

Centre Armand-Frappier Santé Biotechnologie

**CARACTERISATION DE NOUVEAUX GENES
D'AUTOTRANSPORTEUR ET DU FIMBRIAE ET LEURS ROLES DANS
LA VIRULENCE DES SOUCHES D'*E. COLI* PATHOGENES EXTRA-
INTESTINALES**

Par
Hajer Habouria

Thèse présentée pour l'obtention du grade de Philosophiae Doctor (Ph.D.)
en sciences de Biologie

Jury d'évaluation

Président du jury et examineur interne	Charles Calmettes INRS- Centre Armand-Frappier Santé et Biotechnologie
Examineur externe	Michel Frenette Université Laval Faculté des sciences et de génie, Département de biochimie, de microbiologie et de bio-informatique
Examineur externe	Ahmad Saleh Université Laval Faculté des sciences et de génie, Département de biochimie, de microbiologie et de bio-informatique
Directeur de recherche	Charles M. Dozois INRS- Centre Armand-Frappier Santé Biotechnologie

A mon père..
A ma mère..
qui m'ont toujours soutenu et encouragé
quels que soient les choix et les voies que j'ai empruntés

REMERCIEMENTS

Le présent travail représente le fruit d'efforts continus et la contribution précieuse de plusieurs personnes. Les mots à eux seuls ne suffisent à leur exprimer ma profonde gratitude.

Je tiens en premier lieu à remercier les membres du jury pour avoir accepté d'évaluer ce travail.

Je tiens ensuite à exprimer toute ma reconnaissance envers mon directeur de recherche, Pr. **Charles Dozois** pour m'avoir accueillie au sein de son laboratoire. Je le remercie pour son encadrement scientifique, pour ses nombreux conseils constructifs, pour son indéfectible disponibilité et pour ses encouragements. Ce fut pour moi un réel privilège d'être parmi ton équipe et de profiter de ton expertise et ton savoir-faire.

Je tiens à remercier Pr. **Charles Calmettes**, professeur au Centre INRS- AFSB, pour avoir aimablement accepté de présider ce jury de thèse.

Je tiens aussi à remercier Pr. **Michel Frenette**, professeur à l'Université Laval, Faculté des sciences et de génie Département de biochimie, de microbiologie et de bio-informatique, pour avoir accepté d'examiner ce travail

Je suis également reconnaissante envers le Pr. **Ahmad Saleh**, professeur à l'Université Laval, Faculté des sciences et de génie Département de biochimie, de microbiologie et de bio-informatique, de sa disponibilité afin d'examiner ce travail.

Je tiens à remercier Sébastien Houle pour sa grande patience face à mes nombreuses questions, pour ses conseils scientifiques et pour sa contribution à l'avancement de ce projet. Je voudrais aussi exprimer ma profonde gratitude à Ahlem Zaghmi pour ses conseils scientifiques, son encouragement et sa présence chaleureuse. Je tiens aussi à remercier David Bernard pour ses conseils scientifiques. Je voudrais aussi remercier mes chers collègues de laboratoire actuels comme anciens, Amélie, Daniel, Hicham, Katerina, Julie, Hossein, Pravil, Rémi, Paula pour tous les bons moments que nous avons vécus ensemble. Un remerciement spécial à mes précieuses amies Soumia Lahmidi, Asma Lamine, Marwa Khabir, Amina Barakati pour votre gentillesse et votre support tout au long de ma thèse.

Sans oublier d'adresser mes vifs remerciements à mes chers amis qui ont permis de rendre ces années à l'INRS- AFSB plus vivantes et harmonieuses Asma, Ghizlaine, Ahmad, Slimane, Mondher, Mohamed, Mustafa, Hassan qui m'ont encouragée et soutenue tout au long de ces années.

Je tiens à remercier l'indispensable Michel Courcelles pour sa bonne humeur, ses nombreux conseils, ses services tant dans l'accessibilité aux publications et l'emprunt des livres. Un remerciement spécial à Josette Bourdages, Anne Philippon, Geneviève Courtois, Sylvie Richard pour leur assistance dans les démarches administratives.

Je tiens aussi à remercier les organismes subventionnaires CRSNG, CRIPA-FRQNT et le Ministère de l'Enseignement Supérieur et de la Recherche Scientifique en Tunisie, pour le financement de ce projet de recherche.

Enfin, toute ma gratitude va envers mes chers parents, Mohamed et Jamila Habouria, pour leur sacrifice et amour inconditionnel. Vous m'avez appris la persévérance et à rester positive quel que soit les conditions. Je vous serai éternellement reconnaissante pour votre encouragement, votre soutien et votre générosité débordante. Je remercie également mes chères sœurs avec qui j'ai passé de bons moments. Je vous remercie pour votre soutien sans faille, votre affection et d'avoir grandement contribué à ma réussite.

RÉSUMÉ

Les infections bactériennes causées par des souches d'*Escherichia coli* pathogènes extra-intestinales sont à l'origine de nombreuses maladies chez l'homme et les animaux. Ces pathogènes possèdent de nombreux facteurs de virulence, comme les fimbriae et les autotransporteurs, leur permettent de coloniser les cellules de l'hôte. Plusieurs fonctions ont été attribuées à ces protéines qui sont souvent essentielles pour l'établissement d'une infection. Dans ce projet de thèse, nous avons pu caractériser de nouveaux facteurs de virulence qui sont présent chez la souche *E. coli* pathogène aviaire (APEC) QT598. Parmi lesquels on trouve un fimbriae P-like (PL) porté par un plasmide de virulence et trois autotransporteurs (AT) de la famille des sérines protéase (SPATEs) dont un se trouve sur le plasmide de virulence (serine-protease hemagglutinin autotransporter, Sha) et les autres se trouvent sur un îlot de pathogénicité (tandem autotransporter genes, TagB et TagC). L'objectif principale de ce projet est, donc, de caractériser ces nouveaux facteurs en particulier le fimbriae PL et l'AT Sha, et de déterminer leurs rôles dans la virulence chez les souches ExPEC.

Dans la première partie, nous avons tenté de déterminer le rôle potentiel de l'AT Sha en effectuant des analyses *in silico*. Les résultats ont montré que Sha possède une structure typique d'un AT (un peptide signal, un domaine passager portant le site actif sérine protéase et un domaine β) avec un site de clivage modifié. De plus, nous avons montré qu'il appartient au classe II SPATEs avec une homologie de séquence de 66% d'identité avec Tsh, 60% d'identité avec TleA et 54% d'identité avec Vat. Ensuite, afin de déterminer son rôle dans la virulence, nous avons cloné le gène *sha* dans une souche *E. coli* K-12 non-pathogène. Ceci nous a permis d'observer deux phénotypes différents. Le premier était un phénotype d'adhérence, le clone Sha étant capable d'agglutiner différents types de sang, de former un biofilm, de former des auto-agrégations et d'adhérer aux cellules rénales humaines (HEK-293) et de la vessie (5637). Le deuxième phénotype était de type cytotoxique comme mis en évidence sur des cellules vésicales humaines. Toutefois, lors des essais *in vivo*, l'absence du gène *sha* n'a pas diminué la capacité de la souche QT598 à coloniser la vessie et les reins durant l'infection urinaire des souris CBA/J. Seul le mutant dépourvu de tous les SPATEs était moins capable de coloniser les reins. Cela peut suggérer le rôle cumulatif ou la redondance des SPATEs dans la survie et l'adaptation de la souche lors des infections urinaires.

Dans la seconde partie, les analyses *in silico* ont montré que le fimbriae PL est constitué d'un groupe de dix gènes : un gène de régulation (*plfB*) et neuf gènes qui codent pour des

protéines de structure et d'assemblage du fimbriae (*plfA*, *plfH*, *plfC*, *plfD*, *plfJ*, *plfK*, *plfE*, *plfF* et *plfG*). De plus, nous avons montré que la souche UPEC UMEA-3703-1 possède aussi les gènes qui codent pour le fimbriae PL. Ces gènes sont très semblables aux gènes présents chez la souche QT598 (environ 98% d'identité avec les différentes sous-unités) à l'exception de l'adhésine PflG_{UMEA-3703-1} qui est différente (environ 45% d'identité avec PflG_{QT598}). Ceci nous a permis de les classer en classe I pour PflG_{UMEA-3703-1} et classe II pour PflG_{QT598}. Ensuite, nous avons décidé de caractériser ces nouveaux variants de fimbriae PL puisqu'ils partagent une forte homologie et de comparer leur phénotype. Nous avons, donc, cloner les gènes dans *E. coli* K-12. Les deux clones étaient capables d'exprimer les fimbriae à leur surface, d'agglutiner différents types d'érythrocytes, de former un biofilm et d'adhérer aux cellules rénales humaines et de la vessie. Lors des essais *in vivo*, l'absence de l'opéron *plf* ne semble pas diminuer la capacité des souches QT598 ou UMEA-3703-1 à coloniser la vessie et les reins. Toutefois, la capacité compétitive du mutant était significativement moins importante au niveau des reins de souris CBA/J. Dans l'ensemble, les fimbriae PL représentent un nouveau type de fimbriae qui peut contribuer à l'augmentation de l'adhérence aux cellules de l'hôte.

Mots-clés : APEC, UPEC, SPATEs, Autotransporteurs, Fimbriae PL, Cytotoxicité, Adhésines, Infection du tractus urinaire.

ABSTRACT

Bacterial infections caused by extra-intestinal pathogenic *Escherichia coli* are responsible for many diseases in humans and animals. These pathogens express many virulence factors, like fimbriae and autotransporters, allowing them to colonize host cells. Many functions have been described for these proteins for host colonization and survival. In this thesis, we characterized new virulence factors that are present in the APEC strain QT598. Among them, a P-like (PL) fimbrial system found on a virulence plasmid and three autotransporters of the SPATE family. One of which is on the virulence plasmid (serine-protease hemagglutinin autotransporter, Sha) and the others are found on a pathogenicity island (tandem autotransporter genes, TagB and TagC). The main objective of this study was to characterize these new factors, in particular PL fimbriae and Sha, including their roles in the virulence of ExPEC strains.

In the first section, we determined the potential role of Sha by performing *in silico* analyses. The results have shown that Sha has a typical structure of an AT (signal peptide, passenger domain carrying the active site serine protease and a β domain) with a modified cleavage site. Moreover, we have demonstrated that Sha belongs to class II SPATE with a strong homology to Tsh (66% of identity), TleA (60% of identity) and Vat (54% of identity). Next, in order to determine its role in virulence, we cloned the gene in a non-pathogenic *E. coli* K-12 strain. We observed two distinct phenotypes. The first was an adhesion phenotype, in which the Sha clone was able to agglutinate different types of blood, form a biofilm, autoaggregate, and adhere to human kidney and bladder cells. The second phenotype was displayed a cytotoxic effect, which was determined on the bladder cells. However, the absence of the *sha* gene did not reduce the ability of the QT598 strain to colonize the bladder and kidneys during urinary tract infection in CBA/J mice. Only the mutant lacking all five identified SPATEs of the strain showed an impact on the capacity of the QT598 strain to colonize the kidneys. This may suggest the cumulative and/or redundant role of SPATEs in survival and fitness of QT598 strain during urinary tract infections.

For the second part, *in silico* analyses showed us that the PL fimbrial (*plf*) gene cluster contains 10 genes predicted to encode one regulatory gene (*plfB*) and 9 structural/assembly (*plfA*, *plfH*, *plfC*, *plfD*, *plfJ*, *plfK*, *plfE*, *plfF* and *plfG*). Furthermore, we identify another variant of PL fimbriae in the UPEC strain UMEA-3703-1 that share identity to equivalent proteins present in the *plf* gene cluster of QT598 (98% of identity with PlfC). However, the PlfG adhesins of the two strains were very distinct which allowed us to establish new classes: class I for PlfG_{UMEA-3703-1} and class II for PlfG_{QT598} adhesins. Next, we attempted to characterize these new variants of the PL

fimbriae by cloning them into *E. coli* K-12 strain, which allowed us to observe filamentous structures on the surface of the bacteria. The two clones were able to agglutinate different types of blood, form a biofilm at different temperatures, and adhere to human kidney and bladder cells. In *in vivo* assays, deletion of the *plf* operon did not appear to have an impact on the ability of either strain QT598 or UMEA3703-1 to colonize the bladder and kidneys of mice (mono-infection). However, the competitive fitness of the mutant was significantly lower in the kidney of CBA/J mice. Overall, PL fimbriae represent a new type of fimbriae that may contribute to adherence to host cells and tissues.

Keywords: APEC, UPEC, SPATE, Autotransporters, PL Fimbriae, Cytotoxicity, Adhesins, Urinary tract infection.

TABLE DES MATIÈRES

REMERCIEMENTS.....	IV
RÉSUMÉ	VI
ABSTRACT	VIII
TABLE DES MATIÈRES	XI
LISTE DES FIGURES.....	XIII
LISTE DES TABLEAUX	XIII
LISTE DES ABRÉVIATIONS.....	XIV
1 INTRODUCTION.....	2
2 REVUE DE LITTÉRATURE	5
2.1 <i>ESCHERICHIA COLI</i>	6
2.1.1 Généralités.....	6
2.1.2 Classifications	7
2.1.3 <i>E. coli</i> Pathogènes Intestinales.....	13
2.1.4 <i>E. coli</i> Pathogènes Extra-Intestinales (ExPEC)	18
2.1.5 Risque zoonotique.....	29
2.2 FACTEURS DE VIRULENCE CHEZ LES ExPEC	30
2.2.1 Fimbriae.....	34
2.2.2 Autotransporteurs	46
3 ARTICLE DE REVUE	49
4 HYPOTHÈSE ET OBJECTIFS DE RECHERCHES	81
5 PREMIER ARTICLE	85
6 DEUXIÈME ARTICLE	111
7 DISCUSSION GÉNÉRALE	136
7.1 AUTOTRANSPORTEUR SHA	138
7.1.1 Caractérisation des nouveaux membres de la famille des SPATEs	138
7.1.2 Rôle de l'AT Sha dans la virulence	141
7.2 FIMBRIAE PL.....	144
7.2.1 Caractérisation d'un nouveau membre de la famille des π -fimbriae.....	144
7.2.2 Rôle du Fimbriae PL dans la virulence.....	148
7.3 ROLE DE L'AUTOTRANSPORTEUR SHA ET LE FIMBRIAE PL DANS LA VIRULENCE DES SOUCHES D' <i>E. COLI</i> PATHOGENES EXTRA-INTESTINALES	150
8 CONCLUSION ET PERSPECTIVES	154
9 REFERENCES.....	159

10 ANNEXE I : FIGURES SUPPLEMENTAIRES DE L'ARTICLE 1	174
11 ANNEXE II : FIGURES SUPPLEMENTAIRES DE L'ARTICLE 2	188
12 ANNEXE III	193
13 ANNEXE IV.....	194
LISTE DE PUBLICATIONS	195

LISTE DES FIGURES

FIGURE 1: ARBRE PHYLOGENETIQUE REPRESENTANT LES DIFFERENTS GROUPES DU GENRE <i>ESCHERICHIA</i>	9
FIGURE 2: ARBRE PHYLOGENETIQUE REPRESENTANT LES DIFFERENTS PHYLO-GROUPES D' <i>E. COLI</i>	11
FIGURE 3: ARBRE PHYLOGENETIQUE MONTRANT LE PHYLO-GROUPE G PAR RAPPORT AUX AUTRES PHYLO-GROUPES D' <i>E. COLI</i>	12
FIGURE 4: MECANISMES D'INTERACTION DES DIFFERENTS PATHOTYPES D' <i>E. COLI</i> PATHOGENES INTRA-INTESTINALES.....	16
FIGURE 5 : MECANISMES D'INTERACTION DES AIEC ET DES EAHEC.....	18
FIGURE 6: PATHOGENESE DE L'INFECTION DE TRACTUS URINAIRE PAR LES UPEC.....	24
FIGURE 7: COLONISATION DE L'APPAREIL RESPIRATOIRE PAR LES APEC.....	27
FIGURE 8 : REPRESENTATION SCHEMATIQUE DES DIFFERENTS SEROGROUPES D' <i>E. COLI</i> PATHOGENES EXTRA-INTESTINALES.....	30
FIGURE 9: EVOLUTION BACTERIENNE PAR ACQUISITION DES FACTEURS DE VIRULENCE	33
FIGURE 10: ORGANISATION GENETIQUE DE L'OPERON <i>PAP</i>	38
FIGURE 11: BIOGENESE DES FIMBRIAE P.....	40
FIGURE 12: METHYLATION DES SITES DE LIAISON DE LA PROTEINE LRP PAR L'ENZYME DAM.....	42
FIGURE 13: PHASE OFF DE L'OPERON <i>PAP</i> : LA FIXATION DE LRP SUR LE SITE <i>GATC</i> ^{PROX} INHIBE LA TRANSCRIPTION DE L'OPERON <i>PAP</i>	42
FIGURE 14: PHASE ON DE L'OPERON <i>PAP</i>	43
FIGURE 15: PHASE OFF DE L'OPERON <i>PAP</i> : INHIBITION DE LA TRANSCRIPTION PAR LE REGULATEUR H-NS.....	43
FIGURE 16: PHASE OFF DE L'OPERON <i>PAP</i> : INHIBITION DE LA TRANSCRIPTION PAR LE REGULATEUR <i>PAPR</i>	44
FIGURE 17: REPRESENTATION SCHEMATIQUE DU PLASMIDE DE VIRULENCE <i>PEC598</i> RETROUVE CHEZ LA SOUCHE <i>QT598</i>	137
FIGURE 18: ORGANISATION GENETIQUE DE LA REGION UNIQUE DE 25KB RETROUVEE DANS LE PLASMIDE DE VIRULENCE <i>PEC598</i>	138

LISTE DES TABLEAUX

TABLEAU 1: CLASSIFICATION D' <i>E. COLI</i> SELON LE MANUEL DE BACTERIOLOGIE SYSTEMATIQUE DE BERGEY.....	6
--	---

LISTE DES ABRÉVIATIONS

A/T/C/G	Adénine / Thymine / Cytosine / Guanine
AAFs	Aggregative adherence fimbriae
ADN	Acide désoxyribonucléique
AIEC	<i>E. coli</i> adhérents et invasifs / Adherent-invasive <i>E. coli</i>
AMPc	Adénosine monophosphate cyclique
APEC	<i>E. coli</i> pathogènes aviaires / Avian pathogenic <i>E. coli</i>
ARN	Acide ribonucléique
ARNm	ARN messenger
ARNt	ARN de transfert
AT	Autotransporteur
ATP	Adénosine-5'-triphosphate
BHE	Barrière hématoencéphalique
CNF1	Cytotoxic necrotizing factor 1
CRP / CAP	cAMP receptor protein / Catabolite activator protein
CU	Chaperone –usher
DAEC	<i>E. coli</i> à adhérence diffuse / diffusely adherent <i>E. coli</i>
Dam	DNA adenine methylase
<i>E. coli</i>	<i>Escherichia coli</i>
EAEC	<i>E. coli</i> entéroaggrégatives / enteroaggregative <i>E. coli</i>
EAHEC	<i>E. coli</i> entéroaggrégative-hémorragiques/ enteroaggregative haemorrhagic <i>E. coli</i>
EAST1	EAEC heat-stable enterotoxin 1
ECOR	<i>E. coli</i> reference collection
EHEC	<i>E. coli</i> entérohémorragiques / enterohaemorrhagic <i>E. coli</i>
EIEC	<i>E. coli</i> entéroinvasives / enteroinvasive <i>E. coli</i>
EPEC	<i>E. coli</i> entéropathogènes / enteropathogenic <i>E. coli</i>
ETEC	<i>E. coli</i> entérotoxigènes / enterotoxigenic <i>E. coli</i>
ExPEC	<i>E. coli</i> pathogènes extra-intestinales / Extra-intestinal pathogenic <i>E. coli</i>
fimbriae PL	Fimbriae P-like
GbO3	Globotriosylcéramides
GbO4	Globotétraosylcéramides

GbO5	globopentaosylcéramides
GMPc	Guanosine monophosphate cyclique
Hbp	Hemoglobin protease
H-NS	Histone-like nucleoid structuring protein
HUS	Syndromes hémolytiques urémiques
IBC	Communautés bactériennes intracellulaires
InPEC	<i>E. coli</i> pathogènes intestinales / Intestinal pathogenic <i>E. coli</i>
ITU	Infection du tractus urinaire
Lésions A/E	Lésions d'attachement et d'effacement
LPS	Lipopolysaccharide
Lrp	Leucine responsive regulatory protein
LT	Toxines thermolabiles
MLEE	Multilocus enzyme electrophoresis
MLST	Multi-locus sequence typing
NMEC	<i>E. coli</i> associé à la méningite néonatale / Neonatal meningitis <i>E. coli</i>
PAI	Ilots de pathogénicité / Pathogenicity island
Pap	Pyelonephritis-associated pili
PCR	Réaction de polymérisation en chaîne / Polymerase chain reaction
PKA	Protéines kinases A
Prs	Pap-related sequence
Sha	Serine protease hemagglutinin autotransporter
SNP	Single nucleotide polymorphism
SPATE	Serine protease autotransporters of <i>Enterobacteriaceae</i>
ST	Sequence type
Stx	Shiga-toxines
Tsh	Temperature-sensitive hemagglutinin
UPEC	<i>E. coli</i> uropathogènes / uropathogenic <i>E. coli</i>
Vat	Vacuolating autotransporter toxin

UNITÉ DE MESURE

µg	Microgramme
µL	Microlitre

µm	Micromètre
µM	Micromolaire
Kb	Kilobase
kDa	Kilodalton
L	Litre
M	Molaire
mg	Milligramme
mL	Millilitre
mM	Millimolaire
nm	Nanomètre
pb	Paires de bases
rpm	Rotation par minute
UFC	Unité formant colonie

INTRODUCTION

1 INTRODUCTION

Escherichia coli (*E. coli*) est une bactérie bacille Gram-négatif appartenant à la classe des *Gamma-proteobacteria* et à la famille des *Enterobacteriaceae* (Boone *et al.*, 2005). Cette espèce est naturellement présente dans la flore intestinale des humains et des animaux en tant que bactérie commensale (Blount, 2015). Cependant, certaines souches peuvent devenir pathogènes si elles acquièrent des facteurs de virulence particuliers ou si le système immunitaire de l'hôte se trouve affaibli (pathogènes opportunistes), provoquant ainsi différentes pathologies. On distingue alors des pathogènes intestinaux (InPEC pour intestinal pathogenic *E. coli*) et des pathogènes extra-intestinaux (ExPEC pour extra-intestinal pathogenic *E. coli*) (Kaper *et al.*, 2004).

Les ExPEC peuvent coloniser le tractus intestinal de façon asymptomatique sans causer de maladies. Cependant, dès qu'ils atteignent un autre site, ils sont capables de provoquer des infections comme l'infection du tractus urinaire qui constitue un motif fréquent de consultation médicale et d'utilisation d'antibiotique (Flores-Mireles *et al.*, 2015, Foxman, 2014). Les responsables sont généralement des souches *E. coli* uropathogènes (uropathogenic *E. coli* UPEC) (Ronald, 2003) qui sont capables de coloniser la vessie et de causer une cystite (Lobel *et al.*, 2007, Mobley *et al.*, 2009) ou encore de se propager vers les reins et causer une pyélonéphrite qui peut aboutir dans certains cas en septicémie (Lobel *et al.*, 2007, Mobley *et al.*, 2009).

Les ExPEC peuvent aussi provoquer des infections extra-intestinales chez les animaux comme les volailles et causer une colibacillose aviaire (Fratamico *et al.*, 2018). Il s'agit d'une maladie contagieuse due le plus souvent à des souches d'*E. coli* pathogènes aviaires (avian pathogenic *E. coli* ou APEC) de plusieurs sérotypes incluant fréquemment les O1:K1, O2:K1 et O78:K80, qui sont hautement pathogènes. Les souches APEC utilisent une stratégie d'infection dont les points clés sont les suivants : (1) multiplication au niveau du tractus respiratoire supérieur, (2) colonisation des poumons et des sacs aériens, (3) migration vers les organes internes comme le cœur, le foie et la rate, et (4) développement d'une septicémie (Dho-Moulin *et al.*, 1999, Fratamico *et al.*, 2018).

Afin de survivre et de se multiplier dans l'hôte, les APEC et les UPEC doivent surmonter les premières lignes de défense du système immunitaire grâce à l'expression de plusieurs facteurs de virulence, parmi lesquels les toxines, les systèmes d'acquisition de fer, la capsule et certains types de lipopolysaccharides qui peuvent participer à l'évasion des défenses de l'hôte. On retrouve aussi les adhésines fimbriaires (fimbriae) ou afimbriaires qui sont responsables de

l'adhésion aux cellules épithéliales, les autotransporteurs qui peuvent jouer le rôle d'adhésine et/ou de cytotoxine qui provoquent des dommages à l'hôte (Dho-Moulin *et al.*, 1999, Fratamico *et al.*, 2018, Kaper *et al.*, 2004, Rodriguez-Siek *et al.*, 2005). De plus, des études ont montré que les souches ExPEC sont proches génétiquement, ce qui peut expliquer la présence des mêmes facteurs de virulence entre les souches ExPEC humaines et animales et par conséquent, le potentiel d'un risque zoonotique par transmission des animaux à l'homme et vice versa (Melha Mellata, 2013, M Mellata *et al.*, 2018).

Malgré les considérables progrès dans la lutte contre ces pathogènes, l'antibiothérapie demeure l'unique solution efficace. Toutefois, l'utilisation à long terme des antibiotiques pourrait conduire à l'antibiorésistance des souches, ce qui rend le traitement difficile (De Kraker *et al.*, 2011, Melha Mellata, 2013). De plus, en agriculture et en production animale, l'utilisation des antibiotiques est de plus en plus réglementée et il est donc important de trouver d'autres stratégies de traitement afin d'élaborer des nouveaux tests de diagnostic et d'améliorer la prophylaxie contre ce pathogène. Pour cela, les travaux se sont orientés vers l'étude des facteurs de virulence qui sont présents chez ces souches.

Ce projet de thèse vise à caractériser des nouveaux facteurs de virulence, un fimbriae de type P-Like et un autotransporteur de la famille SPATE Sha, et à comprendre davantage leur importance dans la virulence des souches ExPEC. Pour cela, nous avons d'une part, caractérisé ces facteurs en utilisant une approche d'analyse *in silico* ainsi qu'étudié leurs rôles en utilisant des tests phénotypiques, et d'autre part étudié leur expression et leur implication dans le processus pathologique *in vivo* en utilisant des modèles animaux. Ceci va nous permettre d'approfondir les connaissances par rapport à ces nouveaux facteurs de virulence.

REVUE DE LITTÉRATURE

2 REVUE DE LITTÉRATURE

2.1 *Escherichia coli*

2.1.1 Généralités

Escherichia coli est une bactérie à Gram-négatif non sporulante qui appartient à la classe des *Gamma-proteobacteria*, à la famille des *Enterobacteriaceae* (**tableau 1**) (Boone *et al.*, 2005, Whitman, 2015). La première description d'*E. coli* remonte à la fin du XIX^e siècle par le pédiatre allemand Theodor Escherich (1857-1911) qui a isolé des bactéries à partir des selles d'un enfant souffrant de diarrhée et les a nommées *Bacterium coli* commune. En 1919, Castellani et Chalmers ont proposé le nom *E. coli* en l'honneur du Dr Escherich, mais ce nom ne fut adopté qu'en 1954 (Castellani *et al.*, 1919, Whitman, 2015).

E. coli est l'espèce bactérienne la mieux connue et la plus étudiée vue les méthodes efficaces de manipuler son génome et sa croissance rapide, elle est en effet considérée comme un modèle d'étude fondamental de physiologie et métabolisme bactérien (Blount, 2015, Rasko *et al.*, 2008). C'est un bacille ayant un diamètre de 1.1 à 1.5 micromètre (μm) et mesurant entre 2.0 et 6.0 μm de long. Il se distingue par la présence de flagelles péritriches qui lui confèrent une certaine mobilité, et son métabolisme qui est à la fois respiratoire et fermentaire (Whitman, 2015).

Tableau 1: Classification d'*E. coli* selon le manuel de bactériologie systématique de Bergey.

Adapté de (Boone *et al.*, 2005).

Règne	<i>Bacteria</i>
Embranchement	<i>Proteobacteria</i>
Classe	<i>Gamma-proteobacteria</i>
Ordre	<i>Enterobacteriales</i>
Famille	<i>Enterobacteriaceae</i>
Genre	<i>Escherichia</i>
Espèce	<i>coli</i>

Par ailleurs, *E. coli* se définit comme une espèce versatile grâce à sa capacité d'adaptation et de survie vis-à-vis de différentes niches écologiques. Son réservoir naturel est l'intestin des humains et de certains animaux (mammifères, volailles et reptiles) avec une densité de 10^6 à 10^9 bactéries par mg de matière fécale. Elle est également présente dans les sols, les eaux, les plantes et les aliments (Blount, 2015, Leimbach *et al.*, 2013, Van Elsas *et al.*, 2011). À ce titre, sa présence dans les eaux ou dans les aliments reflète une contamination d'origine fécale et éventuellement une contamination par d'autres bactéries pathogènes (Elmund *et al.*, 1999).

Généralement, *E. coli* est considérée comme une bactérie commensale vivant en symbiose avec son hôte. Elle favorise ainsi l'équilibre de la microflore intestinale en empêchant l'invasion des pathogènes (Blount, 2015) et en produisant certains nutriments comme les vitamines K et B2 (des éléments indispensables au métabolisme de l'hôte) (Bacher *et al.*, 2000, Bentley *et al.*, 1982). Toutefois, la plasticité de son génome en termes d'acquisition et de perte de gènes peut rendre la bactérie pathogène : reconnaissance et adhérence plus efficace, invasion des tissus de l'hôte et sécrétion de toxines. Ceci peut être considéré comme un avantage sélectif par rapport à la souche commensale afin de coloniser d'autres niches écologiques et provoquer des maladies (Blount, 2015, Kaper *et al.*, 2004, Ochman *et al.*, 2000). Plusieurs techniques ont été utilisées afin de classer les différentes souches d'*E. coli* pathogènes selon leur groupe sérologique, phylogénétique, leur mode d'interaction avec l'hôte et les symptômes cliniques engendrés par l'infection. Une description de chaque classification sera présentée dans la prochaine section.

2.1.2 Classifications

2.1.2.1 Classification sérologique

La classification sérologique, fondée par Kauffmann en 1947 puis développée par Frits et Ida Ørskov (Kauffmann, 1947, Frits Ørskov *et al.*, 1992), consiste à classer les différentes souches d'*E. coli* selon leurs déterminants antigéniques, à savoir les antigènes somatiques O, les antigènes flagellaires H, les antigènes capsulaires K et parfois les antigènes fimbriaires F (voir section Fimbriae pour plus de détail) (Forsythe, 2020, Frits Ørskov *et al.*, 1992).

L'antigène somatique O dérive du mot allemand « *Ohne* ou *Ohne heuch* » qui signifie « sans » et fait référence aux formes non mobiles des bactéries sur les boîtes de gélose (non-flagelated ou non-swarming). Il fait partie des trois composants du lipopolysaccharide (LPS) : le lipide A, l'oligosaccharide central et l'antigène O (arrangé en unités répétées de glucosides) qui

sont indispensables, dans certains cas, à la résistance au sérum. L'antigène H dérive aussi du mot allemand « *Hauch* » qui signifie « souffle » et se réfère aux bactéries mobiles. Il est associé aux flagellines qui entrent dans la structure du flagelle (Forsythe, 2020, Frits Ørskov *et al.*, 1992, Poolman *et al.*, 2016). Enfin, l'antigène capsulaire K pour « *Kapsel* » ou capsule qui est généralement de nature polysaccharidique. Lorsqu'il est présent, il contribue avec d'autres facteurs (incluant l'antigène O) à la protection de la bactérie contre l'effet du complément. Certaines capsules peuvent avoir une structure similaire aux récepteurs présents à la surface des cellules hôtes, comme la capsule K1 qui possède une structure similaire aux glycoprotéines sialylées se retrouvant à la surface des cellules neuronales, ce qui permet à la bactérie d'échapper à la reconnaissance par le système immunitaire (Forsythe, 2020, Fratamico *et al.*, 2018, Frits Ørskov *et al.*, 1992).

L'antigène O seul définit le sérotype d'une souche (O1, O2, etc.) et la combinaison entre l'antigène O, l'antigène H et l'antigène K définit le sérotype d'une souche O:H:K incluant O1:H1, O111:H8, O1:K, etc. Le typage se fait par un test d'agglutination avec un antisérum. Cette technique est sensible et a permis d'identifier, jusqu'à présent, près de 180 antigènes O, 56 antigènes H et environ 80 antigènes K. Les différentes combinaisons entre les antigènes O, H et K ont permis d'identifier plus de 200 sérotypes, citons notamment les souches commensales (O1:K1:H1, O1:K1:H7 et O2:K1:H4), les souches qui causent des infections intestinales (O157:H7, O1:H20, O1:H7, O1:H2 et O1:H1) (Delannoy *et al.*, 2017, Fratamico *et al.*, 2018), et celles provoquant les infections urinaires, les méningites ou encore les septicémies (O1:K1:H7, O2:K1:H4, O2:K1:H5, O2:K1:H6 et O2:K1:H7) (Forsythe, 2020, Harry L. T. Mobley, 2009, Kauffmann, 1947, Frits Ørskov *et al.*, 1992, Poolman *et al.*, 2016).

2.1.2.2 Classification phylogénétique

Le genre *Escherichia* regroupe plusieurs espèces, dont *E. hermannii*, *E. fergusonii*, *E. vulneris*, *E. blattae*, et *E. albertii* qui diffèrent par leurs propriétés physicochimiques (Boone *et al.*, 2005). Grâce aux méthodes avancées de biologie moléculaire, des modifications ont été apportées au genre *Escherichia*. Selon la classification récente, le genre regroupe *E. coli* (ou *E. coli* sensu stricto), *E. albertii*, *E. fergusonii* et cinq clades cryptiques d'*Escherichia* (I–V). Le clade I est considéré comme une sous-espèce d'*E. coli*, les clades III et IV sont considérés comme des « sous-espèces » d'une nouvelle espèce des *Escherichia* et les clades II et V sont considérés comme deux nouvelles espèces (Fratamico *et al.*, 2018). Par ailleurs, *E. albertii* est l'espèce la plus divergente du genre *Escherichia* tandis que *E. fergusonii* est étroitement lié à *E. coli* sensu

stricto (Fratamico et al., 2018). Les autres espèces précédemment décrites (*E. hermannii*, *E. vulneris*, *E. blattae*) ont été reclassées du genre puisqu'elles partageaient un nombre minimal de similarités génétiques avec les autres espèces.

Dans cette thèse, l'acronyme *E. coli* sera utilisé en vue de désigner l'espèce classique ou *E. coli* sensu stricto (**figure 1**) (Fratamico et al., 2018).

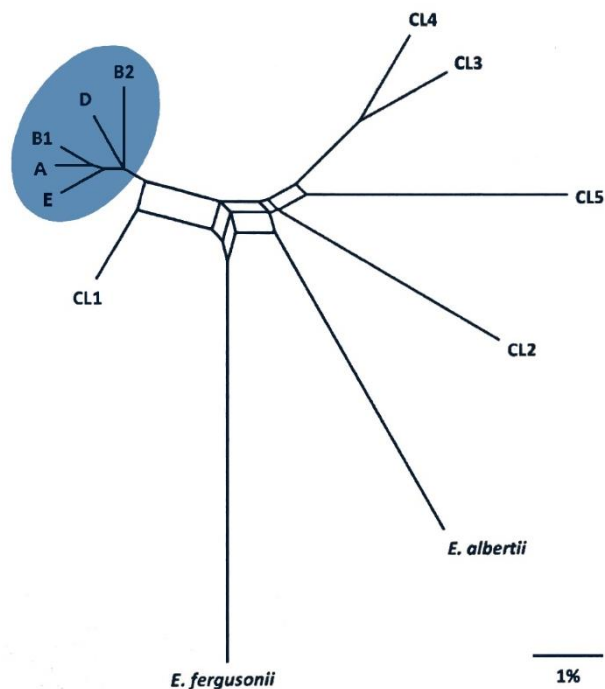


Figure 1: Arbre phylogénétique représentant les différents groupes du genre *Escherichia*.

E. coli sensu stricto est marqué en bleu, cinq clades cryptiques (CL1, CL2, CL3, CL4, et CL5), *E. albertii* et *E. fergusonii*. L'arbre a été construit en utilisant SplitsTree et la méthode de neighbour net en comparant le « core genome » et le polymorphisme d'un seul nucléotide (SNP) de différents groupes du genre *Escherichia*. Figure adaptée de (Fratamico et al., 2018).

Les différentes souches d'*E. coli* ont été classées par la suite en fonction de leur degré de parenté, en utilisant les données de « multilocus enzyme electrophoresis », ou MLEE, qui consiste à évaluer la mobilité électrophorétique des principales enzymes impliquées dans le métabolisme d'une espèce (Stanley et al., 2003). Le profil électrophorétique obtenu dépend essentiellement du poids moléculaire et de la charge globale de l'enzyme. La différence de mobilité des enzymes testées est généralement associée à la variation ou à la mutation génétique

au niveau du locus (Stanley *et al.*, 2003). Cette technique a été utilisée pour classer les différentes souches d'*E. coli* provenant de différents endroits géographiques de la collection « *E. coli* reference collection » (ECOR) afin de mettre en évidence la diversité génotypique de l'espèce. Elle a permis entre autres de classer les différentes souches en quatre groupes (A, B1, B2 et D) (Clermont *et al.*, 2013, Fratamico *et al.*, 2018).

Ensuite, grâce à l'utilisation des techniques plus spécifiques comme la « multi-locus sequence typing » (MLST) et la « méthode de Clermont », d'autres groupes ont été identifiés sur un nombre de souches plus important (A, B1, B2, C, D, E, F) (**figure 2**) (Clermont *et al.*, 2013). Les méthodes de MLST et de Clermont sont des adaptations de la méthode MLEE, elles permettent de détecter un plus grand nombre de variations que la MLEE (Clermont *et al.*, 2013). En effet, la MLST consiste au séquençage de régions de plusieurs loci codant pour des gènes de ménage (housekeeping genes) après les avoir amplifiés par une réaction de polymérisation en chaîne (PCR). Les résultats du séquençage correspondent alors à un profil caractéristique de la souche traitée qui sera soumis à une base de données avec un code et une "sequence type" (ST) (Clermont *et al.*, 2013). Quant à la méthode de Clermont, elle est basée sur la présence et l'absence des gènes *chuA*, *yjaA* et *arpA* ainsi que le fragment d'ADN TspE4.C2 (un gène codant pour l'enzyme lipide estérase) après les avoir amplifiés par PCR triplex/quadruplex. Cette méthode permet d'assigner correctement les différentes souches d'*E. coli* en comparaison avec la MLST.

Les données de l'arbre phylogénétique de différents phylogroupes ont montré que les groupes B2, F et D ont divergé en premier, suivis du groupe E, et plus récemment les groupes A, B1 et C (**figure 2**) (Clermont *et al.*, 2013, Clermont *et al.*, 2019). Les souches des groupes B2 et D sont majoritairement des pathogènes extra-intestinaux avec une diversité importante chez le groupe B2 (qui comprend 9 sous-groupes). Les souches des groupes A et B1, quant à elles, sont majoritairement des souches commensales, mais cela n'empêche pas certaines souches qui causent des infections intestinales de faire partie des groupes A, B1, D et E. Les souches du groupe C sont proches, mais distinctes, du groupe B1 (Clermont *et al.*, 2013, Clermont *et al.*, 2019).

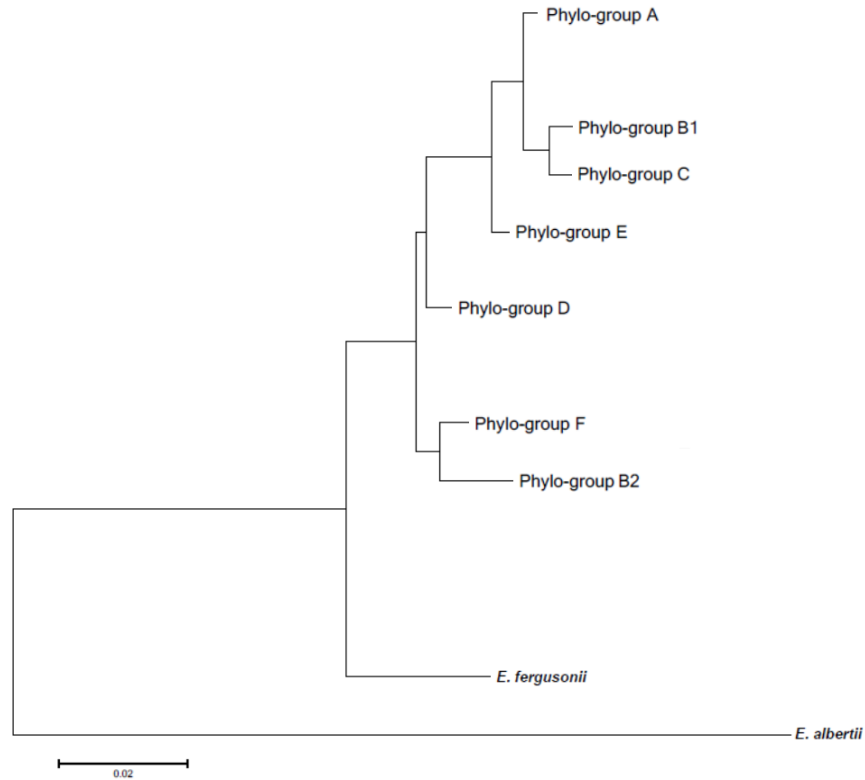


Figure 2: Arbre phylogénétique représentant les différents phylo-groupes d’*E. coli*.

L’arbre a été construit en comparant une partie de séquence génomique de 13 gènes (*aes*, *icd*, *pabB*, *polB*, *putP*, *trpA*, *trpB*, *adk*, *fumC*, *gyrB*, *mdh*, *purA*, *recA*) par le programme PhyML. Figure modifiée de (Clermont *et al.*, 2013).

Par ailleurs, avec le développement des techniques de séquençage, d’autres méthodes plus spécifiques ont été utilisées afin de classer les différentes souches d’*E. coli*, telles que le séquençage du génome (Whole genome sequence), la méthode de « ClermonTyping » qui est une méthode de PCR *in silico*, et le polymorphisme d’un seul nucléotide (single nucleotide polymorphism, SNP) (Clermont *et al.*, 2019). Ces différentes techniques ont permis de proposer un nouveau groupe « groupe G » qui se présente sous forme de groupe intermédiaire entre les phylogroupes F et B2 (**figure 3**). Les souches qui lui appartiennent présentent une résistance étendue aux antibiotiques et se retrouvent majoritairement chez le bétail (Clermont *et al.*, 2019).

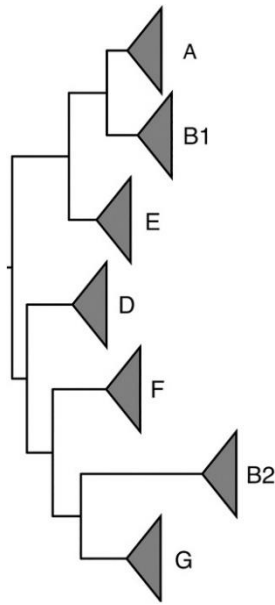


Figure 3: Arbre phylogénétique montrant le phylo-groupe G par rapport aux autres phylo-groupes d'*E. coli*.

L'arbre a été construit en comparant le « core genome » de 112 souches de groupe G avec 20 souches de chaque phylo-groupe A, B1, E, D, B2 et F en utilisant le programme Harvest suite of tools. Figure adaptée de (Clermont *et al.*, 2019).

2.1.2.3 Classification par pathotype

Les *E. coli* pathogènes peuvent être classés en pathotypes, faisant référence ainsi aux souches qui possèdent des facteurs de virulence spécifiques et qui sont capables de provoquer un type spécifique de maladie (Fratamico *et al.*, 2018). Nous distinguons deux groupes d'*E. coli* pathogènes, les *E. coli* pathogènes intestinales et extra-intestinales :

- Les *E. coli* pathogènes intestinales (InPEC) provoquent des gastro-entérites dont les symptômes les plus connus sont des diarrhées qui peuvent parfois être sanglantes, associées à des vomissements, des douleurs abdominales et de la fièvre (Fratamico *et al.*, 2018, Kaper *et al.*, 2004).
- Les *E. coli* pathogènes extra-intestinales (ExPEC) peuvent provoquer des infections urinaires, des méningites néonatales et des septicémies chez les humains et même chez les animaux (Dozois *et al.*, 2003, Fratamico *et al.*, 2018, Kaper *et al.*, 2004, Thomas A Russo *et al.*, 2003).

Ces deux groupes d'*E. coli* seront décrits dans les sections 2.1.3. et 2.1.4., respectivement.

2.1.3 *E. coli* Pathogènes Intestinales

Les *E. coli* Pathogènes Intestinales sont classés en 8 pathotypes en fonction des signes cliniques engendrés et des facteurs de pathogénicité exprimés. Les différents pathotypes seront décrits succinctement dans cette section.

- ***E. coli* entéro-pathogènes ou enteropathogenic *E. coli* (EPEC)**

Les infections aux EPEC représentent la cause majeure des diarrhées infantiles et des vomissements chez les enfants dans les pays en développement. L'infection par les EPEC entraîne des lésions d'attachement et d'effacement (A/E) qui se caractérisent par une adhésion étroite des bactéries à la surface des cellules épithéliales par l'intermédiaire de fimbriae « bundle-forming pili », une destruction locale des microvillosités et un réarrangement du cytosquelette pour former une structure en piédestal (**figure 4.a**). La majorité des facteurs de virulence des EPEC sont portés par un îlot de pathogénicité de 35 kb appelé locus d'effacement des entérocytes (LEE) qui est également présent chez d'autres souches comme les *E. coli* entérohémorragiques (Kaper *et al.*, 2004, Ochoa *et al.*, 2011). Les EPEC peuvent aussi causer des infections intestinales chez les animaux comme les lapins, les bovins, les chevaux, les chiens et les chats (Fratamico *et al.*, 2018, Kaper *et al.*, 2004, Moxley *et al.*, 1986, Ochoa *et al.*, 2011).

- ***E. coli* entérohémorragiques ou enterohaemorrhagic *E. coli* (EHEC)**

Les EHEC, connus aussi par le sérotype O157:H7, sont des agents pathogènes responsables de diarrhées sanglantes associées à des syndromes hémolytiques urémiques (HUS) et à une colite hémorragique (Karmali *et al.*, 1983, Nataro *et al.*, 1998). La transmission de l'infection se fait principalement par la consommation d'aliments contaminés, comme la viande hachée crue ou pas assez cuite ainsi que la consommation de lait cru. Le pouvoir pathogène des EHEC se manifeste par leur capacité à résister aux acides gastro-intestinaux, à adhérer aux cellules épithéliales, provoquant ainsi des lésions d'attachement et d'effacement (A/E) (**figure 4.b**), et par leur capacité à produire des Shiga-toxines (Stx) (Gouali *et al.*, 2013, Kaper *et al.*, 2004).

- ***E. coli* entérotoxigènes ou enterotoxigenic *E. coli* (ETEC)**

Les ETEC sont responsables des diarrhées aqueuses chez les enfants d'au moins trois ans dans les pays en voie de développement. Ils sont aussi responsables de 20 à 75% de ce que l'on appelle diarrhée du voyageur pour les personnes qui se déplacent dans les pays en voie de

développement (Fratamico *et al.*, 2018). La transmission de l'infection se fait principalement par la consommation d'aliments ou eaux contaminés par ETEC. Une fois dans l'intestin, les ETEC peuvent sécréter de toxines thermolabiles (LT) (oligomères AB₅) et/ou de toxines thermostables (STa ou STb) qui reconnaissent des récepteurs spécifiques présents à la surface des cellules épithéliales de la partie proximale de l'intestin grêle. Ceci provoque une augmentation des taux d'AMPc et/ou de GMPc intracellulaires et une activation des protéines kinases PKA et PKGII. Suite à cette activation, les canaux ioniques comme CFTR seront phosphorylés, ce qui entraîne une libération d'ions chlorure et la perte d'eau à travers la lumière intestinale provoquant ainsi la diarrhée (**figure 4.c**) (Croxen *et al.*, 2010, Janda *et al.*, 2006).

