top Curr Chem (J) 375, 14.
- Wu, Y., Aandahl, R.Z., Tanaka, M.M., 2015. Dynamics of bacterial insertion sequences: can transposition bursts help the elements persist? BMC evolutionary biology 15, 288.
- Yadav, S.P., 2007. The wholeness in suffix -omics, -omes, and the word om. Journal of biomolecular techniques : JBT 18, 277.
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T.L., 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC bioinformatics 13, 134.
- Zeileis, A., Grothendieck, G., 2005. zoo: S3 Infrastructure for Regular and Irregular Time Series. Journal of Statistical Software 14, 1-27.
- Zinser, E.R., Schneider, D., Blot, M., Kolter, R., 2003. Bacterial evolution through the selective loss of beneficial Genes. Trade-offs in expression involving two loci. Genetics 164, 1271-1277.

ANNEXE 1 MÉTHODE EMPLOYÉE POUR L'IDENTIFICATION DES IS DANS LES GÉNOMES D'*A. SALMONICIDA* SSP. *SALMONICIDA*

MISE EN CONTEXTE

Lors de la rédaction de l'article présenté au Chapitre 3, il m'est apparu nécessaire de comparer le nombre d'IS de différents génomes d'*A. salmonicida* ssp. *salmonicida*. Cette procédure a été raffinée et utilisée lors de la rédaction du Chapitre 4 (Discussion). Puisque la méthode n'est incluse dans aucune publication, mais qu'elle est toutefois nécessaire à la compréhension des résultats présentés dans le Chapitre 4, elle est donc présentée dans cette annexe.

MÉTHODE

Séquences

Toutes les IS disponibles dans la base de données ISFinder sous le terme de recherche « *Aeromonas salmonicida* » (<http://www-is.biotoul.fr/>, consulté le 21 novembre 2016), excluant les transposons, ont été incluses dans l'analyse. De plus, les IS suivantes ont été extraites de leur séquence respective (en parenthèse) et elles ont été ajoutées à cette liste :

- ISAS4A, ISAS5B et ISAS10A du pAsa5 d'A449 (NC_009350) (voir Figure 4.2)
- ISAS11 du pAsa1 de référence (NC_004338)
- ISKpn3 de pAsa4c (KT033470)
- ISAS35 de pAsa9 (KY555070)
- IS26, ISEc9, IS4321 et IS6100 de pSN254b (KJ909290)

Les composantes du génome d'A449 ont été définies comme suit : le chromosome (RefSeq Genbank : NC_009348), les plasmides pAsa4 (NC_009349), pAsa5 (NC_009350), pAsa1 (NC_004923), pAsa2 (NC_004925) et pAsa3 (NC_004924). Le génome de 01-B526 a été défini de la façon suivante : le chromosome, assemblé selon les paramètres présentés dans

Tanaka *et al.* (2017) (A. Vincent, comm. personnelle), pAsa5 (accession Genbank : KY555069), pAsa9 (KY555070), les pAsa1, pAsa2 et pAsa3 d'A449, ainsi que le pAsa11 de référence (NC_004338) (Tanaka *et al.*, 2017). Les petits plasmides d'autres souches ont été utilisés pour 01-B526 en sachant qu'il n'y a pas d'anomalie dans le profil plasmidique de cette dernière (Trudel *et al.*, 2013). Les contigs de JF2267 (NZ_MII001000001.1) et de 01-B522 (NZ_MIIM01000001.1) ont été obtenus du répertoire WGS (*whole genome shotgun*) du NCBI.

Logiciels et scripts

Le programme lastal version 801 a été utilisé pour faire un alignement entre les IS (les séquences de référence) et les génomes d'A449, de 01-B526, de JF2267 et de 01-B522. Un script en R (work_IS_count.R), disponible à <http://github.com/katherinetanaka/scripts> a été utilisé pour trier et annoter les données. Le reste de l'analyse a été fait manuellement.

ANALYSE

Génomes complètement assemblés

Les alignements de plus de 95% de l'IS de référence ont été considérés comme des IS complètes. Ces données ont été comptabilisées dans le Tableau 4.1. Les alignements partiels des IS [1] ont été étudiés afin de trouver des paires de fragments potentiels. Les candidats devaient avoir une somme des longueurs totalisant plus de 95% de l'IS de référence, correspondre à des alignements rapprochés et sur le même brin de l'IS de référence et posséder des positions rapprochées sur le génome investigué. Lorsqu'une paire potentielle était trouvée, un alignement d'IS complet [2] dans la région séparant les deux fragments était recherché dans le tableau d'alignement. Dans un tel cas, l'IS fragmentée [1] était identifiée comme étant interrompue par l'autre IS complète [2]. C'est de cette façon que les copies d'IS_{AS5} interrompues par IS_{AS9} ont été identifiées pour A449 (voir section 4.2.1.2)

Génomes sous forme de contigs

Les bris de contigage rendent impossible l'identification d'IS complètes si elles sont présentes en plusieurs copies dans le génome. La présence d'une IS (marquée par un + dans

le Tableau 4.1) a été définie de la façon suivante : un alignement entre l'une des deux extrémités d'une IS et l'une des deux extrémités d'un contig. Dans certains cas, aucun alignement (dénnoté par 0) ou un seul alignement sur toute la longueur de l'IS (dénnoté par 1) a été détecté.