- ***E. coli* entéroaggrégatives ou enteroaggrégative *E. coli* (EAEC)**

Les *E. coli* entéroaggrégatives sont responsables de diarrhées aqueuses et parfois mucoïdes chez les enfants et les adultes dans les pays développés et en voie de développement (Hicks *et al.*, 1996). Le pouvoir pathogène des EAEC se caractérise par une adhésion bactérienne à la surface des cellules épithéliales du côlon grâce aux fimbriae AAFs (aggregative adherence fimbriae). Ils forment ensuite un biofilm, suivi de la sécrétion de toxines dans le milieu, ce qui entraîne l'induction d'une réponse inflammatoire (**figure 4.d**). Parmi les toxines sécrétées par les EAEC, nous distinguons les autotransporteurs comme Pet, Sat et Pic, qui possèdent des propriétés cytotoxiques, les entérotoxines comme EAST1 (pour EAEC heat-stable enterotoxin), ainsi que l'entérotoxine ShET1. Celle-ci est codée par le gène *set* auparavant identifié chez *Shigella flexneri*, d'où son nom. La sécrétion de ces toxines dans le milieu peut provoquer une nécrose hémorragique, un raccourcissement des microvillosités et la formation d'abcès cryptiques (Jensen *et al.*, 2014).

- ***E. coli* entéroinvasives ou enteroinvasive *E. coli* (EIEC)**

Les souches EIEC sont responsables de syndromes dysentériques qui se caractérisent par une fièvre, des crampes abdominales sévères, des malaises et des diarrhées aqueuses accompagnées d'un peu de sang (Van den Beld *et al.*, 2012). Au niveau génétique et biochimique, ce groupe d'*E. coli* est très similaire à celui des *Shigella* (Janda *et al.*, 2006). Les souches EIEC utilisent une stratégie d'infection dont les points clés sont les suivants : adhérence puis invasion des cellules épithéliales du côlon, suivi d'une lyse des endosomes qui les contiennent. Ceci entraîne la libération des bactéries dans le cytoplasme où elles pourront se multiplier et

subséquemment se propager aux autres cellules adjacentes de manière similaire (**figure 4.e**). (Janda *et al.*, 2006, Kaper *et al.*, 2004).

- ***E. coli* à adhérence diffuse ou diffusely adherent *E. coli* (DAEC)**

Les DAEC sont associées à des diarrhées aqueuses chez les enfants entre un et cinq an(s). Ce groupe est connu par son phénotype d'adhésion qui se diffuse aux cellules épithéliales humaines Hep-2, et il implique plusieurs facteurs de virulence comme l'autotransporteur « adhesin involved in diffuse adherence, AIDA-1, » les adhésines fimbriaires (Dr et F1845) et les afimbriaires (Afa). Les DAEC ont un effet cytotoxique sur les cellules intestinales. En effet, l'adhérence des bactéries aux cellules épithéliales par l'intermédiaire de fimbriae Dr, induit l'activation des voies de signalisation ce qui provoque la croissance des projections cellulaires en forme de doigt qui va s'enrouler autour de la bactérie (**figure 4.f**) (Kaper *et al.*, 2004, Servin, 2005).

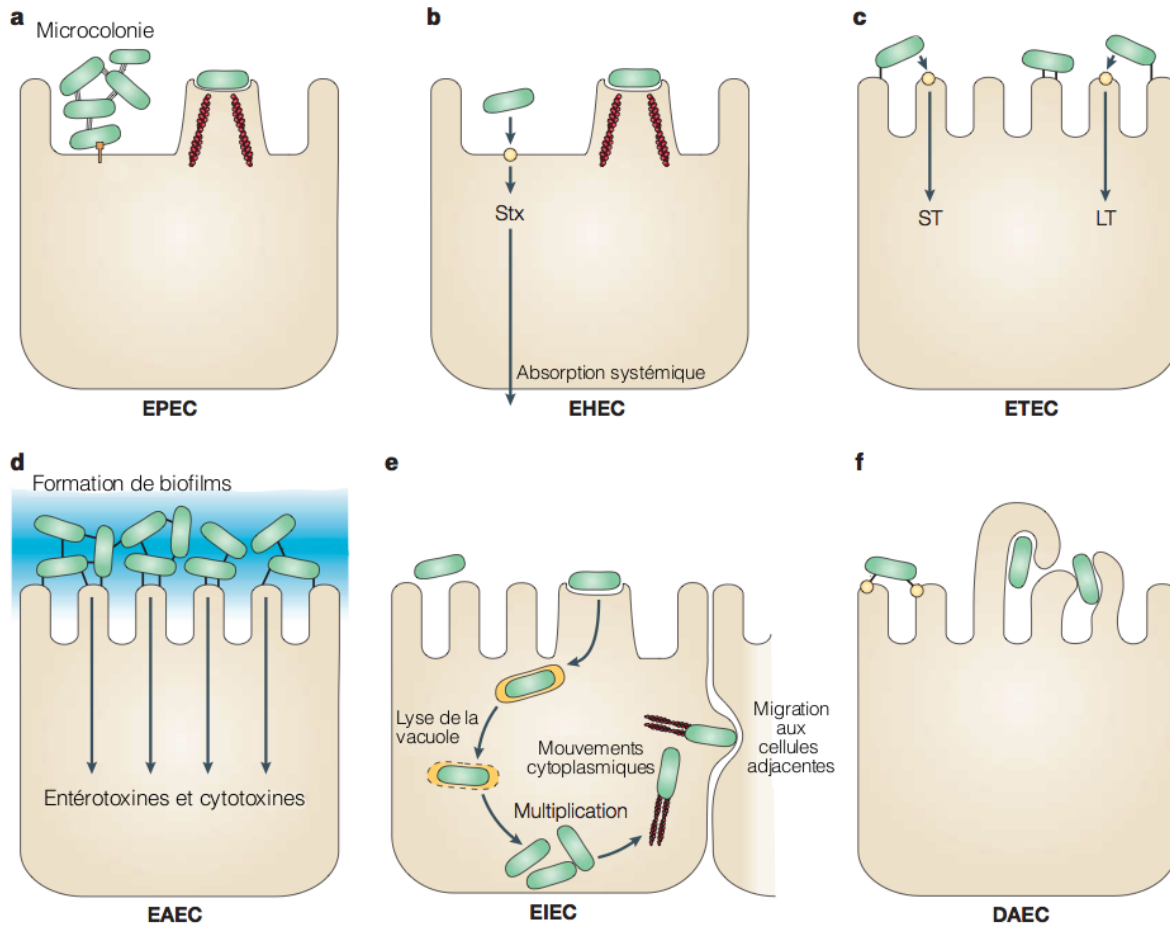


Figure 4: Mécanismes d'interaction des différents pathotypes d'*E. coli* pathogènes intra-intestinales.

- a)** Les EPEC provoquent des lésions d'attachement et d'effacement caractérisées par une adhérence étroite aux cellules épithéliales par la formation de micro-colonies, une destruction locale des microvillosités et un réarrangement du cytosquelette en forme de piédestal.
- b)** Les EHEC provoquent aussi des lésions d'attachement et d'effacement au niveau du côlon accompagnées d'une production de Shiga toxine (Stx).
- c)** Les ETEC provoquent une diarrhée aqueuse par la sécrétion de deux entérotoxines (ST et LT).
- d)** Les EAEC forment des biofilms et sécrètent des entérotoxines et des cytotoxines.
- e)** Les EIEC envahissent la cellule hôte et induisent une réponse inflammatoire destructrice.
- f)** Les DAEC induisent la croissance des projections cellulaires en forme de doigt qui enveloppent les bactéries. Figure modifiée de (Kaper *et al.*, 2004).

- ***E. coli* adhérents et invasifs ou adherent-invasive *E. coli* (AIEC)**

Les AIEC ont été décrits pour la première fois en 1999 (Boudeau *et al.*, 1999). Ils sont associés à la maladie de Crohn et aux diarrhées canines. Les AIEC peuvent adhérer et envahir les cellules intestinales via le fimbriae de type 1, ils peuvent aussi survivre et se multiplier dans les macrophages sans induire le mort cellulaire. Suite à une infection par les AIEC, une grande quantité de cytokine « tumoral necrosis factor ou TNF α » sera libérée dans le milieu, induisant ainsi une inflammation intestinale caractéristique de la maladie de Crohn (**figure 5.a**) (Croxen *et al.*, 2013, Kaper *et al.*, 2004, Miquel, 2010).

- ***E. coli* entéroaggrégative-hémorragique ou enteroaggregative haemorrhagic *E. coli* (EAHEC)**

Un nouveau pathotype de sérotype O104:H4 a été récemment identifié en Europe. Il s'agit d'un hybride entre les souches EHEC et EAEC nommée *E. coli* entéroaggrégative-hémorragique possédant la Shiga toxine 2 (caractéristique des EHEC), le plasmide pAA (caractéristique des EAEC) et l'antigène O102 codé par le gène *wzx*₁₀₄ (Fratamico *et al.*, 2018). Ces souches sont responsables de l'épidémie connue sous le nom du syndrome hémolytique et urémique (SHU) qui fit son apparition en Europe en 2011 avec plus de 4000 cas rapportés. Les symptômes se caractérisent par de sévères diarrhées sanglantes et parfois des troubles neurologiques qui peuvent entraîner la mort des sujets atteints (Frank *et al.*, 2011, Fratamico *et al.*, 2018, Scheutz *et al.*, 2011). Le pouvoir pathogène des EAHEC se caractérise par une adhésion bactérienne à la surface des cellules épithéliales grâce aux fimbriae AAFs. Ils forment ensuite un biofilm, suivi de la sécrétion de toxines dans le milieu comme les autotransporteurs Pet et Pic, ainsi que les Shiga-toxines (Stx) (**figure 5.b**) (Croxen *et al.*, 2013, Paletta *et al.*, 2020).

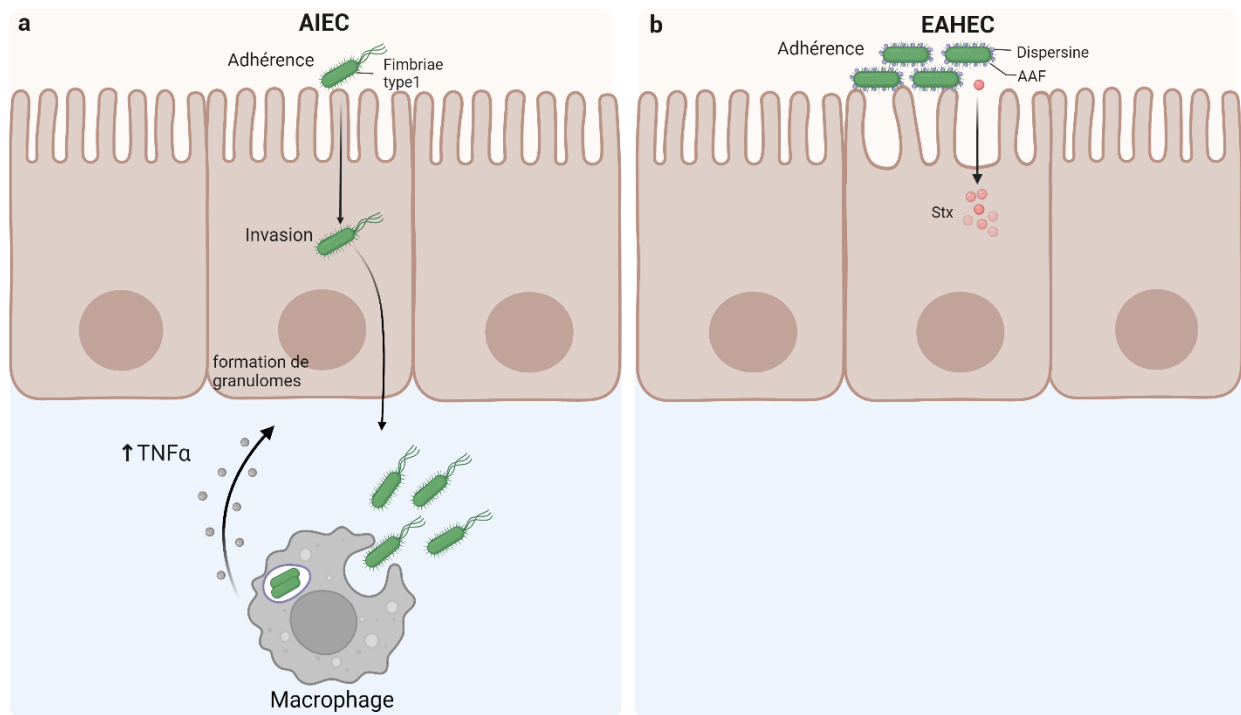


Figure 5 : Mécanismes d'interaction des AIEC et des EAHEC.

a. L'interaction des AIEC avec les cellules épithéliales permettant l'internalisation, la survie et la réplication des AIEC, ils peuvent aussi survivre et se multiplier dans les macrophages sans induire de mort cellulaire. Les macrophages infectés libèrent une grande quantité de cytokine « TNFα » dans le milieu, ce qui va induire une inflammation intestinale. Modifié de (Croxen *et al.*, 2013).

b. Le mécanisme d'interaction des EAHEC est similaire à celui des EAEC et EHEC. L'adhérence aux cellules épithéliales se fait par l'intermédiaire des fimbriae AAFs. Ils sécrètent aussi de toxines comme les Shiga-toxines (Stx) dans le milieu ce qui induit une inflammation intestinale. Figure modifiée de (Croxen *et al.*, 2013, Paletta *et al.*, 2020) et crée avec BioRender.

2.1.4 *E. coli* Pathogènes Extra-Intestinales (ExPEC)

Les différentes souches d'*E. coli* qui causent des infections extra-intestinales sont appelées ExPEC (T. A. Russo *et al.*, 2000). Elles se comportent comme des souches commensales lorsqu'elles se trouvent dans l'intestin, cependant, dès qu'elles colonisent d'autres organes, elles provoquent plusieurs maladies (infection urinaire, méningite, septicémie, etc.) qui peuvent entraîner la mort des sujets atteints (J. R. Johnson *et al.*, 2002). Ce groupe comprend les *E. coli* associées à la méningite néonatale et à la septicémie (MNEC), les *E. coli* associées à des infections du tractus urinaire ou UPEC (T. A. Russo *et al.*, 2000), ainsi que les pathogènes

aviaires APEC qui sont capables d'infecter les volailles et de causer une colibacillose aviaire (Dho-Moulin *et al.*, 1999, Rodriguez-Siek *et al.*, 2005).

La présente thèse se concentrant sur l'étude des souches ExPEC, ces dernières seront décrites plus en détail dans les prochaines sections.

2.1.4.1 *E. coli* associé à la méningite néonatale (NMEC)

La méningite néonatale est une inflammation des méninges chez les nouveau-nés, c'est-à-dire les enveloppes qui entourent leur moelle épinière et leur cerveau (El-Naggar *et al.*, 2019). Elle résulte le plus souvent d'une infection qui peut entraîner la mort ou des séquelles neurologiques d'autant plus importantes que le traitement aura été retardé (Durrmeyer *et al.*, 2012, Leperchois-Loritte, 2019). Celle-ci peut être d'origine virale, bactérienne, fongique ou parasitaire. Dans le cas d'une infection bactérienne, *E. coli* (NMEC) représente la cause la plus fréquente de méningite néonatale associée à Gram négatif après les streptocoques du groupe B (SGB) (S. Bonacorsi *et al.*, 2005a, S Bonacorsi *et al.*, 2001). Les souches isolées à partir des cas de méningite néonatale se différencient des autres souches d'*E. coli* par la présence de l'antigène capsulaire K1 (S. Bonacorsi *et al.*, 2005a) qui, suite à un mimétisme antigénique avec les glycoprotéines sialylées, facilite la survie dans le sang et le franchissement de la barrière hémato-encéphalique (BHE). De plus, les NMEC de sérotype K1 appartiennent essentiellement aux groupes B2 et D (groupes des pathogènes extra-intestinaux) et sont généralement associé aux sérogroupes O1, O8, O16, O45 et surtout O7 et O18 qui sont capables d'induire une bactériémie (E. Bingen, 2001, S. Bonacorsi *et al.*, 2005a, Kim, 2016, Wijetunge *et al.*, 2015).

Par ailleurs, les infections liées aux NMEC peuvent survenir à tous les stades de grossesse et même après la naissance des nouveau-nés, on parle alors d'une infection précoce ou materno-fœtales lorsqu'elle se révèle durant les premiers jours de la vie, ou alors d'une infection tardive lorsqu'elle se révèle au-delà du quatrième jour de vie (Edouard Bingen, 1999). La transmission de l'infection se fait généralement par voie hématogène à travers le cordon ombilical, par voie ascendante lors d'une infection urinaire, ou bien elle peut survenir lors de l'accouchement à travers le col de l'utérus infecté, dans ce cas le nouveau-né sera immédiatement colonisé par *E. coli* (Denis *et al.*, 2016, Leperchois-Loritte, 2019). Toutefois, la contamination post-natale est aussi possible à partir de l'environnement si les règles d'hygiène ne sont pas respectées (les personnels de maternité, les biberons, les matériels de perfusion, etc.) (Leperchois-Loritte, 2019).

Les manifestations cliniques apparaissent généralement quelques heures voire quelques jours après la naissance et peuvent inclure des signes non spécifiques tels que la fièvre, la perte d'appétit, les vomissements, ou encore des signes spécifiques tels que des signes neurologiques (convulsions et troubles de la conscience) (Barrington *et al.*, 2007, Stéphane Bonacorsi *et al.*, 2018, Leperchois-Loritte, 2019), respiratoires (geignement, détresse respiratoire) et cutanés (éruption, purpura) (Barrington *et al.*, 2007, Denis *et al.*, 2016). À ces signes, le nourrisson devrait être traité immédiatement après diagnostic de la ponction lombaire et un traitement aux antibiotiques doit être administré pour traiter l'infection, car tout délai de traitement augmente le risque d'issue négative (Barrington *et al.*, 2007, Stéphane Bonacorsi *et al.*, 2018, El-Naggar *et al.*, 2019).

- **Épidémiologie**

Les infections causées par les NMEC sont la principale cause de morbidité et mortalité chez les nouveau-nés. Leur incidence varie entre 4 et 8 cas pour 1000 naissances vivantes et ce risque augmente considérablement chez les prématurés en raison de l'immaturation de leur système immunitaire (Edouard Bingen, 1999, S. Bonacorsi *et al.*, 2005a, S Bonacorsi *et al.*, 2001). Aux États-Unis, les infections liées aux NMEC représentent entre 20 à 40% des cas de méningite néonatale avec un taux de mortalité de 8%, et environ 30 à 50% des survivants développent des séquelles neurologiques graves (Fratamico *et al.*, 2018). Au Canada, l'infection néonatale par *E. coli* représente une cause majeure de morbidité néonatale, selon une étude menée entre 2010 et 2016 (El-Naggar *et al.*, 2019).

- **Physiopathologie des infections aux NMEC**

Le franchissement de la BHE par les NMEC nécessite une bactériémie élevée et prolongée, soit un seuil supérieur à 10^3 UFC/mL de sang (conséquence à une translocation bactérienne suite à une infection urinaire ou digestive) (E. Bingen, 2001, S. Bonacorsi *et al.*, 2005a, Durrmeyer *et al.*, 2012, Kim, 2003). L'étape suivante consiste en l'adhésion des bactéries aux cellules endothéliales, qui est favorisée par la présence du fimbriae de type 1, codé par l'opéron *fim*, et celle du fimbriae S, codé par l'opéron *sfa*, en association avec la protéine de la membrane externe OmpA (Kim, 2003, Kim, 2016). Ces facteurs de virulence vont interagir avec différents récepteurs glycoprotéiques et glycolipidiques, tels que la protéine CD48, la protéine gp96, ou les groupements glycosidiques *N*-acétylglucosamine (GlcNAc) et α -sialyl-2.3- β -Gal de glycoprotéines présentes sur les cellules des capillaires des méninges ou sur les laminines de la

membrane basale et de l'épithélium des plexus choroïdes (E. Bingen, 2001, Kim, 2003, Kim, 2016). Ensuite, le mécanisme d'endocytose va être activé, entraînant ainsi l'internalisation des bactéries du pôle apical au pôle basolatéral, puis l'exocytose des bactéries se produira à travers la membrane basale des cellules endothéliales (S Bonacorsi *et al.*, 2001, Kim, 2003). Ce processus nécessite, d'une part, l'activation des protéines impliquées dans le processus d'invasion telles que la toxine cytotoxique necrotizing factor 1 (CNF1) et la protéine d'invasion IbeA (Kim, 2003, Kim, 2016), et, d'autre part, l'activation des voies de signalisation de la cellule hôte (ex. les protéines Rho), ce qui provoque le réarrangement du cytosquelette, la formation d'extensions membranaires et l'entrée de l'agent pathogène, sans pour autant causer de dégâts cellulaires (S Bonacorsi *et al.*, 2001, Kim, 2003, Kim, 2016).

2.1.4.2 *E. coli* uropathogènes (UPEC)

On parle d'infection du tractus urinaire (ITU) lorsqu'il y a contamination des voies urinaires (vessie, urètre, reins) par des microorganismes pathogènes. Dans 80% des cas, les ITU sont dues par *E. coli* uropathogènes (UPEC) (Ronald, 2003). D'autres espèces pathogènes, moins fréquentes, peuvent causer une ITU : streptocoques du groupe B, *Klebsiella pneumoniae*, *Proteus mirabilis*, *Staphylococcus aureus*, *Staphylococcus saprophyticus*, *Enterococcus faecalis*, *Pseudomonas aeruginosa*, et des champignons (généralement des espèces de *Candida*) (Foxman, 2014, Nielubowicz *et al.*, 2010, Ronald, 2003).

Des études ont montré que les UPEC qui causent une ITU chez l'homme appartiennent essentiellement aux groupes B2 et D (groupes des pathogènes extra-intestinaux) et sont associés aux sérogroupes O1, O2, O4, O6, O7, O8, O9, O15, O16, O18, O21, O22, O25, O50, O75, O83, O101 et O102 (Ida Ørskov *et al.*, 1985, Poolman *et al.*, 2016). Ils se différencient des *E. coli* commensaux par l'acquisition de nombreux facteurs de virulence, notamment les adhésines (fimbriae), les toxines (CNF1, Hly, etc.), les autotransporteurs, les systèmes d'acquisition du fer, les facteurs associés à la production du biofilm et d'autres facteurs de virulence qui sont généralement codés par des gènes situés sur les îlots de pathogénicité (PAI), conférant ainsi la capacité aux bactéries d'envahir les cellules hôtes (Croxen *et al.*, 2010).

Par ailleurs, l'infection de l'appareil urinaire regroupe différents types de maladies en fonction de la zone infectée (Lobel *et al.*, 2007). Une forme asymptomatique est possible malgré la détection dans les urines d'un niveau bactérien similaire à une infection urinaire, on parle dans ce cas d'une bactériurie asymptomatique (infection limitée aux urines) (Lobel *et al.*, 2007). En contrepartie, les UPEC peuvent coloniser la vessie et causer une cystite (infection des urines et

épithéliums vésicaux) qui se définit par des mictions fréquentes accompagnées par une sensation de brûlure, d'une hématurie et des douleurs ou inconfort supra-pubien (Lobel *et al.*, 2007, Mobley *et al.*, 2009). Elles peuvent ensuite se propager vers les voies supérieures jusqu'aux reins et causer une pyélonéphrite (bassin et parenchyme rénal) caractérisée par un malaise général accompagné par des douleurs abdominales, des nausées, des vomissements et de la fièvre (Lobel *et al.*, 2007, Mobley *et al.*, 2009).

L'ITU peut être une infection compliquée ou simple (Lobel *et al.*, 2007, Tony Mazzulli, 2012). Lorsque celle-ci survient chez des sujets ne présentant aucun facteur de risque, l'infection est souvent bénigne et elle est considérée comme simple ou non compliquée (Lobel *et al.*, 2007, Tony Mazzulli, 2012). Elle est dite compliquée si au contraire le sujet atteint présente un ou plusieurs facteurs de risques (les personnes ayant un tractus urinaire anormal, les immunodéprimés, les diabétiques, les femmes enceintes, etc.). Un traitement antibiotique est administré immédiatement après diagnostic de la culture d'urine afin d'éviter que l'infection se propage vers les reins (Lobel *et al.*, 2007, Tony Mazzulli, 2012).

- **Épidémiologie**

Les infections urinaires représentent un motif fréquent de consultation médicale qui aboutit généralement à un traitement antibiotique et des coûts estimés d'environ 3,5 milliards de dollars chaque année aux États-Unis (Flores-Mireles *et al.*, 2015, Foxman, 2014). *E. coli* est responsable d'environ 85 à 95% des cas de cystite non compliquée (environ 130 à 175 millions de cas dans le monde) et 90% des cas de pyélonéphrite (environ 5,4 millions de cas dans le monde) (Fratamico *et al.*, 2018, Thomas A Russo *et al.*, 2003). Une particularité épidémiologique de ces infections est de survenir préférentiellement chez les femmes. Ceci peut s'expliquer par le fait que l'urètre des femmes est plus court que celui des hommes, ce qui rend l'accès à la vessie plus facile pour les bactéries (Lobel *et al.*, 2007). En plus, il est à proximité de l'orifice anal qui est une source susceptible de contamination par des bactéries de la flore intestinale, y compris *E. coli* (Lobel *et al.*, 2007). On estime qu'environ 60% des femmes souffrent d'une infection urinaire au cours de leur vie et environ 30 à 50% ont un risque d'avoir une infection récurrente (Foxman, 2014). Toutefois, chez les enfants, le risque d'avoir une infection urinaire est plus élevé chez les garçons non circoncis que chez les garçons circoncis et les filles. Une étude a montré qu'environ 20.7% des garçons non circoncis de moins de trois mois ont eu une infection urinaire comparativement à ceux circoncis (2,4 %). Chez les filles de moins de trois mois, le risque d'avoir une infection est estimé à environ 7,5 % (Robinson *et al.*, 2014). Cette proportion inverse de celle

observée chez les adultes peut s'expliquer par le fait qu'*E. coli* colonise le prépuce des garçons non circoncis qui se propage facilement vers l'urètre et cause ainsi une ITU (Stephane Bonacorsi *et al.*, 2005b, Dubrovsky *et al.*, 2012).

- **Physiopathologie des infections aux UPEC**

Les ITU causées par les UPEC sont généralement des infections par voie ascendante (Mobley *et al.*, 2009). Elles se développent à travers la périurétrale jusqu'à la vessie et/ou les reins et se déroulent en plusieurs étapes (**figure 6.a**).

La première étape consiste à coloniser l'urètre distal et l'orifice vaginal par les bactéries de la flore intestinale, entre autres les UPEC (Mobley *et al.*, 2009). Puis se produit leur migration par l'intermédiaire des flagelles vers la lumière de la vessie où elles vont se lier via l'adhésine FimH, des fimbriae type 1, aux récepteurs glycoprotéiques (mannosylés) exprimés à la surface des cellules vésicales (comme les uroplakines) (Mulvey *et al.*, 1998). Cet attachement permet alors à la bactérie de sécréter la toxine CNF1 (Dufies *et al.*, 2021, Flatau *et al.*, 1997), dont la fonction est de catalyser la désamidation de la glutamine en position 63 en acide glutamique chez les protéines Rho (régulateurs du cytosquelette d'actine appartenant à la superfamille des protéines GTPasiques Ras) (Dufies *et al.*, 2021, Flatau *et al.*, 1997). La modification apportée à ce niveau va inhiber l'hydrolyse du GTP en GDP. Cette inhibition va entraîner une activation permanente des GTPases Rho, dont la dérégulation provoquera un réarrangement du cytosquelette d'actine, la formation de projections membranaires et l'internalisation bactérienne (Dufies *et al.*, 2021, Mulvey *et al.*, 1998).

Une fois à l'intérieur des cellules, les bactéries sont capables de se multiplier et de former des communautés bactériennes intracellulaires (IBC) (Anderson *et al.*, 2004, Justice *et al.*, 2004). Ces communautés sont capables de former une matrice d'exopolysaccharides entourant chaque bactérie, de façon similaire à une matrice de biofilm, ce qui confère une résistance vis-à-vis des stress et des antibiotiques (Anderson *et al.*, 2004, Justice *et al.*, 2004). La multiplication intracellulaire des IBC conduit à sa maturation, et à ce stade, les bactéries peuvent se détacher de l'IBC et se diriger vers le lumen de la vessie pour démarrer un nouveau cycle d'infection (**figure 6.b**). Ceci va provoquer une inflammation ou une cystite (Anderson *et al.*, 2004, Justice *et al.*, 2004).

La conséquence la plus grave d'une cystite est la propagation des bactéries vers les voies urinaires supérieures par un reflux vésico-urétéral. Dans ce cas, l'urine rétrograde depuis la vessie vers les uretères facilitant ainsi la dissémination des bactéries vers les reins (Mobley *et*

al., 2009). Une fois que les UPEC ont atteint les reins, elles vont se lier aux bassinets et aux parenchymes rénaux par l'intermédiaire des fimbriae comme les fimbriae P (qui reconnaissent les récepteurs digalactoside présents à la surface des cellules épithéliales rénales) (**figure 6.c**) (Kaper *et al.*, 2004, Korhonen *et al.*, 1986, Lobel *et al.*, 2007). Les bactéries vont aussi sécréter des toxines comme l'hémolysine et certains autotransporteurs cytotoxiques (Sat et Vat) ce qui va provoquer une inflammation rénale ou une pyélonéphrite aiguë (Kaper *et al.*, 2004). Si l'infection n'est pas traitée, les bactéries qui se trouvent dans les tubules rénaux peuvent traverser l'épithélium tubulaire proximal, rejoindre facilement la circulation sanguine (bactériémie) et conduire à une septicémie (Moblely *et al.*, 2009).

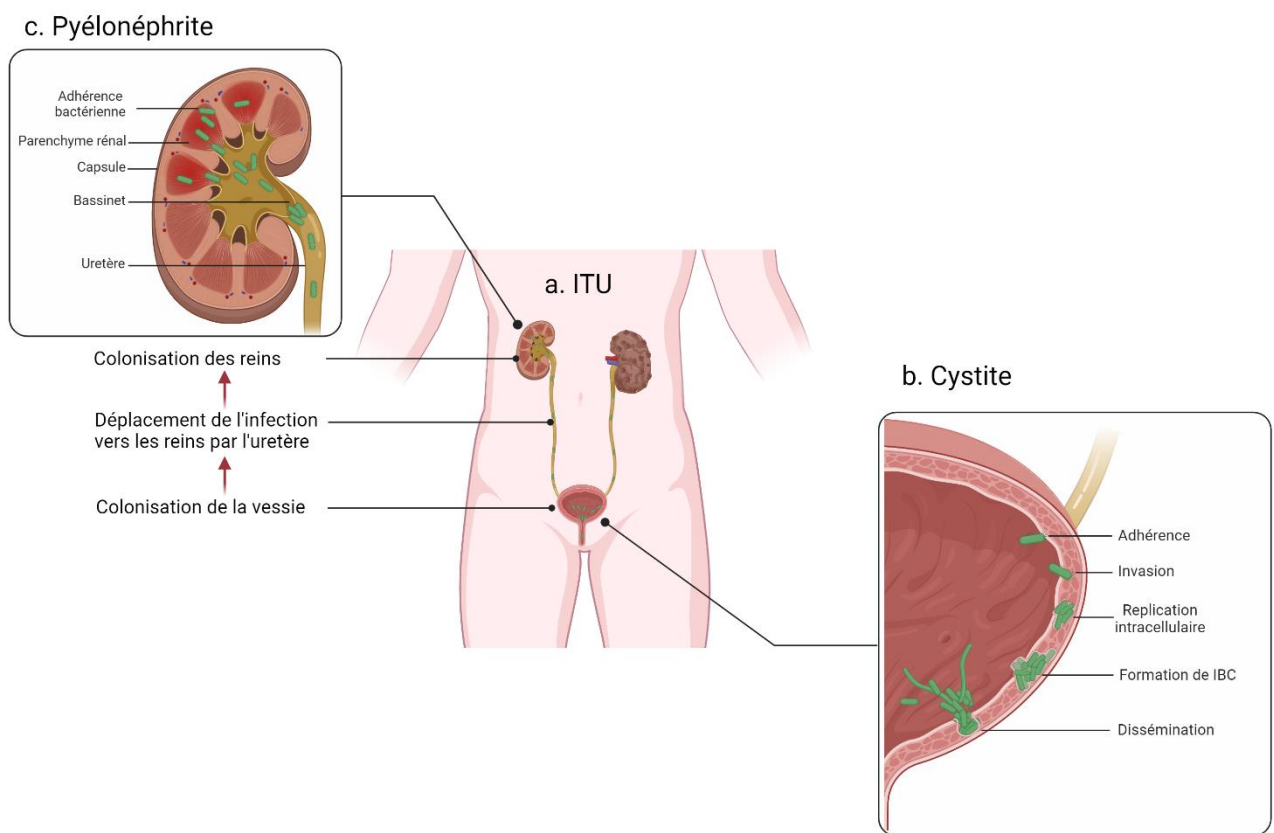


Figure 6: Pathogenèse de l'infection de tractus urinaire par les UPEC.

a) L'infection urinaire commence lorsque les UPEC colonisent la région périurétrale puis l'urètre. Ensuite, grâce aux flagelles, les bactéries vont coloniser la vessie et les reins.

b) Au niveau de la vessie, l'expression de fimbriae de type 1 et d'autres facteurs de virulence aboutit à l'adhérence puis à l'invasion des cellules vésicales. Une fois à l'intérieur, les bactéries commencent à se multiplier afin de former un IBC.

La maturation de l'IBC entraîne une dissémination bactérienne vers d'autres cellules, ce qui va entraîner une cystite. La rétrograde de l'urine vers les uretères va faciliter la migration des bactéries vers les reins.

c) Au niveau des reins, les bactéries vont adhérer aux cellules rénales grâce aux fimbriae et vont aussi sécréter de toxines, ce qui va provoquer une pyélonéphrite aiguë. Si l'infection n'est pas traitée, les bactéries peuvent rejoindre la circulation sanguine et causer une septicémie. Figure modifiée de (Kaper *et al.*, 2004, Rosen *et al.*, 2007) et créée avec BioRender.

2.1.4.3 E. coli pathogènes aviaires (APEC)

Les APEC font aussi partie des groupes des ExPEC (Fratamico *et al.*, 2018). Ils affectent généralement les volailles à savoir les poulets, les dindes et les canards (Melha Mellata *et al.*, 2003) et sont à l'origine de différentes formes de maladies regroupées sous le terme de « colibacilloses aviaires » (Dho-Moulin *et al.*, 1999). Par ailleurs, les souches APEC sont généralement présentes dans l'intestin des oiseaux et font partie intégrante de leur flore commensale. Certaines sont en revanche capables de provoquer une infection qui peut entraîner la mort de l'animal (infection primaire) et d'autres sont des opportunistes susceptibles de déclencher une infection lorsque le système immunitaire de l'hôte est affaibli (infection secondaire) (Dho-Moulin *et al.*, 1999, Kariyawasam *et al.*, 2019, Nolan *et al.*, 2003, Stordeur *et al.*, 2002). Par ailleurs, les souches APEC qui ont été isolées des cas de colibacilloses appartiennent essentiellement aux groupes A, B1, B2 et D (Melha Mellata, 2013) et sont fréquemment associées aux sérogroupes O1, O2 et O78. On trouve aussi d'autres sérogroupes comme O3, O4, O6, O8, O11, O15, O18, O21, O35, O50, O64, O74, O75, O87, O88, O95, O100, O103, O109, O115, O116, O119, O132, O141 et O152 qui ont été décrits par d'autres travaux (Dho-Moulin *et al.*, 1999, Kariyawasam *et al.*, 2019, Kunert Filho *et al.*, 2015, Saif *et al.*, 2011, Stordeur *et al.*, 2002).

Le pouvoir pathogène des souches APEC repose sur leur propriété à coloniser les voies respiratoires de l'animal et à surmonter les défenses de l'hôte (Alber *et al.*, 2020), ceci grâce à la présence de plusieurs facteurs de virulence qui sont généralement portés par des plasmides de virulence de type ColV (Huja *et al.*, 2015, Timothy J Johnson *et al.*, 2006) (leur nom vient de la présence de la microcine « colicine V » qui a été identifiée sur ces plasmide de virulence). Parmi ces facteurs, on peut citer les toxines (ex. *hlyF*) (Guabiraba *et al.*, 2015), les autotransporteurs (ex. *tsh*) (Dozois *et al.*, 2000, Provence *et al.*, 1994), la protéine résistante au sérum *iss* (increased serum survival) (Nolan *et al.*, 2003), la protéase de la membrane externe (*ompT*) (Dho-Moulin *et*

al., 1999, Timothy J Johnson *et al.*, 2006), la protéine de transfert conjugatif (*traT*) (Guabiraba *et al.*, 2015, Timothy J Johnson *et al.*, 2006), et les opérons *iroBCDEN* et *iucABCD* qui codent respectivement pour les salmochélines et les aérobactines (systèmes d'acquisition de fer) (Caza *et al.*, 2008, Dozois *et al.*, 2003). En plus, on peut trouver les fimbriae (ex. type1, P) (La Ragione *et al.*, 2000) et l'antigène capsulaire K1 qui sont généralement associés à des infections systémiques chez la volaille (Dho-Moulin *et al.*, 1999, Guabiraba *et al.*, 2015, Nolan *et al.*, 2003, Stordeur *et al.*, 2002).

- **Épidémiologie**

Les poulets d'élevage et les volailles destinées à la consommation sont fréquemment contaminés par *E. coli* pathogène puisque ces animaux peuvent porter la bactérie dans leurs intestins (Saif *et al.*, 2011). Ceux qui sont infectés peuvent ensuite transmettre la maladie aux autres animaux sains, provoquant ainsi d'importantes pertes économiques dans le secteur avicole (Melha Mellata, 2013, Saif *et al.*, 2011). Plus de 80% des cas déclarés de colibacillose ont été associés aux sérogroupes O1, O2 et O78 qui sont hautement pathogènes (Huja *et al.*, 2015, Kathayat *et al.*, 2021). Ils peuvent infecter les poules de chair et les poules pondeuses tout au long du cycle d'élevage, ce qui entraîne une diminution de la production de viande (environ 2%), des œufs (environ 20%), un retard de croissance, un taux de mortalité élevé qui peut atteindre jusqu'à 53,5% chez les jeunes poulets et un motif fréquent de saisie à l'abattoir (entre 36% à 43%) (Guabiraba *et al.*, 2015, Kathayat *et al.*, 2021, Melha Mellata, 2013).

- **Physiopathologie des infections aux APEC**

Le mécanisme d'infection des souches APEC n'est pas bien connu. Cependant, il est admis que la voie d'entrée principale du pathogène est le tractus respiratoire par inhalation des poussières de litière contaminées par *E. coli* ou encore par contact direct avec les animaux malades, y compris avec leurs sécrétions ou leurs excréments (Brugère-Picoux, 2017, Stordeur *et al.*, 2002). La première étape de colonisation consiste alors à adhérer aux cellules épithéliales du pharynx et de la trachée via les adhésines (comme les fimbriae type 1 et Stg) (Guabiraba *et al.*, 2015, Melha Mellata *et al.*, 2003, Stordeur *et al.*, 2002). Les bactéries vont ensuite coloniser rapidement les poumons et les sacs aériens à cause de leur structure fragile et leur faible vascularisation (pénurie des cellules immunitaire) ce qui va provoquer des lésions d'aérosacculite (Melha Mellata *et al.*, 2003). Puis, ils vont rejoindre la circulation sanguine pour coloniser les organes internes, à savoir le cœur (péricardite), le foie (périhépatite) et la rate, et enfin, développer une septicémie (**figure 7**) (Stordeur *et al.*, 2002).

La résistance bactérienne à l'effet des phagocytes constitue une étape primordiale pour la survie des souches APEC et l'établissement d'une infection (Alber *et al.*, 2020). Cet effet est médié par les différentes structures bactériennes comme la capsule, le lipopolysaccharide, la production de sidérophores, de toxines, et les protéines de la membrane externe (Guabiraba *et al.*, 2015, Nolan *et al.*, 2003, Stordeur *et al.*, 2002).

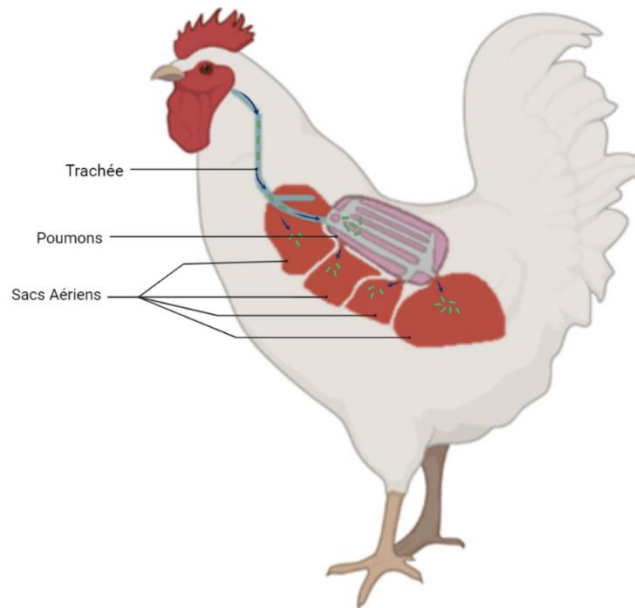


Figure 7: Colonisation de l'appareil respiratoire par les APEC.

Après inhalation des poussières contaminées par *E. coli*, les bactéries infectent les voies respiratoires supérieures (trachée et pharynx) puis les voies respiratoires profondes (poumons et sacs aériens). Si l'infection n'est pas traitée, les APEC vont coloniser les organes internes et développer une septicémie. Les flèches (en bleu) indiquent la voie d'entrée des APEC (en vert). Figure modifiée de (Kathayat *et al.*, 2021) et crée avec BioRender.

Contrairement aux autres infections décrites précédemment, la colibacillose aviaire prend plusieurs formes (localisée, génitale, chronique, etc.) et les signes cliniques de cette maladie varient considérablement en fonction de l'âge de l'animal et des éventuelles lésions produites par l'infection (Saif *et al.*, 2011). On distingue alors :

Septicémie et complexe respiratoire chronique : c'est la forme principale de la colibacillose aviaire. Elle affecte particulièrement les poulets de chair, les faisans, les canards et les dindes et est associée à un taux élevé de morbidité et de mortalité chez les animaux entre 2 à 12 semaines

(Dho-Moulin *et al.*, 1999, Stordeur *et al.*, 2002). Elle se développe suite à une infection aux mycoplasmes, à une bronchite infectieuse, à un accident de vaccination ou à un taux élevée d'ammoniac dans l'air (Stordeur *et al.*, 2002). Les animaux atteints présentent des symptômes respiratoires non spécifiques (toux, râles, éternuements, jetage nasal et sinusites, respiration accélérée et irrégulière) et des lésions inflammatoires plus ou moins exsudatives au niveau des sacs aériens, du foie, du cœur et de la cavité abdominale (péritonite) (Saif *et al.*, 2011, Stordeur *et al.*, 2002).

Omphalite ou infection du sac vitellin : l'infection peut se transmettre à l'œuf verticalement lors de son passage par le cloaque infecté ou, rarement, lors d'une infection du tractus génital (salpingite ou ovarite) ou encore en tombant sur une litière sale (Saif *et al.*, 2011, Stordeur *et al.*, 2002). Les bactéries vont ensuite pénétrer et contaminer la membrane vitelline ce qui va entraîner une infection appelée omphalite (Saif *et al.*, 2011). Cette forme de la maladie entraîne un taux de mortalité élevé chez les poussins âgés de moins d'une semaine. Une aérosacculite et une péricardite sont parfois associées à cette forme chez ceux qui passent le cap de 3 semaines. Le traitement antibiotique réussit à garder des poussins en vie mais il ne guérit pas l'infection (Stordeur *et al.*, 2002).

Dermatite nécrotique : elle se manifeste par l'apparition de lésions sous la peau principalement dans la partie inférieure de l'abdomen, cette forme n'entraîne pas la mort de l'animal, mais est responsable de pertes économiques substantielles (Stordeur *et al.*, 2002).

Salpingites et ovarites : il s'agit d'une inflammation de l'oviducte qui achemine les ovocytes de l'ovaire vers l'utérus. Cette forme de la maladie entraîne la mort de l'animal et est souvent associée à des infections respiratoires causées par *E. coli* (Melha Mellata, 2013, Stordeur *et al.*, 2002)

Swollen head disease : c'est une forme d'inflammation aiguë localisée au niveau des cellules de la peau et du tissu sous-cutané de la tête et des régions périorbitaires (Kabir, 2010). Cette forme de colibacillose peut se développer en quelques semaines après une infection initiale (par exemple, suite à une infection virale par pneumovirus, paramyxovirus ou coronavirus) et les animaux qui présentent ces symptômes meurent rapidement (Stordeur *et al.*, 2002).

Maladie de Hjärre (ou Granulomes à *E. coli*) : est une forme particulière de colibacillose qui se développe sous forme des granulomes dans plusieurs organes (foie, caecum, duodénum et le mésentère). Cette forme de la maladie entraîne la mort de l'animal suite à la rupture de ces granulomes (Saif *et al.*, 2011, Stordeur *et al.*, 2002).

2.1.5 Risque zoonotique

Les zoonoses en termes de l'Organisation mondiale de la santé, sont des maladies infectieuses qui se transmettent naturellement de l'animal à l'homme (Organization, 1992). Par ailleurs, certaines de ces zoonoses sont les conséquences de l'exposition à un agent biologique (bactéries, champignons, parasite, virus, etc.) dans le cadre d'une activité professionnelle (les éboueurs, les agriculteurs, les éleveurs, les vétérinaires, etc.) (Abadia, 2008). La transmission de l'infection se fait alors après un contact avec un animal qu'il ait été malade ou non (les animaux d'élevage, de compagnie ou sauvages), les produits qui en sont issus (la viande, le lait, les œufs, etc.) et/ou son environnement (abattoirs, eaux usées, déchets, etc.) (Abadia, 2008). Les effets des zoonoses sur la santé humaine sont très variables parfois bénignes, parfois mortels. Ils dépendent généralement de la nature du pathogène en cause et de certains facteurs individuels (femmes enceintes, personnes âgées, immunodéprimées, etc.) (Abadia, 2008).

Les zoonoses les plus courantes sont celles d'origine bactérienne comme dans le cas des souches APEC qui peuvent coloniser la vessie et les reins et causer une ITU (dans des modèles d'infection ascendante de tractus urinaire chez les souris). Les analyses phylogénétiques effectuées, ont montré que les différentes souches ExPEC sont proches génétiquement. Cela peut expliquer pourquoi certaines souches APEC peuvent causer une ITU puisqu'elles possèdent les mêmes facteurs virulence que les souches UPEC et/ou NMEC, et parfois peuvent partager le même sérotype ou le groupe phylogénétique (**figure 8**) (Jørgensen *et al.*, 2019, Kariyawasam *et al.*, 2019, Kathayat *et al.*, 2021). Ces similarités suggèrent que les souches APEC pourraient présenter un potentiel risque zoonotique sur la santé humaine.

La seule thérapie efficace contre ces pathogènes est l'antibiothérapie. Toutefois, l'utilisation intensive d'antibiotiques dans le secteur d'élevage favorise l'apparition des souches antibiorésistantes ce qui peut entraîner un problème de santé publique (Kathayat *et al.*, 2021, Melha Mellata, 2013, M Mellata *et al.*, 2018). La gestion de ce problème représente un moyen de réduire le risque zoonotique de transmission des agents pathogènes ainsi que le maintien de la santé humaine. L'exploitation des nouvelles stratégies (comme le développement des vaccins) pourrait limiter l'utilisation des antibiotiques, ce qui amoindrirait le risque de résistance microbienne et apporterait aux éleveurs une alternative à l'utilisation de ces antibiotiques.

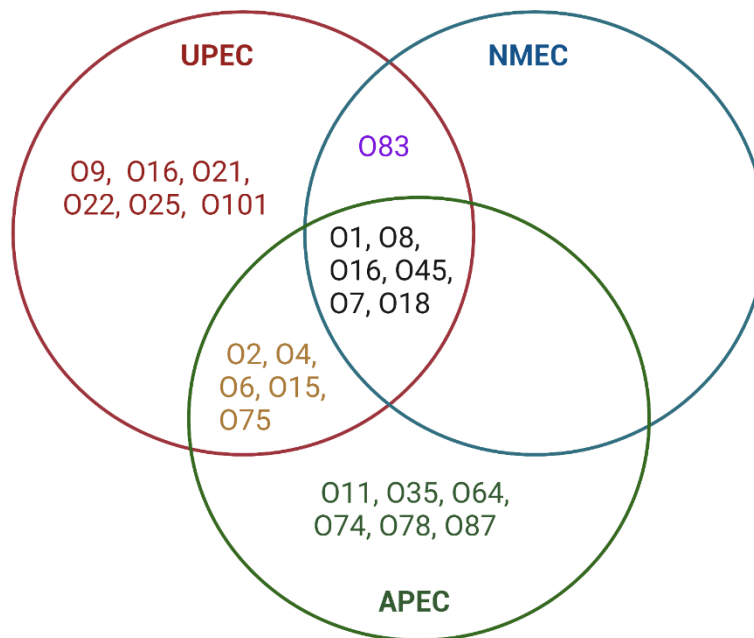


Figure 8 : Représentation schématique des différents sérogroupes d'*E. coli* pathogènes extra-intestinales.

Chaque cercle de couleur représente un pathotype : UPEC en rouge ; NMEC en bleu ; APEC en vert. Les sérogroupes qui sont en commun entre les UPEC, les NMEC et les APEC sont représentés en noire ; les sérogroupes qui sont en commun entre les UPEC et les NMEC sont représentés en mauve ; les sérogroupes qui sont en commun entre les UPEC et les APEC sont représentés en jaune. Figure modifiée de (Dho-Moulin *et al.*, 1999, Poolman *et al.*, 2016).

2.2 Facteurs de virulence chez les ExPEC

Il est admis que l'acquisition de certains facteurs de virulence entraîne une grande plasticité dans le génome bactérien en facilitant l'adaptation des pathogènes aux différentes niches écologiques et en favorisant leur action dans l'invasion des tissus de l'hôte. Cette diversité génétique est généralement due au transfert horizontal des gènes de virulence qui sont portés par des éléments génétiques mobiles (sont caractérisés par leur capacité à se déplacer), jouant ainsi un rôle dans la diversification des génomes (Kaper *et al.*, 2004, Rankin *et al.*, 2011). Parmi ces éléments on peut citer :

- **Les plasmides de virulence**

Les plasmides sont des molécules d'ADN circulaire ou parfois on les retrouve intégrés dans le génome bactérien (épisode). Leur réplication se fait de manière autonome mais ils dépendent entièrement de la machinerie bactérienne (enzymes, précurseurs métaboliques...). Ils peuvent se transmettre régulièrement aux bactéries filles durant la division cellulaire ou vers d'autres bactéries par le phénomène de conjugaison (Rankin *et al.*, 2011). Ainsi, ces plasmides de virulence confèrent aux bactéries qui les hébergent des propriétés supplémentaires qui sont nécessaires à l'établissement d'une infection chez l'organisme hôte. Parmi les principaux plasmides de virulence décrits dans la littérature, on peut citer les plasmides de types ColV. Ils se trouvent généralement chez les APEC et certaines ExPEC qui présentent de grandes similitudes génétiques (Huja *et al.*, 2015, Timothy J Johnson *et al.*, 2006) (**Figure 9**).

- **Les bactériophages**

Les bactériophages ou les phages sont des virus ubiquitaires qui infectent spécifiquement les bactéries. Lorsque les phages sont tempérés, ils peuvent transférer les gènes de virulence aux bactéries en leur conférant de nouvelles propriétés biologiques, on parle alors du phénomène de transduction (Rankin *et al.*, 2011, Wagner *et al.*, 2002). Comme tous les virus, les phages ont besoin d'un hôte pour se propager. La première étape de leur cycle commence par la reconnaissance et la fixation du virus aux récepteurs spécifiques sur la membrane bactérienne, suivie d'une injection et intégration du génome viral à celui de l'hôte. Une fois intégré, il va rester longtemps en état de dormance, sous forme de prophage, avant d'être excisé du génome bactérien suite à une induction et entamer un cycle lytique (Gandon, 2016, Rankin *et al.*, 2011). Lors de l'excision, le prophage va se répliquer à l'intérieur de la bactérie pour former des particules virales ce qui entraîne la lyse de la membrane bactérienne et la libération des virions qui peuvent infecter une autre bactérie et recommencer le cycle (Gandon, 2016, Rankin *et al.*, 2011). Toutefois, le virus peut encapsider un fragment d'ADN bactérien qui code pour un ou des facteurs de virulence et les transmettre à une autre bactérie. Un exemple bien connu est celui de la souche EHEC EDL933 qui possède dans son génome 18 prophages différents, incluant les prophages CP-933V et BP-933W qui codent respectivement pour les toxines Stx1 et Stx2 (Boyd *et al.*, 1998, Perna *et al.*, 2001, Whittam *et al.*, 1993) (**Figure 9**).

- **Les séquences d'insertion**

Les séquences d'insertion sont des courtes séquences d'ADN encadrées par des séquences répétées et inversées, qui permettent leur reconnaissance. L'intégration de ces séquences dans le génome peut modifier l'expression de certains gènes et donc modifier le phénotype de la bactérie (Mahillon *et al.*, 1998). Un exemple de séquences d'insertion est celui des gènes codant pour les toxines STa chez les ETEC qui peuvent être insérés soit dans les plasmides, soit dans le chromosome (Kaper *et al.*, 2004) (**Figure 9**).

- **Les îlots de pathogénicité**

Les îlots de pathogénicité (PAI pour pathogenicity islands) sont des régions génomiques mobiles regroupant différents facteurs de virulence. Ces îlots sont généralement présents dans les génomes des souches pathogènes (Hacker *et al.*, 2000, Paolozzi *et al.*, 2021). Ils sont souvent insérés à proximité d'un ARNt du génome bactérien. Ils possèdent des séquences répétées (*attL* et *attR*) flanquant leurs extrémités gauche et droite et entre ces extrémités on trouve des gènes qui codent pour l'intégrase et les facteurs de virulence (Hacker *et al.*, 2001, Hacker *et al.*, 2000, Paolozzi *et al.*, 2021). L'exemple le plus connu des PAIs est celui des gènes codant pour les fimbriae P (*pap* et *prs*) chez la souche UPEC J96. Les fimbriae P sont portés par deux PAI dont les longueurs sont de 170 kb (PAI-I : porteur de l'opéron *pap* et l'hémolysine α (*hly*)) et de 110 kb (PAI-II : porteur de l'opéron *prs* avec l'opéron *hly* et la toxine (*cnf1*)) (Sarowska *et al.*, 2019, Schmidt *et al.*, 2004) (**Figure 9**).

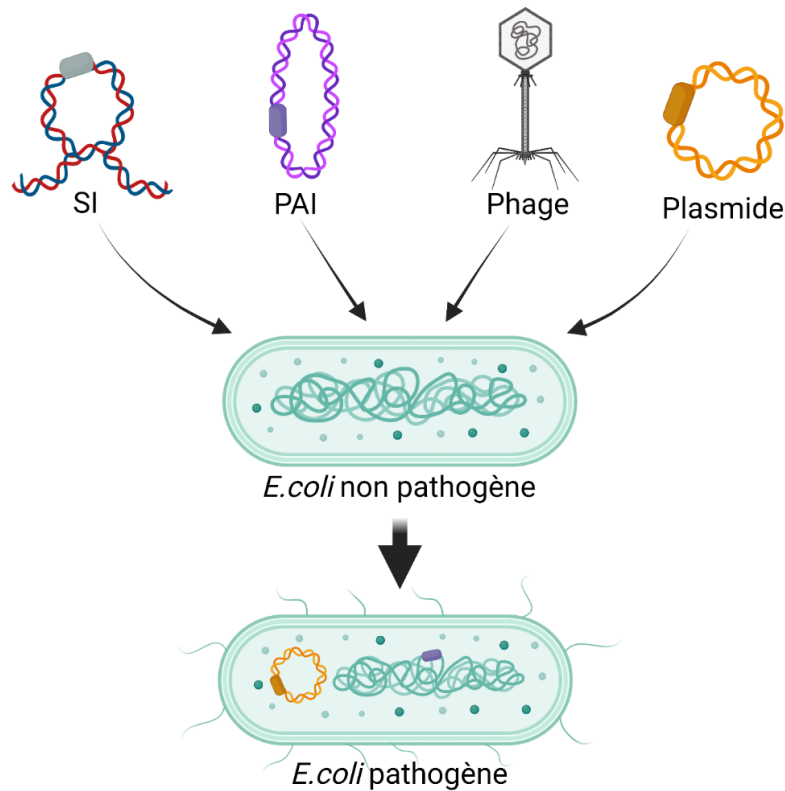


Figure 9: Evolution bactérienne par acquisition des facteurs de virulence

L'acquisition de certains facteurs de virulence portés par des éléments génétiques mobiles (plasmides de virulence, bactériophages, PAIs, séquences d'insertion, etc.) permet à la bactérie pathogène bactéries de mieux s'implanter, survivre et se multiplier dans l'organisme hôte. Figure modifiée de (Kaper *et al.*, 2004) et créée avec BioRender.

Vue la grande similitude génétique entre les souches ExPEC, plusieurs facteurs de virulence ont été identifiés et partagés à l'intérieur de ce groupe. On peut citer :

- Les flagelles qui facilitent le mouvement des bactéries (Fratamico *et al.*, 2018, Macnab, 1992).
- La sécrétion de toxines qui entraînent la formation des pores dans la membrane plasmique des cellules hôtes et la libération des nutriments dans le milieu extracellulaire (Fratamico *et al.*, 2018).
- La résistance au sérum qui est un facteur important pour la survie des pathogènes. Cet effet est généralement médié par différentes structures bactériennes comme la capsule, le lipopolysaccharide et les protéines résistantes au sérum (Fratamico *et al.*, 2018).

- Les systèmes d'acquisition de fer ou sidérophores qui permettent aux bactéries de séquestrer le fer à partir des cellules de l'hôte, un élément nécessaire à la survie bactérienne (Garenaux et al., 2011).
- Les adhésines qui donnent aux bactéries une grande capacité à adhérer aux tissus de l'hôte comme les fimbriae P (Thanassi *et al.*, 2007).
- Les autotransporteurs qui peuvent agir soit comme des adhésines soit comme des toxines (Pokharel *et al.*, 2019).

Les fimbriae P et les autotransporteurs seront décrits plus en détail dans les prochaines sections.

2.2.1 Fimbriae

2.2.1.1 Généralités

L'étape primordiale de la colonisation par *E. coli* pathogène dépend essentiellement de l'adhérence aux cellules eucaryotes (humaines et animales), ce qui leur évite d'être expulsés par les fluides corporels. Ces pathogènes peuvent sécréter plusieurs types d'adhésines ou fimbriae qui permettent à la bactérie de reconnaître des récepteurs spécifiques présents à la surface des cellules hôtes. Ceci joue un rôle important dans le tropisme d'une bactérie vis-à-vis d'un tissu ou organe spécifique (Chahales *et al.*, 2015, Thanassi *et al.*, 2007). Ces adhésines sont des protéines qui se trouvent à la surface externe, permettant à la bactérie d'agir à distance tout en gardant l'intégrité et la structure protectrice de la membrane externe (Chahales *et al.*, 2015). Ce sont des filaments fins et rigides couvrant la surface bactérienne et mesurant entre 2 à 8 nanomètre (nm) de diamètre et environ 0.2 à 2 µm de longueur (Thanassi *et al.*, 2007). Leur structure stéréochimique permet à la bactérie d'adhérer aux récepteurs des cellules cibles (qui sont souvent de nature glycoprotéiques ou glycolipidiques), aux protéines de la matrice extracellulaire, aux autres bactéries pour former des agrégations bactériennes, ou à une surface abiotique (Soto *et al.*, 1999, Thanassi *et al.*, 2007). Plusieurs rôles ont été attribués à ces fimbriae, notamment l'adhérence, l'invasion, la formation de biofilm et la motilité cellulaire. Ils sont composés d'un polymère de protéines appelées sous-unités majeures et au sommet, on trouve des sous-unités mineures auxquelles est fixée l'adhésine fimbriaire (Miquel, 2010, Thanassi *et al.*, 2007).

Généralement, les souches d'*E. coli* peuvent avoir plusieurs gènes fimbriaires distincts (Welch *et al.*, 2002). Toutefois, elles ne peuvent pas exprimer tous les fimbriae à la fois. Ils sont régulés de manière qu'un seul fimbriaire est exprimé au moment opportun (Nowicki *et al.*, 1984).

Cette régulation est appelée « variation de phase : ON /OFF » qui permet d'activer ou de réprimer l'expression de ces fimbriae (pour plus de détails, voir section régulation du fimbriae P).

2.2.1.2 Historique

Les premières descriptions de ces fimbriae remontent en 1955 par Duguid et ses collaborateurs. Ils ont noté la présence d'appendices courts et filamenteux couvrant le corps bactérien et ayant une activité hémagglutinante. Ils ont attribué, par la suite, le terme « fimbriae » à ces structures, qui en latin signifie fils ou fibre (J Po Duguid *et al.*, 1955). En 1959, le terme « pili » (en latin signifie poils ou cheveux) a été introduit par Brinton et ses collaborateurs pour décrire ces mêmes appendices non flagellaires d'*E. coli*. Suite à la découverte du phénomène de conjugaison, le terme F pili a été attribué aux appendices « sexuels » de conjugaison qui sont responsables du transfert de matériel génétique d'une bactérie à une autre (Brinton Jr, 1959).

Afin de distinguer entre les appendices sexuels et ceux de l'adhérence, Ottow *et al.* ont proposé en 1975 l'utilisation du terme fimbriae pour désigner les appendices impliqués dans l'adhérence et le terme pili pour désigner les appendices impliqués dans le phénomène de conjugaison (Ottow, 1975). Cependant, cette proposition n'était pas adoptée par tous les chercheurs et jusqu'à présent les deux termes fimbriae et pili sont utilisés de manière interchangeable pour désigner les appendices qui se trouvent à la surface bactérienne.

2.2.1.3 Classification

Plusieurs classifications ont été proposées pour les fimbriae, la première a été proposée par l'équipe de Duguid en se basant sur la morphologie des fimbriae et leur capacité à agglutiner les érythrocytes en présence ou en absence de mannose (JP Duguid, 1959). On distingue dans cette classification sept groupes principaux : les fimbriae de type 1 (mannose sensible), type 2 (chez *S. enterica*), type 3 (mannose résistant en présence de l'acide tannique), type 4 (hémagglutination mannose résistant sans l'acide tannique), type 5 (chez *Pseudomonas*), type 6 (chez *Klebsiella*) et les fimbriae F (chez les souches *E. coli* F⁺ et Hfr) (Thanassi *et al.*, 2007). Cette classification a été adoptée par d'autres chercheurs afin de classer les nouveaux fimbriae d'*E. coli*, comme le fimbriae de type 1C dont la morphologie ressemble au fimbriae de type 1 mais n'agglutinant pas les érythrocytes, les fimbriae P qui reconnaissent l'antigène P présent à la surface des érythrocytes, et les fimbriae Dr qui reconnaissent l'antigène CD55 du groupe sanguin Dr (Chahales *et al.*, 2015, Thanassi *et al.*, 2007).

Au cours des années qui suivirent, d'autres classifications ont été proposées parmi lesquelles on peut citer la classification sérologique de l'antigène F (pour fimbriae) qui a été développée par l'équipe de Ørskov *et al.* en 1983 en utilisant des anticorps monoclonaux. D'après cette classification, les fimbriae de type 1 et 1C ont été nommés F1A et F1C. On peut aussi trouver différents variants au sein de la même classe. L'exemple de fimbriae P illustre bien cette diversité, plusieurs groupes ont été identifiés allant de F7.1, F7.2 jusqu'à F16, puis F40 et F48 et il y a aussi les fimbriae P-like comme F165.1 (Orskov *et al.*, 1983).

La classification la plus récente et la mieux étudiée est la classification selon la voie de biosynthèse des fimbriae, on distingue 5 groupes : les fimbriae de la voie « chaperone – usher (CU) », pili type IV, curli, système de sécrétion type III et système de sécrétion type IV (Chahales *et al.*, 2015, Thanassi *et al.*, 2007). La majorité des fimbriae présents chez *E. coli* sont assemblés par la voie CU qui implique la présence d'une protéine chaperon et une protéine usher. La protéine chaperon périplasmique se lie aux différentes sous-unités fimbriaires, facilitant ainsi leur repliement et leur passage vers l'extérieur par l'intermédiaire de la protéine d'ancrage « usher » (Busch *et al.*, 2012). La voie de biosynthèse de CU peut être subdivisée en 6 clades majeurs qui sont α , β , γ (divisé en sous-clades γ_1 , γ_2 , γ_3 , γ_4), κ , π et σ (Nuccio *et al.*, 2007). Cette classification basée sur des analyses phylogénétiques tient compte de la conservation de domaines spécifiques aux sous-unités « usher », par exemple les fimbriae dont l'usher est PapC, PrfC, PixC, SfpC, YbgQ, YfcUT, YqiG, ECs0743 et ECs3221 appartiennent au clade π -fimbriae, tandis que les fimbriae dont l'usher est AufC, ECs0594, ECs2110, FimD, FmlC, FocD, SfaF, SfmD, YcbS et YdeT appartiennent au clade γ_1 -fimbriae (Busch *et al.*, 2012, Nuccio *et al.*, 2007). Ici, nous discuterons principalement des fimbriae de type P qui appartiennent aux groupes π -fimbriae.

2.2.1.4 Fimbriae P

Les fimbriae P ont été décrits pour la première fois en 1981 chez les souches UPEC qui sont responsables des ITU supérieures chez les humains (Källenius *et al.*, 1981). Ils sont exprimés préférentiellement au niveau des reins où ils vont se lier aux cellules épithéliales rénales (bassinets et parenchymes rénaux) ce qui va provoquer une pyélonéphrite aiguë (Kaper *et al.*, 2004, Korhonen *et al.*, 1986, Lobel *et al.*, 2007). Ces fimbriae sont retrouvés chez 70 à 90% des souches isolées de cas de pyélonéphrites, dans environ 30 à 50% des souches isolées de cas de cystites et dans environ 10 à 30% des souches fécales. Ainsi, ces fimbriae provoquent des lésions tissulaires et permettent d'induire une réponse inflammatoire accompagnée d'une

augmentation de la production d'interleukine 6 dans un modèle d'infection urinaire de primates (Michael S Donnenberg *et al.*, 1996, Mainil, 2003, Roberts *et al.*, 1994). Toutefois, chez le modèle murin, les fimbriae P jouent un rôle important dans l'établissement d'une infection urinaire, mais ne provoquent pas de lésions tissulaires assez importantes (Michael S Donnenberg *et al.*, 1996, Mainil, 2003, Mobley *et al.*, 1993). Par ailleurs, les souches APEC possèdent aussi les gènes codant pour les fimbriae P dans leur génome. On les retrouve généralement chez les souches isolées de cas de septicémie (entre 20 à 25%) (Dho-Moulin *et al.*, 1999, Dozois *et al.*, 1995, Melha Mellata *et al.*, 2003). Des études ont montré que les fimbriae P ont été transférés horizontalement aux souches APEC, ce qui pourrait augmenter leur pouvoir pathogène lors de l'établissement d'une infection. Cependant, leurs rôles dans la virulence ne sont pas encore entièrement élucidés. Il semble qu'ils pourraient jouer un rôle tardif puisqu'ils n'ont pas montré un effet sur l'adhérence aux cellules du pharynx et de la trachée (Dozois *et al.*, 1995, Pourbakhsh *et al.*, 1997).

- **Déterminisme génétique**

Les gènes codant pour les fimbriae P sont regroupés en opéron nommé *pap* (pyelonephritis-associated pili) ou *prs* (pap-related sequence) en fonction de l'adhésine portée par le fimbriae. Dans les deux cas, l'opéron fimbriaire est constitué de deux gènes de régulation (*papB* et *papI*), six gènes codant pour les protéines de structure (les sous-unités mineures PapG, PapE, PapF, PapH, PapK, et la sous-unité majeure PapA) et trois gènes codant pour les protéines d'assemblage : les chaperons (PapD et PapJ) et la protéine d'ancrage ou usher (PapC) (**figure 10**). Ces gènes sont co-transcrits en un seul ARNm polycystronique (à l'exception de *papI*) puis traduits pour donner les différentes sous-unités fimbriaires (Chahales *et al.*, 2015, Hospenthal *et al.*, 2017, Mainil, 2003, Thanassi *et al.*, 2007).

L'opéron fimbriarie est généralement porté par des îlots de pathogénicité en association avec d'autres gènes de virulence. C'est le cas par exemple du génome de la souche UPEC J96 qui contient deux copies de l'opéron *pap*. Le premier opéron (porteur de l'adhésine PapGI) se trouve sur l'îlot de pathogénicité PAI-I_{J96} avec l'hémolysine α (*hly*), et le deuxième (porteur de l'adhésine PapGIII/PrsG) se trouve sur PAI-II_{J96} en association avec l'hémolysine α (*hly*) et CNF1 (*cnf1*) (Schmidt *et al.*, 2004). Cela peut montrer la diversification génétique de la bactérie.

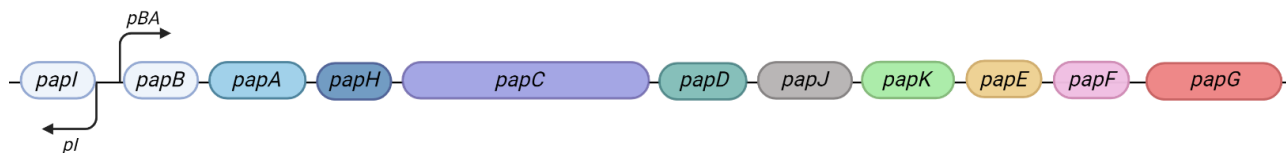


Figure 10: Organisation génétique de l'opéron *pap*.

Les couleurs correspondent aux protéines illustrées dans la figure 11. Les gènes *papB* et *papI* codent pour les gènes de régulation, ils possèdent chacun un promoteur représenté par une flèche noire ; *papA*, code pour la sous-unité majeure ; *papG* code pour l'adhésine ; *papD* code pour la protéine chaperon et *papC* code pour l'usher. Figure modifiée de (Chahales *et al.*, 2015) et créée avec BioRender.

La sous-unité PapH a pour rôle de stabiliser le fimbriae à la protéine d'ancrage. La sous-unité majeure PapA qui est présente à environ 1000 copies va former la partie basale du fimbriae. Les sous-unités mineures PapF, PapE et PapK vont former un appendice qui va porter l'adhésine PapG (Chahales *et al.*, 2015, Hospenthal *et al.*, 2017).

L'adhésine PapG est responsable de la reconnaissance des récepteurs glycoprotéiques membranaires de type Gal-(α 1-4)-Gal qui se trouve sur les cellules de l'hôte. Quatre allèles ont été identifiés et ont été classés en PapG classe I, PapG classe II, PapG classe III (ou PrsG) et PapG classe IV (Manning *et al.*, 2001, Sarowska *et al.*, 2019, Strömberg *et al.*, 1990). Les gènes codant pour des fimbriae P porteurs d'une adhésine Pap GI, PapGII ou papGIV sont regroupés sous le nom « *pap* » et ceux qui codent pour des fimbriae P porteurs d'une adhésine PapGIII/PrsG sont regroupés sous le nom « *prs* » (Mainil, 2003). Par ailleurs, la prévalence de PapGI chez les souches pathogènes est très rare, tandis que PapGII et PapGIII sont fréquemment rencontrés chez les souches UPEC. PapGIV a été identifiée chez des souches fécales mais son rôle dans la virulence est encore mal connu (Manning *et al.*, 2001).

Il a été démontré que PapG classe I reconnaît le récepteur Gal-(α 1-4)-Gal-(β 1-4)-Glc-, ou globotriosylcéramides (GbO3) (Lund *et al.*, 1987, Strömberg *et al.*, 1990), la classe II reconnaît le récepteur GalNAc-(β 1-3)-Gal-(α 1-4)-Gal-(β 1-4)-Glc-, ou globotétraosylcéramides (GbO4) et la classe III reconnaît le récepteur GalNAc-(α 1-3)-GalNAc-(β 1-3)-Gal-(α 1-4)-Gal-(β 1-4)-Glc-, ou globopentaosylcéramides (GbO5) ainsi que l'antigène Forssmann (Strömberg *et al.*, 1990). Ces récepteurs sont généralement présents sur les globules rouges des humains et des animaux (porc, pigeon, volaille, chèvre, chien), ainsi que sur les cellules épithéliales du tractus urinaire

chez l'homme (vessie pour GbO3, rein pour GbO4) et les chiens (GbO5) (Johanson *et al.*, 1993, Mainil, 2003, Strömberg *et al.*, 1990).

- **Biogenèse**

Après synthèse cytoplasmique, les différentes sous-unités fimbriaires sont reconnues par leur peptide signal et sont recrutées par le système de sécrétion SecYEG, ce qui permet leur translocation vers la membrane interne (Thanassi *et al.*, 2007). Une fois dans le périplasme, les sous-unités fimbriaires adoptent un repliement de type immunoglobuline incomplet (ou Ig-like) qui est composé par 6 brins β antiparallèles (de A à F). L'absence de brin G en C-terminal des sous-unités va créer un sillon hydrophobe et va les rendre instables, ce qui facilite leur agrégation et par conséquent leur dégradation par la protéase DegP (**figure 11**) (Busch *et al.*, 2012, Lillington *et al.*, 2014, Miquel, 2010, Thanassi *et al.*, 2007).

L'interaction des sous-unités avec la chaperonne PapD dès leur entrée dans le périplasme va stabiliser et aider leur repliement. En effet, PapD adopte aussi un repliement de type Ig et est composé de deux domaines de 7 brins β antiparallèles (A1 à G1 et A2 à G2). Le brin β G1 et la boucle F1-G1 de PapD interagissent avec le sillon hydrophobe de différentes sous-unités d'assemblage (PapG, PapE, PapF, PapK, PapA et PapH) et complète correctement le repliement de son domaine Ig-like (Busch *et al.*, 2012, Thanassi *et al.*, 2007). Ce mécanisme est appelé complémentation par brin donneur (Thanassi *et al.*, 2007). Ainsi, la première étape de la biogenèse commence lorsque l'adhésine PapG se lie au PapD pour former un complexe PapD-PapG. Ce complexe est recruté en premier vers l'usher PapC qui forme un pore de 2 à 3 nm dans la membrane externe. Il va se lier sur le premier site de fixation du côté N-terminal de l'usher PapC, puis va passer au deuxième site de fixation du côté C-terminal. À ce niveau, une modification de conformation spatiale a lieu pour permettre la translocation de PapG à travers le pore jusqu'à la surface de la bactérie (Busch *et al.*, 2012, Forsythe, 2020, Thanassi *et al.*, 2007). Ces étapes vont se répéter avec les différentes sous-unités jusqu'à la formation du fimbriae (**figure 11**). On trouve alors, au sommet, l'adhésine PapG qui possède un domaine lectine impliqué dans la reconnaissance des glycoprotéines spécifiques, puis les sous-unités mineures (PapE, PapF et PapK), suivies par les sous-unités majeures PapA qui vont former la partie basale du fimbriae. Cette partie mesure entre 5 et 7 μm de longueur et elle est composée d'environ 1000 sous-unités polymérisées en hélice droite. Enfin, le complexe PapD-PapH va se lier sur le domaine « Plug » de l'usher PapC, ce qui provoque la terminaison de l'assemblage du fimbriae

(Busch *et al.*, 2012, Chahales *et al.*, 2015, Hospenthal *et al.*, 2017, Lillington *et al.*, 2014, Thanassi *et al.*, 2007).

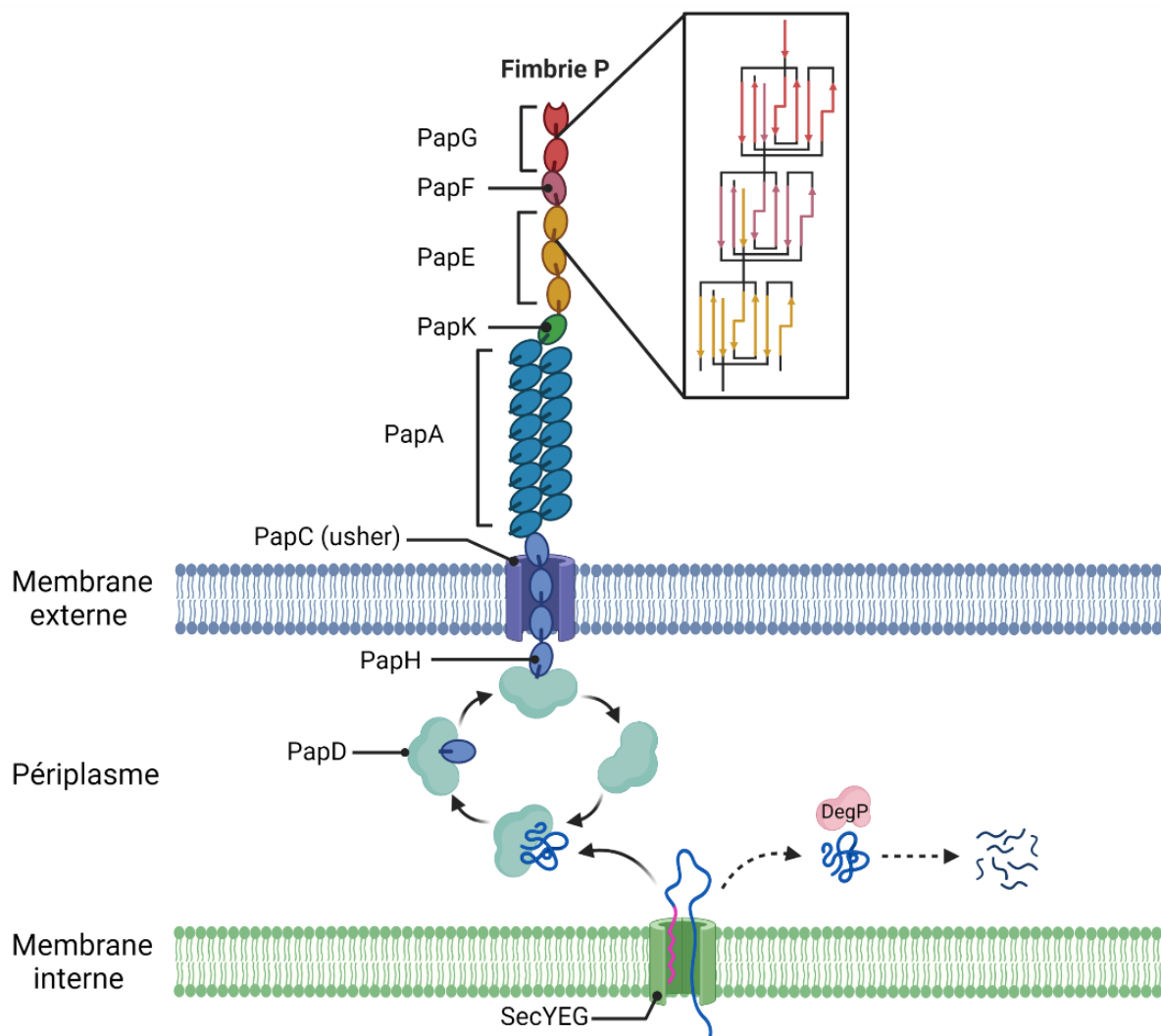


Figure 11: Biogenèse des fimbriae P.

L'assemblage commence par le recrutement des sous-unités vers le périplasma par le système Sec (YEG). Celles qui ne se sont pas liées correctement au PapD, seront dégradées par la protéase DegP. La protéine chaperon PapD se lie aux différentes sous-unités d'assemblage, ce qui facilite leur translocation dans la membrane externe par l'intermédiaire de PapC. Le complexe PapD-PapG se lie en premier au PapC, ce qui initie l'assemblage du fimbriae. Puis, les autres sous-unités seront incorporées : PapF, PapE, PapK et PapA. Enfin, une fois que le fimbriae se trouve dans le milieu externe, il va assumer sa conformation quaternaire pour exercer sa fonction. Figure modifiée de (Hospenthal *et al.*, 2017, Lillington *et al.*, 2014, Thanassi *et al.*, 2007) et créée avec BioRender.

- **Régulation des fimbriae P**

La régulation des fimbriae P chez les *E. coli* pathogènes dépend essentiellement de la variation de phase. C'est un mécanisme épigénétique qui conduit à l'alternance entre la Phase-ON où le fimbriae P est exprimé et la Phase-OFF où il est réprimé (Van Der Woude *et al.*, 2004). Ce phénomène met en jeu plusieurs niveaux de régulation (transcriptionnel, post-transcriptionnel et post-traductionnel) face aux diverses interactions environnementales (température, source de carbone...), ce qui peut déclencher soit la Phase-ON soit la Phase-OFF. L'activation de cette voie serait donc une réponse appropriée pour permettre à la bactérie d'augmenter son adhérence vis-à-vis des cellules cibles (Graveline *et al.*, 2011, Holden *et al.*, 2007, Van Der Woude *et al.*, 2004, Zamora *et al.*, 2020).

Plusieurs régulateurs entrent en compétition au niveau transcriptionnel pour activer ou réprimer la transcription de l'opéron. Parmi ceux-ci, on peut citer les régulateurs PapI et PapB, le régulateur universel histone-like nucleoid structuring protein (H-NS) (White-Ziegler *et al.*, 2000), le Catabolite activator protein (CAP, ou cAMP receptor protein (CRP)), le DNA adenine methylase (Dam) et le leucine-responsive regulatory protein (Lrp) (Graveline *et al.*, 2011, Van Der Woude *et al.*, 2004).. Généralement, le rôle de ces régulateurs a été attribué suite à une inhibition observée de l'expression des fimbriae P chez les souches mutantes, ce qui indique que ces composés sont impliqués dans la régulation (Graveline *et al.*, 2011, Van Der Woude *et al.*, 2004).

Tout d'abord, l'opéron *pap* est sous le contrôle de ses propres gènes de régulation *papI* et *papB*. Le gène *papI* est régulé par son propre promoteur *pI* qui est transcrit en sens inverse du reste de l'opéron *pap*. Quant au gène *papB*, il est régulé par le promoteur *pBA* dont la transcription génère l'unique ARNm polycistronique mentionné précédemment, et qui sera traduit pour donner les différentes sous-unités fimbriaires qui seront par la suite assemblées (Holden *et al.*, 2007, Van Der Woude *et al.*, 2004). La région intergénique de 416-pb située entre les gènes *papI* et *papB* possède six sites de liaison de la protéine Lrp, désignés de 1 à 6. Les sites 1 et 6 sont séparés d'environ 32 pb tandis que les sites 2 et 5 sont séparés d'environ 102 pb. Ces deux sites contiennent deux séquences d'ADN désignées GATC proximal (1-3) et GATC distal (4-6) par rapport au promoteur *pBA* qui peuvent être méthylés par DAM au niveau de la position N6 des adénines. Cette modification est réversible, empêchant ainsi la fixation des autres régulateurs comme la protéine Lrp à un des deux sites (**figure 12**) (Holden *et al.*, 2007, Van Der Woude *et al.*, 2004).

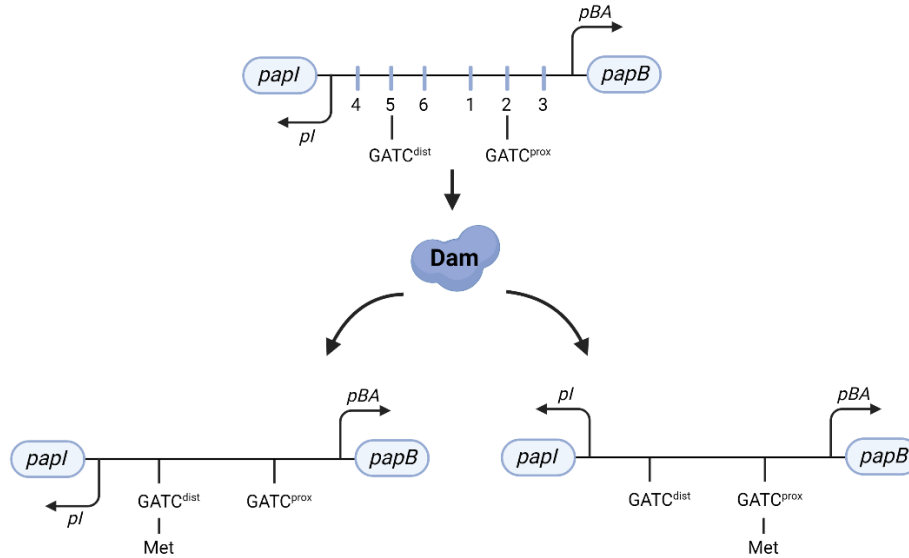


Figure 12: Méthylation des sites de liaison de la protéine Lrp par l'enzyme DAM.

Les sites de liaison de la protéine Lrp sont désignés de 1 à 6 et contiennent les deux séquences GATC^{prox} et GATC^{dist} qui peuvent être méthylées par DAM. Figure modifiée de (Van Der Woude *et al.*, 2004) et crée avec BioRender.

Quant à la protéine Lrp, elle a tendance à se lier aux sites GATC^{prox} de la région promotrice, ce qui bloque la méthylation de ce site et la fixation de l'ARN polymérase au niveau du promoteur pBA. Cela a pour conséquence l'inhibition de la transcription de l'opéron *pap* (**figure 13**) (Graveline *et al.*, 2011, Holden *et al.*, 2007, Van Der Woude *et al.*, 2004).



Figure 13: Phase OFF de l'opéron pap : la fixation de LRP sur le site GATC^{prox} inhibe la transcription de l'opéron pap.

Figure modifiée de (Van Der Woude *et al.*, 2004) et crée avec BioRender.

Toutefois, la présence de régulateur PapI dans le cytoplasme est nécessaire pour réguler la transcription du gène *papB* et pour former le complexe PapI-Lrp. La formation de ce complexe va favoriser sa fixation sur le site GATC^{dist} et, par conséquent, la transcription de l'opéron *pap*.

Cette activation est maintenue grâce à d'autres régulateurs comme PapB qui active à son tour la transcription du gène *papI*. Ainsi, PapB va former des multimères de 8 à 10 sous-unités qui vont se fixer au sillon mineur de l'ADN localisé entre le promoteur *pI* et le site de liaison de la protéine CAP. En parallèle, la protéine CAP va former un complexe avec l'AMPc pour se fixer sur le site qui se trouve à 215 pb en amont de promoteur *pBA*. Cela permet d'activer la transcription des deux promoteurs *pI* et *pBA* et de maintenir la phase ON (**figure 14**) (Graveline *et al.*, 2011, Holden *et al.*, 2007, Van Der Woude *et al.*, 2004).

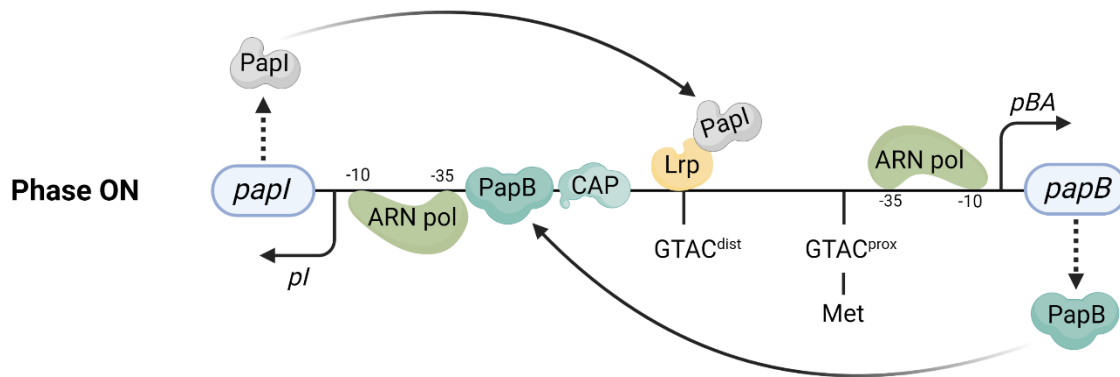


Figure 14: Phase ON de l'opéron *pap*.

Figure modifiée de (Van Der Woude *et al.*, 2004) et créée avec BioRender.

Le régulateur universel H-NS intervient aussi dans la régulation de l'expression de l'opéron *pap*. Son effet répresseur a été montré à basse température (inférieure à 26°C). En effet, il se fixe à la région intergénique entre *papI* et *papB*, empêchant ainsi la méthylation des sites $GATC^{prox}$ et $GATC^{dist}$, et par conséquent, l'inhibition de la transcription du promoteur *pBA*. Cela permet de bloquer la transcription avant même de passer à la phase OFF (**figure 15**) (White-Ziegler *et al.*, 2000, Zamora *et al.*, 2020).

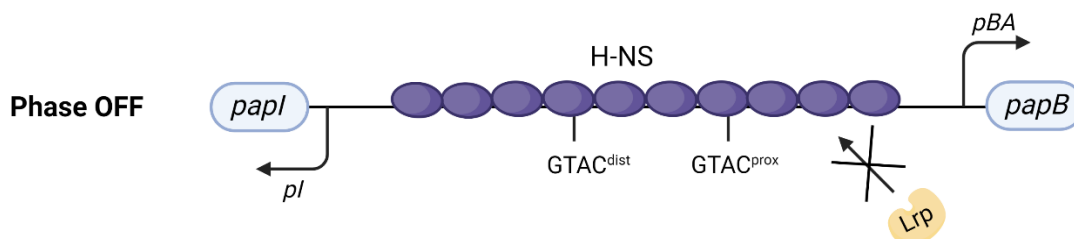


Figure 15: Phase OFF de l'opéron *pap* : inhibition de la transcription par le régulateur H-NS.

Figure modifiée de (Zamora *et al.*, 2020) et créée avec BioRender.

Également chez les UPEC, il a été démontré que le petit ARN nommé P-fimbriae regulator, *papR*, contrôle négativement l'expression des fimbriae P en dégradant l'ARNm de *papI*. La présence de ce régulateur dans le cytoplasme empêche la fixation de la protéine Lrp aux sites $\text{GATC}^{\text{dist}}$, et par conséquent, la répression des fimbriae P (Khandige et al., 2015) (**figure 16**).

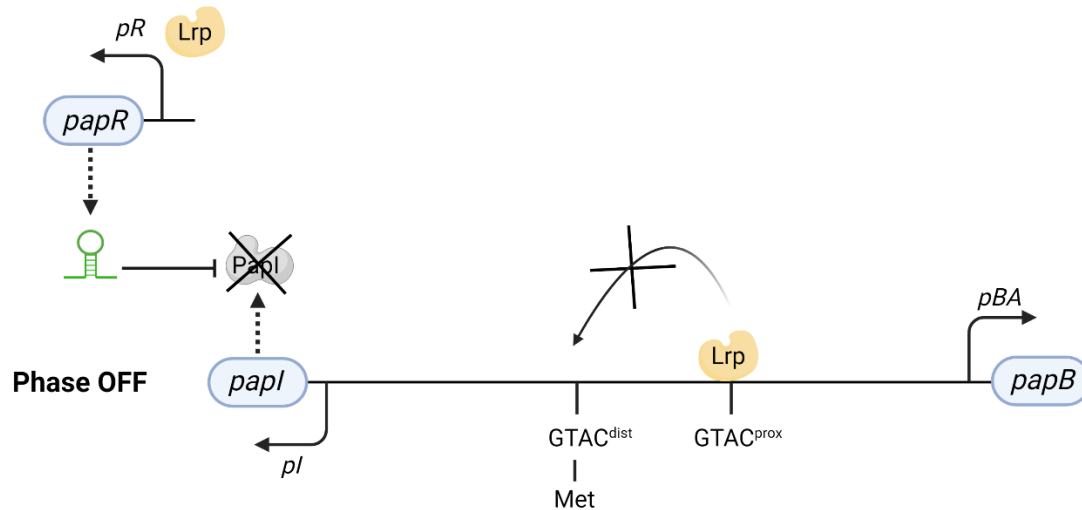


Figure 16: Phase OFF de l'opéron *pap* : inhibition de la transcription par le régulateur *papR*.

Figure modifiée de (Khandige et al., 2015) et créée avec BioRender.

Le système à 2 composants CpxAR régule également l'expression des fimbriae P à plusieurs niveaux. Au niveau traductionnel, la surexpression de certaines sous-unités fimbriaires dans le cytoplasme va activer la protéine senseur CpxA (possédant une activité phosphatase/kinase) qui va à son tour activer la protéine CpxR par phosphorylation. La protéine CpxR phosphorylée (CpxR-P) va entrer en compétition avec le complexe PapI-Lrp sur les sites $\text{GATC}^{\text{dist}}$, ce qui va inhiber la transcription du promoteur *pBA* (Hernday et al., 2004).

Après transcription de l'opéron *pap*, la modulation du nombre des sous-unités fimbriaires déjà existantes dans le cytoplasme se fait par des modifications post-traductionnelles qui constituent un moyen efficace pour la cellule de répondre à divers signaux. Ces modifications peuvent changer la conformation des protéines pour augmenter ou bloquer l'interaction entre les sous-unités fimbriaires ou modifier le nombre des sous-unités amenées au périplasma (Thanassi et al., 2007). Parmi ces régulateurs, on peut citer la protéase DegP qui joue un rôle important dans la dégradation des agrégats des sous-unités fimbriaires ou celles qui n'ont pas lié correctement la chaperon PapD, cela empêchera leur accumulation « létale » dans le périplasma bactérien (Thanassi et al., 2007). Aussi, il y a l'enzyme disulfide oxidoreductase A (Dsb A) qui

joue un rôle important dans l'assemblage du fimbriae en catalysant la formation de ponts disulfures au niveau de la protéine chaperon PapD et de l'adhésine PapG. Cette modification est nécessaire pour stabiliser la structure quaternaire de ces protéines et pour maintenir la liaison entre les différentes sous-unités fimbriaires (pour PapD) ou entre PapG et les récepteurs glycosylés qui sont présents au niveau des cellules de l'hôte (Thanassi *et al.*, 2007).

2.2.1.5 Autres fimbriae

- **Fimbriae type 1**

Les fimbriae de type 1 sont des fimbriae qui ressemblent structurellement au fimbriae P. Ils sont présents dans la majorité des souches pathogènes et non pathogènes d'*E. coli* (environ 70% à 80% des souches possèdent l'opéron *fim* dans leur génome). Chez les souches UPEC, les fimbriae de type 1 jouent un rôle important dans le développement d'une ITU chez l'homme en adhérant aux résidus mannosylés retrouvés sur les cellules endothéliales de la vessie. Ils permettent aussi la formation de biofilm et l'adhérence des bactéries aux différentes surfaces biotiques et abiotiques (Klemm *et al.*, 2004, Schembri *et al.*, 2001).

Les fimbriae de type 1 sont des filaments fins et rigides couvrant la surface bactérienne et mesurant entre 3 à 7 nm de diamètre et environ 0.1 à 2 µm de longueur. Ils sont codés par l'opéron *fimAICDFGH* qui se trouve sur le chromosome bactérien. Les gènes *fimA*, *fimF*, *fimG* et *fimH* codent pour les protéines de structures (*fimA* code pour la sous-unité majeure et *fimH* code pour l'adhésine), le gène *fimC* code pour la protéine chaperon, le gène *fimD* code pour le usher et le rôle du gène *fimI* n'est pas assez clair (Klemm *et al.*, 2004, Schembri *et al.*, 2001).

L'expression de cet opéron dépend de l'orientation de son promoteur *fimS* qui est constitué de deux petites séquences inversées répétées nommées région inverse gauche (RIG) et région inverse droite (RID). Ce promoteur est sous le contrôle des recombinases FimB et FimE qui déterminent l'orientation de l'élément chromosomique. Phase ON : si le promoteur porté par l'élément est dans l'orientation de la transcription. Phase OFF : si le promoteur porté par l'élément est dans l'orientation inverse de la transcription (Nielubowicz *et al.*, 2010).

- **Fimbriae Pix**

Les fimbriae Pix sont des fimbriae P-like. Ils sont codés par l'opéron *pix* (constitué de 8 gènes *pixBAHCDJFG*) qui a été identifié pour la première fois sur un îlot de pathogénicité chez la souche UPEC X2194 (Lügering *et al.*, 2003).

Les fimbriae Pix sont codés par l'opéron *pixBAHCDJFG*. Le gène *pixB* est un gène de régulation, les gènes *pixA*, *pixpH*, *pixJ*, *pixpF* et *pixG* codent pour les protéines de structures (*pixA* code pour la sous-unité majeure et *pixG* code pour l'adhésine), le gène *pixD* code pour la protéine chaperon, le gène *pixC* code pour le usher (Lügering *et al.*, 2003).

La prévalence de ces fimbriae chez les souches ExPEC est très faible, seulement 8% des souches UPEC et 4% des souches fécales contiennent les gènes codant pour ces fimbriae. Ils permettent l'adhérence des bactéries aux lignées cellulaires du cancer du col de l'utérus (HeLa) (Lügering *et al.*, 2003). Cependant, leur rôle dans la virulence n'est pas bien élucidé (Michael Donnenberg, 2013).

- **Fimbriae Sfp**

Les fimbriae Sfp sont aussi des fimbriae P-like qui ont été identifiées chez la souche EHEC O157:NM (Brunner *et al.*, 2001) sur le plasmide de virulence de type « sorbitol-fermenting ». Ce plasmide est présent chez certaines souches EHEC O157 sorbitol-positives, d'où le nom Sfp pour « sorbitol-fermenting EHEC O157 fimbriae, plasmid-encoded » (Brunner *et al.*, 2001). Les fimbriae Sfp sont codées par l'opéron *sfpAHCDJFG*. Les gènes *sfpA*, *sfpH*, *sfpJ*, *sfpF* et *sfpG* codent pour les protéines de structures (*sfpA* code pour la sous-unité majeure et *sfpG* code pour l'adhésine), le gène *sfpD* code pour la protéine chaperon, le gène *sfpC* code pour le usher (Brunner *et al.*, 2001).

Le rôle des fimbriae Sfp dans la virulence n'est pas bien élucidé. Toutefois, il a été montré qu'il est capable d'agglutiner le sang humain et d'adhérer aux cellules épithéliales de l'intestin (Bielaszewska *et al.*, 2009, Brunner *et al.*, 2001).

2.2.2 Autotransporteurs

Depuis leur découverte, la famille des autotransporteurs (AT) représente une des plus importantes familles de protéines sécrétées par les bactéries Gram négatif avec plus de 1000 membres qui ont été identifiées (Benz *et al.*, 2011, Pokharel *et al.*, 2019). Elles sont sécrétées par le système de sécrétion de type V qui est répandu chez les bactéries pathogènes (Pokharel *et al.*, 2019). Ce système est subdivisé en cinq catégories :

- Système de sécrétion (Va), représenté par les ATs monomériques.
- Système de sécrétion à deux partenaires (Vb), composé de deux protéines séparées (TpsA constitue la protéine sécrétée et TpsB le domaine translocateur).

- Système de sécrétion de type Vc représenté par les ATs trimériques.
- Système de sécrétion de type Vd est un hybride entre Va et Vb.
- Système de sécrétion de type Ve représenté par les intimines et les invasines.

Les ATs de la famille sérine protéase (SPATE) constituent un groupe de protéines extracellulaires produites par la plupart des souches d'*E. coli* pathogènes. Ce sont des ATs monomériques sécrétés par le système de sécrétion de type Va (Dautin, 2010, Pokharel *et al.*, 2019). Ils sont constitués par

- Un peptide signal du côté N-terminale.
- Un « domaine passager » portant le site actif sérine-protéase « Ser, His, Asp ».
- Une région de jonction appelée « domaine de liaison » nécessaire pour le repliement et la translocation du domaine passager.
- Un domaine C-terminal (domaine β) qui se replie et forme un tonneau β dans la membrane externe (Dautin, 2010).

Les ATs de la famille des SPATEs sont caractérisés par un peptide signal assez long de plus de 49 acides aminés, un domaine β très conservé qui se trouve à l'extrémité C-terminale (277 acides aminés chez tous les ATs) et un site catalytique conservé (GDSGS) qui se trouve à l'extrémité N-terminale du domaine passager (entre les résidus 250 à 270). De plus, la majorité des ATs sont clivés après la translocation de leur domaine passager à travers la membrane externe par un mécanisme autoprotéolytique (Dautin, 2010, Henderson *et al.*, 2001, Pokharel *et al.*, 2019).

Les connaissances actuelles sur les ATs de la famille SPATE sont discutées en détail dans la revue 1.

ARTICLE DE REVUE

3 ARTICLE DE REVUE

Serine Protease Autotransporters of the *Enterobacteriaceae* (SPATEs): Out and about and chopping it up

Auteurs : Pravil Pokharel^{1,2#}, Hajer Habouria^{1,2#}, Hicham Bessaiah^{1,2}, and Charles M. Dozois^{1,2,3}

1 Institut national de recherche scientifique (INRS) - Centre Armand-Frappier Santé Biotechnologie, Laval, Quebec, Canada

2 Centre de recherche en infectiologie porcine et avicole (CRIPA), Saint-Hyacinthe, Quebec, Canada

3 Institute Pasteur International Network

These two authors contributed equally as primary authors of this research

Objectif de la revue :

L'objectif de l'article revue est de rassembler l'ensemble des connaissances actuelles sur les ATs de la famille SPATE. Initialement, nous avons introduit la famille des SPATEs qui appartient à une famille des ATs monomériques sécrétées par le système de sécrétion Va des bactéries Gram négatif. Par la suite, nous avons abordé la biogenèse et la sécrétion des ATs de la famille des SPATEs. Nous avons ensuite présenté les différents membres de cette famille et évalué leurs rôles dans la virulence des souches ExPEC. Enfin, nous avons détaillé les différents mécanismes de régulation des différents gènes de cette famille.

Contribution des auteurs :

En qualité de co-premier auteur, j'ai structuré le texte et rédigé l'ensemble de cette revue. J'ai également élaboré l'ensemble des figures et procédé à l'ensemble des corrections suggérées par mon directeur de recherche et mes collaborateurs.

Pravil Pokharel a contribué au plan de la revue ainsi qu'à sa rédaction. Il a également participé à la préparation des figures, coordonné et assemblé les corrections de notre directeur de recherche et de nos collaborateurs.

Hicham Bessaiah a contribué à la révision du manuscrit ainsi qu'à l'assemblage des corrections de notre directeur de recherche et de nos collaborateurs.

Le Professeur Charles M. Dozois a contribué à la conception de la revue. Il a révisé, corrigé et soumis le manuscrit au journal.

Serine Protease Autotransporters of the *Enterobacteriaceae* (SPATEs): Out and about and chopping it up

Auteurs : Pravit Pokharel^{1, 2 #}, **Hajer Habouria**^{1, 2 #}, Hicham Bessaiah^{1,2}, and Charles M. Dozois^{1,2,3}

1 Institut national de recherche scientifique (INRS) - Centre Armand-Frappier Santé Biotechnologie, Laval, Quebec, Canada

2 Centre de recherche en infectiologie porcine et avicole (CRIPA), Saint-Hyacinthe, Quebec, Canada

3 Institute Pasteur International Network

These two authors contributed equally as primary authors of this research

Journal : Microorganisms.

2019, 7(12), 594|

Soumis : 25 October 2019, **Accepté** : 13 November 2019, **Publié** : 21 November 2019

DOI : <https://doi.org/10.3390/microorganisms7120594>.



Review

Serine Protease Autotransporters of the *Enterobacteriaceae* (SPATEs): Out and About and Chopping It Up

Pravil Pokharel ^{1,2,†}, Hajer Habouria ^{1,2,†}, Hicham Bessaiah ^{1,2} and Charles M. Dozois ^{1,2,3,*}

¹ Institut National de Recherche Scientifique (INRS)-Centre Armand-Frappier Santé Biotechnologie, Laval, QC H7V 1B7, Canada; micropravil@gmail.com (P.P.); hajer.hbr2@gmail.com (H.H.); hicham.bessaiah@iaf.inrs.ca (H.B.)

² Centre de Recherche en Infectiologie Porcine et Avicole (CRIPA), Saint-Hyacinthe, QC J2S 2M2, Canada

³ Institut Pasteur International Network, Laval, QC H7V 1B7, Canada

* Correspondence: charles.dozois@iaf.inrs.ca

† These authors contributed equally to this work.

Received: 25 October 2019; Accepted: 13 November 2019; Published: 21 November 2019



Abstract: Autotransporters are secreted proteins with multiple functions produced by a variety of Gram-negative bacteria. In *Enterobacteriaceae*, a subgroup of these autotransporters are the SPATEs (serine protease autotransporters of *Enterobacteriaceae*). SPATEs play a crucial role in survival and virulence of pathogens such as *Escherichia coli* and *Shigella* spp. and contribute to intestinal and extra-intestinal infections. These high molecular weight proteases are transported to the external milieu by the type Va secretion system and function as proteases with diverse substrate specificities and biological functions including adherence and cytotoxicity. Herein, we provide an overview of SPATEs and discuss recent findings on the biological roles of these secreted proteins, including proteolysis of substrates, adherence to cells, modulation of the immune response, and virulence in host models. In closing, we highlight recent insights into the regulation of expression of SPATEs that could be exploited to understand fundamental SPATE biology.

Keywords: serine protease autotransporters of *Enterobacteriaceae* (SPATE); autotransporters; cytotoxins; adhesins; *E. coli*; gene regulation

1. Introduction

Bacteria have acquired a capacity to export and secrete proteins and other molecules to the cell surface in order to interact with the extracellular environment. The transport of proteins to the cell surface is achieved through a number of highly specialized protein secretion systems that release them into the extracellular milieu. In Gram-negative bacteria, which have an inner and outer membrane that contains a periplasmic space, secretion can be a two-step process involving export to the periplasmic space, and in some cases, subsequent secretion through the outer membrane. Autotransporter (AT) proteins comprise a large family with more than 1000 members that have been characterized [1]. AT proteins represent the largest family of secreted polypeptides in Gram-negative bacteria and serine protease autotransporters of *Enterobacteriaceae* (SPATEs) are a subclass of AT proteins that contain a protease domain belonging to the trypsin-like family, which typically contains a serine in the catalytic motif [1]. In recent years, considerable information has been obtained about how these proteins are assembled and secreted. Here, we review the latest findings on the AT secretion system with a recent model for transporting SPATE cargo out of the bacterial cell and in-depth updates of members of SPATEs including studies on genomic distribution, gene regulation, classification, and fate of the protein during in vitro or in vivo host interaction.

2. The Autotransporter Secretion Pathway

AT secretion through the outer membrane is mediated by the type V secretion system (T5SS) or AT secretion pathway. The T5SS pathway has been subdivided into five subtypes: (i) T5SS of monomeric ATs is classed as type Va secretion; (ii) two-partner secretion is classed as type Vb secretion; (iii) trimeric AT secretion is classed as type Vc secretion [2]; (iv) secretion of ATs homologous to both type Va and type Vb is described as type Vd [3]; and (v) secretion of intimins and invasins is classed as subtype Ve [4]. SPATEs are monomeric ATs that are secreted by the type Va secretion pathway.

The figure below depicts the major differences between these subtypes, which includes the variations in alignments of different domains (Figure 1). In type Va ATs, release of the N-terminal passenger domain is assisted by a C-terminal translocation domain or autoprocessed and liberated into the external milieu (explained in detail below) [1]. Type Vb is a split variant of the type Va system as the passenger domain and translocation domain are located in different polypeptide chains, and the translocated domain contains periplasmic polypeptide transport-associated (POTRA) motifs. As such, the type Vb class has also been described as a two-partner secretion system [5]. The type Vc class comprises ATs that form trimers and are also called trimeric autotransporter adhesins [2]. Type Vd ATs differ from type Va due to the presence of additional periplasmic domains between the passenger domain and the translocation domain, which is homologous to the periplasmic domains present in type Vb proteins [3]. Likewise, in type Ve ATs, the domains have a reverse order, wherein the passenger domain is at the C-terminal and translocation domain is N-terminal [4].

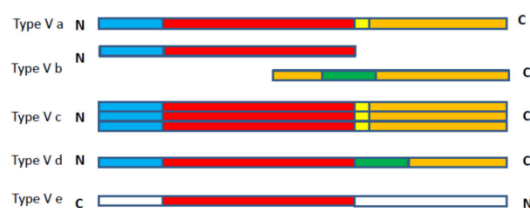


Figure 1. Scheme presenting domain organization among the subclasses of type-V bacterial autotransporter proteins. The labeling includes the conserved domains, colored blocks correspond to: Signal peptide (blue), passenger domain (red), polypeptide transport associated (POTRA) domain (green), linker domain (yellow) and translocation domain (orange). Adapted from [1,4].

Understanding the biogenesis of the SPATEs is of great interest for the isolation, purification, and characterization of these proteins. Over the last two decades, a diversity of predicted AT proteins, including SPATEs, have been identified through the sequencing of many bacterial genomes and through bioinformatics analysis. However, only a few SPATEs have been more extensively studied with regards to their biological functions and structural characterization. The crystal structure of the passenger domain of three SPATEs has been determined: EspP from an *Escherichia coli* O157:H7 strain [6], Hbp (also called Tsh) from an extra-intestinal pathogenic *E. coli* (ExPEC) strain [7] and Pet from enteroaggregative *E. coli* (EAEC) strains [8]. Based upon these crystal structures, general models of structure and translocation have been proposed, although, whether such models derived from only a few SPATE structures collectively represent all other SPATEs remains to be determined.

The general structure of AT proteins, including SPATEs, comprises three functional domains: The signal peptide, which mediates the Sec-dependent transport of the protein into the periplasm; the N-terminal passenger domain (also called the α -domain), which is the mature protein that is exposed at the surface of the outer membrane and/or released extracellularly; and the pore forming carboxyl-terminal translocator domain (also called as β -barrel), which provides the channel through which the passenger domain is translocated to the surface of the outer membrane [9]. Initial proposals of ATs as autonomously secreted proteins have been rejected due to recent findings, indicating a role for accessory proteins located in the inner membrane [10], the periplasm [11] and the outer membrane [10] which facilitate or mediate translocation of AT proteins to the cell surface.

ATs are exported into the periplasmic space through the Sec-dependent pathway [1], and export can occur co-translationally or follow autotransporter synthesis into the cytoplasm [12]. Immediate export following translation could improve export by preventing premature folding and degradation in the cytoplasm. Upon reaching the periplasm via the Sec-translocon and cleavage of the signal sequence, AT proteins are then protected by conserved periplasmic chaperones such as Skp, SurA, and DegP and directed toward the β -barrel assembly machinery (Bam) complex, which catalyzes the insertion and assembly of the outer membrane protein (Figure 2) [13,14]. A “Hybrid barrel-model” (Figure 2) has been proposed to explain the translocation of the passenger domain through the outer membrane. It has been shown that passenger domain secretion does not appear to use ATP, but that vectorial folding of the C-terminal of the passenger domain may contribute the necessary energy required for transmembrane passage and folding [15].

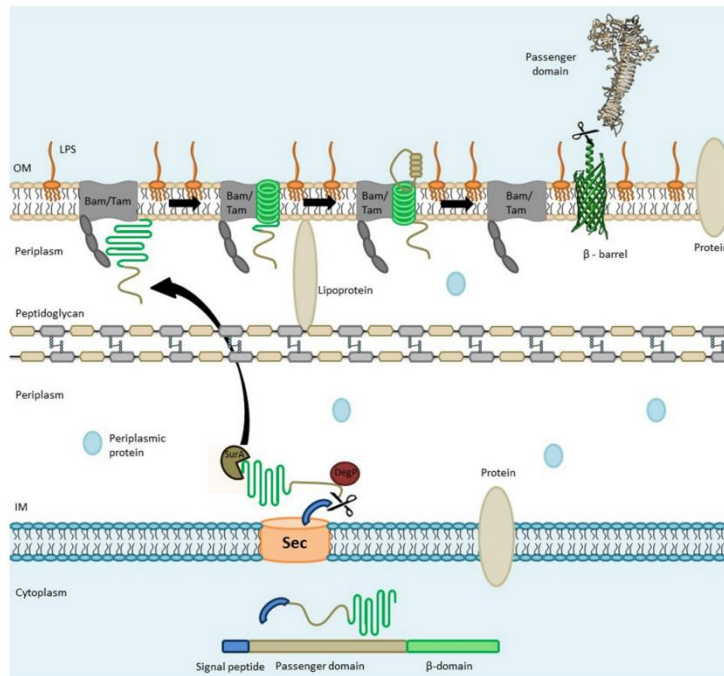


Figure 2. Schematic overview of Autotransporter (AT) processing, export, and secretion: The Signal peptide (SP) mediated by the Sec apparatus, guides translocation of the autotransporter to the periplasmic space. In the periplasm, the AT is kept in a “translocation competent state” by recruiting chaperones such as Skp, SurA, and DegP. Further, the Bam complex assists in the integration of the β -domain into the outer membrane and promotes the translocation of the passenger domain across the outer membrane through a hybrid-barrel mechanism wherein the AT β -barrel and Bam/Tam protein domains interact. Periplasmic chaperones such as SurA and DegP deliver an unfolded autotransporter to BamA/TamA. POTRA domains (gray in beaded structure) are contact sites for the AT protein to be transported. A hybrid-barrel is then formed by insertion of the AT β -strand through the gate region between strands 1 and 16 of the BamA/TamA barrel. Barrel expansion results in pore opening and the passenger domain can then protrude through the hybrid barrel. Subsequently, the passenger domain is released from the hybrid structure and may remain on the bacterial cell surface or can be cleaved for release from the bacterial cell. Adapted from [13,14].

2.1. Sec-Dependent Export of AT Proteins through the Inner Membrane

The N-terminal signal peptide of all autotransporters mediates insertion through the inner membrane through a Sec-dependent mechanism that is common to many other exported or secreted proteins. Generally, the N-terminal signal sequences in proteins show a tripartite organization of n, h, and c regions which corresponds to N-terminal, hydrophobic, and cleavage sites [16,17]. Analysis of signal sequence mutants revealed that a hydrophobic, H region is an essential part for targeting and membrane insertion [16].

The signal peptide sequence has the function of targeting or protein translocation to the inner membrane. In *E. coli*, protein export through the Sec pathway can involve two distinct pathways: (i) The SecB/SecA pathway wherein the chaperone SecB, prevents premature aggregation or folding, keeping the protein in a “translocation—competent state” and leads to transfer to Sec A [18], (ii) the SRP (Signal Recognition Particle) pathway in which the SRP nucleoprotein complex mediates co-translational targeting by interacting with a highly hydrophobic signal sequence following translation from ribosome towards the translocon [19].

2.2. AT Protein Transit through the Periplasm

The mechanism of secretion from the periplasm and the transitional state of ATs while localized in the periplasm is still debated. In fact, in the periplasmic space, these proteins are prone to immature folding or aggregation and degradation by periplasmic proteases. Misfolding or degradation of ATs can be prevented apparently either by prolonged interaction with the Sec-translocon or by interaction with periplasmic chaperones [11], or both. The periplasmic localization of ATs is likely to be very transient, and translocation to the outer membrane may occur rapidly following export and processing of the signal peptide through the cytoplasmic membrane.

2.3. Transport of ATs through the Outer Membrane (The Hybrid-Barrel Model)

Proteins of the Omp85 superfamily such as BamA promote the insertion and folding of β -barrel outer-membrane proteins (OMPs) including AT proteins across the bacterial outer membrane [14,20]. BamA possesses five periplasmic POTRA domains which are believed to recognize substrates mediated by β -strand augmentation [21,22]. In addition to BamA, TamA has been implicated in the translocation of autotransporters [10]. TamA has structural similarities to BamA and also contains domains associated with catalytic functions like insertase and chaperone foldase activity [23]. Thus, a similar model of AT translocation has been proposed for both the BamA and TamA translocation systems.

Despite the lack of a satisfactory model for autotransporter delivery at the outer membrane, the hybrid barrel model (Figure 2) provides a plausible mechanistic model which is based on interactions with the open 16-stranded BamA and TamA barrels. The unzipped strands of these proteins can incorporate β -strands of autotransporter by β -augmentation, creating a hybrid-barrel of the AT protein with BamA/TamA. The resulting hybrid-barrel would form a pore through which the AT passenger domain would be translocated. Subsequent to passenger domain translocation through the outer membrane, a final unzipping of the hybrid complex would separate the two barrels, releasing the assembled autotransporter laterally into the outer membrane and returning BamA/ TamA to its free, uncoupled state [23,24].

2.4. Passenger Domain Cleavage

At the bacterial surface, the fate of AT proteins can be dependent on the specific proteins themselves as well as the physiological conditions or the environmental niche. Some ATs remain associated with the outer membrane surface whereas others, such as the Pet and EspP SPATES, show autocatalytic activity within the β -barrel leading to cleavage of the linker and release of the passenger domain from the bacterial cell surface [25,26]. Cleavage of the passenger domain from the β -domain can take place by various mechanisms. In the case of the SPATES, the cleavage site is generally conserved.

2.5. A Cleavage Site is Located in the “Linker Domain” of SPATE Proteins

The linker domain encompasses an invariant 14-residue segment that spans the passenger domain and translocation domain junction in the SPATEs, and studies have shown that the cleavage site is conserved in this domain. Analysis of the (¹⁰²¹EVNNLNKRMGDL¹⁰³²) sequence motif of EspP showed that the passenger domain is cleaved after the first asparagine residue [27,28]. The mutation in the linker peptide resulted in impaired passenger domain cleavage of EspP [28]. Similar mutations impaired passenger domain cleavage and passenger domain translocation of another SPATE, Tsh [29]. These findings suggest that the linker domain and sequence plays an important role in processing of the passenger domain. However, some SPATEs lack a twin asparagine sequence. For instance, RpeA, from a rabbit enteropathogenic *E. coli* (EPEC) strain lacks the twin asparagine residues within its linker domain and it was reported that this protein was not released into the supernatant [30]. By contrast, a recently identified SPATE called Sha (Serine protease hemagglutinin autotransporter) lacks the twin asparagine residues, but was released into the culture supernatant [31], suggesting motifs other than the twin asparagine site may be recognized for cleavage of certain SPATEs.

3. SPATEs

Members of the SPATE family are autotransporter proteins from a variety of enterobacterial species that all contain a consensus serine protease motif, and they have most notably been described from pathogenic *Escherichia coli* and *Shigella* spp. Although some other SPATEs have also been described in other enterobacteria including *Serratia marcescens* [32], *Salmonella bongori* [33], *Citrobacter rodentium* [34], and *Edwardsiella tarda* [35].

Other conserved architecture in SPATEs include: (1) A highly conserved secretion domain/translocation domain called a β -domain. Overall, protein homology ranges from 25% to 55% but in the case of the β -domain, homology ranges from 60% to 90% identity [36]; (2) A conserved serine protease motif (consensus GDSGSP where S is the catalytic serine) at similar positions in their N-terminal passenger domain between residues 250–270 [25,26,37,38]; (3) The serine protease motif of SPATEs does not have a role in the cleavage of the passenger domain from the β -domain [33]; (4) The passenger domain of the SPATEs are cleaved from the β -domain from a conserved cleavage site between the two asparagines [25,39]; (5) All SPATEs have unusually long signal sequences (>50 amino acids) that have been shown to facilitate post-translational targeting [40]; (6) In contrast to the classical autotransporter IgA1 protease, none of the SPATEs can cleave IgA1; (7) SPATEs are highly immunogenic proteins having specific phenotypes [41].

3.1. Classification of SPATEs

SPATEs are subdivided into class 1 and class 2 on the basis of structural and functional properties. The class 1 SPATEs are cytopathic/cytotoxic toxins (eliciting cellular changes such as cytoplasmic shrinkage, loss of membrane integrity and activation of apoptosis), likely caused by cleavage of cytoskeletal proteins such as spectrin/fodrin [41]. Further, class 1 SPATEs are believed to show cytotoxic activity primarily through targeting of intracellular substrates. On the other hand, class 2 SPATEs, the larger phylogenetic cluster, comprise O-glycoproteases that cleave mucin and other O-glycoproteins present not only on epithelial cells but also on the surface of hematopoietic cells [42]. Thus, due to mucinolytic activity, class 2 SPATEs can impart a subtle competitive advantage in mucosal colonization [41]. Sat, Pet, EspP, EspC, and SigA SPATEs belong to class 1 whereas Pic, PicU, Tsh/Hbp, Vat, EatA, and SepA belong to class 2 SPATEs.

3.2. Distribution of SPATEs among Intestinal and Extra-Intestinal Pathogenic *E. coli*

Some SPATEs are present in one or more pathotypes of *E. coli* (Figure 3). Generally, the following SPATEs are found in various groups: EspP (extracellular serine protease plasmid (pO157-encoded) from enterohemorrhagic *E. coli* (EHEC) [26], Pet (plasmid-encoded toxin) from enteroaggregative

E. coli (EAEC) [25], Pic (protein involved in intestinal colonization) from EAEC and uropathogenic *E. coli* (UPEC) and *Shigella* [39], EspC (EPEC secreted protein C) from enteropathogenic *E. coli* (EPEC), EatA (ETEC autotransporter A) from enterotoxigenic *E. coli* (ETEC) [43], Tsh (temperature-sensitive hemagglutinin)/Hbp (hemoglobin protease) mainly in avian pathogenic *E. coli* (APEC) and some ExPEC (MNEC-UPEC) [37,44], Sat (secreted autotransporter toxin) from UPEC [45], and Vat (vacuolating autotransporter toxin), TagBC (Tandem autotransporter genes B and C), and Sha (Serine-protease hemagglutinin autotransporter) from APEC and UPEC strains [31,46].

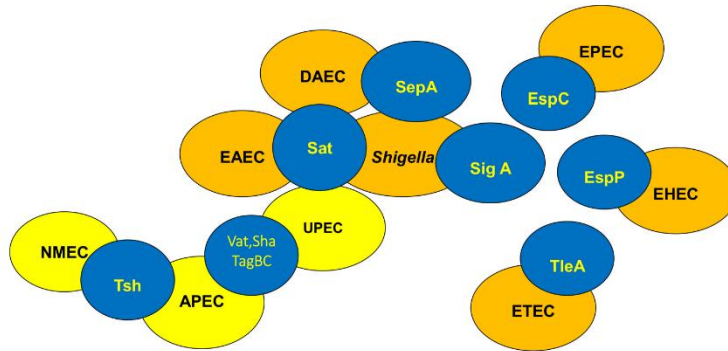


Figure 3. Distribution of SPATEs among intestinal and extra-intestinal pathogenic *E. coli*. SPATEs have been found not only in all recognized intestinal *E. coli* pathotypes (highlighted in orange) but also in extra-intestinal *E. coli* pathotypes (highlighted in yellow). The recent availability of many more bacterial genome sequences, has led to identification of new and previously described SPATE proteins among both human and animal pathogens, including *Salmonella*, *Citrobacter*, and *Edwardsiella*, as well as some commensal *E. coli* strains. (APEC, Avian pathogenic *E. coli*; DAEC, Diffuse Adhering *E. coli*; UPEC, Uropathogenic *E. coli*; NMEC, Neonatal meningitis *E. coli*; EAEC, Enteroaggregative *E. coli*; EHEC, Enterohemorrhagic *E. coli*; EPEC; Enteropathogenic *E. coli*, ETEC, Enterotoxigenic *E. coli*).

3.3. Allelic Variation

Phylogenetic comparisons of SPATE sequences available in the National Center for Biotechnology Information (NCBI) databases demonstrate that these proteins share high homology among their fellow members because of the high degree of amino acid identity in the C-terminal β -domains [47]. If we consider only the passenger domain, the percentage of homology drops considerably and signifies that different members contain distinct regions and have differing biological functions (Figure 4). In addition, some predicted SPATEs are related allelic variants of some of the characterized SPATEs. Despite demonstrating some closer identity to some of the characterized SPATEs, only functional testing of these proteins can confirm their specific bioactivities.

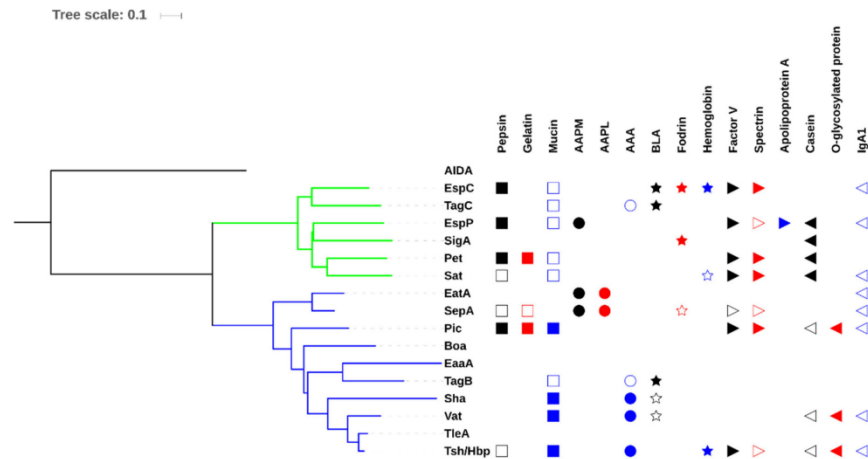


Figure 4. Evolutionary relationships of SPATEs based on passenger domain sequences and presentation of known protease substrates. The evolutionary history of passenger domains of characterized SPATEs was inferred using the neighbor-joining method [48]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method [49] and are in the units of the number of amino acid substitutions per site. The analysis involved 17 protein sequences. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6 [50]. Multiple sequence alignment was performed by Clustal W and the tree was constructed using the Mega6 software with PhyML/bootstrapping and iTOL [51]. Cluster of cytotoxic SPATEs (class 1) are in green branches while immunomodulator SPATEs (class 2) are in blue branches. SPATE protein sequences are available in NCBI database as follows: EspC, GenBank Accession No. AAC44731; EspP, NP_052685; SigA, AF200692; Pet, SJK83553; Sat, AAG30168; EatA, CAI79539, SepA, Z48219; Sha, MH899684; Pic, ALT57188; Boa, AAW66606; TagB and TagC, MH899681; EaaA, AAF63237; Vat, AY151282; TleA, KF494347; Tsh/Hbp, AF218073; AIDA, ABS20376; AAPM: MeOSuc-Ala-Ala-Pro-Met-pNA, AAPL: Suc-Ala-Ala-Pro-Leu-pNA, AAA: N-Succinyl-Ala-Ala-Ala-p-nitroanilide, BLA: N-Benzoyl-L-arginine 4-nitroanilide. The cleaved substrates, substrates that were negative for cleavage, and those untested substrates are represented by filled symbol, unfilled symbol and no symbol respectively. This comparison shows that some of the activities are phylogenetically distributed.

3.4. Confusion Due to Improper Annotation of Uncharacterized SPATE Encoding Proteins

Due to the presence of conserved domains and similarities in protein sequences among SPATEs, there are some cases where SPATE protein sequences were misnamed or given the name of a related, but distinct SPATE. For example, Vat (GenBank Accession No. AAN78874) was annotated as Tsh/Hbp in the UPEC CFT073 genome although this protein shares 78% identity with Tsh/Hbp [33,52,53]. Similarly, in the genome of *E. coli* PCN033, which was isolated from a pig with meningitis, a putative SPATE gene (GenBank Accession No. AKK51062) is named as EspC, despite this protein only having 59% identity with EspC (GenBank Accession No. WP_109867760). So, for annotations in some enterobacterial genomes, it is unfortunate that numerous uncharacterized SPATEs are incorrectly labeled as specific characterized SPATEs, despite sharing a limited degree of identity, particularly in the absence of demonstration of any biological or experimentally confirmed activities.

4. SPATEs Demonstrate a Diversity of Biological Activities Associated with Virulence

Since the discovery of the first SPATE, Tsh, two decades ago, research has focused on characterizing SPATEs by determining their biological activities and substrate specificities and in some cases, structural

properties have been studied through crystallography. Serine protease activity is due to the GDSCG motif and this activity is inhibited by phenylmethane sulfonyl fluoride (PMSF), but not by the metalloprotease inhibitor, Ethylene diamine tetra acetic acid (EDTA) [26]. The virulence properties of SPATEs may in part be attributed to proteolytic activity. However, it is now clear that there is a staggering diversity in biological substrates and modes of action. In the next sections, we focus on specific characteristics and functions of various SPATEs (summarized in Table 1) and their possible roles in enterobacterial pathogenesis.

Table 1. Summary of characteristics of different SPATEs.

SPATEs	Organism ^a	Biological Functions	References
EatA	ETEC	Enterotoxin	[54]
EspC	EPEC	Cytotoxin, Enterotoxin Cleavage of fodrin, hemoglobin, pepsin, coagulation factor V, translocator components (EspA/EspD) of T3SS Cell rounding and cell detachment	[41,55–58]
Pet	EAEC	Mucosal cytotoxicity, Cleavage of spectrin, pepsin, factor V	[25,41,59–61]
Pic	Shigella, EAEC	Serum resistance Mucinase, Hemagglutination Colonization, Cleavage of gelatin, factor V, O-glycans: PSGL-1, CD44, CD45, CD93 and CX3CL1	[39,41,42,52,62,63]
EspP	EHEC, STEC	Cleaved pepsin, factor V, apolipoprotein, complement factors: C3/C3b and C5	[26,41,64–66]
Tsh/Hbp	APEC	Hemagglutinin, Binding to Caco-2 cells and to EMPs (laminin, fibronectin, and collagen IV) and heme Cleavage of mucin, factor V and O-glycosylated proteins in leukocyte	[37,41,63,67]
Sha	APEC, UPEC	Autoaggregation, hemagglutination, biofilm formation, proteolytic activity on synthetic peptide: N-Succinyl-Ala-Ala-Ala-p-nitroanilide, adherence and cytopathic effects on bladder epithelial cell line	[31]
TleA		Binding to Caco-2 cells Cleavage of bovine submaxillary mucin, leukocyte surface glycoproteins CD45 and P-selectin glycoprotein ligand 1	[68]
Vat	APEC, UPEC	Vacuolating cytotoxin, Agglutinate leukocyte Cleavage of O-glycosylated proteins in leukocyte	[46,63]
Sat	UPEC	Vacuolating cytotoxin on HK-2, HEp-2 and Vero monkey kidney cells Cleavage of casein, factor V and spectrin	[41,45,69]
SepA	<i>Shigella flexneri</i>	Intestinal inflammation, proteolytic activity toward synthetic peptides: Suc-Ala-Ala-Pro-Phe-pNA, Suc-Val-Pro-Phe-pNA and Suc-Phe-Leu-Phe-pNA	[70,71]
SigA	<i>Shigella flexneri</i>	Cytotoxin, Cleavage of casein, recombinant human α II spectrin Cell rounding and cell detachment	[72,73]
Boa	<i>Salmonella bongori</i>	Unknown	[74]
TagBC	UPEC, APEC	Autoaggregation, proteolytic effect on synthetic peptide: N-Benzoyl-L-arginine 4-nitroanilide cytopathic effect on human bladder cell lines	[31]

^a Bacteria known to produce these SPATEs. APEC, Avian pathogenic *E. coli*; DAEC, Diffuse Adhering *E. coli*; STEC, Shiga toxin-producing *E. coli*; UPEC, Uropathogenic *E. coli*; NMEC, Neonatal meningitis *E. coli*; EAEC, Enteroaggregative *E. coli*; EHEC, Enterohemorrhagic *E. coli*; EPEC, Enteropathogenic *E. coli*; ETEC, Enterotoxigenic *E. coli*.

4.1. EaaA/EaaC

The EaaA and EaaC SPATEs (GenBank Accession No. Protein AAF63237.1, Nucleotide AF151091) were first identified from *Escherichia coli* reference strain ECOR-9, a human commensal intestinal isolate. These two SPATE-encoding genes were found to be associated with genes encoding non-immune immunoglobulin binding proteins that were also associated with prophages [75]. The EaaA and EaaC SPATEs are very similar 1335 aa proteins, sharing 99% identity (only 8 aa substitutions). Other than identification of these genes in ECOR-9, *eaa* gene sequences were also shown to be present in other ECOR strains (ECOR-2, ECOR-5, and ECOR-12) belonging to phylogenetic group A, but these SPATE sequences were not identified in a variety of clinical isolates [76]. However, screening of genomic sequence databases indicates EaaA/EaaC sequences are present in a diversity of *E. coli* strains (at least 50 entries of highly similar proteins identified from uniprot.org). Other than identification of the sequences in different strains, no phenotypic or biochemical properties of Eaa SPATEs have been investigated thus far.

4.2. EatA

EatA for ETEC autotransporter A (GenBank Accession No. CAI79539, Q84GK0; AY163491.2) is secreted by some ETEC strains and has been shown to contribute to virulence in the rabbit ileal loop model of infection, since an *eatA* mutant demonstrated less marked and slowed fluid accumulation [43]. The *eatA* sequence was identified on a plasmid, pCS1, in *E. coli* strain H10407. The EatA protein shares over 80% identity to the SepA SPATE from *Shigella flexneri*. EatA was found to degrade a bacterial adhesin EtpA, which could reduce intestinal colonization, but in parallel increased access of ETEC toxins at the host cell surface [54]. EatA was also shown to be highly immunogenic and could contribute to ETEC virulence by degrading the MUC2 protein at the small intestinal mucous layer, which could further promote access of ETEC toxins to epithelial cell surfaces [77]. The EatA protein was also shown to be present in most ETEC (over 70%) [78] and has been identified less commonly in some EAEC strains (4.1% of isolates) [79]. EatA has vaccine potential for prevention of ETEC, as it generated a high antibody response and protection against ETEC intestinal infection [77].

4.3. EspC

The EspC (EPEC secreted protein C) passenger domain is a 110 kDa protein (GenBank Accession No. AAC44731, U69128.1), and one of the first proteins reported to be secreted by EPEC [38]. As with other SPATEs, although EspC shares some sequence homology to IgA proteases, it cannot cleave IgA nor is the catalytic serine GDSGDG motif required for release of the passenger domain from the β -domain into the external milieu [38]. It was also shown that EspC is not involved in EPEC generation of attaching and effacing (A/E) lesions, nor is it required for adherence or invasion of tissue culture cells. EspC demonstrated enterotoxic activity and increased tissue PD (potential difference) and Isc (short circuit current) of rat jejunum mounted in Ussing chambers [80]. EspC enterotoxic activity was nullified by pre-incubation with an antiserum against another SPATE, Pet [80]. Similar to Pet, EspC produced cytotoxic effects on cultured epithelial cells but with three times higher dose (120 $\mu\text{g}/\text{mL}$) than Pet. The actin cytoskeleton was disrupted, resulting in cell contraction and cell detachment. This overall effect was caused by the serine protease motif of EspC [55]. Also, similar to Pet, EspC cleaved an intracellular target, α -fodrin but the cleavage sites were different. Pet cleaves fodrin within the calmodulin binding domain between M¹¹⁹⁸ and V¹¹⁹⁹ [61]. EspC cleavage of fodrin occurred outside of the calmodulin binding domain [55]. Once inside the cells, kinetics of protein degradation indicate that purified EspC cleaves fodrin at two sites (within the 11th repetitive unit between Q¹²¹⁹ and L¹²²⁰ and within the 9th repetitive unit between D⁹³⁸ and L⁹³⁹) which then results in disruption of focal adhesion including dephosphorylation and degradation of paxillin and FAK; leading to cell rounding and detachment [58]. However, entry to a target cell (cytosol) is critical for EspC cytotoxicity. Though internalization of purified EspC by pinocytosis was shown by [81], it cannot be considered as

a natural physiological phenomenon for infection, as it took 8 h of incubation for insertion while EPEC infection delivers EspC into the cells after 30 min. Interestingly, EspC is secreted into the milieu by the T5SS and then incorporated into the T3SS translocon for entry into host cells [82]. Further, EspC has a relevant role in cell death induced by EPEC. EspC is able to induce apoptosis and necrosis in epithelial cells [56] and apoptosis could be the first event which can manifest to increased necrosis. Also, EspC was shown to interfere with the caspase cascade required for induction of apoptosis which was partially dependent on serine protease activity.

A number of biological targets recognized by EspC which are relevant to its diarrheagenic activity have been identified. Purified EspC cleaved human hemoglobin at an optimum pH between 5 and 6 [57]; although correlation of this biological property for EPEC virulence is yet to be established. EspC has also been shown to cleave other substrates like pepsin, glycoprotein coagulation factor V, and spectrin [41]. Apart from cleaving different biological substrates, EspC also cleaves bacterial components of the secretion system. EspC was found to target EspA/EspD which are translocator components of the Type III secretion system (T3SS) and control of pore formation and cytotoxicity by T3SS for the host cell [83]. The T3SS acts as a molecular syringe, comprised of pore-forming translocator proteins EspB and EspD which insert into the host cell plasma membrane and EspA which forms a hollow structure connecting the T3SS needle into the cell [84]. This result indicates that control of pore formation by EspC can support bacterial colonization, by mediating a controlled release of effector proteins from the T3SS to limit host cell death, since this could increase the immune response and potential clearance of EPEC at an early stage of infection.

4.4. *Pet*

Pet (Plasmid encoded toxin) is a 104 kDa enterotoxin produced by EAEC (GenBank Accession No. SJK83553), which has been found to increase jejunal potential difference (PD) and *Isc* (short circuit currents) accompanied by mucosal damage, exfoliation of cells and development of crypt abscesses [25]. *Pet* is encoded on the pAA plasmid and comprises a 52 aa N-terminal signal peptide and the secreted passenger domain (amino acids 53–1018), which cleaves from the β -domain between N¹⁰¹⁸ and N¹⁰¹⁹ [25].

A host-specific factor is required for proper folding of the *Pet* autotransporter. Interestingly, clones of *E. coli* HB101 produced both folded and misfolded variants of *Pet* but the wild-type EAEC only produced the properly folded active *Pet*, suggesting that the accessory protein from EAEC may be absent or non-functional in strain HB101 [85]. This observation shows that correct protein folding is not required for AT secretion but that accessory proteins are necessary for folding ATs in the right conformation [11,13,86]. X-ray structure of the *Pet* passenger domain was resolved and when compared with the most similar SPATE, EspP (50% sequence identity); *Pet* harbors a β -pleated sheet from residues 181–1900 whereas EspP has a coiled loop. Further, the *Pet* passenger domain showed more β -sheets between residues 135–143 compared to EspP [8]. These β -helices are presumed to confer functionality to the protein.

Pet produced changes in host cytoskeletal architecture in both HEP-2 and HT29 epithelial cells characterized by time and dose dependent cell elongation followed by cell rounding and detachment from the substrate, which was dependent on serine protease activity [60]. Cellular morphological changes were visible after 2 h of incubation with *Pet* (25 μ g/mL). *Pet* also contributes to pathology at the mucosa which is characterized by dilation of crypt opening, extrusion of colonic enterocytes, development of intercrypt crevices and loss of apical mucus from goblet cells as a result of contraction of interlinking cytoskeleton integrity and loss of actin stress fibers and focal contacts [59]. These cytopathic effects may be due to internalization in host cells by the vesicular system as *Pet* activity completely vanished following incubation with brefeldin A [87]. *Pet* interaction with host cells requires a two-hour time lag leading to cell damage and requires a sequence of events: (1) Binding to the cell. (2) Entry into the cell by clathrin-dependent receptor-mediated endocytosis. (3) Entry into early endosomes. (4) Passage to the Golgi apparatus from endosomes. (5) Retrograde vesicular transport from the

Golgi complex to the ER. (6) Delivery to the cytosol through the ER-associated degradation (ERAD) pathway [61,88]. Once internalized, loss of actin microfilaments takes place due to breakdown of cellular spectrin [60,89]. Similar to Pet, EspC also produces cytotoxic activities on epithelial cells, although the dose of EspC was three times higher and a longer incubation time was required to produce a similar result [55]. Subsequently, it was shown that the d2 subdomain of the passenger domain is required for Pet internalization by recognizing the Pet host cell receptor, cytokeratin 8 [90]; the d1 subdomain is the largest domain having a serine protease motif and was incapable of binding the cell surface without the aid of d2 subdomain [90,91]. Pet preferentially cleaves α -fodrin between M¹¹⁹⁸ and V¹¹⁹⁹ residues within the calmodulin-binding domain of fodrin's 11th repetitive unit [61,92]. Pet-mediated cleavage of α -fodrin (spectrin) has been suggested to induce enterocyte death via apoptosis [93,94]. Together, this phenotype of Pet could explain the cellular alteration during EAEC pathogenesis [92]. Recent studies have shown that the spice, curcumin, can also affect the secretion of Pet on the bacterial surface and subsequent internalization into the epithelial cells [95].

4.5. Pic

Protein involved in colonization (Pic) is a 109.8 kDa extracellular protein secreted by both EAEC (GenBank Accession No. ALT57188, AF097644.1), and *Shigella flexneri* 2a [39] and also atypical EPEC [96]. Pic catalyzed mucin degradation and has also been shown to confer serum resistance and hemagglutination [39]. Further, Pic also cleaved gelatin, but did not demonstrate any activity against human immunoglobulins. In addition to Pic identified in diarrheagenic *E. coli*, another highly similar SPATE named PicU (96% amino acid identity to Pic) has been identified in some uropathogenic isolates. PicU was functionally similar to Pic from EAEC, degraded mucin and contributed to colonization during urinary tract infection (UTI) [52,62]. Interestingly, mucinase activity was not only important as a virulence factor but also may contribute to nutrient availability, since the *pic* mutant was less able to grow when compared to the wild-type strain [97]. The Pic mucinase (from all groups EAEC, UPEC, and *Shigella flexneri*) is responsible for increased secretion of mucus in the intestinal lumina of rat ileal loops by increasing mucus production in goblet cells even though this activity was independent of the serine protease motif [98]. This secretory activity of Pic favors the formation of biofilm by EAEC, a hallmark of EAEC infection. The mucolytic activity of Pic not only contributed to damage of the intestinal mucosal layer, but also cleaved mucin-type O-glycans of the immune system, including PSGL-1, CD44, CD45, CD93, and CX3CL1 [42,63]. Further, Pic significantly reduced complement activation by cleaving complement cascade factors- C3, C4 and C2 [99]. Downregulation of complement activation by Pic may contribute to EAEC, *Shigella* and UPEC infections. To add to its virulence potential, Pic was also shown to induce polymorphonuclear leucocytes/neutrophil (PMN) activation and programmed T-cell death [42].

PicU was identified in a UPEC strain and also demonstrates mucinase activity that may contribute to UTI pathogenesis [62]. To support this observation, a *picU* mutant of *E. coli* CFT073 was less able to colonize compared to wild-type parent although differences were not statistically significant. In addition, pepsin and coagulation factor V are other cleavage substrates for PicU [62].

Interestingly, a SPATE in *Citrobacter rodentium* similar to Pic and PicU, named PicC (79% identity at amino acid level), was shown to demonstrate mucinolytic activity [100]. The *picC* mutant actually outcompeted the wild-type and elicited more colitis. The PicC protease may therefore be an important immune regulator that could function to decrease stimulation of the host immune system during infection [100].

4.6. EspP

The extracellular serine protease plasmid (pO157-encoded) (EspP) was isolated from the culture supernatant of EHEC O157: H7 EDL933 associated with hemolytic uremic syndrome (HUS) (GenBank Accession No. NP_052685; CAA66144, X97542.1) [26] and Shiga toxin-producing *Escherichia coli* (STEC) [64,101]. PssA (protease secreted by STEC) is the homologue of EspP which differs by a single

amino acid change and was found to be cytotoxic for Vero cells [102]. EspP is generally associated with STEC, EHEC and atypical EPEC [103]. In silico analysis suggests that the signal sequence of EspP is cleaved between residues A⁵⁵ and A⁵⁶ [26] and the 104 kDa passenger domain contains a chaperone motif and a 30 AA linker domain that connects the β -domain and the passenger domain [104]. It was later shown that biogenesis and export of EspP was not self-mediated but also required additional periplasmic chaperones SurA and DegP, which aided in the proper folding of the AT passenger domain during translocation through the Bam complex [105].

EspP proteins have been classified as four distinct alleles, namely EspP α , EspP β , EspP γ , and EspP δ , where EspP α is associated with highly virulent EHEC O157: H7 and major non-O157 EHEC and can contribute to biofilm formation by forming macroscopic rope-like polymers which were refractory to antibiotics and showed adhesive and cytopathic effects [106]; EspP γ cleaved pepsin and human coagulation factor V, although EspP β and EspP δ were either not secreted or proteolytically inactive [64]. Interestingly, EspP α was more prevalent in human isolates (84%) than in environmental isolates (47%) but EspP γ was more prevalent in the environment (40%) than from human sources (11%) [64,107]. The diversity of EspP alleles with different proteolytic activities necessitates specific subtyping for the screening of *espP* genes. The crystal structure of the EspP passenger domain was solved at 2.5 Å, and revealed a large β -helical stalk and a globular subdomain with the catalytic triad [6] and shared overall structure to the previously crystallized Hbp protein [7]. However, in contrast to Hbp, the active site of EspP α is slightly wider, deeper and more exposed, suggesting that it likely interacts with larger substrates [6].

Esp was shown to cleave substrates such as coagulation factor V [26], porcine pepsin A [26], apolipoprotein [65], major complement factor proteins C3/C3b but not factors H and I [66]. A valid argument for prolonged hemorrhage due to EspP α is strengthened by the ability of EspP α to cleave various serpins (serine protease inhibitors) from human plasma which are involved in blood coagulation [108]. Cleavage was specific, targeting only procoagulatory serpins such as α 2-AP and α 1-PI [108]. Like EspP α , EspI, another SPATE from STEC also cleaved α 2-AP and α 1-PI [65], but unlike EspP α , EspI mediated cleavage was not complete due to the formation of an inhibitory serpin-enzyme-complex [108]. Further, EspP was found to stimulate electrogenic ion transport in human colonoid monolayers, although this activity was Ca²⁺ dependent but independent of serine protease activity [109]. Taken together, the role of EspP in blood coagulation, pathophysiology and immunomodulation can contribute to pathogenesis of EHEC.

4.7. Tsh/Hbp

The first characterized SPATE, temperature-sensitive hemagglutinin (Tsh) was identified on a ColV-type plasmid in APEC strain χ 7122 [37,110]. The *tsh* gene (GenBank Accession No. AF218073) encodes a protein of 1377 amino acids with a molecular weight of approximately 148-kDa. It is composed of a leader sequence that is cleaved between residues A⁵² and A⁵³, a 106-kDa passenger domain that encompasses the serine protease motif (S₂₅₉) extending from residues 53 to 1100, and a β -barrel domain of 33-kDa that extends from residues 1101 to 1377 [37]. Based on the sequence homology, Tsh was also reported in some UPEC strains [52]. However, Heimer et al. were actually screening for *vat* not *tsh*, and they erroneously named the *vat* gene in CFT073 and other UPEC as *tsh*. However, it has been reported in one study that some human ExPEC strains from newborn meningitis (11 to 50 percent) also contain *tsh* located on large plasmids similar to ColV plasmids [111]. A Tsh-like protein sharing 60% aa identity with *E. coli* Tsh was also reported in *Edwardsiella tarda*, a fish pathogen [35]. Despite the 60% identity of this SPATE from *Edwardsiella tarda* to Tsh, the bioactivity of this protein and its role in virulence fish may be quite distinct.

The production of Tsh in *E. coli* K-12 was found to be higher at low temperature (26 °C) and it conferred the capacity to agglutinate chicken erythrocytes in a mannose resistant manner. However, at higher temperatures, Tsh was released into the supernatant medium and this agglutination activity was lost leading the name, temperature-sensitive hemagglutinin [37]. Interestingly,

the hemagglutination phenotype was also observed for Tsh with sheep [112], bovine, pig, turkey, rabbit, horse, and dog [31] and human erythrocytes [63]. Furthermore, Tsh promoted adherence to Caco-2 cells and to extracellular matrix proteins such as laminin, fibronectin, and collagen IV [112]. Sequence analyses indicated that Tsh is a homologue to IgA proteases (56% of similarity) of *Hemophilus influenzae* and *Neisseria gonorrhoea* [113], but was unable to cleave IgA. Tsh was shown to cleave mucin, factor V [41] and O-glycosylated proteins such as CD43, CD44, CD45, CD93, CD162, and CX3CL1 in vitro [42,63], suggesting diverse roles in cellular and immune functions. Tsh seems to also contribute directly to APEC infection, as its presence accelerates the progression of the infection and could lead to the development of lesions and deposition of fibrin in avian air sacs [37,112]. In addition to these functions, Tsh was shown to have potential enterotoxin function to induce fluid accumulation in a rabbit ligated ileal loop assay [114], although the significance of this role for Tsh in mammalian enteric disease is unknown.

Hbp (hemoglobin protease) is a near-identical variant of Tsh that was isolated from a patient with an intra-abdominal wound infection, and only differs by two amino acids (Q209K and A842T) (GenBank Accession No. AJ223631) [44]. Like *tsh*, the *hbp* gene is located on a virulence plasmid, pColV-K30. Hbp was shown to cleave hemoglobin and acquire heme in an iron depleted host niche [44]. Heme captured by Hbp could also promote growth of *Bacteroides fragilis*. *B. fragilis* was shown to contain a specific receptor that recognizes the heme-Hbp complex and it is capable of exploiting heme liberated from host complexes and cause intra-abdominal abscesses in patients [67].

4.8. TleA

Tsh-like ETEC autotransporter or TleA (GenBank Accession No. KF494347) is an AT that was identified in an ETEC strain [68]. It is a 4110-bp gene that encodes a 1369-amino-acid precursor. Sequence analyses have shown that TleA contains a signal peptide from residues 1 to 52, a passenger domain from residues 53 to 1092 with a serine protease motif GDSGS from residues 257 to 261 and a β -barrel domain from residues 1093 to 1369. The alignment of the passenger domain sequences of the other members of SPATEs indicated that TleA is a class 2 SPATE and shares 97% identity with the Tsh autotransporter [68]. TleA may have a role in intestinal colonization and immunomodulation as it was shown to degrade bovine submaxillary mucin and leukocyte surface glycoproteins CD45 and P-selectin glycoprotein ligand 1. Further, nonadherent *E. coli* HB101 expressing TleA conferred the capacity to adhere to Caco-2 cells, while such adherence was not observed in the wild-type ETEC strain 1766a [68].

4.9. Vat

Vacuolating autotransporter toxin (Vat) is a 140 kDa class II SPATE that was found to be encoded on a pathogenicity island from APEC *E. coli* Ec222, and was shown to contribute to respiratory tract infection, cellulitis and septicemia in poultry [46]. The *vat* gene is located on a pathogenicity island (VAT-PAI) between the *proA* and *yagLI* genes (GenBank Accession No. AY151282). The *vat* pathogenicity island in Ec222 contains 33 ORFs wherein *vat* is ORF27. Interestingly, ORF26 share 44% aa identity with the PapX regulator from UPEC strains [46]. The *vat* predicted gene product also shares 97% identity with an AT present in UPEC strain CFT073, which was annotated mistakenly as Tsh/Hbp. However, the new annotation has been corrected as Vat-ExPEC [33,52,53]. Vat shares 77.5% identity to Tsh/Hbp of APEC [44,110], which explains why Vat has sometimes been mislabeled as Tsh for certain strains or genome sequences.

Unlike Tsh/Hbp, the secreted 111.8-kDa Vat passenger domain was unable to cleave casein-based substrate [46]. Compared to the *vat* mutant, APEC Ec222 exhibited a robust cytotoxic effect on chicken embryonic fibroblast (CEF) cells. Cytotoxic activity involved vacuole formation in cells that was visualized from 2–24 h after exposure [46]. In a cellulitis infection model, chickens challenged with *E. coli* Ec222 developed cellulitis, whereas none of the chickens developed cellulitis when infected with the *vat* mutant [46]. In UPEC, Vat elicits an antibody response in some urosepsis patients, and

the titer of Vat-specific IgG was higher in the plasma of patients compared to the titer found in controls and patients infected with *vat*-negative UPEC strains [115]. Vat contributed to fitness in UPEC during murine systemic infection [116]. Further, a role for Vat in combination with other SPATEs was demonstrated for fitness in murine kidney colonization and cytotoxicity [31]. Vat has also been shown to mediate agglutination of erythrocytes and cleavage of O-glycoproteins in vitro [63].

Vat (Vat-AIEC) from an adherent invasive *E. coli* (AIEC) associated with Crohn's disease, is 97% similar to Vat from APEC Ec222 and contains a modification in the serine catalytic domain to GDSGSP instead of the conserved ATSGSP motif present in Vat [117]. Vat-AIEC acts as a mucinase and inactivation of *vat* reduced gut colonization by one-log in a mouse model [117].

4.10. *Sha*

A recently identified SPATE-encoding gene, *sha* (serine-protease hemagglutinin autotransporter), is located on a distinct region of a ColV-type plasmid [31]. The *sha* gene (GenBank Accession No. MH899684) was more common among APEC (present in 20% of 299 APEC strains) than in UPEC (0.9% of 697 strains). *Sha* is more closely related to Tsh (43% identity) and Vat proteins (38% identity) than to other SPATEs. Similar to Vat and Tsh, *Sha* showed elastase-like activity by cleaving N-Succinyl-Ala-Ala-Ala-p-nitroanilide. In addition, *Sha* increased adherence to both human and avian epithelial cells, whereas Tsh only increased adherence to bladder cells, and Vat only increased adherence to kidney cells. *Sha* also demonstrated hemagglutination for erythrocytes of a variety of animal species (sheep, bovine, pig, dog, chicken, turkey, rabbit, horse, and human), contributed to increased biofilm formation, and delayed cytotoxicity (release of LDH after 12 h from bladder epithelial cells). As similar carbohydrates may be present on erythrocyte surfaces, it is not surprising that *Sha* as well as Tsh and Vat autotransporters demonstrated extensive hemagglutination activity for a variety of erythrocytes. We believe that the contrasting phenotype when *Sha* is expressed in high copy vector—agglutination with erythrocytes by the protein present on the bacterial surface as well as the cytopathic effects of released protein in culture supernatant could be explained by the presence of protein both the bacterial surface and released in the supernatant (Figure 5). Although, loss of *sha* didn't affect competitive fitness during colonization of the urinary tract of female mice, *sha* expression was upregulated six-fold in infected bladder compared to culture in lysogeny broth (LB) [31].

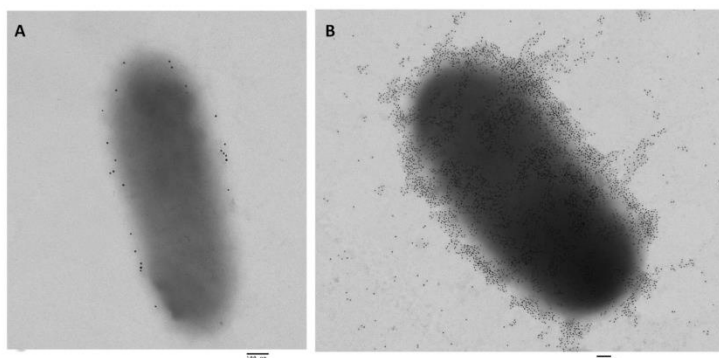


Figure 5. Transmission electron micrographs of *E. coli* BL21 expressing the Sha autotransporter [31] immunolabelled with 10-nm-diameter gold particles (A) The negative control, adsorbed serum was used for the primary incubation with *E. coli* BL21 empty vector and shows very low background of gold particles. (B) Gold-labelling with AT-specific antibodies shows that when *Sha* is constitutively expressed, proteins are localized on the surface as well as released in the supernatant. So, the protein released can gain access to host cell targets, whereas the protein associated with bacterial cells can mediate agglutination, autoaggregation, and adherence to host cells. Bar 100 nm.

4.11. Sat

A Secreted autotransporter toxin (Sat) was identified in UPEC strain CFT073 [45] and later also described in *Shigella*, EAEC, DEAC and neonatal septicemia *E. coli* strains [118]. The 3885-bp *sat* gene (GenBank Accession No. AAG30168) is located on a pathogenicity island that also carries a *pap* fimbrial gene cluster. The *sat* gene product is a 1295-amino-acid precursor with a molecular weight of 142 kDa [45]. Sat was unable to agglutinate human erythrocytes and did not cleave glycoproteins CD43 and CD162 from Jurkat cells [63]. Sat was also shown to cleave casein, coagulation factor V and nonerythroid spectrin, but not pepsin or mucin [41,45]. Like Vat, Sat also exhibited cytotoxicity on epithelial cells, including HK-2, HEP-2 and Vero cells, characterized by vacuole formation, autophagy and cell detachment [45,69,119]. Cytotoxicity included disruption of actin and other cytoskeletal and nuclear proteins and was dependent on the serine protease active site [120]. Sat is internalized by an unknown mechanism, and shown to be localized specifically to the cytoskeletal fraction of bladder and kidney epithelial cells [120].

The autophagy in HeLa epithelial cells triggered by Sat led to disruption of the F-actin cytoskeleton [119]. Sat also modified tight junction-associated proteins ZO-1, ZO-2 and occludins in human polarized epithelial intestinal Caco-2 and TC7 cells, resulting in increased paracellular permeability [121]. In addition, Sat induced a strong immune response in a murine model of ascending urinary tract infection, although a *sat* mutant colonized urine, bladder and kidneys as well as the wild-type strain [45]. However, kidneys of mice infected with the wild-type strain showed dissolution of the glomerular membrane and vacuolation of proximal tubules and these lesions were absent in kidneys infected with the *sat* mutant [69]. In addition, Sat from a Diffusely adhering *E. coli* (DAEC) strain triggered pronounced fluid accumulation and villous necrosis in rabbit ileal tissue [122]. Taken together, Sat plays an important role as a cytotoxin in the pathogenicity of urinary tract infection as well as in intestinal infections. Still, the role of Sat as an enterotoxin is paradoxical, since Sat is also present in the probiotic strain *Escherichia coli* Nissle 1917 even though it was shown to be a functioning protease and was expressed during colonization of the mouse intestine [123].

4.12. SepA

Shigella extracellular protein A (SepA) was identified in *Shigella flexneri*, which causes diarrhea (shigellosis) in humans [70] and also in some enteroaggregative *E. coli* (EAEC) strains [103]. The *sepA* gene (GenBank Accession No. Z48219) is located on the 200 kb virulence plasmid pWR100 and encodes a 1366 a.a. precursor of 146 kDa [70]. SepA exhibited protease activity of some synthetic peptides such as Suc-Val-Pro-Phe-pNA [71], although no protease activity was found on natural substrates such as gelatin, IgA1 [70], angiotensin-I, egg lysozyme [71], fibronectin, mucin, pepsin, factor V, spectrin, or fodrin [41]. Unlike other class-2 SPATEs, SepA did not cleave O-linked glycoproteins from leucocytes such as CD43, CD44, CD45, CD93, CD162, and CX3CL1 [63]. Deletion of *sepA* did not affect entry of *S. flexneri* into epithelial cells or cell-to-cell spread; however, the *sepA* mutant showed a decrease in fluid accumulation and inflammation in the rabbit ileal loop model compared to the wild type [70]. Furthermore, in a human explant model, the *sepA* mutant demonstrated reduced mucosal damage and a significant reduction in desquamation of intestinal epithelial barrier [124]. SepA also induced disruption of the apical pole of a polarized epithelial barrier and facilitated invasion of intestinal cells by *Shigella* [125]. This was associated with a decrease in LIMK1 leading to increased accumulation of colifin, a protein involved in actin dynamics [125,126].

4.13. SigA

Shigella IgA-like protease homologue or SigA is another SPATE identified in *Shigella flexneri* 2a [72] (GenBank Accession No. AF200692). The *sigA* gene is located on the *she* pathogenicity island and encodes a 103 kDa protein [72,74]. SigA imparts cytotoxic and enterotoxic effects and was shown to cleave casein [72] and recombinant human α II spectrin (α -fodrin) [73], indicating that it might

contribute to patho-physiological manifestation of *Shigella*. SigA shares 56% identity to the Pet. A *S. flexneri* 2a sigA mutant had 30% reduction in fluid accumulation in a rabbit ileal loop model. It was also shown that immunogenic SigA can bind to HEp-2 cells and induce cell rounding and detachment, phenotypes similar to purified Pet toxin from enteroaggregative *E. coli*, suggesting that SigA could play a key role in the virulence of *S. flexneri* [72,73]. Owing to the immunogenic properties of SigA [127], a computational approach has been used to identify potential epitopes for generation of a peptide vaccine against *Shigella* [128]. Due to their high immunogenicity, the potential of SigA and other SPATEs as vaccine targets is warranted.

4.14. Boa

Although most SPATEs identified to date were characterized from different pathotypes of *E. coli* and *Shigella* spp., Boa represents the only SPATE identified in *Salmonella* spp. Boa is present in *Salmonella bongori*, a *Salmonella* species associated mainly with reptiles, but which have been reported to infect some animals and humans. Surprisingly, no other SPATEs have been identified in all other *Salmonella enterica* serovars. It has been suggested that Boa may have been acquired by horizontal gene transfer from another Enterobacteria such as *E. coli* [74]. The *boa* gene encodes a 1384 protein (GenBank Accession No. AAW66606, FR877557) predicted to have a long signal peptide of 57 residues, a passenger domain of 1050 residues containing a serine protease GDSCS motif from residues 262 to 266 and a β -barrel domain of 277 residues [74,129]. To date, the biological properties of the Boa protein or its potential role for *Salmonella bongori* pathogenesis have not been determined.

4.15. TagBC

Recently, two new SPATEs referred to as “tandem autotransporter genes, *tagB* and *tagC*, located adjacent to each other on a genomic island between the conserved *E. coli* genes *yjdI* and *yjdK* in an APEC O1:K1 strain were identified [31]. Genome analysis and screening for *tagB* and *tagC* (GenBank Accession No. MH899681) indicate that these SPATE genes are present in APEC as well as UPEC strains [31]. Among 697 UPEC isolates, *tagB* sequences were present in 70 isolates (10%), whereas *tagC* sequences were present in 80 isolates. While amongst 299 APEC strains, *tagB* sequences were present in 14 isolates (4.7%) and *tagC* sequences were present in 21 isolates (7%). Interestingly, all of the *tagB* or *tagC*-positive APEC isolates were exclusively from infections in turkeys. Both proteins showed trypsin-like activity and efficiently cleaved N-Benzoyl-L-arginine 4-nitroanilide, similarly to the activity of EspC protein. Further, TagB and TagC were autoaggregating, hemagglutinating, could promote adherence to the HEK 293 renal and 5637 bladder cell lines as well as cytotoxic to human bladder cell lines when expressed in *E. coli* K-12 (with early release of LDH after 5 h). Neither of them required a functional serine protease motif for secretion [31]. In spite of these in vitro phenotypes, loss of TagBC did not have any appreciable effect on virulence or fitness of the mutant in bladder and kidneys of infected mice.

5. Regulation of Expression of SPATEs

Although numerous SPATEs have been characterized and their roles in infection and cell toxicity have been reported, determination of the mechanisms of regulation of SPATEs has been limited. In general, most SPATEs are thermally regulated and are better expressed under conditions similar to the host infection sites and upon contact with host cells, in tissue culture medium and at neutral to alkaline pH (7 to 9). EspP, EspC, Tsh, Pic, SigA, SepA are thermoregulated, although mechanisms of regulation remain to be elucidated [26,37,39,57,113]. Expression of the *vat-AIEC* gene was upregulated when grown at pH 7.5 with bile salts and mucin, which are similar to conditions in the distal ileal segment of the GI tract [117]. Expression of EspP in culture supernatant was higher in lysogeny broth (LB) than in minimal essential medium (MEM) as well as higher at 37 °C than at 20 °C [26,130]. Further, when EHEC 5236/96 (O26:H11) was grown in contact with human intestinal epithelial HCT-8 cells, *espP* expression was upregulated more than 35-fold [131]. Likewise, transcription of *pet* expression was

increased in tryptone-containing medium, which might be of clinical significance for the milk-drinking pediatric population that can be infected with EAEC. Similarly, *tsh* and *vat* expression was shown to be upregulated in minimal medium when compared to rich LB medium [31]. Despite determining what conditions increase expression of some SPATEs, defining which regulatory mechanisms control SPATE expression has been limited. Specific aspects of regulation of expression of some SPATEs are presented below.

5.1. Regulation by LER

The locus of enterocyte effacement (LEE) encoded regulator (Ler) regulates the LEE pathogenicity island of EPEC and EHEC which produce attaching and effacing lesions on host intestinal epithelial cells. Ler activates the transcription of various LEE operons [132]. Apart from regulating operons associated with the LEE and its Type 3 secretion system, Ler also strongly activates the *espC* promoter (by 31-fold) and hence increases the production of the EspC SPATE in EPEC [133].

5.2. Regulation by H-NS

Histones are small, abundant, highly conserved proteins that have been recognized as DNA binding proteins. They play a role in compacting DNA into the nucleosome, the main structures to form chromosomes, in eukaryotic cell nuclei [134,135]. In bacteria, such as *E. coli*, some proteins have been described as histone-like proteins. They may not share the same functions as compacting prokaryotic DNA but it was shown that they play a role in the regulation of genes by competing for binding to their promoter, and these genes could be associated with virulence, osmoregulation, pH and temperature sensing [136]. Based on sequence homology, four major groups of histone-like proteins were described: histone-like proteins *Escherichia coli* U93 (HU), histone-like nucleoid structuring proteins (H-NS), integration host factors (IHF), and factors for inversion stimulation (FIS). Among these, a role for H-NS has been reported for regulation of different AT encoding genes including regulation of the SPATE *Vat*.

H-NS has been shown to regulate different trimeric autotransporters [137–139]. Further, H-NS repressed the expression of *vat*; in UPEC strain CFT073, as a Δhns mutant was shown to secrete a significantly higher level of *Vat* than the wild type strain [115]. Further, sequence analysis of the *vat* promoter region predicted the presence of three potential H-NS binding sites in the *vat* promoter region [115]. It seems H-NS could also potentially directly regulate other SPATE genes, as there are predicted putative H-NS binding sites present in SPATE gene promoter regions (Table 2). However, further experimental evidence will be required to confirm whether or not regulation of expression of SPATEs by H-NS is a common phenomenon.

Table 2. The potential H-NS binding sites on the promoter region of SPATEs as predicted by Virtual Footprint Software ^a.

SPATEs	Potential H-NS Binding Sites
<i>boa</i>	⁻⁹⁷ GCAATAAAC ⁻⁸⁸ (-), ⁻⁹⁶ GCAATAAAAT ⁻⁸⁷ (-), ⁻⁸⁰ GCTATAAAAA ⁻⁷¹ (-)
<i>sigA</i>	⁻¹⁷⁹ TGGTTAGATA ⁻¹⁷⁰ (-), ⁻¹⁷⁰ GTGATTGATT ⁻¹⁶¹ (-), ⁻¹⁹ CCGATATTTC ⁻¹⁰ (-)
<i>pic</i>	⁻¹⁵⁹ CAGATAAAAC ⁻¹⁵⁰ (+), ⁻¹⁰⁹ TGCATTAATG ⁻¹⁰⁰ (-), ⁻³⁵ GGGATATAAA ⁻²⁶ (-)
<i>sepA</i>	⁻¹⁷⁶ ATGATAAAAA ⁻¹⁶⁷ (+), ⁻³⁵ AAGATTAATT ⁻²⁶ (-)
<i>tsh/hbp</i>	⁻¹⁶⁴ CACATAAAGT ⁻¹⁵⁵ (-), ⁻²⁸ AAAATAAAAT ⁻¹⁹ (-), ⁻¹⁰ GTAATTA AAA ⁻¹ (+)
<i>espC</i>	⁻³⁰⁰ ACCATTAAAA ⁻²⁹¹ (+), ⁻²⁹⁹ CCATTAAAAAT ⁻²⁹⁰ (+), ⁻¹¹¹ GCCACAAACT ⁻¹⁰² (-)
<i>espP</i>	⁻²⁸⁰ TCGATTGTTA ⁻²⁷¹ (-), ⁻⁹⁶ CAGATAAATG ⁻⁸⁷ (-), ⁻⁴⁶ CTGATACATT ⁻³⁷ (+)
<i>pet</i>	⁻¹⁷⁷ ATGATTAATT ⁻¹⁶⁸ (+), ⁻⁴² AGGATTAAGA ⁻³³ (-), ⁻²⁴ TCAATAAATG ⁻¹⁵ (+)
<i>sat</i>	⁻¹⁷⁷ ACGATCAATT ⁻¹⁶⁸ (+), ⁻¹⁶⁶ ACGATCAATT ⁻¹⁵⁷ (+), ⁻²⁴ TCAATAAATG ⁻¹⁵ (+)
<i>eatA</i>	⁻⁸⁸ GCTACTATT ⁻⁷⁹ (+), ⁻⁷¹ ACAATAAATG ⁻⁶² (+), ⁻⁴⁰ TCCACACAAC ⁻³¹ (-)

Table 2. Cont.

SPATEs	Potential H-NS Binding Sites
<i>aaaA</i>	⁻³¹⁴ ACCATACAGC ⁻³⁰⁵ (-), ⁻¹²⁴ GCGGTAAAAA ⁻¹¹⁵ (-)
<i>tagB</i>	⁻³⁰⁴ ACGAAAAAAA ⁻²⁹⁵ (-), ⁻¹⁶¹ CTGATAAATA ⁻¹⁵² (-), ⁻¹²⁸ TCGATAAATG ⁻¹¹⁹ (+)
<i>tagC</i>	⁻²⁵⁶ GCAATTAATA ⁻²⁴⁷ (+), ⁻⁶² TCGCTATATT ⁻⁵³ (+), ⁻⁵⁶ ACTATAAATA ⁻⁴⁷ (-)
<i>sha</i>	⁻¹⁸⁷ CCCACAAATC ⁻¹⁷⁸ (-), ⁻⁴⁸ TCCTTATATT ⁻³⁹ (+), ⁻³² TCAATAGATA ⁻²³ (-)
<i>vat</i>	⁻²⁹⁶ TCCATATATC ⁻²⁸⁷ (+), ⁻²⁹⁵ TGGATATATG ⁻²⁸⁶ (-), ⁻¹⁰⁷ GCTATATAAT ⁻⁹⁸ (-)

^a Virtual Footprint software [140] was used for in silico analysis of different SPATEs promoter region for putative regulatory H-NS binding sites and additional experimentation is required to confirm a role for H-NS in regulation of SPATE gene regulation. Pattern matching tool Virtual Footprint used specific position weight matrices (PWMs) to generate the top three potential binding sites upstream to the start codon with high scoring matches. (+/- strand).

5.3. Regulation of *Vat* by the MarR-Related Protein *VatX*

The multiple antibiotic resistance regulator (MarR) family are proteins that regulate the expression of many genes involved in resistance to multiple antibiotics including tetracycline, chloramphenicol, β -lactams, nalidixic acid, penicillins, fluoroquinolones, toxic substances, organic solvents, oxidative stress agents and pathogenic factors [115,141–145]. In *E. coli*, MarR is located in the chromosome in the *mar* locus and consists of an operator *marO* and two divergent transcriptional units *marC* and *marRAB* [146]. Besides, the possible role of H-NS as a negative regulator of *Vat*, an open reading frame was located downstream of the *vat* gene designated as ORF26 in the VAT-PAI from Ec222 [46] and c0392 in CFT073 [53]. This ORF shares 44% amino acid identity to the protein PapX (P pilus-associated transcriptional regulatory) from CFT073 [115,147]. PapX has been described as a member of the MarR family, it regulates flagella by binding to the *flhDC* promoter region [115,147]. Considering identity to PapX, the ORF was named *VatX*. The *vatX* gene is present and adjacent to the *vat* gene in many strains. The *VatX* protein contains a MarR protein family (PFAM) domain (PF01047) and a helix-turn-helix motif that is characteristic of DNA binding proteins and was classified in a different clade of MarR family regulators, and is more closely related to the PapX, SfaX, and FocX fimbria-associated regulators. In UPEC strain CFT073, overexpression of *VatX* increased expression of *Vat* 3-fold compared to wild-type levels. Interestingly, in the absence of H-NS, the *vat* and *vatX* genes were both co-transcribed, suggesting *VatX* may compete with H-NS to promote expression of *vat* [115].

5.4. Co-Regulation of SPATEs by CRP and FIS Proteins

The cyclic AMP receptor protein (CRP) is a global transcription factor which is required by *E. coli* in carbon metabolism [148,149]. CRP binds as a homodimer to a 16-bp DNA binding site, is allosterically activated by cAMP binding and interacts with RNA polymerase as either an activator or a repressor of transcription initiation [150]. CRP has been identified as a key transcription factor for the *pet* gene in EAEC strain 042 [151]. In addition, transcription from the *pet* promoter was found to be co-dependent on CRP and Fis regulators and the synergy between these regulators was due to the non-optimal position for transcription initiation of CRP and the additional required binding of the Fis regulator [151]. Fis is a versatile transcription factor and is involved in site-specific recombination events, organization of local DNA topology in bacterial chromosomes, and as a global transcription factor [152,153]. It binds DNA as a homodimer that recognizes a degenerate 15-bp binding sequence often found in many promoter regions [154]. Through a co-activation mechanism, Fis was also shown to co-regulate SPATE genes with CRP including the *sat* gene from UPEC and *sigA* from *Shigella sonnei* [155].

Clearly, further investigation of regulation of SPATE expression by either DNA-binding proteins, transcriptional factors and mechanisms controlling expression of different SPATE proteins is needed. Further, potential cross-talk or hierarchical production of different SPATEs needs to be considered and to what extent other virulence proteins affect the trafficking and secretion of SPATEs, if they are

using the same secretion machinery such as BAM/TAM systems for their export and biogenesis at the bacterial cell surface.

6. Some SPATEs Can Also Mediate Degradation of Bacterial Protein Targets

The perspective of considering SPATEs uniquely as virulence proteins specifically produced to damage host cells and promote infection has been called into question in recent years. A newly described role for some SPATEs has been identified to be the degradation of bacterial proteins and secretion systems. Protease activity of SPATEs has mainly been determined for host proteins, however, importantly, some SPATEs can play an important role in infection through targeting of other bacterial substrates. The EatA SPATE from ETEC degrades the ETEC EtpA adhesin which can lead to reduced intestinal colonization [54]. Reduced colonization may be an advantage for ETEC to quickly deliver enterotoxin without causing extensive damage and may reduce inflammation and decrease the host immune response – acting as hit and run strategy of the bacteria. Similar activity was shown by EspC, wherein EspC reduced secretion of EspA and EspD, which are effector proteins of the EPEC T3SS [83]. In this case, EspC contributed to reducing pore formation and cytotoxicity by degrading EspA and EspD. A similar interchangeable activity was found for EspP [83]. Therefore, these SPATEs can also contribute to EPEC/EHEC infections by degrading other bacterial secreted proteins secreted by the T3SS before they contact the host cell. Another example of interference with other bacterial virulence factors is EspP α present in EHEC, which has been shown to degrade the pore-forming repeats-in-toxin (RTX)-protein hemolysin Ehx (EHEC- Hly) [156] in its free and vesicle-bound form [131]. EspP mediated cleavage was specific to the Ehx hydrophobic domain, which is crucial for interaction with the host cell membrane and pore formation [157]. Similarly, the reduction in the level of active pore-forming toxin may also alter the host immune response and contribute to intestinal colonization and pathogenesis.

7. Conclusions

Improved technologies and affordability of DNA sequencing and other methods to investigate bacterial/host interactions has greatly advanced our understanding of many aspects of bacterial pathogenesis in recent years. The large amount of data from genome sequences has also resulted in the determination of new putative SPATEs, which may also contribute to host colonization and infection. Two decades after the characterization of the first SPATE, much progress has been made in the study of SPATEs by combining *in silico* bioinformatics, bacteriology, molecular biology, biochemistry and cellular biology and pathogenesis studies. Within the SPATE family, all members may share a similar global structure, but each SPATE may mediate distinct functional properties and substrate specificities and be associated with certain bacterial pathotypes, the type of the disease they cause, and host animal species they infect. One would expect that the variety of host and tissues niches could result in diversification of SPATE functions. Further, these pathogens can also differ in their interactions with the host, with some SPATE-producing pathogens remaining extracellular during infection, whereas others may invade cells or survive in host phagocytes. Many questions remain to be answered about the molecular basis for substrate recognition and specificity of SPATEs and also fundamentally understanding how these proteins are regulated. It is also unclear to what extent these SPATEs may modulate or alter proteins and the bacterial surface and effect pathogenesis. Moreover, the SPATE proteins are also frequently present in certain enterobacterial pathogens commonly associated with both enteric and systemic diseases in humans and animal hosts. As these proteins are highly immunogenic, it stands to reason that the SPATEs or their conserved epitopes can be targeted for potential approaches to develop vaccines to prevent such diseases.

Author Contributions: Conceptualization, P.P. and H.H.; formal analysis, P.P. and H.H.; writing—review and editing, P.P., H.H., H.B. and C.M.D.; visualization, P.P. and H.H.; supervision, C.M.D.; funding acquisition, C.M.D.

Funding: Funding for this work was supported by NSERC Canada Discovery Grants 2014-06622 and 2019-06642, and funding from the CRIPA, FRQ-NT center of research excellence.

Acknowledgments: We thank Alma Lilián Guerrero Barrera, professor at Universidad Autónoma de Aguascalientes, Mexico for providing us the primary antibodies against SPATes.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Henderson, I.R.; Navarro-Garcia, F.; Desvaux, M.; Fernandez, R.C.; Ala'Aldeen, D. Type V protein secretion pathway: The autotransporter story. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 692–744. [[CrossRef](#)] [[PubMed](#)]
- Linke, D.; Riess, T.; Autenrieth, I.B.; Lupas, A.; Kempf, V.A. Trimeric autotransporter adhesins: Variable structure, common function. *Trends Microbiol.* **2006**, *14*, 264–270. [[CrossRef](#)] [[PubMed](#)]
- Salacha, R.; Kovačić, F.; Brochier-Armanet, C.; Wilhelm, S.; Tommassen, J.; Filloux, A.; Voulhoux, R.; Bleves, S. The *Pseudomonas aeruginosa* patatin-like protein PlpD is the archetype of a novel Type V secretion system. *Environ. Microbiol.* **2010**, *12*, 1498–1512. [[PubMed](#)]
- Leo, J.C.; Grin, I.; Linke, D. Type V secretion: Mechanism (s) of autotransport through the bacterial outer membrane. *Phil. Trans. R. Soc. B* **2012**, *367*, 1088–1101. [[CrossRef](#)] [[PubMed](#)]
- Jacob-Dubuisson, F.; Locht, C.; Antoine, R. Two-partner secretion in Gram-negative bacteria: A thrifty, specific pathway for large virulence proteins. *Mol. Microbiol.* **2001**, *40*, 306–313. [[CrossRef](#)] [[PubMed](#)]
- Khan, S.; Mian, H.S.; Sandercok, L.E.; Chirgadze, N.Y.; Pai, E.F. Crystal structure of the passenger domain of the *Escherichia coli* autotransporter EspP. *J. Mol. Biol.* **2011**, *413*, 985–1000. [[CrossRef](#)]
- Otto, B.R.; Sijbrandi, R.; Luirink, J.; Oudega, B.; Heddle, J.G.; Mizutani, K.; Park, S.-Y.; Tame, J.R. Crystal structure of hemoglobin protease, a heme binding autotransporter protein from pathogenic *Escherichia coli*. *J. Biol. Chem.* **2005**, *280*, 17339–17345. [[CrossRef](#)]
- Meza-Aguilar, J.D.; Fromme, P.; Torres-Larios, A.; Mendoza-Hernández, G.; Hernandez-Chiñas, U.; Campos, C.A.E.; Fromme, R. X-ray crystal structure of the passenger domain of plasmid encoded toxin (Pet), an autotransporter enterotoxin from enteroaggregative *Escherichia coli* (EAEC). *Biochem. Biophys. Res. Commun.* **2014**, *445*, 439–444. [[CrossRef](#)]
- Veiga, E.; Sugawara, E.; Nikaido, H.; de Lorenzo, V.; Fernández, L.A. Export of autotransported proteins proceeds through an oligomeric ring shaped by C-terminal domains. *EMBO J.* **2002**, *21*, 2122–2131. [[CrossRef](#)]
- Selkrig, J.; Mosbahi, K.; Webb, C.T.; Belousoff, M.J.; Perry, A.J.; Wells, T.J.; Morris, F.; Leyton, D.L.; Totsika, M.; Phan, M.-D. Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat. Struct. Mol. Biol.* **2012**, *19*, 506. [[CrossRef](#)]
- Ruiz-Perez, F.; Henderson, I.R.; Leyton, D.L.; Rossiter, A.E.; Zhang, Y.; Nataro, J.P. Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of Enterobacteriaceae. *J. Bacteriol.* **2009**, *191*, 6571–6583. [[CrossRef](#)]
- Sijbrandi, R.; Urbanus, M.L.; Corinne, M.; Bernstein, H.D.; Oudega, B.; Otto, B.R.; Luirink, J. Signal recognition particle (SRP)-mediated targeting and Sec-dependent translocation of an extracellular *Escherichia coli* protein. *J. Biol. Chem.* **2003**, *278*, 4654–4659. [[CrossRef](#)] [[PubMed](#)]
- Purdy, G.E.; Fisher, C.R.; Payne, S.M. IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp, and SurA. *J. Bacteriol.* **2007**, *189*, 5566–5573. [[CrossRef](#)] [[PubMed](#)]
- Sauri, A.; Soprova, Z.; Wickström, D.; de Gier, J.-W.; Van der Schors, R.C.; Smit, A.B.; Jong, W.S.; Luirink, J. The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology* **2009**, *155*, 3982–3991. [[CrossRef](#)] [[PubMed](#)]
- Peterson, J.H.; Tian, P.; Ieva, R.; Dautin, N.; Bernstein, H.D. Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17739–17744. [[CrossRef](#)]
- Michaelis, S.; Hunt, J.; Beckwith, J. Effects of signal sequence mutations on the kinetics of alkaline phosphatase export to the periplasm in *Escherichia coli*. *J. Bacteriol.* **1986**, *167*, 160–167. [[CrossRef](#)]
- Martoglio, B.; Dobberstein, B. Signal sequences: More than just greasy peptides. *Trends Cell Biol.* **1998**, *8*, 410–415. [[CrossRef](#)]
- Randall, L.; Hardy, S. SecB, one small chaperone in the complex milieu of the cell. *Cell. Mol. Life Sci. Cmls* **2002**, *59*, 1617–1623. [[CrossRef](#)]
- Valent, Q.A.; Scotti, P.A.; High, S.; de Gier, J.W.L.; von Heijne, G.; Lentzen, G.; Wintermeyer, W.; Oudega, B.; Luirink, J. The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. *EMBO J.* **1998**, *17*, 2504–2512. [[CrossRef](#)]

20. Bredemeier, R.; Schlegel, T.; Ertel, F.; Vojta, A.; Borissenko, L.; Bohnsack, M.T.; Groll, M.; Von Haeseler, A.; Schleiff, E. Functional and phylogenetic properties of the pore-forming β -barrel transporters of the Omp85 family. *J. Biol. Chem.* **2007**, *282*, 1882–1890. [[CrossRef](#)]
21. Fleming, P.J.; Patel, D.S.; Wu, E.L.; Qi, Y.; Yeom, M.S.; Sousa, M.C.; Fleming, K.G.; Im, W. BamA POTRA domain interacts with a native lipid membrane surface. *Biophys. J.* **2016**, *110*, 2698–2709. [[CrossRef](#)] [[PubMed](#)]
22. Albrecht, R.; Schütz, M.; Oberhettinger, P.; Faulstich, M.; Bermejo, I.; Rudel, T.; Diederichs, K.; Zeth, K. Structure of BamA, an essential factor in outer membrane protein biogenesis. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**, *70*, 1779–1789. [[CrossRef](#)] [[PubMed](#)]
23. Gruss, F.; Zähringer, F.; Jakob, R.P.; Burmann, B.M.; Hiller, S.; Maier, T. The structural basis of autotransporter translocation by TamA. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1318. [[CrossRef](#)] [[PubMed](#)]
24. Noinaj, N.; Kuszak, A.J.; Balusek, C.; Gumbart, J.C.; Buchanan, S.K. Lateral opening and exit pore formation are required for BamA function. *Structure* **2014**, *22*, 1055–1062. [[CrossRef](#)]
25. Eslava, C.; Navarro-García, F.; Czczulin, J.R.; Henderson, I.R.; Cravioto, A.; Nataro, J.P. Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect. Immun.* **1998**, *66*, 3155–3163.
26. Brunder, W.; Schmidt, H.; Karch, H. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157: H7 cleaves human coagulation factor V. *Mol. Microbiol.* **1997**, *24*, 767–778. [[CrossRef](#)]
27. Dautin, N.; Barnard, T.J.; Anderson, D.E.; Bernstein, H.D. Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism. *EMBO J.* **2007**, *26*, 1942–1952. [[CrossRef](#)]
28. Dautin, N.; Bernstein, H.D. Residues in a conserved α -helical segment are required for cleavage but not secretion of an *Escherichia coli* serine protease autotransporter passenger domain. *J. Bacteriol.* **2011**, *193*, 3748–3756. [[CrossRef](#)]
29. Kostakioti, M.; Stathopoulos, C. Role of the α -helical linker of the C-terminal translocator in the biogenesis of the serine protease subfamily of autotransporters. *Infect. Immun.* **2006**, *74*, 4961–4969. [[CrossRef](#)]
30. Leyton, D.L.; Adams, L.M.; Kelly, M.; Sloan, J.; Tauschek, M.; Robins-Browne, R.M.; Hartland, E.L. Contribution of a novel gene, rpeA, encoding a putative autotransporter adhesin to intestinal colonization by rabbit-specific enteropathogenic *Escherichia coli*. *Infect. Immun.* **2007**, *75*, 4664–4669. [[CrossRef](#)]
31. Habouria, H.; Pokharel, P.; Maris, S.; Garénaux, A.; Bessaiah, H.; Houle, S.; Veyrier, F.J.; Guyomard-Rabenirina, S.; Talarmin, A.; Dozois, C.M. Three new serine-protease autotransporters of Enterobacteriaceae (SPATEs) from extra-intestinal pathogenic *Escherichia coli* and combined role of SPATEs for cytotoxicity and colonization of the mouse kidney. *Virulence* **2019**, *10*, 568–587. [[CrossRef](#)] [[PubMed](#)]
32. Ohnishi, Y.; Beppu, T.; Horinouchi, S. Two genes encoding serine protease homologues in *Serratia marcescens* and characterization of their products in *Escherichia coli*. *J. Biochem.* **1997**, *121*, 902–913. [[CrossRef](#)] [[PubMed](#)]
33. Parham, N.J.; Pollard, S.J.; Desvaux, M.; Scott-Tucker, A.; Liu, C.; Fivian, A.; Henderson, I.R. Distribution of the serine protease autotransporters of the Enterobacteriaceae among extraintestinal clinical isolates of *Escherichia coli*. *J. Clin. Microbiol.* **2005**, *43*, 4076–4082. [[CrossRef](#)] [[PubMed](#)]
34. Vijayakumar, V.; Santiago, A.; Smith, R.; Smith, M.; Robins-Browne, R.M.; Nataro, J.P.; Ruiz-Perez, F. Role of class 1 serine protease autotransporter in the pathogenesis of *Citrobacter rodentium* colitis. *Infect. Immun.* **2014**, *82*, 2626–2636. [[CrossRef](#)]
35. Hu, Y.-H.; Zhou, H.-Z.; Jin, Q.-W.; Zhang, J. The serine protease autotransporter Tsh contributes to the virulence of *Edwardsiella tarda*. *Vet. Microbiol.* **2016**, *189*, 68–74. [[CrossRef](#)]
36. Dautin, N. Serine protease autotransporters of enterobacteriaceae (SPATEs): Biogenesis and function. *Toxins* **2010**, *2*, 1179–1206. [[CrossRef](#)]
37. Dozois, C.M.; Dho-Moulin, M.; Brée, A.; Fairbrother, J.M.; Desautels, C.; Curtiss, R. Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the Tsh genetic region. *Infect. Immun.* **2000**, *68*, 4145–4154. [[CrossRef](#)]
38. Stein, M.; Kenny, B.; Stein, M.A.; Finlay, B.B. Characterization of EspC, a 110-kilodalton protein secreted by enteropathogenic *Escherichia coli* which is homologous to members of the immunoglobulin A protease-like family of secreted proteins. *J. Bacteriol.* **1996**, *178*, 6546–6554. [[CrossRef](#)]
39. Henderson, I.R.; Czczulin, J.; Eslava, C.; Noriega, F.; Nataro, J.P. Characterization of Pic, a Secreted Protease of *Shigella flexneri* and Enteroaggregative *Escherichia coli*. *Infect. Immun.* **1999**, *67*, 5587–5596.
40. Peterson, J.H.; Szabady, R.L.; Bernstein, H.D. An unusual signal peptide extension inhibits the binding of bacterial presecretory proteins to the signal recognition particle, trigger factor, and the SecYEG complex. *J. Biol. Chem.* **2006**, *281*, 9038–9048. [[CrossRef](#)]

41. Dutta, P.R.; Cappello, R.; Navarro-García, F.; Nataro, J.P. Functional comparison of serine protease autotransporters of Enterobacteriaceae. *Infect. Immun.* **2002**, *70*, 7105–7113. [[CrossRef](#)] [[PubMed](#)]
42. Ruiz-Perez, F.; Wahid, R.; Faherty, C.S.; Kolappaswamy, K.; Rodriguez, L.; Santiago, A.; Murphy, E.; Cross, A.; Sztein, M.B.; Nataro, J.P. Serine protease autotransporters from *Shigella flexneri* and pathogenic *Escherichia coli* target a broad range of leukocyte glycoproteins. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12881–12886. [[CrossRef](#)] [[PubMed](#)]
43. Patel, S.K.; Dotson, J.; Allen, K.P.; Fleckenstein, J.M. Identification and molecular characterization of EatA, an autotransporter protein of enterotoxigenic *Escherichia coli*. *Infect. Immun.* **2004**, *72*, 1786–1794. [[CrossRef](#)]
44. Otto, B.R.; Van Dooren, S.J.; Nuijens, J.H.; Luirink, J.; Oudega, B. Characterization of a hemoglobin protease secreted by the pathogenic *Escherichia coli* strain EB1. *J. Exp. Med.* **1998**, *188*, 1091–1103. [[CrossRef](#)] [[PubMed](#)]
45. Guyer, D.M.; Henderson, I.R.; Nataro, J.P.; Mobley, H.L. Identification of sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol. Microbiol.* **2000**, *38*, 53–66. [[CrossRef](#)] [[PubMed](#)]
46. Parreira, V.; Gyles, C. A novel pathogenicity island integrated adjacent to the thrW tRNA gene of avian pathogenic *Escherichia coli* encodes a vacuolating autotransporter toxin. *Infect. Immun.* **2003**, *71*, 5087–5096. [[CrossRef](#)] [[PubMed](#)]
47. Ruiz-Perez, F.; Nataro, J.P. Bacterial serine proteases secreted by the autotransporter pathway: Classification, specificity, and role in virulence. *Cell. Mol. Life Sci.* **2014**, *71*, 745–770. [[CrossRef](#)]
48. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425.
49. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* **1992**, *8*, 275–282. [[CrossRef](#)]
50. Tamura, K.; Stecher, G.; Peterson, D.; Filipinski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)]
51. Ciccarelli, F.D.; Doerks, T.; Von Mering, C.; Creevey, C.J.; Snel, B.; Bork, P. Toward automatic reconstruction of a highly resolved tree of life. *Science* **2006**, *311*, 1283–1287. [[CrossRef](#)] [[PubMed](#)]
52. Heimer, S.R.; Rasko, D.A.; Lockatell, C.V.; Johnson, D.E.; Mobley, H.L. Autotransporter genes pic and tsh are associated with *Escherichia coli* strains that cause acute pyelonephritis and are expressed during urinary tract infection. *Infect. Immun.* **2004**, *72*, 593–597. [[CrossRef](#)] [[PubMed](#)]
53. Welch, R.A.; Burland, V.; Plunkett, G.; Redford, P.; Roesch, P.; Rasko, D.; Buckles, E.; Liou, S.-R.; Boutin, A.; Hackett, J. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 17020–17024. [[CrossRef](#)] [[PubMed](#)]
54. Roy, K.; Kansal, R.; Bartels, S.R.; Hamilton, D.J.; Shaaban, S.; Fleckenstein, J.M. Adhesin degradation accelerates delivery of heat-labile toxin by enterotoxigenic *E. coli*. *J. Biol. Chem.* **2011**, *286*, 29771–29779. [[CrossRef](#)]
55. Navarro-García, F.; Canizalez-Roman, A.; Sui, B.Q.; Nataro, J.P.; Azamar, Y. The serine protease motif of EspC from enteropathogenic *Escherichia coli* produces epithelial damage by a mechanism different from that of Pet toxin from enteroaggregative *E. coli*. *Infect. Immun.* **2004**, *72*, 3609–3621. [[CrossRef](#)]
56. Serapio-Palacios, A.; Navarro-García, F. EspC, an autotransporter protein secreted by enteropathogenic *Escherichia coli*, causes apoptosis and necrosis through caspase and calpain activation, including direct procaspase-3 cleavage. *MBio* **2016**, *7*, e00479-16. [[CrossRef](#)]
57. Elisa Drago-Serrano, M.; Gavilanes Parra, S.; Angel Manjarrez-Hernández, H. EspC, an autotransporter protein secreted by enteropathogenic *Escherichia coli* (EPEC), displays protease activity on human hemoglobin. *FEMS Microbiol. Lett.* **2006**, *265*, 35–40. [[CrossRef](#)]
58. Navarro-García, F.; Serapio-Palacios, A.; Vidal, J.E.; Salazar, M.I.; Tapia-Pastrana, G. EspC promotes epithelial cell detachment by enteropathogenic *Escherichia coli* via sequential cleavages of a cytoskeletal protein and then focal adhesion proteins. *Infect. Immun.* **2014**, *82*, 2255–2265. [[CrossRef](#)]
59. Henderson, I.R.; Hicks, S.; Navarro-García, F.; Elias, W.P.; Philips, A.D.; Nataro, J.P. Involvement of the Enterotoxigenic *Escherichia coli* Plasmid-Encoded Toxin in Causing Human Intestinal Damage. *Infect. Immun.* **1999**, *67*, 5338–5344.
60. Navarro-García, F.; Sears, C.; Eslava, C.; Cravioto, A.; Nataro, J.P. Cytoskeletal effects induced by pet, the serine protease enterotoxin of enteroaggregative *Escherichia coli*. *Infect. Immun.* **1999**, *67*, 2184–2192.
61. Canizalez-Roman, A.; Navarro-García, F. Fodrin CaM-binding domain cleavage by Pet from enteroaggregative *Escherichia coli* leads to actin cytoskeletal disruption. *Mol. Microbiol.* **2003**, *48*, 947–958. [[CrossRef](#)] [[PubMed](#)]

62. Parham, N.J.; Srinivasan, U.; Desvaux, M.; Foxman, B.; Marrs, C.F.; Henderson, I.R. PicU, a second serine protease autotransporter of uropathogenic *Escherichia coli*. *FEMS Microbiol. Lett.* **2004**, *230*, 73–83. [[CrossRef](#)]
63. Ayala-Lujan, J.L.; Vijayakumar, V.; Gong, M.; Smith, R.; Santiago, A.E.; Ruiz-Perez, F. Broad spectrum activity of a lectin-like bacterial serine protease family on human leukocytes. *PLoS ONE* **2014**, *9*, e107920. [[CrossRef](#)] [[PubMed](#)]
64. Brockmeyer, J.; Bielaszewska, M.; Fruth, A.; Bonn, M.L.; Mellmann, A.; Humpf, H.-U.; Karch, H. Subtypes of the plasmid-encoded serine protease EspP in Shiga toxin-producing *Escherichia coli*: Distribution, secretion, and proteolytic activity. *Appl. Environ. Microbiol.* **2007**, *73*, 6351–6359. [[CrossRef](#)] [[PubMed](#)]
65. Schmidt, H.; Zhang, W.-L.; Hemmrich, U.; Jelacic, S.; Brunder, W.; Tarr, P.; Dobrindt, U.; Hacker, J.; Karch, H. Identification and characterization of a novel genomic island integrated at selC in locus of enterocyte effacement-negative, Shiga toxin-producing *Escherichia coli*. *Infect. Immun.* **2001**, *69*, 6863–6873. [[CrossRef](#)] [[PubMed](#)]
66. Orth, D.; Ehrlenbach, S.; Brockmeyer, J.; Khan, A.B.; Huber, G.; Karch, H.; Sarg, B.; Lindner, H.; Würzner, R. EspP, a serine protease of enterohemorrhagic *Escherichia coli*, impairs complement activation by cleaving complement factors C3/C3b and C5. *Infect. Immun.* **2010**, *78*, 4294–4301. [[CrossRef](#)]
67. Otto, B.R.; Van Dooren, S.J.; Dozois, C.M.; Luirink, J.; Oudega, B. *Escherichia coli* hemoglobin protease autotransporter contributes to synergistic abscess formation and heme-dependent growth of *Bacteroides fragilis*. *Infect. Immun.* **2002**, *70*, 5–10. [[CrossRef](#)]
68. Gutiérrez, D.; Pardo, M.; Montero, D.; Oñate, A.; Farfán, M.J.; Ruiz-Pérez, F.; Del Canto, F.; Vidal, R. TleA, a Tsh-like autotransporter identified in a human enterotoxigenic *Escherichia coli* strain. *Infect. Immun.* **2015**, *83*, 1893–1903. [[CrossRef](#)]
69. Guyer, D.M.; Radulovic, S.; Jones, F.-E.; Mobley, H.L. Sat, the secreted autotransporter toxin of uropathogenic *Escherichia coli*, is a vacuolating cytotoxin for bladder and kidney epithelial cells. *Infect. Immun.* **2002**, *70*, 4539–4546. [[CrossRef](#)]
70. Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C. SepA, the major extracellular protein of *Shigella flexneri*: Autonomous secretion and involvement in tissue invasion. *Mol. Microbiol.* **1995**, *17*, 123–135. [[CrossRef](#)]
71. Benjelloun-Touimi, Z.; Tahar, M.S.; Montecucco, C.; Sansonetti, P.J.; Parsot, C. SepA, the 110 kDa protein secreted by *Shigella flexneri*: Two-domain structure and proteolytic activity. *Microbiology* **1998**, *144*, 1815–1822. [[CrossRef](#)] [[PubMed](#)]
72. Al-Hasani, K.; Henderson, I.R.; Sakellaris, H.; Rajakumar, K.; Grant, T.; Nataro, J.P.; Robins-Browne, R.; Adler, B. The sigA gene which is borne on the shepathogenicity island of *Shigella flexneri* 2a encodes an exported cytopathic protease involved in intestinal fluid accumulation. *Infect. Immun.* **2000**, *68*, 2457–2463. [[CrossRef](#)] [[PubMed](#)]
73. Al-Hasani, K.; Navarro-Garcia, F.; Huerta, J.; Sakellaris, H.; Adler, B. The immunogenic SigA enterotoxin of *Shigella flexneri* 2a binds to HEP-2 cells and induces fodrin redistribution in intoxicated epithelial cells. *PLoS ONE* **2009**, *4*, e8223. [[CrossRef](#)] [[PubMed](#)]
74. Yen, Y.T.; Kostakioti, M.; Henderson, I.R.; Stathopoulos, C. Common themes and variations in serine protease autotransporters. *Trends Microbiol.* **2008**, *16*, 370–379. [[CrossRef](#)]
75. Sandt, C.H.; Hill, C.W. Four Different Genes Responsible for Nonimmune Immunoglobulin-Binding Activities within a Single Strain of *Escherichia coli*. *Infect. Immun.* **2000**, *68*, 2205–2214. [[CrossRef](#)]
76. Restieri, C.; Garriss, G.; Locas, M.-C.; Dozois, C.M. Autotransporter-encoding sequences are phylogenetically distributed among *Escherichia coli* clinical isolates and reference strains. *Appl. Environ. Microbiol.* **2007**, *73*, 1553–1562. [[CrossRef](#)]
77. Kumar, P.; Luo, Q.; Vickers, T.J.; Sheikh, A.; Lewis, W.G.; Fleckenstein, J.M. EatA, an immunogenic protective antigen of enterotoxigenic *Escherichia coli*, degrades intestinal mucin. *Infect. Immun.* **2014**, *82*, 500–508. [[CrossRef](#)]
78. Del Canto, F.; Valenzuela, P.; Cantero, L.; Bronstein, J.; Blanco, J.E.; Blanco, J.; Prado, V.; Levine, M.; Nataro, J.; Sommerfelt, H. Distribution of classical and nonclassical virulence genes in enterotoxigenic *Escherichia coli* isolates from Chilean children and tRNA gene screening for putative insertion sites for genomic islands. *J. Clin. Microbiol.* **2011**, *49*, 3198–3203. [[CrossRef](#)]
79. Andrade, F.B.; Abreu, A.G.; Nunes, K.O.; Gomes, T.A.; Piazza, R.M.; Elias, W.P. Distribution of serine protease autotransporters of Enterobacteriaceae in typical and atypical enteroaggregative *Escherichia coli*. *Infect. Genet. Evol.* **2017**, *50*, 83–86. [[CrossRef](#)]

80. Mellies, J.L.; Navarro-Garcia, F.; Okeke, I.; Frederickson, J.; Nataro, J.P.; Kaper, J.B. espC pathogenicity island of enteropathogenic *Escherichia coli* encodes an enterotoxin. *Infect. Immun.* **2001**, *69*, 315–324. [[CrossRef](#)]
81. Vidal, J.E.; Navarro-García, F. Efficient translocation of EspC into epithelial cells depends on enteropathogenic *Escherichia coli* and host cell contact. *Infect. Immun.* **2006**, *74*, 2293–2303. [[CrossRef](#)] [[PubMed](#)]
82. Vidal, J.E.; Navarro-García, F. EspC translocation into epithelial cells by enteropathogenic *Escherichia coli* requires a concerted participation of type V and III secretion systems. *Cell. Microbiol.* **2008**, *10*, 1975–1986. [[CrossRef](#)] [[PubMed](#)]
83. Guignot, J.; Segura, A.; Van Nhieu, G.T. The serine protease EspC from enteropathogenic *Escherichia coli* regulates pore formation and cytotoxicity mediated by the type III secretion system. *PLoS Pathog.* **2015**, *11*, e1005013. [[CrossRef](#)] [[PubMed](#)]
84. Garmendia, J.; Frankel, G.; Crepin, V.F. Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: Translocation, translocation, translocation. *Infect. Immun.* **2005**, *73*, 2573–2585. [[CrossRef](#)]
85. Nemeč, K.N.; Scaglione, P.; Navarro-García, F.; Huerta, J.; Tatulian, S.A.; Teter, K. A host-specific factor is necessary for efficient folding of the autotransporter plasmid-encoded toxin. *Biochimie* **2010**, *92*, 171–177. [[CrossRef](#)]
86. Wagner, J.K.; Heindl, J.E.; Gray, A.N.; Jain, S.; Goldberg, M.B. Contribution of the periplasmic chaperone Skp to efficient presentation of the autotransporter IcsA on the surface of *Shigella flexneri*. *J. Bacteriol.* **2009**, *191*, 815–821. [[CrossRef](#)]
87. Navarro-García, F.; Canizalez-Roman, A.; Luna, J.; Sears, C.; Nataro, J.P. Plasmid-Encoded Toxin of Enteroaggregative *Escherichia coli* is Internalized by Epithelial Cells. *Infect. Immun.* **2001**, *69*, 1053–1060. [[CrossRef](#)]
88. Navarro-García, F.; Canizalez-Roman, A.; Burlingame, K.E.; Teter, K.; Vidal, J.E. Pet, a non-AB toxin, is transported and translocated into epithelial cells by a retrograde trafficking pathway. *Infect. Immun.* **2007**, *75*, 2101–2109. [[CrossRef](#)]
89. Dutta, P.R.; Sui, B.Q.; Nataro, J.P. Structure-function analysis of the enteroaggregative *Escherichia coli* plasmid-encoded toxin autotransporter using scanning linker mutagenesis. *J. Biol. Chem.* **2003**, *278*, 39912–39920. [[CrossRef](#)]
90. Nava-Acosta, R.; Navarro-García, F. Cytokeratin 8 is an epithelial cell receptor for Pet, a cytotoxic serine protease autotransporter of Enterobacteriaceae. *MBio* **2013**, *4*, e00838-13. [[CrossRef](#)]
91. Chavez-Dueñas, L.; Serapio-Palacios, A.; Nava-Acosta, R.; Navarro-García, F. The subdomain 2 of the autotransporter Pet is the ligand site for recognizing Pet receptor on the epithelial cell surface. *Infect. Immun.* **2016**, *84*, 2012–2021. [[CrossRef](#)] [[PubMed](#)]
92. Villaseca, J.M.; Navarro-García, F.; Mendoza-Hernández, G.; Nataro, J.P.; Cravioto, A.; Eslava, C. Pet toxin from enteroaggregative *Escherichia coli* produces cellular damage associated with fodrin disruption. *Infect. Immun.* **2000**, *68*, 5920–5927. [[CrossRef](#)] [[PubMed](#)]
93. Martin, S.J.; O'Brien, G.A.; Nishioka, W.K.; McGahon, A.J.; Mahboubi, A.; Saido, T.C.; Green, D.R. Proteolysis of fodrin (non-erythroid spectrin) during apoptosis. *J. Biol. Chem.* **1995**, *270*, 6425–6428. [[CrossRef](#)] [[PubMed](#)]
94. Wang, K.K.; Posmantur, R.; Nath, R.; McGinnis, K.; Whitton, M.; Talanian, R.V.; Glantz, S.B.; Morrow, J.S. Simultaneous degradation of α II- and β II-spectrin by caspase 3 (CPP32) in apoptotic cells. *J. Biol. Chem.* **1998**, *273*, 22490–22497. [[CrossRef](#)]
95. Sanchez-Villamil, J.I.; Navarro-García, F.; Castillo-Romero, A.; Gutierrez-Gutierrez, F.; Tapia-Pastrana, G. Curcumin blocks cytotoxicity of enteroaggregative and enteropathogenic *Escherichia coli* by blocking Pet and EspC proteolytic release from bacterial outer membrane. *Front. Cell. Infect. Microbiol.* **2019**, *9*, 334. [[CrossRef](#)]
96. Abreu, A.G.; Abe, C.M.; Nunes, K.O.; Moraes, C.T.; Chavez-Dueñas, L.; Navarro-García, F.; Barbosa, A.S.; Piazza, R.M.; Elias, W.P. The serine protease Pic as a virulence factor of atypical enteropathogenic *Escherichia coli*. *Gut Microbes* **2016**, *7*, 115–125. [[CrossRef](#)]
97. Harrington, S.M.; Sheikh, J.; Henderson, I.R.; Ruiz-Perez, F.; Cohen, P.S.; Nataro, J.P. The Pic protease of enteroaggregative *Escherichia coli* promotes intestinal colonization and growth in the presence of mucin. *Infect. Immun.* **2009**, *77*, 2465–2473. [[CrossRef](#)]
98. Navarro-García, F.; Gutierrez-Jimenez, J.; Garcia-Tovar, C.; Castro, L.A.; Salazar-Gonzalez, H.; Cordova, V. Pic, an autotransporter protein secreted by different pathogens in the Enterobacteriaceae family, is a potent mucus secretagogue. *Infect. Immun.* **2010**, *78*, 4101–4109. [[CrossRef](#)]

99. Abreu, A.G.; Fraga, T.R.; Granados Martínez, A.P.; Kondo, M.Y.; Juliano, M.A.; Juliano, L.; Navarro-García, F.; Isaac, L.; Barbosa, A.S.; Elias, W.P. The serine protease Pic from enteroaggregative *Escherichia coli* mediates immune evasion by the direct cleavage of complement proteins. *J. Infect. Dis.* **2015**, *212*, 106–115. [[CrossRef](#)]
100. Bhullar, K.; Zarepour, M.; Yu, H.; Yang, H.; Croxen, M.; Stahl, M.; Finlay, B.B.; Turvey, S.E.; Vallance, B.A. The serine protease autotransporter pic modulates citrobacter rodentium pathogenesis and its innate recognition by the host. *Infect. Immun.* **2015**, *83*, 2636–2650. [[CrossRef](#)]
101. Brockmeyer, J.; Spelten, S.; Kuczus, T.; Bielaszewska, M.; Karch, H. Structure and function relationship of the autotransport and proteolytic activity of EspP from Shiga toxin-producing *Escherichia coli*. *PLoS ONE* **2009**, *4*, e6100. [[CrossRef](#)] [[PubMed](#)]
102. Djafari, S.; Ebel, F.; Deibel, C.; Krämer, S.; Hudel, M.; Chakraborty, T. Characterization of an exported protease from Shiga toxin-producing *Escherichia coli*. *Mol. Microbiol.* **1997**, *25*, 771–784. [[CrossRef](#)] [[PubMed](#)]
103. Boisen, N.; Ruiz-Perez, F.; Scheutz, F.; Krogfelt, K.A.; Nataro, J.P. High prevalence of serine protease autotransporter cytotoxins among strains of enteroaggregative *Escherichia coli*. *Am. J. Trop. Med. Hyg.* **2009**, *80*, 294–301. [[CrossRef](#)] [[PubMed](#)]
104. Velarde, J.J.; Nataro, J.P. Hydrophobic residues of the autotransporter EspP linker domain are important for outer membrane translocation of its passenger. *J. Biol. Chem.* **2004**, *279*, 31495–31504. [[CrossRef](#)]
105. Ruiz-Perez, F.; Henderson, I.R.; Nataro, J.P. Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut Microbes* **2010**, *1*, 339–344. [[CrossRef](#)]
106. Xicohtencatl-Cortes, J.; Saldaña, Z.; Deng, W.; Castañeda, E.; Freer, E.; Tarr, P.I.; Finlay, B.B.; Puente, J.L.; Girón, J.A. Bacterial macroscopic rope-like fibers with cytopathic and adhesive properties. *J. Biol. Chem.* **2010**, *285*, 32336–32342. [[CrossRef](#)]
107. Khan, A.B.; Naim, A.; Orth, D.; Grif, K.; Mohsin, M.; Prager, R.; Dierich, M.P.; Würzner, R. Serine protease espP subtype α , but not β or γ , of Shiga toxin-producing *Escherichia coli* is associated with highly pathogenic serogroups. *Int. J. Med Microbiol.* **2009**, *299*, 247–254. [[CrossRef](#)]
108. Weiss, A.; Joerss, H.; Brockmeyer, J. Structural and functional characterization of cleavage and inactivation of human serine protease inhibitors by the bacterial SPATE protease EspP α from enterohemorrhagic *E. coli*. *PLoS ONE* **2014**, *9*, e111363. [[CrossRef](#)]
109. Tse, C.; In, J.; Yin, J.; Donowitz, M.; Doucet, M.; Foulke-Abel, J.; Ruiz-Perez, F.; Nataro, J.; Zachos, N.; Kaper, J. Enterohemorrhagic *E. coli* (EHEC)—Secreted Serine Protease EspP Stimulates Electrogenic Ion Transport in Human Colonoid Monolayers. *Toxins* **2018**, *10*, 351. [[CrossRef](#)]
110. Provence, D.L.; Curtiss, R. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic *Escherichia coli* strain. *Infect. Immun.* **1994**, *62*, 1369–1380.
111. Nicholson, B.A.; West, A.C.; Mangiamale, P.; Barbieri, N.; Wannemuehler, Y.; Nolan, L.K.; Logue, C.M.; Li, G. Genetic characterization of ExPEC-like virulence plasmids among a subset of NMEC. *PLoS ONE* **2016**, *11*, e0147757. [[CrossRef](#)] [[PubMed](#)]
112. Kostakioti, M.; Stathopoulos, C. Functional analysis of the Tsh autotransporter from an avian pathogenic *Escherichia coli* strain. *Infect. Immun.* **2004**, *72*, 5548–5554. [[CrossRef](#)] [[PubMed](#)]
113. Stathopoulos, C.; Provence, D.L.; Curtiss, R. Characterization of the Avian Pathogenic *Escherichia coli* Hemagglutinin Tsh, a Member of the Immunoglobulin A Protease-Type Family of Autotransporters. *Infect. Immun.* **1999**, *67*, 772–781. [[PubMed](#)]
114. Maluta, R.P.; Gatti, M.S.V.; Joazeiro, P.P.; De Paiva, J.B.; Rojas, T.C.G.; Silveira, F.; Houle, S.; Kobayashi, R.K.T.; Dozois, C.M.; Dias da Silveira, W. Avian extraintestinal *Escherichia coli* exhibits enterotoxigenic-like activity in the in vivo rabbit ligated ileal loop assay. *Foodborne Pathog. Dis.* **2014**, *11*, 484–489. [[CrossRef](#)]
115. Nichols, K.B.; Totsika, M.; Moriel, D.G.; Lo, A.W.; Yang, J.; Wurpel, D.J.; Rossiter, A.E.; Strugnell, R.A.; Henderson, I.R.; Ulett, G.C. Molecular characterization of the vacuolating autotransporter toxin in uropathogenic *Escherichia coli*. *J. Bacteriol.* **2016**, *198*, 1487–1498. [[CrossRef](#)]
116. Subashchandrabose, S.; Smith, S.N.; Spurbeck, R.R.; Kole, M.M.; Mobley, H.L. Genome-wide detection of fitness genes in uropathogenic *Escherichia coli* during systemic infection. *PLoS Pathog.* **2013**, *9*, e1003788. [[CrossRef](#)]
117. Gibold, L.; Garenaux, E.; Dalmaso, G.; Gallucci, C.; Cia, D.; Mottet-Auselo, B.; Fais, T.; Darfeuille-Michaud, A.; Nguyen, H.T.T.; Barnich, N. The Vat-AIEC protease promotes crossing of the intestinal mucus layer by Crohn's disease-associated *Escherichia coli*. *Cell. Microbiol.* **2016**, *18*, 617–631. [[CrossRef](#)]

118. Tapader, R.; Chatterjee, S.; Singh, A.; Dayma, P.; Haldar, S.; Pal, A.; Basu, S. The high prevalence of serine protease autotransporters of Enterobacteriaceae (SPATEs) in *Escherichia coli* causing neonatal septicemia. *Eur. J. Clin. Microbiol. Infect. Dis.* **2014**, *33*, 2015–2024. [[CrossRef](#)]
119. Liévin-Le Moal, V.; Comenge, Y.; Ruby, V.; Amsellem, R.; Nicolas, V.; Servin, A.L. Secreted autotransporter toxin (Sat) triggers autophagy in epithelial cells that relies on cell detachment. *Cell. Microbiol.* **2011**, *13*, 992–1013. [[CrossRef](#)]
120. Maroncle, N.M.; Sivick, K.E.; Brady, R.; Stokes, F.-E.; Mobley, H.L. Protease activity, secretion, cell entry, cytotoxicity, and cellular targets of secreted autotransporter toxin of uropathogenic *Escherichia coli*. *Infect. Immun.* **2006**, *74*, 6124–6134. [[CrossRef](#)]
121. Guignot, J.; Chaplais, C.; Coconnier-Polter, M.H.; Servin, A.L. The secreted autotransporter toxin, Sat, functions as a virulence factor in Afa/Dr diffusely adhering *Escherichia coli* by promoting lesions in tight junction of polarized epithelial cells. *Cell. Microbiol.* **2007**, *9*, 204–221. [[CrossRef](#)] [[PubMed](#)]
122. Taddei, C.R.; Fasano, A.; Ferreira, A.J.; Trabulsi, L.R.; Martinez, M.B. Secreted autotransporter toxin produced by a diffusely adhering *Escherichia coli* strain causes intestinal damage in animal model assays. *FEMS Microbiol. Lett.* **2005**, *250*, 263–269. [[CrossRef](#)] [[PubMed](#)]
123. Toloza, L.; Giménez, R.; Fábrega, M.J.; Alvarez, C.S.; Aguilera, L.; Cañas, M.A.; Martín-Venegas, R.; Badia, J.; Baldomà, L. The secreted autotransporter toxin (Sat) does not act as a virulence factor in the probiotic *Escherichia coli* strain Nissle 1917. *BMC Microbiol.* **2015**, *15*, 250. [[CrossRef](#)] [[PubMed](#)]
124. Coron, E.; Flamant, M.; Aubert, P.; Wedel, T.; Pedron, T.; Letessier, E.; Galmiche, J.P.; Sansonetti, P.J.; Neunlist, M. Characterisation of early mucosal and neuronal lesions following *Shigella flexneri* infection in human colon. *PLoS ONE* **2009**, *4*, e4713. [[CrossRef](#)] [[PubMed](#)]
125. Maldonado-Contreras, A.; Birtley, J.R.; Boll, E.; Zhao, Y.; Mumy, K.L.; Toscano, J.; Ayehunie, S.; Reinecker, H.-C.; Stern, L.J.; McCormick, B.A. *Shigella* depends on SepA to destabilize the intestinal epithelial integrity via cofilin activation. *Gut Microbes* **2017**, *8*, 544–560. [[CrossRef](#)]
126. Huang, T.Y.; DerMardirossian, C.; Bokoch, G.M. Cofilin phosphatases and regulation of actin dynamics. *Curr. Opin. Cell Biol.* **2006**, *18*, 26–31. [[CrossRef](#)]
127. Scorza, F.B.; Colucci, A.M.; Maggiore, L.; Sanzone, S.; Rossi, O.; Ferlenghi, I.; Pesce, I.; Caboni, M.; Norais, N.; Di Cioccio, V. High yield production process for *Shigella* outer membrane particles. *PLoS ONE* **2012**, *7*, e35616.
128. Oany, A.R.; Pervin, T.; Mia, M.; Hossain, M.; Shahnaij, M.; Mahmud, S.; Kibria, K. Vaccinomics approach for designing potential peptide vaccine by targeting *Shigella* spp. serine protease autotransporter subfamily protein SigA. *J. Immunol. Res.* **2017**, *2017*, 6412353. [[CrossRef](#)]
129. Fookes, M.; Schroeder, G.N.; Langridge, G.C.; Blondel, C.J.; Mammia, C.; Connor, T.R.; Seth-Smith, H.; Vernikos, G.S.; Robinson, K.S.; Sanders, M. *Salmonella bongori* provides insights into the evolution of the *Salmonellae*. *PLoS Pathog.* **2011**, *7*, e1002191. [[CrossRef](#)]
130. Ebel, F.; Deibel, C.; Kresse, A.U.; Guzman, C.A.; Chakraborty, T. Temperature- and medium-dependent secretion of proteins by Shiga toxin-producing *Escherichia coli*. *Infect. Immun.* **1996**, *64*, 4472–4479.
131. Brockmeyer, J.; Aldick, T.; Soltwisch, J.; Zhang, W.; Tarr, P.I.; Weiss, A.; Dreisewerd, K.; Müthing, J.; Bielaszewska, M.; Karch, H. Enterohaemorrhagic *Escherichia coli* haemolysin is cleaved and inactivated by serine protease EspPα. *Environ. Microbiol.* **2011**, *13*, 1327–1341. [[CrossRef](#)] [[PubMed](#)]
132. Friedberg, D.; Umanski, T.; Fang, Y.; Rosenshine, I. Hierarchy in the expression of the locus of enterocyte effacement genes of enteropathogenic *Escherichia coli*. *Mol. Microbiol.* **1999**, *34*, 941–952. [[CrossRef](#)] [[PubMed](#)]
133. Elliott, S.J.; Sperandio, V.; Girón, J.A.; Shin, S.; Mellies, J.L.; Wainwright, L.; Hutcheson, S.W.; McDaniel, T.K.; Kaper, J.B. The Locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect. Immun.* **2000**, *68*, 6115–6126. [[CrossRef](#)] [[PubMed](#)]
134. Drlica, K.; Rouviere-Yaniv, J. Histone-like proteins of bacteria. *Microbiol. Rev.* **1987**, *51*, 301.
135. Workman, J.; Kingston, R. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* **1998**, *67*, 545–579. [[CrossRef](#)]
136. Anuchin, A.; Goncharenko, A.; Demidenok, O.; Kaprelyants, A. Histone-like proteins of bacteria (review). *Appl. Biochem. Microbiol.* **2011**, *47*, 580–585. [[CrossRef](#)]

137. Allsopp, L.P.; Beloin, C.; Ulett, G.C.; Valle, J.; Totsika, M.; Sherlock, O.; Ghigo, J.-M.; Schembri, M.A. Molecular characterization of UpaB and UpaC, two new autotransporter proteins of uropathogenic *Escherichia coli* CFT073. *Infect. Immun.* **2012**, *80*, 321–332. [[CrossRef](#)]
138. Allsopp, L.P.; Beloin, C.; Moriel, D.G.; Totsika, M.; Ghigo, J.-M.; Schembri, M.A. Functional heterogeneity of the UpaH autotransporter protein from uropathogenic *Escherichia coli*. *J. Bacteriol.* **2012**, *194*, 5769–5782. [[CrossRef](#)]
139. Totsika, M.; Wells, T.J.; Beloin, C.; Valle, J.; Allsopp, L.P.; King, N.P.; Ghigo, J.-M.; Schembri, M.A. Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic *Escherichia coli*. *Appl. Environ. Microbiol.* **2012**, *78*, 2179–2189. [[CrossRef](#)]
140. Münch, R.; Hiller, K.; Grote, A.; Scheer, M.; Klein, J.; Schobert, M.; Jahn, D. Virtual Footprint and PRODORIC: An integrative framework for regulon prediction in prokaryotes. *Bioinformatics* **2005**, *21*, 4187–4189. [[CrossRef](#)]
141. George, A.; Levy, S. Gene in the major cotransduction gap of the *Escherichia coli* K-12 linkage map required for the expression of chromosomal resistance to tetracycline and other antibiotics. *J. Bacteriol.* **1983**, *155*, 541–548. [[PubMed](#)]
142. Cohen, S.P.; McMurry, L.; Hooper, D.; Wolfson, J.; Levy, S. Cross-resistance to fluoroquinolones in multiple-antibiotic-resistant (Mar) *Escherichia coli* selected by tetracycline or chloramphenicol: Decreased drug accumulation associated with membrane changes in addition to OmpF reduction. *Antimicrob. Agents Chemother.* **1989**, *33*, 1318–1325. [[CrossRef](#)] [[PubMed](#)]
143. George, A.M.; Levy, S.B. Amplifiable resistance to tetracycline, chloramphenicol, and other antibiotics in *Escherichia coli*: Involvement of a non-plasmid-determined efflux of tetracycline. *J. Bacteriol.* **1983**, *155*, 531–540. [[PubMed](#)]
144. Asako, H.; Nakajima, H.; Kobayashi, K.; Kobayashi, M.; Aono, R. Organic solvent tolerance and antibiotic resistance increased by overexpression of marA in *Escherichia coli*. *Appl. Environ. Microbiol.* **1997**, *63*, 1428–1433. [[PubMed](#)]
145. Ariza, R.; Cohen, S.; Bachhawat, N.; Levy, S.; Demple, B. Repressor mutations in the marRAB operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. *J. Bacteriol.* **1994**, *176*, 143–148. [[CrossRef](#)]
146. Cohen, S.P.; Hächler, H.; Levy, S. Genetic and functional analysis of the multiple antibiotic resistance (mar) locus in *Escherichia coli*. *J. Bacteriol.* **1993**, *175*, 1484–1492. [[CrossRef](#)]
147. Simms, A.N.; Mobley, H.L. PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic *Escherichia coli*. *Infect. Immun.* **2008**, *76*, 4833–4841. [[CrossRef](#)]
148. Gosset, G.; Zhang, Z.; Nayyar, S.; Cuevas, W.A.; Saier, M.H. Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *J. Bacteriol.* **2004**, *186*, 3516–3524. [[CrossRef](#)]
149. Green, J.; Stapleton, M.R.; Smith, L.J.; Artymiuk, P.J.; Kahramanoglou, C.; Hunt, D.M.; Buxton, R.S. Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: Adaptation for different ecological niches. *Curr. Opin. Microbiol.* **2014**, *18*, 1–7. [[CrossRef](#)]
150. Won, H.-S.; Lee, Y.-S.; Lee, S.-H.; Lee, B.-J. Structural overview on the allosteric activation of cyclic AMP receptor protein. *Biochim. Biophys. Acta (BBA) Proteins Proteom.* **2009**, *1794*, 1299–1308. [[CrossRef](#)]
151. Rossiter, A.E.; Browning, D.F.; Leyton, D.L.; Johnson, M.D.; Godfrey, R.E.; Wardius, C.A.; Desvaux, M.; Cunningham, A.F.; Ruiz-Perez, F.; Nataro, J.P. Transcription of the plasmid-encoded toxin gene from Enteroaggregative *Escherichia coli* is regulated by a novel co-activation mechanism involving CRP and Fis. *Mol. Microbiol.* **2011**, *81*, 179–191. [[CrossRef](#)] [[PubMed](#)]
152. Koch, C.; Kahmann, R. Purification and properties of the *Escherichia coli* host factor required for inversion of the G segment in bacteriophage Mu. *J. Biol. Chem.* **1986**, *261*, 15673–15678. [[PubMed](#)]
153. Travers, A.; Schneider, R.; Muskhelishvili, G. DNA supercoiling and transcription in *Escherichia coli*: The FIS connection. *Biochimie* **2001**, *83*, 213–217. [[CrossRef](#)]
154. Kahramanoglou, C.; Seshasayee, A.S.; Prieto, A.I.; Ibberson, D.; Schmidt, S.; Zimmermann, J.; Benes, V.; Fraser, G.M.; Luscombe, N.M. Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res.* **2010**, *39*, 2073–2091. [[CrossRef](#)] [[PubMed](#)]
155. Rossiter, A.E.; Godfrey, R.E.; Connolly, J.A.; Busby, S.J.; Henderson, I.R.; Browning, D.F. Expression of different bacterial cytotoxins is controlled by two global transcription factors, CRP and Fis, that co-operate in a shared-recruitment mechanism. *Biochem. J.* **2015**, *466*, 323–335. [[CrossRef](#)] [[PubMed](#)]

156. Schmidt, H.; Beutin, L.; Karch, H. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157: H7 strain EDL 933. *Infect. Immun.* **1995**, *63*, 1055–1061. [[PubMed](#)]
157. Welch, R. RTX toxin structure and function: A story of numerous anomalies and few analogies in toxin biology. In *Pore-Forming Toxins*; Springer: Berlin/Heidelberg, Germany, 2001.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

HYPOTHÈSE ET OBJECTIFS DE RECHERCHES

4 HYPOTHÈSE ET OBJECTIFS DE RECHERCHES

La colibacillose provoquée par les APEC, est une pathologie aviaire très fréquente qui peut entraîner une certaine mortalité et une baisse de la production industrielle. Le pouvoir pathogène de ces souches repose sur leurs propriétés à coloniser l'appareil respiratoire, leur résistance au système immunitaire et leur capacité à produire des effets cytotoxiques. Plusieurs facteurs de virulence potentiels sont associés aux souches APEC comme les adhésines fimbriales, les ATs, les systèmes d'acquisition du fer, l'antigène capsulaire K, la résistance au pouvoir bactéricide du sérum et les toxines.

Nous avons séquencé une souche APEC appelée QT598, isolée d'un cas de colibacillose chez une dinde en France. La particularité de cette souche repose sur la présence d'un plasmide de virulence de type CoIV d'environ 200 kb qui possède tous les gènes typiques de virulence, en plus d'une région unique qui n'est présente que chez certaines souches ExPEC référencées dans la base de donnée Genbank. Le but de ce projet est donc d'étudier la virulence potentielle des nouveaux gènes qui se trouvent dans la région unique, entre autres l'AT Sha et le fimbriae PL *in vitro* et *in vivo*.

Hypothèse

Généralement, la présence des ATs et/ou des fimbriae confère aux bactéries la capacité à coloniser certains organes de l'hôte donc ils peuvent être associés à la virulence des souches ExPEC. Nous pouvons, par conséquent, émettre l'hypothèse suivante : les nouveaux gènes d'AT et du fimbriae sont impliqués dans la virulence de la souche QT598.

Objectifs

L'objectif principal de notre projet est d'étudier l'implication des nouveaux gènes de l'AT Sha et du fimbriae PL dans la virulence des souches ExPEC, et ceci est concrétisé par deux sous objectifs qui sont les suivants :

- Déterminer leurs rôles *in vitro* en utilisant des tests phénotypiques :
 - Effet adhérent et protéolytique / cytopathique de l'AT Sha
 - Effet adhérent des fimbriae PL
- Etudier leurs rôles *in vivo* en utilisant des modèles animaux et déterminer leurs niveaux d'expression dans les vessies des souris après 48h d'infection par RT-qPCR. Pour les essais *in vivo*, deux modèles ont été utilisés : (1) le modèle de mono-infection qui va servir

à déterminer la capacité des souches à coloniser les organes et à déterminer l'expression des gènes *in vivo* ainsi que (2) le modèle de co-infection qui va servir à déterminer si l'une des 2 souches (sauvage ou mutante) aura un avantage compétitif sur l'autre.

PREMIER ARTICLE

5 PREMIER ARTICLE

Three new serine-protease autotransporters (SPATEs) from extra-intestinal pathogenic *Escherichia coli* and combined role of SPATEs for cytotoxicity and colonization of the mouse kidney

Auteurs : Hajer Habouria^{1,2#}, Pravil Pokharel^{1,2#}, Segolène Maris^{1,2}, Amélie Garénaux^{1,2}, Hicham Bessaiah^{1,2}, Sébastien Houle^{1,2}, Frédéric J. Veyrier^{1,4}, Stéphanie Guyomard-Rabenirina^{3,4}, Antoine Talarmin^{3,4}, Charles M. Dozois^{1,2,4}

1INRS-Institut Armand-Frappier, Laval, Québec, Canada

2CRIPA-Centre de recherche en infectiologie porcine et avicole, Saint-Hyacinthe, Québec, Canada

3Institut Pasteur de Guadeloupe, Les Abymes, Guadeloupe, France.

4Institut Pasteur International Network

#These two authors contributed equally as primary authors of this research

Objectif de l'article :

Dans ce premier article, nous avons caractérisé de nouveaux gènes d'ATs de la famille des SPATEs. Ce travail de recherche publié représente la première partie de mon projet de doctorat qui consistait en la caractérisation de l'AT Sha. Nous avons démontré que l'AT Sha est appartient à la famille des SPATEs possédant un site de clivage atypique. L'AT Sha appartient à la classe II des immunomodulateurs et possède une homologie de séquence avec Tsh, TleA et Vat. Nous avons également mis en évidence que AT Sha est capable d'agglutiner le sang de différentes espèces, de se lier aux cellules épithéliales rénales et de la vessie, de former des agrégations et enfin de former un biofilm. De plus, l'AT Sha possède un effet protéolytique de type estérase et un effet cytotoxique sur les cellules épithéliales de la vessie après 12h d'incubation. Nous avons montré également que l'absence du gène *sha* n'a pas affecté le niveau de colonisation de la souche QT598 dans les vessies et les reins des souris dans le modèle de mono-infection. Par ailleurs, la délétion des 5 SPATEs a considérablement réduit la colonisation bactérienne au niveau des reins, ce qui permet de suggérer un rôle cumulatif des SPATE chez la souche QT598 dans le modèle ITU de souris.

Contribution des auteurs :

En tant que co-premier auteur, j'ai effectué l'ensemble des manipulations relatives aux protéines Sha, Tsh et Vat. J'ai procédé au clonage des gènes d'intérêt, purifié les protéines et effectué tous les tests phénotypiques. J'ai aussi analysé l'ensemble des résultats, préparé les figures et écrit le manuscrit. J'ai également coordonné et assemblé les corrections de mon directeur de recherche et de mes collaborateurs.

Pravil Pokharel : a effectué les manipulations relatives aux protéines TagB, TagC et EspC. Il a procédé au clonage des gènes d'intérêt, purifié les protéines et effectué la majorité des tests phénotypiques. Il a aussi analysé l'ensemble des résultats, préparé les figures et écrit le manuscrit. Il a également coordonné et assemblé les corrections de notre directeur de recherche et de nos collaborateurs.

Segolène Maris : a contribué au criblage des 697 souches UPEC.

Amélie Garénaux : a contribué à l'analyse des données de séquençage PacBio de QT598 et à la révision du manuscrit.

Hicham Bessaiah et *Sébastien Houle* : ont participé aux infections des souris, à l'analyse des résultats et la révision du manuscrit.

Frédéric J. Veyrier : a contribué au séquençage de la souche QT598 et à la révision du manuscrit.

Stéphanie Guyomard-Rabenirina et *Antoine Talarmin* : ont contribué à l'isolement des souches cliniques et à la révision du manuscrit.

Professeur Dozois Charles : a contribué à la planification et à la conception de l'étude, à l'analyse des résultats, à la révision, à la correction et enfin à la soumission du manuscrit au journal.

Three new serine-protease autotransporters (SPATEs) from extra-intestinal pathogenic *Escherichia coli* and combined role of SPATEs for cytotoxicity and colonization of the mouse kidney

Auteurs : Hajer Habouria^{1,2#}, Pravil Pokharel^{1,2#}, Segolène Maris^{1,2}, Amélie Garénaux^{1,2}, Hicham Bessaiah^{1,2}, Sébastien Houle^{1,2}, Frédéric J. Veyrier^{1,4}, Stéphanie Guyomard-Rabenirina^{3,4}, Antoine Talarmin^{3,4}, Charles M. Dozois^{1,2,4}

1INRS-Institut Armand-Frappier, Laval, Québec, Canada

2CRIPA-Centre de recherche en infectiologie porcine et avicole, Saint-Hyacinthe, Québec, Canada

3Institut Pasteur de Guadeloupe, Les Abymes, Guadeloupe, France.

4Institut Pasteur International Network

#These two authors contributed equally as primary authors of this research

Nom du journal : Virulence.

Volume 10, 2019 - Issue 1 Pages 568-587 | Reçu 07 février 2019, Accepté 17 mai 2019, Publié en ligne: 14 juin 2019.

DOI <https://doi.org/10.1080/21505594.2019.1624102>

Résumé en français

Les autotransporteurs sérines protéases des *Enterobacteriaceae* (SPATEs, Serine Protease Autotransporters of *Enterobacteriaceae*) sont des protéines sécrétées qui peuvent jouer un rôle dans la virulence en tant que protéases, toxines, adhésines et/ou immunomodulateurs. La souche *E. coli* pathogène extra-intestinale (ExPEC) O1:K1, QT598, isolée d'une dinde, possède cinq gènes d'AT de la famille des SPATEs : *vat*, *tsh* et trois nouveaux gènes non caractérisés qui sont *sha* (Serine-protease hemagglutinin autotransporter), *tagB* et *tagC* (tandem autotransporter genes B et C). Ces nouveaux ATs ont été clonés dans *E. coli* K-12 afin de déterminer leurs rôles dans l'hémagglutination, l'auto-agrégation et la cytotoxicité. Les clones Sha et TagB ont montré une activité d'auto-agrégation et d'hémagglutination, alors que TagB, TagC et Sha ont tous montré une activité cytopathique sur les cellules épithéliales de la vessie. Les gènes *tagB* et *tagC* qui sont codés en tandem sur un îlot de pathogénicité, sont présents chez environ 10 % des souches UPEC et 4,7 % des souches APEC. Le gène *sha* qui est porté par le plasmide de virulence, est présent dans environ 1 % des souches UPEC et 20 % des souches APEC. Afin de déterminer le rôle spécifique des SPATEs chez la souche QT598, nous avons mutés ces 5 gènes puis testés leurs activités. Bien que la présence de l'un de ces ATs ait montré une activité cytopathique sur les cellules de la vessie, la perte des cinq gènes a diminué significativement cet effet. Dans les infections urinaires des souris, le gène *sha* a été régulé positivement et de façon significatif (six fois plus) au niveau de la vessie par rapport à sa croissance dans le milieu de culture *in vitro*. La délétion de *tagBC* ou de *sha* n'a pas réduit la colonisation des bactéries au niveau des voies urinaires. Cependant, la délétion des 5 SPATEs a considérablement réduit la colonisation bactérienne au niveau des reins, ce qui suggère un rôle cumulatif des SPATE chez la souche QT598 dans le modèle ITU de souris.

Mots-clés : Autotransporteurs ; SPATE ; Toxines ; pathogène aviaire ; infection de souris ; la volaille ; autotransporteur protéase de sérine ; uropathogène.

Three new serine-protease autotransporters of *Enterobacteriaceae* (SPATEs) from extra-intestinal pathogenic *Escherichia coli* and combined role of SPATEs for cytotoxicity and colonization of the mouse kidney

Hajer Habouria^{*a,b}, Pravil Pokharel^{*a,b}, Segolène Maris^{a,b}, Amélie Garénaux^{a,b}, Hicham Bessaiah^{a,b}, Sébastien Houle^{a,b}, Frédéric J. Veyrier^{a,c}, Stéphanie Guyomard-Rabenirina^{cd}, Antoine Talarmin^{cd}, and Charles M. Dozois^{a,b,c}

^aInstitut national de recherche scientifique (INRS)-Institut Armand Frappier, Laval, Quebec, Canada; ^bCentre de recherche en infectiologie porcine et avicole (CRIPA); ^cInstitut Pasteur International Network; ^dUnité Environnement Santé, Institut Pasteur de Guadeloupe, Les Abymes, Guadeloupe, France

ABSTRACT

Serine protease autotransporters of *Enterobacteriaceae* (SPATEs) are secreted proteins that contribute to virulence and function as proteases, toxins, adhesins, and/or immunomodulators. An extra-intestinal pathogenic *E. coli* (ExPEC) O1:K1 strain, QT598, isolated from a turkey, was shown to contain *vat*, *tsh*, and three uncharacterized SPATE-encoding genes. Uncharacterized SPATEs: *Sha* (Serine-protease hemagglutinin autotransporter), *TagB* and *TagC* (tandem autotransporter genes *B* and *C*) were tested for activities including hemagglutination, autoaggregation, and cytotoxicity when expressed in *E. coli* K-12. *Sha* and *TagB* conferred autoaggregation and hemagglutination activities. *TagB*, *TagC*, and *Sha* all exhibited cytopathic effects on a bladder epithelial cell line. In QT598, *tagB* and *tagC* are tandemly encoded on a genomic island, and were present in 10% of UTI isolates and 4.7% of avian *E. coli*. *Sha* is encoded on a virulence plasmid and was present in 1% of UTI isolates and 20% of avian *E. coli*. To specifically examine the role of SPATEs for infection, the 5 SPATE genes were deleted from strain QT598 and tested for cytotoxicity. Loss of all five SPATEs abrogated the cytopathic effect on bladder epithelial cells, although derivatives producing any of the 5 SPATEs retained cytopathic activity. In mouse infections, *sha* gene-expression was up-regulated a mean of sixfold in the bladder compared to growth *in vitro*. Loss of either *tagBC* or *sha* did not reduce urinary tract colonization. Deletion of all 5 SPATEs, however, significantly reduced competitive colonization of the kidney supporting a cumulative role of SPATEs for QT598 in the mouse UTI model.

ARTICLE HISTORY

Received 7 February 2019
Revised 15 May 2019
Accepted 17 May 2019

KEYWORDS

Escherichia coli;
Autotransporters; serine
protease autotransporter;
SPATE; Toxins; avian
pathogenic *E. coli*; mouse
infection; uropathogenic
E. coli; poultry

Introduction


Escherichia coli is a common commensal of the gastrointestinal tract of mammals and birds, and is also a versatile pathogen associated with a variety of intestinal and extra-intestinal infections. Pathogenic *E. coli* belong to two main groups: intestinal pathogenic *E. coli*, and extra-intestinal pathogenic *E. coli* (ExPEC) [1,2]. Among ExPEC, the strains have been classified into pathotypes based on the sites of infection or the animal species they have infected, although these different ExPEC subgroups often share certain traits [3–8]. Such pathotypes include neonatal meningitis *E. coli* (NMEC), uropathogenic *E. coli* (UPEC), and avian pathogenic *E. coli* (APEC) [2,9,10]. Avian pathogenic *E. coli* (APEC) are a subset of ExPEC that cause respiratory infections and septicemia in poultry [4,10–12]. The genomes of a number of APEC strains and their virulence plasmids have been sequenced and share

similarities to some human ExPEC isolates and their plasmids [13–18]. The plasticity of the *E. coli* genome has led to the emergence of numerous combinations of genes that can be encoded on genomic islands or harbored on plasmids that can contribute to fitness, adaptability, and virulence of a variety of ExPEC strains [19–21].

APEC and human ExPEC strains share multiple virulence factors that promote survival and colonization of the host during extraintestinal infections. These include fimbriae, iron acquisition systems, autotransporter (AT) proteins, capsular polysaccharides, O-antigens, toxins and secretion systems [1,2,9,11,12]. Most APEC strains also contain conjugative colicin V (ColV) or similar plasmids that encode multiple virulence genes that have been shown to contribute to virulence in poultry [11,22,23], and also to urinary tract infection or systemic infection in rodent models [6,24,25]. The shared battery of virulence genes and the close phylogenetic relatedness of some

CONTACT Charles M. Dozois  charles.dozois@inrs.ca

*These two authors contributed equally as primary authors of this research

 Supplemental data for this article can be accessed here.

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

APEC and human ExPEC strains suggest that some APEC may be potential zoonotic pathogens for humans [6,7,25–28].

Among pathogenic *E. coli* virulence factors, AT proteins comprise a large family that falls into three main categories: SPATEs (Serine Protease Autotransporters of *Enterobacteriaceae*), trimeric AT proteins, and the self-associating autotransporters (SAATs), such as AIDA-1 and Antigen43 (Ag43) [29–31]. AT proteins are exported by the type V secretion system, which can be classified into 5 subgroups: Va for the monomeric autotransporters which includes SPATEs, Vb for the two-partner secretion system, Vc for the trimeric AT, Ve for the ATs that are homologous to both type Va and type Vb, and Vd for the intimins and invasins which have a reverse order of domains [32]. The export process of the AT may also require additional proteins such as the BAM and TAM assembly systems [33,34]. SPATEs consist of three specific domains: (i) a signal peptide which translocates the protein from cytoplasm to periplasm by the Sec-dependant pathway (ii) a functional passenger domain which contains a conserved serine protease motif (GDSGS), and (iii) a β -barrel domain which is localized in the outer membrane acting as a pore-forming domain that translocates the passenger domain [35]. SPATEs have been grouped into two main classes; class 1 SPATEs consist of cytotoxic proteins, whereas class 2 SPATEs represent immunomodulator proteins [32]. Certain SPATEs including the secreted autotransporter toxin (Sat), vacuolating autotransporter protein (Vat), temperature-sensitive hemagglutinin (Tsh), which has also been called hemoglobin protease (Hbp) [36], and protein involved in colonization (Pic) [37] have been previously reported in APEC and human ExPEC.

The SPATEs comprise a diverse group of autotransporter proteins that contribute to the virulence of pathogenic *E. coli* and *Shigella* spp., and other Enterobacteria [2,22,32,37–43]. Some SPATEs were shown to be important virulence factors in disseminated infection of ExPEC due to their proteolytic activity, which can promote the degradation of host cell substrates and elicit an inflammatory response [32,44]. In ExPEC, SPATE proteins have previously been characterized and have been shown to be associated with infections of both humans and other animals including poultry. SPATEs identified in uropathogenic *E. coli* include Sat [44], Vat [45,46] and PicU [41]. The *sat* gene encodes a vacuolating toxin and *sat* sequences were present in 55% of UPEC strains [40] but were not identified in a collection of APEC isolates [47]. PicU is homologous to the Pic protein identified in *Shigella* and enteroaggregative *E. coli* (EAEC) [37]. *PicU* was found in 22% of UPEC isolates [41] and 9% of APEC strains [47]. The Vat autotransporter was first discovered in APEC [45], was present in 60–70% of ExPEC from human infections

[46,48] and 33% of APEC strains [47]. The Vat toxin was shown to contribute to virulence, respiratory infection, and cellulitis in broiler chickens [45]. Both *pic* and *vat* were shown to contribute to the fitness of UPEC in a mouse model of systemic infection [43]. Tsh was the first SPATE identified in *E. coli* [49] and was shown to contribute to the development of respiratory lesions in the air sacs of chickens [22]. The *tsh* gene is located on ColV-type plasmids, was present in 50% of APEC strains [47], is less commonly associated with human ExPEC, but can be associated with certain human ExPEC strains [18,50–52].

In this report, analysis of the genome sequence of an APEC O1 strain, QT598, revealed that it contained 5 distinct SPATEs. Three of these, two chromosomally encoded SPATE genes (we name *tagB* and *tagC*) and a novel plasmid-encoded SPATE gene (*sha*) have not been previously characterized. The remaining two SPATEs were the previously characterized Vat and Tsh proteins. Herein, we have characterized the three novel SPATEs, determined their prevalence among avian and human urinary tract isolates, and investigated the role of these SPATEs for cytotoxicity and in the colonization of the murine urinary tract.

Results

Genomic analysis identifies five predicted SPATEs encoded by *E. coli* strain QT598

Strain QT598 was initially isolated from an infected turkey poult in France as MT156 [53]. It is a phylogenetic group B2 strain belonging to serogroup O1, a common serogroup among ExPEC strains causing infections in both poultry and humans. This APEC strain was sequenced initially because it contains most of the known APEC-associated virulence genes and was previously found to be virulent in one-day-old chicks [54]. QT598 belongs to sequence type (ST) 1385. Other strains belonging to ST1385 include other APEC O1 isolates, a canine urinary isolate, and environmental isolates (<http://enterobase.warwick.ac.uk/>). Interestingly, ST1385 strains are related to other STs including ST91, which contains some strains from human extra-intestinal infections and *E. coli* F54, an O18:K1 human fecal isolate sharing many virulence genes found in ExPEC from neonatal meningitis [55].

The genome of QT598 contains five SPATE-encoding sequences (Figure 1). Two of the SPATE genes, *tsh* and a novel SPATE which we have called *sha* (for serine-protease hemagglutinin autotransporter), are located on a ColV-type plasmid (pEC598). The *vat* gene was also identified on a genomic island. Finally, a genomic region was identified containing two distinct SPATE-encoding sequences in close proximity to each other, which we have

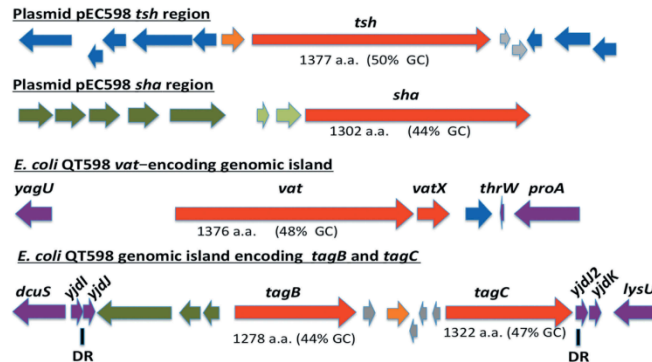


Figure 1. Regions containing the five SPATE-encoding genes in *E. coli* QT598.

The *tsh* and *sha* genes are located on a ColV-type plasmid (pEC598). The *vat*, *tagB*, and *tagC* genes are located on genomic islands. Arrows indicate open reading frames (ORFs). SPATE encoding ORFs and regulatory gene *vatX* are in red. Predicted full amino acid lengths and GC content of the SPATE ORFs are indicated below arrows. Blue ORFs are related to insertion elements, integrases, or mobile elements. Dark green ORFs are predicted fimbrial proteins. Light green ORFs are predicted EAL-domain proteins. Grey ORFs are hypothetical uncharacterized ORFs. Orange ORFs are hypothetical regulatory proteins. Purple ORFs are genes conserved in *E. coli* K-12 that border the SPATE-encoding genomic regions. Direct repeats (DR) are indicated for the region containing the *tag* AT genes.

named *tagB* and *tagC* (for *t*andem *a*utotransporter genes *B* and *C*).

The *tsh* open reading frame on plasmid pEC598 shares highest identity to *tsh* from plasmid pACN001-B (accession number KC853435.1) [56] and similar sequences in the NCBI database, differing in only 1 nucleotide, a Gly₁₁₇₇-Ser₁₁₇₇ substitution. Compared to the characterized Tsh (Hbp) proteins, Tsh from APEC strain χ 7122 [22] and hemoglobin protease (Hbp) from ExPEC strain EB1 [36], Tsh_{QT598} contains 4 and 2 amino acid differences, respectively. In QT598, *tsh* is also flanked by sequences related to transposases and insertional sequences (Figure 1) that also flank *tsh* on other IncFII plasmids [22,36]. The *sha* gene is also located on pEC598 and has a 44% GC content. Sequence analysis of the *sha* gene revealed an open reading frame (ORF) of 3909 bp encoding a predicted precursor protein of 1302 amino acids with an N-terminal domain signal peptide (residues 1–51), a passenger domain (residues 52–1026) (predicted molecular mass of 105.7 kDa) containing a consensus serine protease motif ²⁵⁶GDSGS, and β -barrel domain (residues 1027–1302). Interestingly, the highly conserved SPATE cleavage site of two consecutive asparagines “EVNNLNK”, found between the passenger domain and the β -barrel of most SPATEs [57], is absent in Sha. Sha is more closely related to Tsh and Vat proteins than to other SPATEs (Figure 2). The global alignment of Sha with Tsh_{QT598} has 43% identity/56% similarity with 237 gaps, whereas the global alignment of Sha with Vat_{QT598} is 38% identity/52% similarity with 252 gaps.

The *vat* gene from QT598 is present on a genomic island that includes the *vatX* regulatory gene, and is

located between the *E. coli* conserved genes *yagU* and *proA* adjacent to the *thrW*-tRNA gene (Figure 1). This is a conserved insertion site for *vat*-encoding genomic islands [58]. Vat_{QT598} is a predicted 1376 aa precursor with a single substitution (His₅₃₄-Arg₅₃₄) compared to Vat from UPEC strain CFT073 (accession number AAN78874.1). At least 41 predicted Vat protein sequences from different *E. coli* strains share an identical predicted Vat_{QT598} sequence, indicating it is a common allelic variant of Vat (Supplemental Table 2). These entries include sequences from strains isolated from fecal samples and infections of poultry and two human UTIs that are labeled as “Hbp” or “SepA” proteins in the databank.

The two new chromosomal-encoded SPATE genes were named *tagB* and *tagC*, Tandem autotransporter genes (Tag) because of their tandem co-localization in the genome of QT598 as well as in various other *E. coli* strains such as: multidrug-resistant CTX-M-15-producing ST131 isolate *E. coli* JJ1886 (Accession number CP006784), porcine *E. coli* PCN033 (Accession number CP006632), and *E. coli* Cl5 (Accession number CP011018). The *tagB* and *tagC* SPATE-encoding genes from QT598 are located on a genomic island between the *E. coli* conserved genes *yjdI* and *yjdK* (Figure 1). This genomic island has a mean GC content of 41%, which is considerably lower than the 50% GC of *E. coli*, suggesting horizontal gene transfer. This genomic region is also bordered by direct repeat (DR) sequences that correspond to duplication of *yjdI* sequences bordering each side of the genomic island (Figure 1).

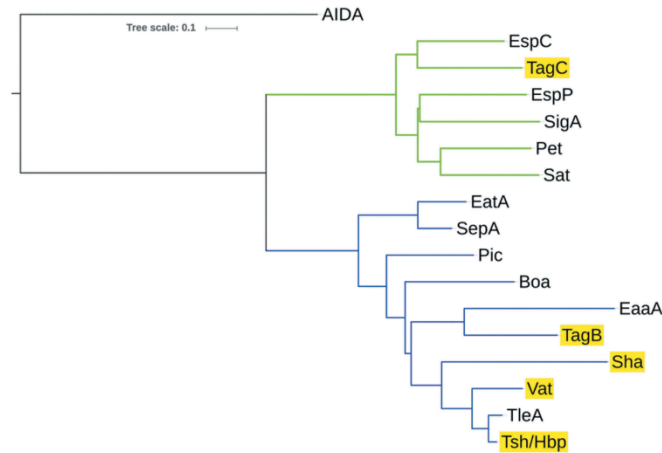


Figure 2. Phylogenetic analysis of new SPATEs identified in the QT598 genome.

The evolutionary history of passenger domains of QT598 SPATEs (highlighted in yellow) as well as other characterized SPATEs was inferred using the Neighbor-Joining method [98]. The optimal tree with the sum of branch length = 8.78918031 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method [99] and are in the units of the number of amino acid substitutions per site. The analysis involved 17 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 723 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 [85]. Multiple sequence alignment was performed by Clustal W, and the tree was constructed using the Mega6 software with PhyML/bootstrapping. A cluster of cytotoxic SPATEs (class 1) comprise the green branches, while immunomodulator SPATEs (class 2) are in blue branches. DNA regions encoding SPATE protein sequences are available in NCBI database as follows: EspC, GenBank Accession No. AAC44731, TagB and TagC, MH899681; EspP, NP_052685; SigA, AF200692; Pet, SJK83553; Sat, AAG30168; EatA, CAI79539, SepA, Z48219; Pic, ALT57188; Boa, AAW66606; EaaA, AAF63237; Sha, MH899684; Vat, MH899682; TleA, KF494347; Tsh/Hbp, MH899683.

Related genomic islands containing similar SPATE encoding genes at this insertion site are present in other *E. coli* genome sequences including antibiotic-resistant strains isolated from the urinary tract and other infections in humans (Supplemental Table 3). The predicted TagB and TagC proteins share the closest identity to the EaaA [59] and EspC [60], respectively (Figure 2). TagB shares 46% identity/63% similarity to EaaA with 84 gaps over its full length. TagC shares 60% identity/74% similarity to EspC with 22 gaps. TagB comprises a predicted signal peptide (residues 1–58), a passenger domain from residues 59–1006 (predicted molecular mass of 101 kDa) containing a consensus serine protease motif ²⁵³GDSGS, and a β -barrel domain from residues 1007–1283. TagC comprises a predicted signal peptide (residue 1–53), a passenger domain from residues 55–1032 (predicted molecular mass of 105.14 kDa) with a consensus serine protease motif ²⁵⁰GDSGS, and β -barrel domain ranging between 1033 and 1309 residues. Both TagB and TagC contain the conserved twin asparagine (N-N) residues in the linker domain connecting passenger and β -barrel domains, EIN¹⁰⁰⁶NLNDRM and EVN¹⁰³²NLNKRM, respectively.

Prevalence of new SPATE genes in human uropathogenic and avian pathogenic *E. coli*

To determine the distribution of the SPATE sequences among *E. coli* clinical isolates, the presence of these three new SPATE sequences as well as *vat*, *sat*, and *tsh* were detected by PCR in a collection of UPEC isolates from Guadeloupe (697 isolates) [61] and from avian pathogenic *E. coli* (299 isolates) [22]. For the UPEC isolates, *tagB* sequences were present in 70 isolates (10%), whereas *tagC* sequences were present in 80 isolates (11.5%). Interestingly, 96.8% (69/70) of the *tagB*-positive isolates were also *tagC*-positive. Furthermore, 68 of the *tagB* isolates belonged to phylogenetic group B2, with one isolate from group B1 and one untypable isolate. The 11 isolates that contained *tagC* but not *tagB* sequences belonged to groups other than B2: B1 (3 isolates), D (3 isolates), F (4 isolates), or A (1 isolate). *Sha* was the least common sequence, and was present in only 6 UTI isolates (0.9%), all of which belonged to group B2 and were also *vat*-positive. Five of the *sha*-positive strains also contained *tagB* and *tagC*, whereas *vat* and *sat* sequences were more common and found in 333 isolates (47.8%) and 217 isolates (31.1%), respectively. The *tsh* gene was present in 41 UPEC isolates (5.9%). In summary, in

UPEC, *tagB* and *tagC* were found together in a subset of strains belonging to phylogenetic group B2, although some strains belonging to other phylogenetic groups were only *tagC* positive.

With regards to the APEC strains, *tagB* sequences were present in 14 isolates (4.7%) and *tagC* sequences were present in 21 isolates (7%). All 14 *tagB*-positive APEC were also *tagC*-positive, and 13/14 of these belonged to phylogenetic group B2. Among these, 10 strains belonged to serogroup O1, 1 was serogroup O78, and 3 were of undetermined serogroup. Interestingly, among the 299 APEC strains that were screened, comprised of 109 from chickens, 175 from turkeys, and 15 from ducks, all of the *tagB* or *tagC*-positive isolates were exclusively from infections in turkeys. Overall, similar to UPEC, *tagB*, and *tagC* were specifically present in a subset of APEC strains, mainly belonging to group B2, although some strains belonging to other phylogenetic groups only contained *tagC* sequences.

The *sha* sequences were present in 61 APEC strains (20%). The majority, 42 strains, belonged to phylogenetic group A, 11 strains belonged to group B1, 5 strains belonged to group B2, and 3 strains belonged to group D. Among these *sha*-containing strains, 35 belonged to serogroup O78, 3 were from serogroup O1, 2 strains each belonged to serogroups O11, O54, O21, and O8, and one belonged to serogroup O55. The remaining 12 strains were from undetermined serogroups. The *sha* gene is, therefore, clearly more prevalent among APEC than UPEC in the subset of strains we analyzed.

Cloning and production of SPATEs in culture supernatants

All five of the predicted SPATE-encoding genes and promoter regions were cloned to determine their expression and for use in a variety of phenotypic tests. Each of the five SPATE genes, when cloned into *E. coli* BL21, produced a high-molecular-weight protein (>100 kDa) in culture supernatants that corresponded to the expected product (Figure 3). In addition, derivatives of strain QT598 wherein these SPATE-encoding genes were inactivated were generated. Analysis of supernatant fractions of QT598, by SDS-PAGE, revealed the presence of SPATE proteins expressed under laboratory conditions, as demonstrated by visualization of bands with a high molecular mass (>100 kDa) secreted in the external milieu. By contrast, no such bands were observed in the supernatant extracts of the SPATE-free, Δ 5ATs, derivative of QT598 (Figure 3). The purity of concentrated supernatant filtrate was also evaluated by silver staining (Supplemental Fig. S1). Protein bands of the newly identified SPATEs were extracted from gels and sampled by mass spectrometry for peptide analysis following trypsin digestion

(Supplemental Fig. S2). For Sha, peptides corresponding to the mature secreted protein spanned from amino acids 52 to 1009. Despite not containing the twin asparagine (N-N) cleavage site present in most SPATEs, peptide profiles suggest the cleavage site from the β -barrel domain likely resides within the 1010–1020 region. This region contains two adjacent polar amino acids, Ser1015 and Asp1016, that may serve as the cleavage site. For TagB, peptide scans suggest that the mature secreted protein spans from amino acids 54 to 1006 based on the twin Asp1006-Asp1007 location. For TagC, peptide scans suggest that the mature secreted protein spans from at least amino acid 60 to 1026 with a predicted twin Asp1032-Asp1033 cleavage site. As expected, peptides corresponding to the predicted amino-terminal signal peptides and the carboxy-terminal predicted β -barrel domains of the Sha, TagB, and TagC SPATEs were not identified from peptide analysis of the secreted proteins (Supplemental Fig. S2).

Cleavage of oligopeptides by SPATE proteins

To determine the protease substrate cleavage specificity of the new SPATEs, we used synthetic polypeptides conjugated with pNA at the C-terminus. Purified proteins from supernatants of each SPATE were incubated with N-Succinyl-Ala-Ala-Ala-p-nitroanilide (elastase substrate), N-Benzoyl-L-arginine 4-nitroanilide (trypsin substrate) and N-succinyl-ala-ala-pro-phe-p-nitroanilide (chymotrypsin substrate) (Sigma-Aldrich, St. Louis, MO, USA). TagB and TagC demonstrated trypsin-like activity and efficiently cleaved N-Benzoyl-L-arginine 4-nitroanilide, similarly to the EspC protein (Figure 4(a)). By contrast, Sha demonstrated significant elastase-like activity toward N-Succinyl-Ala-Ala-Ala-p-nitroanilide, as did the Vat and Tsh proteins (Figure 4(a)). The cleavage activity of high-molecular-weight supernatant fractions from WT strain QT598 and SPATE mutant derivatives was also determined (Figure 4(b)). QT598 demonstrated both trypsin-like and elastase-like activity, whereas the Δ 5ATs mutant had lost these activities. By contrast, a strain which had lost only *tagBC* demonstrated only elastase-like activity conferred by *vat*, *tsh*, and *sha* (Figure 4(b)). In addition, pre-incubation of these supernatants with PMSF eliminated or sharply inhibited oligo-peptide cleavage indicating the activity demonstrated was due to the SPATE proteins produced by the strains (Supplemental Fig. S3).

Multiple alignment of the new autotransporters with other SPATEs places TagC within the class 1, cytotoxic and enterotoxic SPATEs, along with the EspC SPATE from Enteropathogenic *E. coli* (EPEC) (Figure 2). EspC was previously shown to cleave spectrin, Factor V, pepsin and hemoglobin [42,62] and as with TagC and TagB, similarly demonstrated trypsin-like protease activity (Figure 4(a)). The SPATE sharing closest identity to TagB

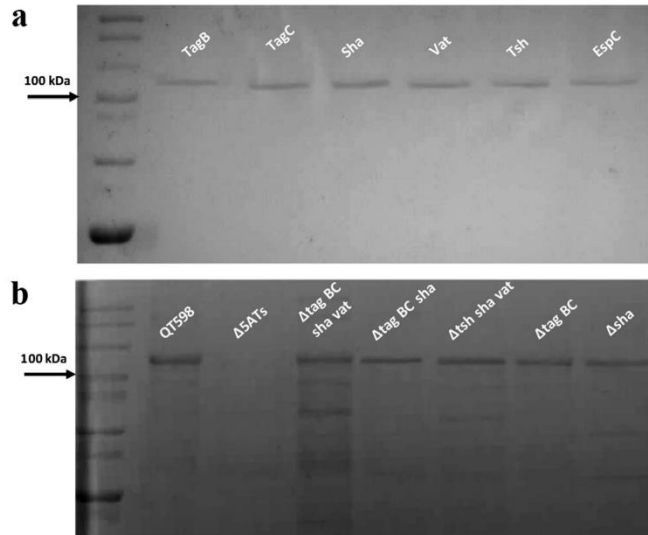


Figure 3. Detection of SPATE proteins by SDS-PAGE. **A.** SDS-PAGE analysis of cloned SPATE genes. **a.** Clones expressing SPATE proteins were produced in the BL21 background with high-copy plasmid pBCsk+. Supernatants were filtered then concentrated through Amicon filters with 50 kDa cutoff. Samples containing 5 μg protein were migrated with protein marker (10–200 kDa) and stained with Coomassie blue (arrow represents 100 kDa size marker). **b.** Detection of SPATES from supernatants of strain QT598 and various SPATE gene mutant derivatives. Supernatants from an overnight culture of the respective mutants were filtered, concentrated and run on SDS-polyacrylamide gels and stained with Coomassie blue to visualize proteins.

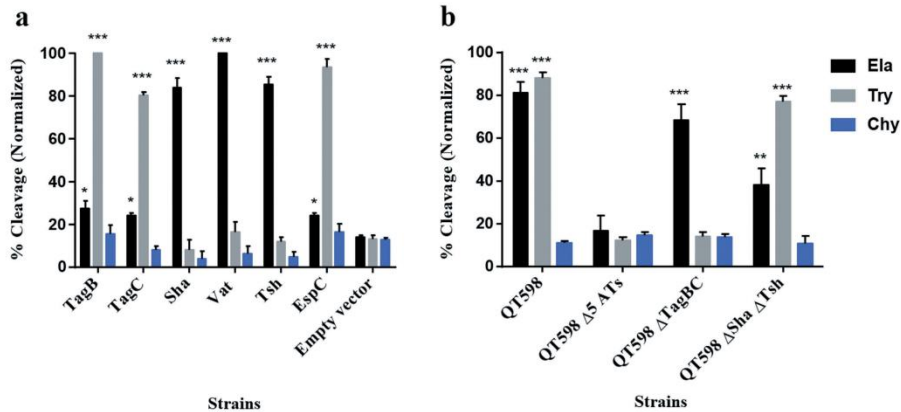


Figure 4. Oligopeptide cleavage profiles of SPATEs. **a.** Enzymatic activity of cloned SPATEs. Five μg of each SPATE-containing supernatant was incubated at 37°C for 3 h with 1mM of synthetic oligopeptide specifically recognized by the following enzymatic activities: Elastase (Ela)-(N-Suc-Ala-Ala-Ala-pNA); Trypsin (Try)-(N-Ben-L-arginine-pNA); or Chymotrypsin (Chy)-(N-Suc-Ala-Ala-Pro-Phe-pNA). Absorbance at 410nm was normalized to the maximum absorbance value. **b.** Enzymatic activity of supernatants from strain QT598 and SPATE gene mutant derivatives. Samples were tested as described above. Data are the means of three independent experiments, and error bars represent the standard errors of the means (* p < 0.05, **p < 0.01, ***p < 0.001 one-way ANOVA with multiple comparisons vs pBCsk+ (A) or QT598Δ5ATs (B)).

is the class 2 SPATE EaaA, identified from commensal *E. coli* ECOR-9 [59] (Figure 2). Sha shares more identity to class 2 SPATEs Tsh/Hbp (66% identity) [36,63], TleA (60% identity) [64] and Vat (56% identity) [45]. As such, Sha is likely to share other properties more similar to Tsh (adhesin, hemagglutinin) and Vat (cytotoxin), and the elastase-like substrate specificity of Sha, also shown for Tsh and Vat, is in line with this.

To predict the 3D structure of the passenger domain of new SPATEs, we used the I-Tasser program to generate a 3D structure model and UCSF chimera to compare structures [65,66]. We found that the Sha protein is also predicted to contain the small domain 2 that was identified in Tsh/Hbp and was considered to be characteristic of class 2 SPATEs [32] (Figure 5). This domain was however absent in predicted models of TagB and TagC (Fig. S4).

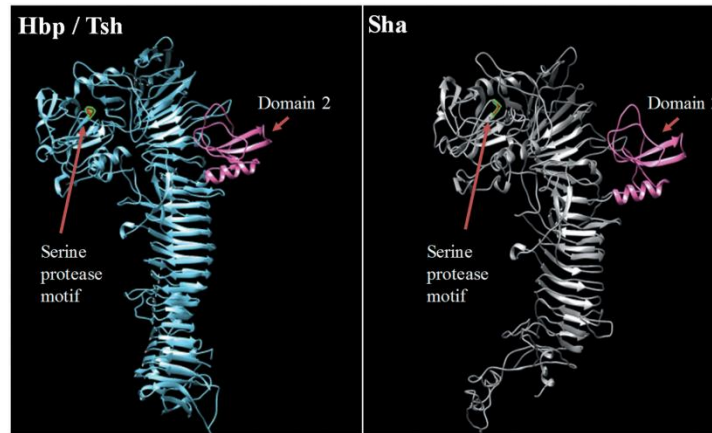


Figure 5. Stereo ribbon diagram showing the predicted three-dimensional structure of the Sha SPATE passenger domain derived from the Hbp/Tsh crystal structure.

Crystal structure of the Heme-binding protein (Hbp) (PDB 1WXR), which is near identical to Tsh, was used to model a homologous structure based on alignment with the Sha protein sequence. The model was generated using the I-TASSER server with 100.0% confidence by the single highest scoring template. Sha is shown to harbor a conserved domain, domain 2 (shown in pink), which is characteristic of class 2 SPATEs.

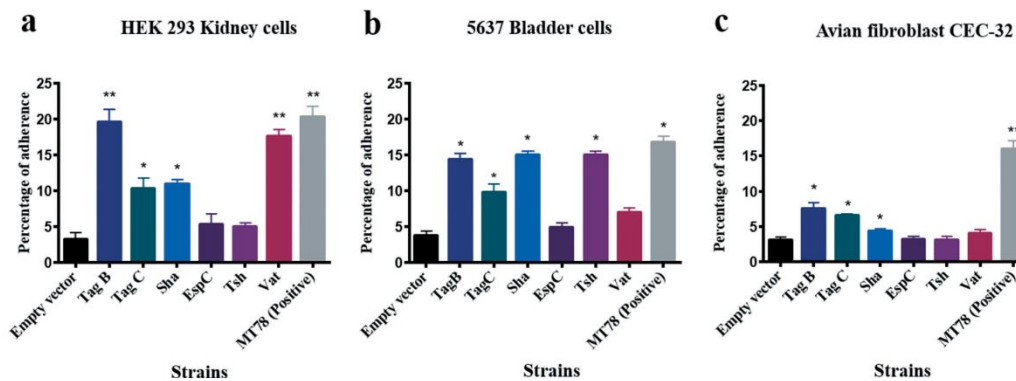


Figure 6. TagB, TagC, and Sha SPATEs promote adherence to the human kidney (HEK-293) and bladder (5637) epithelial, and avian fibroblast (CEC-32) cell lines.

Cell monolayers were infected with *E. coli* *fim*-negative ORN172 expressing SPATE proteins at a multiplicity of infection (MOI) of 10 and incubated at 37°C at 5% CO₂ for 2 h. Adherent bacteria were enumerated by plating on LB agar. Empty vector (pBCsk+) was used as a negative-control and APEC MT78 [80] as a positive control for adherence to cell lines. Data are the averages of three independent experiments. Error bars represent standard errors of the means. (**p* < 0.05, ***p* < 0.01, ****p* < 0.001 vs empty vector by one-way ANOVA).

Increased adherence to epithelial cells is mediated by SPATEs

Some AT proteins can promote adherence to host cells [31,67]. To investigate this with TagB, TagC, and Sha, the different SPATE encoding genes were cloned in the *fim*-negative *E. coli* K-12 strain ORN172 and tested for increased adherence to avian fibroblasts (CEC-32), human kidney (HEK-293) and bladder (5637) epithelial cells (Figure 6). Clones expressing either TagB, TagC, Sha, or Vat adhered significantly to kidney cells, whereas Tsh and EspC did not significantly increase adherence (Figure 6(a)). For the bladder cells, all SPATEs tested except for Vat and EspC significantly increased adherence (Figure 6(b)). By contrast, for avian fibroblasts, only TagB and TagC increased adherence (Figure 6(c)). These cell culture-based results suggest that TagB, TagC, and Sha, as well as Vat and Tsh, may contribute to host cell interactions in the urinary tract, and that TagBC may also contribute to adherence to tissues in poultry. Although some individual SPATEs were shown to increase adherence to host cells, loss of all 5 SPATEs did not reduce adherence of strain QT598 (Fig. S5).

Sha, TagB, Vat, and Tsh are hemagglutinins

Tsh has been previously shown to hemagglutinate chicken and sheep erythrocytes [22,49,68]. To assess whether other SPATEs also show hemagglutinin activity, we verified hemagglutination by the different SPATEs with erythrocytes from a variety of species. Interestingly, Sha, Tsh, and Vat all demonstrated hemagglutinin activities against sheep, bovine, pig, dog, chicken, turkey, rabbit, horse, and human blood (type O and A groups). In addition, TagB and TagC hemagglutinated sheep, bovine and pig erythrocytes, but not human, turkey, rabbit, dog, chicken, or horse erythrocytes (Table 1). However, the titer for TagC was very low for any erythrocytes tested and EspC demonstrated no hemagglutination.

TagB, TagC, and Sha mediate autoaggregation, but only Sha increases biofilm formation

Some AT proteins such as AIDA-1 and Ag43 mediate cell-cell interactions and autoaggregation, which can contribute

to virulence and facilitate host cell adherence [69–71]. We observed that the absorbance of clones expressing Sha, TagB, and TagC dropped rapidly similar to the positive control AIDA-1 (Figure 7(a)). An aggregative adherence pattern was also observed on interaction with bladder epithelial cell culture (Figure 7(b)). As these SPATEs demonstrated autoaggregation, we were interested to know if these plasmids could also confer autoaggregation to ExPEC QT598. However, the introduction of these plasmids did not lead to an autoaggregation phenotype in QT598 (Figure 7(a)).

Proteins involved in autoaggregation can also increase biofilm formation [69]. We, therefore, checked biofilm forming capacity of these new SPATE clones at different temperatures (25°C, 30°C, 37°C, and 42°C). TagB and TagC did not increase biofilm formation. However, Sha, Vat, and Tsh significantly increased biofilm production at lower temperatures (25°C and 30°C), but no significant differences in biofilm production were observed at higher temperatures of 37°C and 42°C (Figure 8).

Assessment of the cytopathic effect of 5 different SPATEs on bladder cells

Since some SPATEs produce cytopathic activity on host cells, we assessed the cytopathic effect of extracts of supernatants of the different SPATEs as well as the supernatant of wild-type strain QT598 and SPATE-free Δ 5ATs mutant, on the human bladder 5637 cell line. Incubation of SPATEs from concentrated filtered supernatant (Figure 3(b)) of strain QT598 with bladder cells triggered a cytopathic effect, characterized by the dissolution of cytoplasm and enlargement of the nucleus (Figure 9(a)), after 5 h of incubation. After 12 h, the majority of the cells were affected and showed similar morphological changes. These phenotypes were absent upon the treatment with the supernatant of the Δ 5AT SPATE-free mutant or with culture media alone.

Further, we assessed the cytopathic effect of individual SPATEs that were from recombinant clones. Following 12 h of interaction, cells incubated with TagB or TagC proteins elicited distinct cytopathic changes including

Table 1. Hemagglutination activities of different SPATEs.

SPATEs	Erythrocytes – Species (titer dilution) ^a								
	Sheep	Bovine	Pig	Chicken	Turkey	Rabbit	Horse	Dog	Human
Vat	6	5	4	3	3	7	4	7	7
Tsh	8	7	6	7	4	7	7	7	7
Sha	6	3	5	6	4	3	3	5	5
TagB	6	3	3	-	-	-	-	-	-
TagC	0–1	1	0–1	-	0–1	-	-	-	-
EspC	-	-	-	-	-	-	-	-	-
pBCsk+ vector	-	-	-	-	-	-	-	-	-

a- Clones expressing SPATEs in *E. coli* ORN172 were adjusted to an O.D._{600nm} of 0.6 and then concentrated 100-fold. Samples were then diluted twofold in microwell plates containing final suspensions of 3% erythrocytes from different species. Titers are the average maximal dilution showing agglutination. Both human A and O blood gave similar titers. 0–1: little or no detectable agglutination. See Methods section for details.

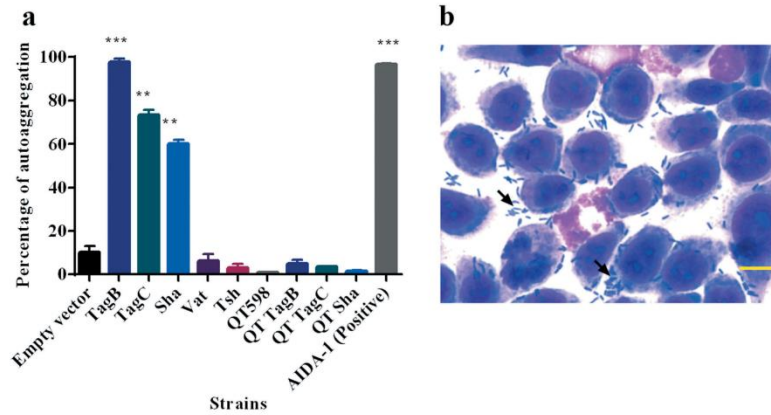


Figure 7. TagB, TagC, and Sha are autoaggregating proteins.

a. Clones of *E. coli fim*-negative strain ORN172 expressing SPATE proteins were grown 18 h and adjusted to OD_{600} of 1.5 and left to rest at 4° C. Samples were taken at 1 cm from the top surface of the cultures after 3 h to determine the change in OD_{600} . Assays were performed in triplicate, and the rate of autoaggregation was determined by the mean decrease in OD after 3 h. The autoaggregation phenotype was absent when plasmids expressing *tagB*, *tagC*, or *sha* were introduced into APEC strain QT598 (QT TagB, QT TagC, QT Sha respectively). Empty vector (pBCsk+) was used as a negative control and the AIDA-1 AT as a positive control for autoaggregation. Error bars represent standard errors of the means (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ compared to empty vector using one-way ANOVA). b. Giemsa stain of the 5637 bladder cell line infected with a *tagB* expressing clone after 2 h demonstrates an aggregative adherence pattern to bladder cells (arrow head). A similar pattern was found for *tagC* and *sha* expressing clones (not shown here). Bar represents 50 μm.

dissolution of cytoplasm, nuclear enlargement and vacuolation in the nucleus, cells treated with Sha were rounded, and Vat-treated cells showed numerous vacuoles within the cytoplasm (Figure 9(b)). By contrast, cells exposed to Tsh did not show distinct morphological changes. These results suggest that TagB, TagC, Sha and Vat proteases demonstrate cytopathic effects that alter bladder cell

morphology, whereas Tsh was less cytopathic to this cell line. We further investigated the cytopathic effect of these SPATEs by measuring the release of lactate dehydrogenase (LDH) after 5 h and 12 h following exposure to supernatant extracts. The release of LDH after 5 h was demonstrated only following exposure to either TagB or TagC (Figure 9(c)). Interestingly, although LDH was not

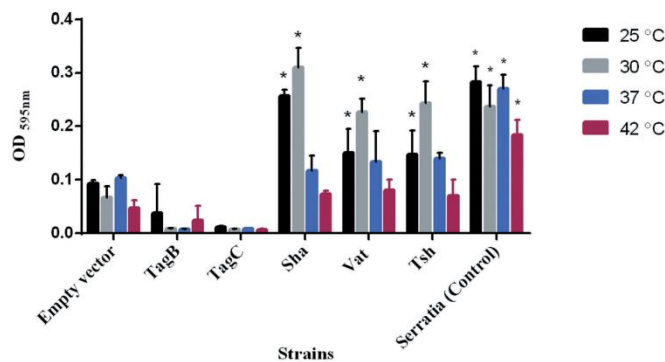


Figure 8. Sha, Vat, and Tsh promote biofilm formation.

Clones of *E. coli fim*-negative strain ORN172 expressing SPATE proteins were grown at different temperatures (25°C, 30°C, 37°C, and 42°C) in polystyrene plate wells for 48 h and then stained with crystal violet. Remaining crystal violet after washing with acetone was measured as absorbance at 595 nm. Data are the means of three independent experiments, and error bars represent standard errors of the means. Empty vector (pBCsk+) was used as a negative control, and a string biofilm producing *Serratia* strain [100] served as a positive control for biofilm formation (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ compared to empty vector using one-way ANOVA).

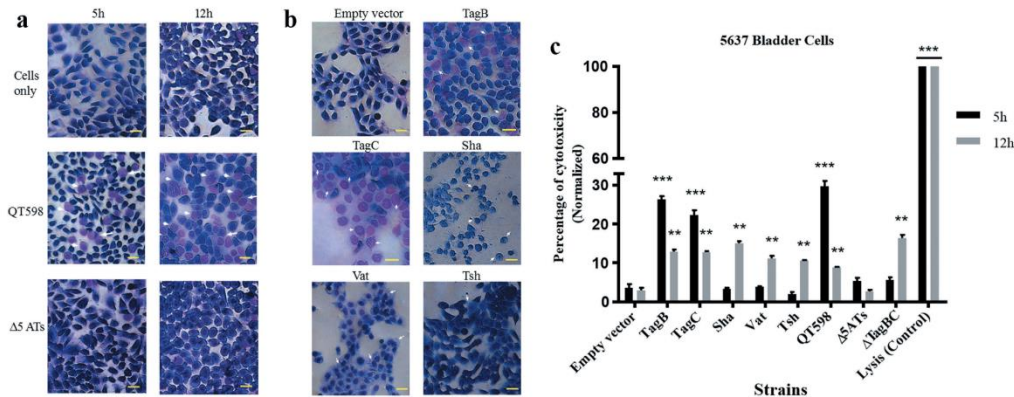


Figure 9. All five different SPATEs from strain QT598 induce cytopathic effects.

a. Concentrated supernatants (30 μ g of protein per well) from wild type QT598 as well SPATE free Δ 5ATs were incubated with human bladder cell line 5637 for 5 h and 12 h. Cytopathic cells (arrowheads) were found in the cells treated with QT598 supernatant while there were no morphological changes in the cells treated with supernatant of Δ 5ATs. b. Concentrated supernatants of *E. coli* BL21 pBCsk+ (30 μ g of protein per well) clones overexpressing different SPATEs were incubated with monolayers of 5637 bladder cell lines for 12 h at 37°C. TagB, TagC, Sha, and Vat showed more cytopathic effect (arrowheads) compared to Tsh and empty vector. c. LDH release by 5637 cells after incubation with culture filtrates of different clones (30 μ g of protein per well) expressing SPATEs or from supernatants from wild-type *E. coli* strain QT598 and Δ 5ATs mutant derivatives with 5637 human bladder cells at 37°C for 5 h and 12 h. Empty vector (pBCsk+) was used as a negative control. Lysis solution was added as a positive control for maximum LDH release. Data are the means of three independent experiments, and error bars represent the standard errors of the means (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ vs empty vector using one-way ANOVA). Bar represents 50 μ m.

detected from samples exposed to Vat, Tsh and Sha after 5 h (Figure 9(c)), some LDH release was observed after 12 h of exposure, suggesting these SPATEs may demonstrate a delayed cytotoxic effect. Hence, all 5 SPATEs elicited some cytotoxicity that corresponded with cytotoxic effects that were observed in cells following Giemsa staining. Of further interest, the concentrated supernatant filtrates from ExPEC strain QT598 showed early toxicity comparable to the concentrated supernatant filtrates from *tagB* or *tagC* expressing clones. However, loss of *tagB* and *tagC* resulted in only a late onset cytotoxic effect at 12 h, and loss of all 5 SPATEs abrogated LDH release at either early or late time points (Figure 9(c)). Taken together, these results indicate TagB and TagC can mediate an early (5 h) cytotoxicity, whereas Vat, Tsh, and Sha mediate late onset (12 h) cytotoxicity, and that these 5 SPATEs collectively mediate the overall cytotoxic effects of ExPEC QT598 on bladder epithelial cells.

Cumulative role of SPATEs for colonization during urinary infection in mice

In order to determine the potential role of SPATEs during UTI, we tested isogenic mutants with deletions of the *tagB*, *tagC*, and *sha* genes in a murine transurethral infection model. We observed no significant differences in colonization of the wild-type compared to *tagB*, *tagC* or *sha*

knockout mutants (Figure 10). To determine the collective role of the SPATEs in UTI for QT598, we deleted all five SPATE encoding genes and did transurethral infections. We again observed no significant difference in the colonization of kidneys or bladder following single-strain infections. By contrast, when a co-infection model was used (1 mouse died after 24 h), the Δ 5ATs mutant was significantly outcompeted by the wild-type by more than 10-fold in kidneys ($p = 0.0037$) (Figure 10)

Sha gene expression is upregulated during infection in the mouse bladder

The level of expression of SPATE encoding genes from samples grown in different culture conditions as well as from infected mouse bladders was determined. All 5 SPATE genes were expressed in LB medium as well as in minimal M63-glycerol medium. Interestingly, the *vat* gene was upregulated by 10-fold in the minimal medium compared to LB. In bladder samples from infected mice, all the SPATE genes were detected and expressed during infection. Interestingly, the *sha* gene was upregulated 6-fold in the bladder (Figure 11). However, expression levels of the four other SPATE genes were not significantly different in bladders when compared to expression during culture in LB.

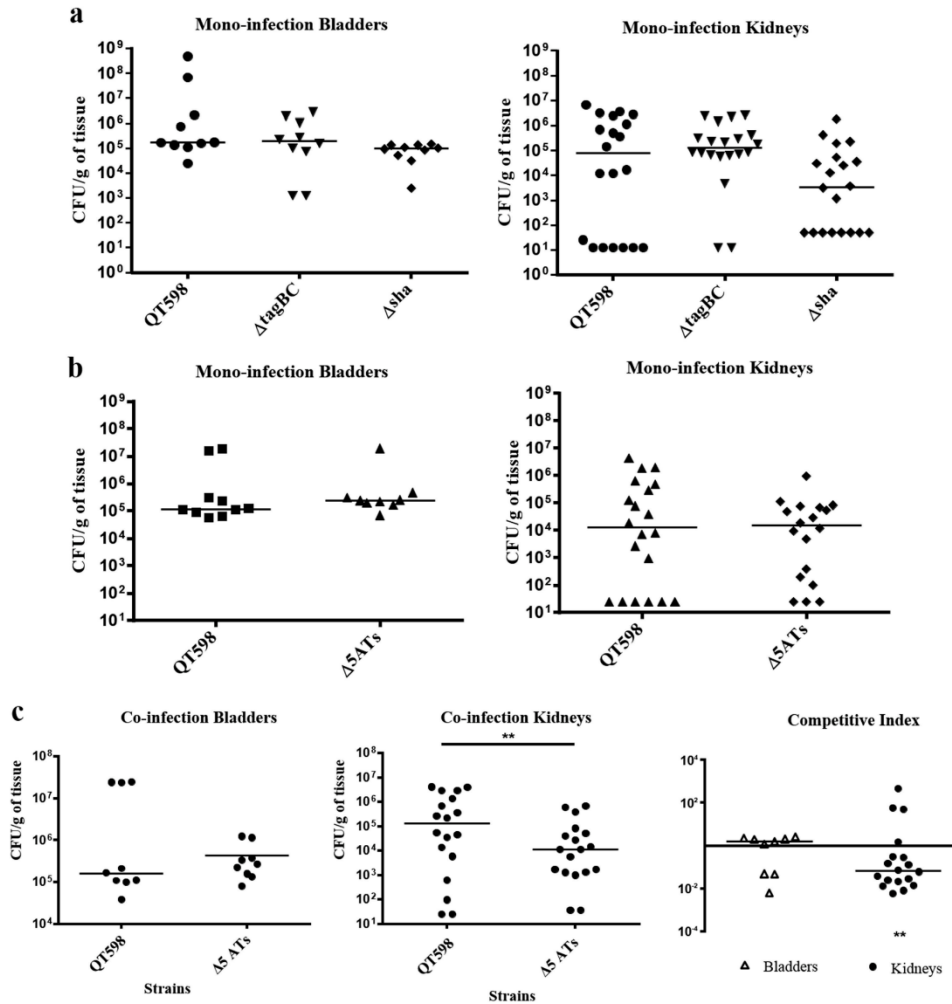


Figure 10. Role of SPATEs for *E. coli* QT598 in the murine model of ascending UTI.

CBA/J mice were challenged transurethrally with QT598 and isogenic strains $\Delta tagBC$, Δsha , or $\Delta 5ATs$ (wherein all 5 SPATE genes are inactivated). Mice were euthanized after 48 h, and bladder and kidneys were harvested for colony counts. a. Single-strain infections to compare wild-type strain QT598 to $\Delta tagBC$, Δsha mutants. There were no significant differences in bacterial numbers in either the bladders or kidneys. b. Single-strain infections to compare wild-type strain QT598 to the $\Delta 5ATs$ mutant. Similarly, there were no significant differences in colonization observed. c. Co-infection experiments between the QT598 Δlac and the $\Delta 5ATs$ mutant. The $\Delta 5ATs$ mutant colonized the bladder at similar levels to the wild-type (Δlac) strain; however, the $\Delta 5ATs$ strain was outcompeted in the kidneys by over 10-fold (Data are means \pm standard errors of the means of 10 mice (* $p < 0.05$, ** $p < 0.01$, Mann-Whitney Test).

Discussion

Some pathogenic *E. coli* produce multiple SPATEs, and this may provide a greater capacity to infect different host species or tissues. Having a combination of SPATEs may also allow a reserve of functions that may include both specific and redundant functions which may importantly be differentially regulated during infection or

colonization. UPEC strain CFT073 has 10 autotransporter proteins including 3 SPATEs: Sat, Vat and PicU [41]. *Shigella flexneri* also contains 3 SPATEs: SepA, Pic and SigA [37,72,73]. Similarly, *Citrobacter rodentium*, a model for A/E lesions of EPEC and Enterohemorrhagic *E. coli* (EHEC) also produces 3 SPATEs [74]. The ExPEC strain QT598 we investigated in this report produces a total of 5

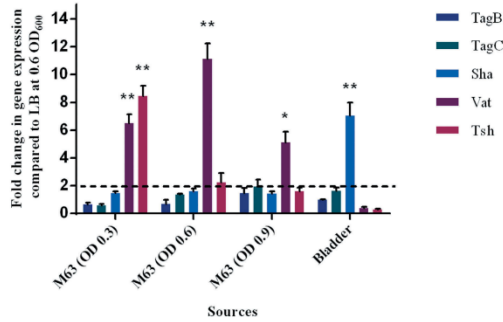


Figure 11. Differential expression of some SPATE genes occurs *in vitro* and in mouse bladder.

qRT-PCR analysis of SPATE gene transcription from QT598 strain grown in different conditions. Growth in rich medium (LB) to OD₆₀₀ of 0.6 was used as a standard and compared to growth in M63 minimal medium (with glycerol as carbon) at different growth phases (OD₆₀₀ of 0.3, 0.6, and 0.9). RNAs were also extracted from infected mouse bladder. Transcription of *tsh* and *vat* genes were significantly increased in minimal medium. Further, the *sha* gene was shown to be significantly upregulated in the mouse bladder. (* $p < 0.05$, ** $p < 0.01$, error bars indicate standard deviations, Student *t*-test). Expression of other SPATE genes was tested under similar conditions. See methods section for details concerning the calculation of gene expression levels. Expressions of other SPATE genes were similar under all conditions tested. The dashed line corresponds to the cutoff for a significant difference in expression.

SPATEs including three previously uncharacterized members: Sha, TagB, and TagC. Each of the 5 SPATEs demonstrated individual as well as shared properties or functions. Sha, located on a ColV-type plasmid (pEC598), is a class 2 SPATE, and it is closely related to Tsh and Vat proteins. We found this new AT has a modified cleavage site lacking twin asparagine (NN) sites between the passenger domain and the β -barrel. This absence of a typical cleavage site was also seen in *rpeA* (Rabbit-specific enteropathogenic *Escherichia coli* (REPEC) plasmid-encoded autotransporter) [75], indicating that the passenger domain is not cleaved from the outer membrane. However, in the case of Sha when cloned in BL21, a band was detected in polyacrylamide gel indicating that the modified cleavage site contributes to the separation of the passenger domain from the β -barrel.

TagB and TagC were found on a genomic island between the conserved *E. coli* genes *yjdI* and *yjdK*. Prevalence of the SPATE sequences among UPEC and APEC demonstrated that *tagB* and *tagC* sequences were present in at least 10% of the UPEC strains and that these genes were largely associated with strains belonging to phylogenetic group B2. Interestingly, genomic islands harboring *tagB* and *tagC* are also present in the genomes of

Table 2. Strains and plasmids used in this study.

Strains	Characteristic(s)	References
ORN172	<i>thr-1 leuB thi-1Δ(argF-lac)U169 xyl-7 ara-13 mtl-2 gal-6 rpsL tonA2 supE44Δ(fimBEACDFGH)::Km pilG1</i>	[101]
BL21	<i>fhuA2 [lon] ompT gal [dcm] Δhds5</i>	New England Biolabs
QT1603	HB101 with AIDA-1 operon	[92]
QT598	APEC O1:K, serum resistant	[53]
MT78	APEC O2:K1	[80]
QT2799	<i>Serratia liquefaciens</i>	ATCC27592
QT4567	QT598 Δ lacZYA	This study
QT4726	QT598 Δ tagBC::kan, Km ^R	This study
QT5187	QT598 Δ tagBC::FRT	This study
QT5188	QT598 Δ tagBC Δ vat::cat, Cm ^R	This study
QT5189	QT598 Δ tagBC Δ vat::cat Δ sha::kan, Cm ^R Km ^R	This study
QT5182	Δ SATs or QT598 Δ tagBC Δ vat:: cat Δ sha:: kan Δ tsh::tetAR(B) Cm ^R Km ^R Tc ^R	This study
QT5190	QT598 Δ sha:: kan Δ tsh:: tetAR(B) Km ^R Tc ^R	This study
QT5191	QT598 Δ tagBC Δ vat:: cat Δ tsh:: tetAR(B) Cm ^R Km ^R Tc ^R	This study
QT5192	QT598 Δ sha:: kan Km ^R	This study
QT5193	QT598 Δ tsh:: tetAR(B) Tc ^R	This study
QT12	ORN172/pBCsk+	This study
QT4750	ORN172/pIJ551, (Expressing <i>vat</i>)	This study
QT4751	ORN172/pIJ552 (Expressing <i>tsh</i>)	This study
QT4767	ORN172/pIJ553 (Expressing <i>sha</i>)	This study
QT5194	BL21/pIJ548 (Expressing <i>tagB</i>)	This study
QT5195	BL21/pIJ549 (Expressing <i>tagC</i>)	This study
QT5197	BL21/pIJ550 (Expressing <i>espC</i>)	This study
QT5198	ORN172/pIJ548 (Expressing <i>tagB</i>)	This study
QT5199	ORN172/pIJ549 (Expressing <i>tagC</i>)	This study
QT5201	ORN172/pIJ550 (Expressing <i>espC</i>)	This study
QT5431	BL21/pIJ551 (Expressing <i>vat</i>)	This study
QT5432	BL21/pIJ552 (Expressing <i>tsh</i>)	This study
QT5433	BL21/pIJ553 (Expressing <i>sha</i>)	This study
Plasmids		
pKD3	Plasmid used for amplification of <i>cat</i> cassette	[89]
pKD4	Plasmid used for amplification of <i>kan</i> cassette	[89]
pKD13	Plasmid used for amplification of <i>kan</i> cassette	[89]
pKD46	λ Red plasmid; Amp ^r	[89]
pCP20	FLP recombinase, Amp ^r	[89]
pBCsk+	Cloning vector; Cm ^r	Stratagene, La Jolla, CA
pIJ548	pBCsk+:: <i>tagB</i>	This study
pIJ549	pBCsk+:: <i>tagC</i>	This study
pIJ550	pBCsk+:: <i>espC</i>	This study
pIJ551	pBCsk+:: <i>vat</i>	This study
pIJ552	pBCsk+:: <i>tsh</i>	This study
pIJ553	pBCsk+:: <i>sha</i>	This study

numerous multi-resistant clinical isolates including members of the *E. coli* ST131 pandemic clone such as *E. coli* JJ1877 [76] and many other CTX-producing clinical isolates from urinary tract infections or sepsis (Supplemental Table 2). Further investigation into the potential role of these newly identified SPATE toxins for the virulence of such human ExPEC is therefore warranted.

Although *tagB* and *tagC* were less prevalent in APEC, the strains carrying those genes were mostly O1 strains belonging to phylogenetic group B2 and were all isolated from diseased turkeys. Although *sha* was only present in 1% of UPEC, it was present in 20% of APEC strains. As with Tsh, Sha is plasmid-encoded and these ColV-type plasmids are present in nearly all APEC but are less common in UPEC [77]. Similar phenotypes were observed for

the Tsh, Vat, and Sha autotransporters, including hemagglutination, adherence, protease activity, biofilm formation, and cytopathic effects. As such, the role of these three SPATEs may also be cumulative for some APEC strains.

We assessed the cumulative role of SPATEs in APEC strain QT598 with cell cytotoxicity assays and infection experiments in the murine UTI model. It has been shown that certain APEC strains are highly similar to human ExPEC and can belong to the same clonal groups and contain similar virulence gene profiles [2,6,9,11,12,14,24,25]. Previous reports have also tested APEC strains in the mouse UTI model [78–81]. QT598 is a serogroup O1 strain, which is a common serogroup of both APEC and human ExPEC strains, and is clonally related to some human ExPEC, further supporting verification of the role of SPATEs for this strain in a UTI model. QT598 was initially isolated from a young turkey poult and was virulent by subcutaneous infection of 1-day-old chicks. We initially tested QT598 strain in a 3-week-old chicken air sac inoculation model. However, the strain only caused very limited disease and was rapidly cleared in this model. It is possible that the strain is only able to infect very young chicks or it may be more specific for infection of turkeys than chickens. It will be of interest to investigate this in the future.

By cloning each of the SPATE encoding genes in *E. coli* K-12, specific activities of individual SPATEs could be determined. Protease cleavage using oligopeptides demonstrated that Sha, as well as Tsh and Vat, demonstrated elastase-like activity, whereas TagB and TagC demonstrated trypsin-like activity, similar to the EspC autotransporter (Figure 4). Oligopeptide degradation was also observed in strain QT598 with TagB and TagC being required for trypsin-like activity, whereas the Class 2 SPATEs contributed to Elastase-like activity. As such, the combination of these SPATEs provides an expanded spectrum of protease activity. However, adding a serine protease inhibitor (PMSF) to the supernatant of the SPATEs neutralized their effect on the cleavage of these oligopeptides. Further, when the catalytic site of the ATs was mutated (serine was replaced by alanine) proteolytic activity was absent, indicating that the serine protease motif is important for the activity of these SPATEs (Supplemental Fig. S3). Previously, Tsh was shown to be proteolytic to substrates including mucin and factor V, whereas EspC could cleave other proteins such as pepsin and spectrin [62]. This suggests that the combination of SPATEs produced by QT598 can target multiple substrates. It will be of interest to more specifically investigate cleavage of different host substrates by the newly characterized SPATEs Sha, TagB, and TagC to try to identify their mechanisms of

interaction with host cells. Adherence to host cell lines was also increased by the production of SPATEs. In particular, TagB, TagC, and Sha increased adherence to both human and avian epithelial cells, whereas Tsh only increased adherence to bladder cells, and Vat only increased adherence to kidney cells (Figure 6). Interestingly, in addition to promoting general adherence to host cells, TagB, TagC, and Sha also mediated bacterial aggregation, suggesting a self-associating phenotype similar to the AIDA-1 autotransporter (Figure 7). Although Tsh and Vat were less effective at general adherence to different epithelial cells, these SPATEs as well as Sha were effective hemagglutinins for erythrocytes of a variety of animal species and also demonstrated increased biofilm formation, whereas production of TagB and TagC only conferred limited hemagglutination activity (Table 2) and no increase in biofilm formation (Figure 8). It is interesting to note that some class 2 SPATEs have been shown to recognize a variety of glycans on leukocyte surfaces [82]. As similar carbohydrates may be present on erythrocyte surfaces, it is not surprising that Sha as well as Tsh and Vat autotransporters demonstrated extensive hemagglutination activity for a variety of erythrocytes. It will be of interest to further determine if these SPATEs can also recognize glycosylated surface receptors on either human or avian leukocytes that may alter host immune function.

Notably, some phenotypes such as autoaggregation, hemagglutination, and biofilm formation were only present when the genes were expressed in high-copy vectors but were absent in the clinical strain. Cloning in a higher copy vector provides a means to constitutively express proteins *in vitro*, whereas these proteins or systems may not always be expressed when in a single copy or on a native plasmid in the clinical strain under *in vitro* growth conditions. Also, in the wild-type clinical strain, there can be a great deal of redundancy of function with multiple proteins (both the autotransporters and other outer membrane proteins and fimbrial adhesins that may similarly mediate hemagglutination to different erythrocytes or adherence to host cells).

Interestingly all of the SPATEs demonstrated cytopathic effects on epithelial cells. However, the class 2 SPATEs, Vat, Tsh, and Sha, only caused delayed cell death after 12 h exposure, compared to the TagB and TagC SPATEs that demonstrated cytotoxicity within 5 h of interaction (Figure 9). Importantly, loss of all 5 of the SPATE encoding genes from QT598 was required to abrogate the cytopathic effect (Figure 9). However, the “slow acting” cytopathic effect may be due to a less directed or non-specific internalization such as pinocytosis for some of the SPATEs and since this effect only occurs after long-term 12-h exposure it may indicate that the cytotoxic capacity of these SPATEs may be limited and that the

roles of such SPATEs may be more specifically linked to other protease functions such as mucinase activity or immunomodulatory roles as has been proposed as the main function of some of the Class-2 SPATEs. Similarly, a slow effect was seen in the case of EspC, when it was internalized slowly into the cells by pinocytosis, and no receptor was required for this process [83]. Further experiments are in progress to elucidate the cytopathic effects of Sha, TagB, and TagC.

Of further note, loss of all 5 SPATEs resulted in decreased fitness in the kidneys of infected mice (Figure 10), whereas loss of individual SPATEs did not have any reduction in virulence or fitness. This suggests that collectively these SPATEs can provide a selective advantage during kidney infection for QT598 in the murine UTI model. The levels of expression of the 5 SPATEs were different depending on the growth conditions – all 5 SPATEs were expressed in LB broth, mass spectrometry results have confirmed that Tsh (48% of coverage) and Vat (22% of coverage) proteins were highly expressed compared to the other SPATE proteins. In minimal M63-glycerol medium, the *vat* gene was upregulated by 10-fold compared to LB, and the *sha* gene was upregulated six-fold in infected bladder compared to culture in LB. These results indicate that SPATE-encoding genes can be subjected to environmental changes that influence their regulation. However, other than the *vat* gene [58], there is very limited information concerning identification of factors that regulate the expression of different SPATEs.

In conclusion, we investigated the role of three new SPATEs in an APEC strain, TagB, TagC, and Sha, which were present in some APEC and UPEC strains. These SPATEs may confer fitness and capabilities to infect multiple hosts. Our findings highlight the potential role of these proteins in virulence and show they significantly contribute to autoaggregation, hemagglutination, adherence to epithelial cell lines and also exhibit cytopathic effects. Furthermore, we have shown, in combination, all five ATs contribute to fitness and colonization of QT598 during urinary tract infection in the mouse model. In future studies, it will be of interest to determine mechanisms of action of the TagB, TagC, and Sha autotransporters, identify specific targets, and determine the effects of these SPATEs on host immune responses.

Materials and methods

Strains, media, and PCR screening

Strains and plasmids are listed in Table 2. The 697 *E. coli* strains isolated from urinary tract infections were clinical isolates from Guadeloupe. Strains were collected by laboratories or hospitals over a period of 17 months and included

community or nosocomial urinary tract infection isolates [61]. The 299 APEC strains were previously described elsewhere [22]. MT156 (QT598) is an APEC serogroup O1 strain that was isolated from the liver of a young turkey poult [53]. *E. coli* DH5 α , ORN172, and BL21 were used for gene cloning and phenotypic tests. Bacteria were grown at 37°C on solid or liquid Luria-Bertani medium (Alpha Bioscience, Baltimore, MD, USA) supplemented with the appropriate antibiotics when required at concentrations of 100 μ g/ml ampicillin, 30 μ g/ml chloramphenicol, or 50 μ g/ml of kanamycin. For mice infection studies, QT598 and mutant derivatives were grown in brain heart infusion broth (Alpha Bioscience). M63-glycerol minimal medium contained the following per liter: 5.3 g KH₂PO₄, 13.9 g K₂HPO₄ · 3H₂O, and 2.0 g (NH₄)₂SO₄. The pH was adjusted to 7.5 with KOH, and the medium was supplemented with 1 mM MgSO₄, 1 mM CaCl₂, 1 mM thiamine, and 0.6% (wt/vol) glycerol.

Multiplex PCR was performed to determine the prevalence of different ATs within clinical isolates using primers listed in Supplemental Table 1. PCR amplification was done in a 25 μ l reaction mixture containing 10 mM of primers, 1X of Taq FroggaMix (FroggaBio, Toronto, ON, Canada), DNA templates and deionized water when necessary. To detect *sat*, *tsh*, *tagB* and *tagC* genes, the reaction mixture was placed in a thermocycler (Eppendorf, Mississauga, ON, Canada) and set for initial denaturation at 95°C for 2 min followed by 30 cycle of denaturation at 95°C for 30 sec, annealing at 58°C for 40 sec and extension at 72°C for 30 sec. A final extension step was added at 72°C for 5 min. In order to detect *sha* gene, PCR was set for initial denaturation at 95°C for 2 min, 35 cycle of denaturation (95°C, 30 sec), annealing (56°C, 40 sec) and extension (72°C, 1 min), and a final extension step at 72°C for 5 min. To detect *vat* gene, PCR was set for initial denaturation at 95°C for 2 min, 30 cycle of denaturation (95°C, 30 sec), annealing (60°C, 40 sec) and extension (72°C, 30 sec), and a final extension step at 72°C for 5 min. The phylogenetic group of strains was determined by multiplexed PCR as described [84]. Amplified samples were separated by electrophoresis on 0.8% agarose gel stained with gel stain (Civic Bioscience, Beloeil, QC, Canada) and DNA was visualized using Gel Doc (Syngene Chemi Genius, Frederick, MD, USA) at 400 nm.

DNA and genetic manipulations

Plasmid DNA was extracted using the EZ DNA Miniprep kit (Bio Basic, Markham, ON, Canada). PCR products and DNA were purified using the EZ-10 Spin Column PCR Product Purification Kit (Bio Basic). DNA for SPATE-encoding genes was amplified using Q5 High Fidelity-

DNA polymerase (New England Biolabs, Whitby, ON, Canada).

Bioinformatic analysis

Presence of AT sequences *in silico* was determined in *E. coli* genomes available from the NCBI database by Blast analyses using the *tagB*, *tagC* and *sha* sequences from the QT598 genome. Phylogenetic analyses of the predicted full-length passenger domain sequence of each SPATE were performed by Clustal W, and the phylogenetic tree was constructed using PhyML/bootstrapping in MEGA6 [85]. The Conserved Domain Database (CDD) and SignalP were used to predict the three domains of the AT proteins [86,87]. The I-TASSER server and chimera were used for three-dimensional (3D) structure [65,66] visualization of the predicted Sha AT protein, and the Protein Data Bank (PDB) server was used to obtain the predicted 3-D structure of the Sha protein for comparison to Hbp [88]. Amino acid sequence comparisons were performed using Clone Manager Suite 7 (SciED, Denver, CO, USA) and online BLAST programs available from the National Center for Biotechnology Information (NCBI).

Construction of plasmids

AT-encoding genes were amplified by PCR using specific primers (listed in Supplemental Table 2). The *tagB*, *tagC*, *sha*, *vat*, and *tsh* genes were amplified from QT598. The *espC* gene was amplified from EPEC strain E2348/69 (Accession number AF297061) [60]. To clone *tagB*, *tagC*, *espC*, *sha*, *tsh*, and *vat*, PCR products contained 15 bp extensions homologous to the pBCsk+ multi-cloning site. Linearized pBCsk+ digested with *XhoI* and *BamHI* was used to clone inserts by fusion reaction with the Quick-fusion cloning kit (Biotool, Houston, TX, USA). All the recombinant plasmids (pIJ548, pIJ549, pIJ550, pIJ551, pIJ552, pIJ553) were first transformed into *E. coli* DH5a then the plasmids were extracted using a Miniprep kit according to the manufacturer's recommendations (Bio Basic) and transformed into *E. coli* BL21 for protein expression into culture supernatants and into the *fim*-negative *E. coli* strain ORN172 for other phenotypic assays.

Mutagenesis of SPATE-encoding genes

Mutants were generated using the lambda red recombinase method [89]. The *tagBC* genes were first replaced with a kanamycin resistance cassette, from plasmid pKD13 with knock out primers: 2094 and 2095 (Supplemental Table 1). PCR products were electroporated into QT598 containing the lambda red recombinase-expressing

plasmid pKD46. Deletion of the *tagB* and *tagC* was confirmed by PCR with screening primers (Supplemental Table 1). The kanamycin resistance cassette, flanked by FLP recombination target (FRT) sequences, was removed by the introduction of plasmid pCP20 expressing the FLP recombinase [90]. Then, the *vat* gene was replaced by a chloramphenicol resistance cassette amplified from pKD4 in the background QT598 Δ *tagBC*:FRT (QT5187) using the same approach. Similarly, kanamycin resistance cassette amplified from plasmid pKD3 was used to replace the *sha* gene in QT598 Δ *tagBC* Δ *vat*::*cat* (QT5188). Finally, the *tsh* gene was replaced with a tetracycline-resistance cassette by allelic exchange as detailed in [22] in QT598 Δ *tagBC* Δ *vat*::*cat* Δ *sha*::*kan* (QT5189) generating the five SPATE genes, Δ 5ATs mutant (QT598 Δ *tagBC* Δ *vat*::*cat* Δ *sha*::*kan* Δ *tsh*::*tetAR(B)*) (QT5182).

Protein preparation and analysis

Culture supernatants, from LB broth cultures of *E. coli* BL21 expressing AT proteins were centrifuged at 7500 \times g for 15 min at 4°C. Supernatants were filtered through 0.22 μ m filters and concentrated through 50 kDa Amicon filters (Millipore Sigma, St. Louis, MO, USA). Protein concentrations were determined using the Pierce Coomassie Plus Assay Reagent kit (Thermo Fisher Scientific, St. Laurent, QC, Canada) and proteins were visualized on SDS-PAGE by Coomassie blue as well as silver staining and identified based on the protein markers (10–200 kDa) (Bio Basic). To identify the proteins, bands were excised from denatured gels. Protein digestion by the trypsin, peptide labeling, and mass spectrometry analyses was performed by the proteomics platform of the Institut de Recherche en Immunologie et en Cancérologie (IRIC) of the Université de Montréal (Montréal, QC, Canada). The data were visualized and analyzed using Scaffold 4 Proteomics software.

Oligopeptide cleavage assays

Synthetic peptide cleavage was performed as previously described in [91] with slight modifications. Briefly, 5 μ g/ml of each SPATES was added to 200 μ l of three different pNA-conjugated oligopeptides: N-Succinyl-Ala-Ala-Ala-p-nitroanilide, N-Benzoyl-L-arginine 4-nitroanilide and N-succinyl-ala-ala-pro-phe-p-nitroanilide (Millipore Sigma) at 1 mM concentration in a buffer containing 0.2 M NaCl, 0.01 mM ZnSO₄, 0.1 M MOPS (3-(N-morpholino) propanesulfonic acid), pH 7.3 were incubated at 37°C for 3 h and absorbance readings were determined at 410 nm. Readings were normalized to the maximum absorbance of positive control. All reactions were performed in triplicate.

Autoaggregation

The autoaggregation test was carried out as described before [92]. Briefly, overnight cultures were adjusted to an OD₆₀₀ of 1.5. A volume of 10 ml of each culture was placed in sterile 20 ml glass tubes. Tubes were then vortexed for 5 s then left at 4°C for 3 h. Samples 1 cm from the top were taken, and the percentage of change in OD₆₀₀ was used as the value for autoaggregation.

Hemagglutination tests

Hemagglutination in microtiter plates was performed as described by [93] with some modifications. Human, chicken, turkey, pig, bovine, canine, rabbit, horse and sheep red blood cells (RBCs) were washed and resuspended in PBS at a final concentration of 3%. The *E. coli* fim-negative K-12 strain ORN172 expressing different SPATEs was grown overnight at 37°C in LB medium, harvested and adjusted to an optical density (O.D.₆₀₀) of 60. Suspensions were serially diluted in 96-well round bottom plates containing 20 µl of PBS mixed with 20 µl of 3% red blood cells and incubated for 30 min at 4°C.

Biofilm assay

Biofilm formation on polystyrene surfaces was assessed in 96-well plates as previously described [67]. Strains were grown at various temperatures (25°C, 30°C, 37°C, and 42°C) for 48 h under static conditions in LB medium. Wells were washed and stained with 0.1% crystal violet (Millipore Sigma) for 15 min, then 200 µl of ethanol-acetone (80:20) solution was added, followed by measuring at an optical density at OD₅₉₅.

Adherence assays

The 5637 human bladder cells (ATCC HTB-9), human embryonic kidney cells HEK-293 (ATCC® CRL-1573™), and the avian fibroblast cell line (CEC-32) were used to determine adherence of *E. coli* ORN172 expressing different SPATEs as described before [94]. The 5637 cells were maintained in RPMI 1640 (Thermo Fisher Scientific) supplemented with 10% heat-inactivated fetal bovine serum (FBS) at 37°C in 5% CO₂, and 2 × 10⁵ cells/well were seeded into 24-well cell culture plates. Cell lines were washed twice with phosphate-buffered saline (pH 7.2) and then incubated at a multiplicity of infection (MOI) of 10 at 37°C for 2 h in RPMI 1640 medium with 10% FBS. Non-adherent bacteria were removed by washing with PBS three times. Cells were then exposed to 1% Triton X-100 for 5 min, and serial dilutions were plated on LB agar plus an antibiotic. For HEK-293 cells, Eagle's Minimum Essential Medium supplemented with 10% of FBS was used. For CEC-32 cells Dulbecco's Modified

Eagle's medium supplemented with 10% of FBS was used. The adherence assays were done in triplicate for each sample.

Determination of cytopathic effects

The 5637 cells were maintained in RPMI 1640 medium (Thermo Fisher Scientific) supplemented with 10% heat-inactivated FBS at 37°C in 5% CO₂, and 2 × 10⁵ cells/well were seeded into eight-well chamber slides (Thermo Fisher Scientific) and allowed to grow to 75% confluence. Thirty µg/ml of each SPATE (final concentration) was added directly to monolayers and incubated for 5 h or 12 h at 37°C with 5% CO₂. Cells were then washed twice with PBS (phosphate-buffered saline), fixed with 70% methanol, and stained with Giemsa stain.

Lactate dehydrogenase (LDH) release assay

Culture supernatants were incubated with monolayers of 5637 cells in RPMI 1640 supplemented with 10% heat-inactivated FBS at 37°C in 5% CO₂ for up to 12 h. Release of LDH was quantified by CytoTox 96® Non-Radioactive Cytotoxicity Assay kit (Promega, Madison, WI, USA) at 5 h and 12 h. A lysis solution (provided in the kit) was added to the non-infected cells to generate the maximum LDH release control from lysed cells.

Ascending urinary tract infection in mice

For single strain infections with the wild-type parent QT598 and isogenic SPATE mutants QT598ΔtagBC and QT598Δsha, 25 µl (10⁹ CFU) were tested in an ascending UTI model adapted from [95] with 10 mice in each group. Similarly, a murine ascending UTI model with 10 mice in each group was used for co-infection, in which a virulent ΔlacZYA derivative of QT598 was co-infected with the Δ5ATs strain, a QT598-derivative lacking all 5 SPATE-encoding genes. Twenty-five µl (10⁹ CFU) of a mixed culture containing equal amounts of each strain were inoculated through a catheter in six-week-old CBA/J female mice. Mice were euthanized after 48 h, and bladders and kidneys were harvested aseptically, homogenized, diluted and plated on MacConkey agar plates. Bladder samples were frozen at -80°C in TRIzol® reagent (Thermo Fisher Scientific) for RNA extraction.

qRT-PCR analysis of SPATE gene expression in vivo and in vitro

Expression of the 5 SPATE-encoding genes was determined after growth in LB medium, minimal M63

medium, and from infected mouse bladders. Total RNAs from bacterial samples were extracted in the EZ-10 Spin Column Total RNA Miniprep Kit (Bio Basic) as described elsewhere [96]. Briefly, to extract RNA from infected bladder, samples were homogenized with TRIzol® reagent (Thermo Fisher Scientific), centrifuged for 30 sec at $12,000 \times g$, the supernatant was incubated with ethanol (95–100%) and transferred into Zymo-Spin™ IICR Column to extract RNA using Direct-zol RNA Miniprep kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's recommendations. All RNA samples were treated with TURBO DNase (Thermo Fisher Scientific). PCR (35 cycles) was used to verify DNA contamination. cDNAs were synthesized by using the Iscript™ Reverse transcription supermix according to the manufacturer's protocol (Bio-Rad Life Science, Mississauga, ON, Canada). The RNA polymerase sigma factor *rpoD* gene was used as a housekeeping control. Primers designed to amplify *rpoD*, *tagB*, *tagC*, *vat*, *tsh*, and *sha* (Supplemental Table 1) were used. For each sample, 50 ng of cDNA and 100 nM concentrations of each primer set were mixed with 10 µl of SsoFast Evagreen supermix (Bio-Rad Life Science) per well. Assays were performed in triplicate on a Corbett Rotorgene (Thermo Fisher Scientific) instrument. All data were normalized to *rpoD* expression levels. Melting-curve analysis was verified to differentiate accumulation of Evagreen-bound DNA and determine that signal was gene-amplification specific and not due to the primer-dimer formation. The data were analyzed by the $2^{-\Delta\Delta CT}$ method [97].

Statistical analyses

Experimental data were expressed as a mean \pm standard error of the mean (SEM) in each group. The means of groups were combined and analyzed by a two-tailed Student *t*-test for pairwise comparisons and analysis of variance (ANOVA) to compare means of more than two populations. For mouse infection experiments, the Mann-Whitney test was used to compare the samples by pairs, and the Kruskal-Wallis test was used to compare groups. A *P* value of <0.05 was considered statistically significant. All data were analyzed with the Graph Pad Prism 7 software (GraphPad Software, San Diego, CA, USA).

Ethics statement

The protocol for mice urinary tract infection was approved by the animal ethics evaluation committee – *Comité Institutionnel de Protection des Animaux* (CIPA No 1608–02) of the INRS-Institut Armand-Frappier.

Nucleotide accession numbers

Sections of the genome of *E. coli* strain QT598 containing the five different SPATE-encoding genes were submitted to NCBI Genbank from the analysis of whole-genome survey sequence. The corresponding accession numbers are *tsh* (MH899683), *sha* (MH899684), *vat* (MH899682), and *tagB* and *tagC* (MH899681) genetic regions.

Acknowledgments

We thank E. Bonnaie, manager at the proteomics platform of Université de Montréal for assistance with mass spectrometry and peptide analysis and Y. Lopez de Los Santos for assistance for 3-D model prediction comparison of autotransporters. Funding for this work was supported by NSERC Canada Discovery Grants 2014-06622 and 2019-06642 (C.M.D.) and a collaborative ACIP grant from the Institut Pasteur International Network (to A.T., S.G.-R., and C.M.D.).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Institut Pasteur [ACIP]; Natural Sciences and Engineering Research Council of Canada [2014-06622] and [2019-06642].

References

- [1] Croxen MA, Finlay BB. Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol.* 2010;8(1):26–38.
- [2] Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nature Rev Microbiol.* 2004;2(2):123.
- [3] Bauchart P, Germon P, Bree A, et al. Pathogenomic comparison of human extraintestinal and avian pathogenic *Escherichia coli*—search for factors involved in host specificity or zoonotic potential. *Microb Pathog.* 2010;49(3):105–115.
- [4] Bélanger L, Garenaux A, Harel J, et al. *Escherichia coli* from animal reservoirs as a potential source of human extraintestinal pathogenic *E. coli*. *FEMS Immunol Med Microbiol.* 2011;62(1):1–10.
- [5] Clermont O, Olier M, Hoede C, et al. Animal and human pathogenic *Escherichia coli* strains share common genetic backgrounds. *Infect Genet Evol.* 2011;11(3):654–662.
- [6] Ewers C, Li G, Wilking H, et al. Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: how closely related are they?. *Int J Med Microbiol.* 2007;297(3):163–176.
- [7] Moulin-Schouleur M, Schouler C, Tailliez P, et al. Common virulence factors and genetic relationships between O18: K1:H7 *Escherichia coli* isolates of human and avian origin. *J Clin Microbiol.* 2006;44(10):3484–3492.

- [8] Nandanwar N, Janssen T, Kuhl M, et al. Extraintestinal pathogenic *Escherichia coli* (ExPEC) of human and avian origin belonging to sequence type complex 95 (STC95) portray indistinguishable virulence features. *Int J Med Microbiol.* 2014;304(7):835–842.
- [9] Johnson JR, Russo TA. Extraintestinal pathogenic *Escherichia coli*: “the other bad *E.coli*”. *J Lab Clin Med.* 2002;139(3):155–162.
- [10] Guabiraba R, Schouler C. Avian colibacillosis: still many black holes. *FEMS Microbiol Lett.* 2015;362(15):fnv118.
- [11] Dho-Moulin M, Fairbrother JM. Avian pathogenic *Escherichia coli* (APEC). *Vet Res.* 1999;30(2–3):299–316.
- [12] Dziva F, Stevens MP. Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic *Escherichia coli* in their natural hosts. *Avian Pathol.* 2008;37(4):355–366.
- [13] Dziva F, Hauser H, Connor TR, et al. Sequencing and functional annotation of avian pathogenic *Escherichia coli* serogroup O78 strains reveal the evolution of *E.coli* lineages pathogenic for poultry via distinct mechanisms. *Infect Immun.* 2013;81(3):838–849.
- [14] Johnson TJ, Kariyawasam S, Wannemuehler Y, et al. The genome sequence of avian pathogenic *Escherichia coli* strain O1: K1: H7 shares strong similarities with human extraintestinal pathogenic *E.coli* genomes. *J Bacteriol.* 2007;189(8):3228–3236.
- [15] Johnson TJ, Nolan LK. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol Mol Biol Rev.* 2009;73(4):750–774.
- [16] Johnson TJ, Wannemuehler YM, Johnson SJ, et al. Plasmid replicon typing of commensal and pathogenic *Escherichia coli* isolates. *Appl Environ Microbiol.* 2007;73(6):1976–1983.
- [17] Mellata M, Touchman JW, Curtiss R. Full sequence and comparative analysis of the plasmid pAPEC-1 of avian pathogenic *E.coli* chi7122 (O78: K80:H9). *PLoS One.* 2009;4(1):e4232.
- [18] Peigne C, Bidet P, Mahjoub-Messai F, et al. The plasmid of *Escherichia coli* strain S88 (O45: K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E.coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. *Infect Immun.* 2009;77(6):2272–2284.
- [19] Cordoni G, Woodward MJ, Wu H, et al. Comparative genomics of European avian pathogenic *E.coli* (APEC). *BMC Genomics.* 2016;17(1):960.
- [20] Leimbach A, Hacker J, Dobrindt U. *E.coli* as an all-rounder: the thin line between commensalism and pathogenicity. *Curr Top Microbiol Immunol.* 2013;358:3–32.
- [21] Mokady D, Gophna U, Ron EZ. Extensive gene diversity in septicemic *Escherichia coli* strains. *J Clin Microbiol.* 2005;43(1):66–73.
- [22] Dozois CM, Dho-Moulin M, Bree A, et al. Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the Tsh genetic region. *Infect Immun.* 2000;68(7):4145–4154.
- [23] Dozois CM, Daigle F, Curtiss R 3rd. Identification of pathogen-specific and conserved genes expressed in vivo by an avian pathogenic *Escherichia coli* strain. *Proc Natl Acad Sci U S A.* 2003;100(1):247–252.
- [24] Johnson TJ, Johnson SJ, Nolan LK. Complete DNA sequence of a ColBM plasmid from avian pathogenic *Escherichia coli* suggests that it evolved from closely related ColV virulence plasmids. *J Bacteriol.* 2006;188(16):5975–5983.
- [25] Krishnan S, Chang AC, Hodges J, et al. Serotype O18 avian pathogenic and neonatal meningitis *Escherichia coli* strains employ similar pathogenic strategies for the onset of meningitis. *Virulence.* 2015;6(8):777–786.
- [26] Johnson TJ, Wannemuehler Y, Johnson SJ, et al. Comparison of extraintestinal pathogenic *Escherichia coli* strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. *Appl Environ Microbiol.* 2008;74(22):7043–7050.
- [27] Mellata M. Human and avian extraintestinal pathogenic *Escherichia coli*: infections, zoonotic risks, and antibiotic resistance trends. *Foodborne Pathog Dis.* 2013;10(11):916–932.
- [28] Rodriguez-Siek KE, Giddings CW, Doetkott C, et al. Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. *Microbiology.* 2005;151(Pt 6):2097–2110.
- [29] Henderson IR, Navarro-García F, Nataro JP. The great escape: structure and function of the autotransporter proteins. *Trends Microbiol.* 1998;6(9):370–378.
- [30] Klemm P, Vejborg RM, Sherlock O. Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol.* 2006;296(4–5):187–195.
- [31] Wells TJ, Totsika M, Schembri MA. Autotransporters of *Escherichia coli*: a sequence-based characterization. *Microbiology.* 2010;156(8):2459–2469.
- [32] Ruiz-Perez F, Nataro JP. Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cell Mol Life Sci.* 2014;71(5):745–770.
- [33] Albenne C, Ieva R. Job contenders: roles of the beta-barrel assembly machinery and the translocation and assembly module in autotransporter secretion. *Mol Microbiol.* 2017;106(4):505–517.
- [34] Bernstein HD. Looks can be deceiving: recent insights into the mechanism of protein secretion by the autotransporter pathway. *Mol Microbiol.* 2015;97(2):205–215.
- [35] Henderson IR, Navarro-García F, Desvaux M, et al. Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev.* 2004;68(4):692–744.
- [36] Otto BR, Van Dooren SJ, Nuijens JH, et al. Characterization of a hemoglobin protease secreted by the pathogenic *Escherichia coli* strain EB1. *J Exp Med.* 1998;188(6):1091–1103.
- [37] Henderson IR, Czczulin J, Eslava C, et al. Characterization of Pic, a Secreted Protease of *Shigella flexneri* and Enterotoxigenic *Escherichia coli*. *Infect Immun.* 1999;67(11):5587–5596.
- [38] Brunder W, Schmidt H, Karch H. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157: H7 cleaves human coagulation factor V. *Mol Microbiol.* 1997;24(4):767–778.
- [39] Eslava C, Navarro-García F, Czczulin JR, et al. Pet, an autotransporter enterotoxin from enterotoxigenic *Escherichia coli*. *Infect Immun.* 1998;66(7):3155–3163.

- [40] Guyer DM, Henderson IR, Nataro JP, et al. Identification of sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol Microbiol.* 2000;38(1):53–66.
- [41] Parham NJ, Srinivasan U, Desvaux M, et al. PicU, a second serine protease autotransporter of uropathogenic *Escherichia coli*. *FEMS Microbiol Lett.* 2004;230(1):73–83.
- [42] Stein M, Kenny B, Stein MA, et al. Characterization of EspC, a 110-kilodalton protein secreted by enteropathogenic *Escherichia coli* which is homologous to members of the immunoglobulin A protease-like family of secreted proteins. *J Bacteriol.* 1996;178(22):6546–6554.
- [43] Subashchandrabose S, Smith SN, Spurbeck RR, et al. Genome-wide detection of fitness genes in uropathogenic *Escherichia coli* during systemic infection. *PLoS Pathog.* 2013;9(12):e1003788.
- [44] Guyer DM, Radulovic S, Jones F-E, et al. Sat, the secreted autotransporter toxin of uropathogenic *Escherichia coli*, is a vacuolating cytotoxin for bladder and kidney epithelial cells. *Infect Immun.* 2002;70(8):4539–4546.
- [45] Parreira VR, Gyles CL. A novel pathogenicity island integrated adjacent to the *thrW* tRNA gene of avian pathogenic *Escherichia coli* encodes a vacuolating autotransporter toxin. *Infect Immun.* 2003;71(9):5087–5096.
- [46] Parham NJ, Pollard SJ, Desvaux M, et al. Distribution of the serine protease autotransporters of the Enterobacteriaceae among extraintestinal clinical isolates of *Escherichia coli*. *J Clin Microbiol.* 2005;43(8):4076–4082.
- [47] Restieri C, Garriss G, Locas MC, et al. Autotransporter-encoding sequences are phylogenetically distributed among *Escherichia coli* clinical isolates and reference strains. *Appl Environ Microbiol.* 2007;73(5):1553–1562.
- [48] Spurbeck RR, Dinh PC Jr., Walk ST, et al. *Escherichia coli* isolates that carry *vat*, *fyuA*, *chuaA*, and *yfcV* efficiently colonize the urinary tract. *Infect Immun.* 2012;80(12):4115–4122.
- [49] Provence DL, Curtiss R. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic *Escherichia coli* strain. *Infect Immun.* 1994;62(4):1369–1380.
- [50] Cyoia PS, Rodrigues GR, Nishio EK, et al. Presence of virulence genes and pathogenicity islands in extraintestinal pathogenic *Escherichia coli* isolates from Brazil. *J Infect Dev Ctries.* 2015;9(10):1068–1075.
- [51] Maluta RP, Logue CM, Casas MR, et al. Overlapped sequence types (STs) and serogroups of avian pathogenic (APEC) and human extra-intestinal pathogenic (ExPEC) *Escherichia coli* isolated in Brazil. *PLoS One.* 2014;9(8):e105016.
- [52] Otto BR, van Dooren SJ, Dozois CM, et al. *Escherichia coli* hemoglobin protease autotransporter contributes to synergistic abscess formation and heme-dependent growth of *Bacteroides fragilis*. *Infect Immun.* 2002;70(1):5–10.
- [53] Marc D, Dho-Moulin M. Analysis of the *fim* cluster of an avian O2 strain of *Escherichia coli*: serogroup-specific sites within *fimA* and nucleotide sequence of *fimI*. *J Med Microbiol.* 1996;44(6):444–452.
- [54] Dozois CM, Fairbrother JM, Harel J, et al. *pap*- and *pil*-related DNA sequences and other virulence determinants associated with *Escherichia coli* isolated from septicemic chickens and turkeys. *Infect Immun.* 1992;60(7):2648–2656.
- [55] Dobrindt U, Agerer F, Michaelis K, et al. Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J Bacteriol.* 2003;185(6):1831–1840.
- [56] Wang X, Wei L, Wang B, et al. Complete genome sequence and characterization of avian pathogenic *Escherichia coli* field isolate ACN001. *Stand Genomic Sci.* 2016;11:13.
- [57] Kostakioti M, Stathopoulos C. Role of the alpha-helical linker of the C-terminal translocator in the biogenesis of the serine protease subfamily of autotransporters. *Infect Immun.* 2006;74(9):4961–4969.
- [58] Nichols KB, Totsika M, Moriel DG, et al. Molecular characterization of the vacuolating autotransporter Toxin in Uropathogenic *Escherichia coli*. *J Bacteriol.* 2016;198(10):1487–1498.
- [59] Sandt CH, Hill CW. Four Different Genes Responsible for Nonimmune Immunoglobulin-Binding Activities within a Single Strain of *Escherichia coli*. *Infect Immun.* 2000;68(4):2205–2214.
- [60] Mellies JL, Navarro-Garcia F, Okeke I, et al. *espC* pathogenicity island of enteropathogenic *Escherichia coli* encodes an enterotoxin. *Infect Immun.* 2001;69(1):315–324.
- [61] Guyomard-Rabenirina S, Malespine J, Ducat C, et al. Temporal trends and risks factors for antimicrobial resistant Enterobacteriaceae urinary isolates from outpatients in Guadeloupe. *BMC Microbiol.* 2016;16(1):121.
- [62] Dutta PR, Cappello R, Navarro-García F, et al. Functional comparison of serine protease autotransporters of Enterobacteriaceae. *Infect Immun.* 2002;70(12):7105–7113.
- [63] Stathopoulos C, Provence DL, Curtiss R. Characterization of the Avian Pathogenic *Escherichia coli* Hemagglutinin Tsh, a Member of the Immunoglobulin A Protease-Type Family of Autotransporters. *Infect Immun.* 1999;67(2):772–781.
- [64] Gutierrez D, Pardo M, Montero D, et al. TleA, a Tsh-like autotransporter identified in a human enterotoxigenic *Escherichia coli* strain. *Infect Immun.* 2015;83(5):1893–1903.
- [65] Yang J, Yan R, Roy A, et al. The I-TASSER Suite: protein structure and function prediction. *Nat Methods.* 2015;12(1):7.
- [66] Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605–1612.
- [67] Genevaux P, Muller S, Bauda P. A rapid screening procedure to identify mini-Tn10 insertion mutants of *Escherichia coli* K-12 with altered adhesion properties. *FEMS Microbiol Lett.* 1996;142(1):27–30.
- [68] Kostakioti M, Stathopoulos C. Functional analysis of the Tsh autotransporter from an avian pathogenic *Escherichia coli* strain. *Infect Immun.* 2004;72(10):5548–5554.
- [69] Sherlock O, Vejborg RM, Klemm P. The TibA adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. *Infect Immun.* 2005;73(4):1954–1963.
- [70] Hasman H, Chakraborty T, Klemm P. Antigen-43-mediated autoaggregation of *Escherichia coli* is blocked by fimbriation. *J Bacteriol.* 1999;181(16):4834–4841.

- [71] Benz I, Schmidt MA. Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic *Escherichia coli*. *Infect Immun*. 1989;57(5):1506–1511.
- [72] Al-Hasani K, Henderson IR, Sakellaris H, et al. The *sigA* gene which is borne on the *she* pathogenicity island of *Shigella flexneri* 2a encodes an exported cytopathic protease involved in intestinal fluid accumulation. *Infect Immun*. 2000;68(5):2457–2463.
- [73] Benjelloun-Touimi Z, Sansonetti PJ, Parsot C. SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Mol Microbiol*. 1995;17(1):123–135.
- [74] Vijayakumar V, Santiago A, Smith R, et al. Role of class 1 serine protease autotransporter in the pathogenesis of *Citrobacter rodentium* colitis. *Infect Immun*. 2014;82(6):2626–2636.
- [75] Leyton DL, Adams LM, Kelly M, et al. Contribution of a novel gene, *rpeA*, encoding a putative autotransporter adhesin to intestinal colonization by rabbit-specific enteropathogenic *Escherichia coli*. *Infect Immun*. 2007;75(9):4664–4669.
- [76] Johnson TJ, Aziz M, Liu CM, et al. Complete genome sequence of a CTX-M-15-producing *Escherichia coli* strain from the H30Rx subclone of sequence type 131 from a patient with recurrent urinary tract infections, closely related to a lethal urosepsis isolate from the patient's sister. *Genome Announc*. 2016;4(3):e00334–16.
- [77] Johnson TJ, Siek KE, Johnson SJ, et al. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. *J Bacteriol*. 2006;188(2):745–758.
- [78] Jakobsen L, Hammerum AM, Frimodt-Møller N. Virulence of *Escherichia coli* B2 isolates from meat and animals in a murine model of ascending urinary tract infection (UTI): evidence that UTI is a zoonosis. *J Clin Microbiol*. 2010;48(8):2978–2980.
- [79] Mellata M, Johnson J, Curtiss III R. *Escherichia coli* isolates from commercial chicken meat and eggs cause sepsis, meningitis and urinary tract infection in rodent models of human infections. *Zoonoses Public Health*. 2018;65(1):103–113.
- [80] Pavanelo DB, Houle S, Matter LB, et al. The periplasmic trehalase affects type 1 fimbriae production and virulence of the extraintestinal pathogenic *E.coli* strain MT78. *Infect Immun*. 2018;IAI:00241.
- [81] Skyberg JA, Johnson TJ, Johnson JR, et al. Acquisition of avian pathogenic *Escherichia coli* plasmids by a commensal *E.coli* isolate enhances its abilities to kill chicken embryos, grow in human urine, and colonize the murine kidney. *Infect Immun*. 2006;74(11):6287–6292.
- [82] Ruiz-Perez F, Wahid R, Faherty CS, et al. Serine protease autotransporters from *Shigella flexneri* and pathogenic *Escherichia coli* target a broad range of leukocyte glycoproteins. *Proc Natl Acad Sci*. 2011;108(31):12881–12886.
- [83] Vidal JE, Navarro-García F. Efficient translocation of EspC into epithelial cells depends on enteropathogenic *Escherichia coli* and host cell contact. *Infect Immun*. 2006;74(4):2293–2303.
- [84] Clermont O, Christenson JK, Denamur E, et al. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylogroups. *Environ Microbiol Rep*. 2013;5(1):58–65.
- [85] Tamura K, Stecher G, Peterson D, et al. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30(12):2725–2729.
- [86] Bendtsen JD, Nielsen H, von Heijne G, et al. Improved prediction of signal peptides: signalP 3.0. *J Mol Biol*. 2004;340(4):783–795.
- [87] Marchler-Bauer A, Lu S, Anderson JB, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res*. 2010;39(suppl_1):D225–D9.
- [88] Kouranov A, Xie L, de la Cruz J, et al. The RCSB PDB information portal for structural genomics. *Nucleic Acids Res*. 2006;34(suppl_1):D302–D5.
- [89] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci*. 2000;97(12):6640–6645.
- [90] Cherepanov PP, Wackernagel W. Gene disruption in *Escherichia coli*: Tc R and Km R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*. 1995;158(1):9–14.
- [91] Benjelloun-Touimi Z, Tahar MS, Montecucco C, et al. SepA, the 110 kDa protein secreted by *Shigella flexneri*: two-domain structure and proteolytic activity. *Microbiology*. 1998;144(7):1815–1822.
- [92] Charbonneau M-È, Berthiaume F, Mourez M. Proteolytic processing is not essential for multiple functions of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence (AIDA-I). *J Bacteriol*. 2006;188(24):8504–8512.
- [93] Provence D, Curtiss R. Role of *crl* in avian pathogenic *Escherichia coli*: a knockout mutation of *crl* does not affect hemagglutination activity, fibronectin binding, or curli production. *Infect Immun*. 1992;60(11):4460–4467.
- [94] Matter LB, Barbieri NL, Nordhoff M, et al. Avian pathogenic *Escherichia coli* MT78 invades chicken fibroblasts. *Vet Microbiol*. 2011;148(1):51–59.
- [95] Hagberg L, Engberg I, Freter R, et al. Ascending, unobstructed urinary tract infection in mice caused by pyelonephritogenic *Escherichia coli* of human origin. *Infect Immun*. 1983;40(1):273–283.
- [96] Porcheron G, Habib R, Houle S, et al. The small RNA RyhB contributes to siderophore production and virulence of uropathogenic *Escherichia coli*. *Infect Immun*. 2014;82(12):5056–5068.
- [97] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻ $\Delta\Delta$ CT method. *Methods*. 2001;25(4):402–408.
- [98] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4(4):406–425.
- [99] Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics*. 1992;8(3):275–282.
- [100] Labbate M, Queck SY, Koh KS, et al. Quorum sensing-controlled biofilm development in *Serratia liquefaciens* MG1. *J Bacteriol*. 2004;186(3):692–698.
- [101] Woodall LD, Russell PW, Harris SL, et al. Rapid, synchronous, and stable induction of type 1 piliation in *Escherichia coli* by using a chromosomal *lacUV5* promoter. *J Bacteriol*. 1993;175(9):2770–2778.

DEUXIÈME ARTICLE

6 DEUXIÈME ARTICLE

A newly identified group of P-like (PL) fimbriae from extra-intestinal pathogenic *Escherichia coli* (ExPEC) encode distinct adhesin subunits and mediate adherence to host cells.

Auteurs : Hajer Habouria^{1,2}, Hicham Bessaiah^{1,2}, Pravil Pokharel^{1,2}, Sabin Dhakal^{1,2}, Ségolène Maris^{1,2}, Julie Buron^{1,2}, Sébastien Houle^{1,2}, and Charles M. Dozois^{1,2}

1INRS-Centre Armand-Frappier Santé Biotechnologie, Laval, Québec, Canada

2CRIPA-Centre de recherche en infectiologie porcine et avicole, Saint-Hyacinthe, Québec, Canada

Objectif de l'article :

Ce travail de recherche a été publié dans le cadre de la deuxième partie de mon projet de doctorat. Celui-ci a consisté en la caractérisation des fimbriae PL chez les ExPEC. Dans cet article, nous avons mis en évidence que ce fimbriae partage une homologie de séquence avec les fimbriae P, Pix et Sfp. Nous avons également montré qu'il est constitué d'un opéron de dix gènes avec un gène de régulation (*plfB*) et neufs gènes qui codent pour des protéines de structure et d'assemblage du fimbriae (*plfA*, *plfH*, *plfC*, *plfD*, *plfJ*, *plfK*, *plfE*, *plfF* et *plfG*). Par ailleurs, nous avons aussi montré qu'il existe cinq allèles différents de l'adhésine *plfG* avec PlfG classe I et classe II qui sont prédominants parmi une variété des souches ExPEC testées. Ensuite, nous avons mis en évidence le rôle des fimbriae PL dans la virulence par une série des tests phénotypiques. Nos résultats montrent que le clone porteur de l'adhésine PlfG II est capable d'agglutiner uniquement les érythrocytes de dinde et les érythrocytes humains, de former un biofilm à différentes températures et d'adhérer aux cellules rénales et de la vessie. D'autre part, le clone porteur de l'adhésine PlfG I est capable d'agglutiner les érythrocytes de 8 espèces différentes, de former un biofilm à des températures qui ne dépasse pas 37°C et d'adhérer aux cellules rénales et de vessie. Au cours des essais *in vivo*, nous avons pu montrer que l'absence de l'opéron *plf* n'a pas diminué le niveau de colonisation de la souche QT598 dans les vessies et les reins des souris infectées. Toutefois, dans le modèle de co-infection des souris, la capacité compétitive du mutant était significativement moins importante au niveau des reins de souris comparativement à la souche sauvage. Cela peut suggérer le potentiel rôle des fimbriae PL dans les infections systémiques comme les infections des voies urinaires.

Contribution des auteurs :

En tant que premier auteur de cet article, j'ai effectué la totalité des expérimentations à savoir (clonage, mutation, tests *in vitro* et *in vivo*), j'ai procédé à l'analyse l'ensemble des résultats, préparé les figures et écrit le manuscrit. Celui-ci a été corrigé par mon directeur de recherche et par mes collaborateurs. J'ai également coordonné et assemblé les corrections de mon directeur de recherche et de mes collaborateurs.

Hicham Bessaiah : a contribué aux infections de souris et à l'interprétation des résultats.

Pravil Pokharel et *Sabin Dhakal* : ont contribué à la correction du manuscrit. Ils ont aussi participé aux infections de souris et à l'interprétation des résultats

Ségolène Maris : a participé au clonage de l'opéron *plf* de la souche UMEA 3703-1.

Sébastien Houle et *Julie Buron* : ont participé aux infections de souris et à l'interprétation des résultats.

Professeur Dozois Charles M. : a contribué à la planification, à la conception de l'étude, à l'analyse des résultats, à la révision, à la correction et enfin à la soumission du manuscrit au journal.

A newly identified group of P-like (PL) fimbriae from extra-intestinal pathogenic *Escherichia coli* (ExPEC) encode distinct adhesin subunits and mediate adherence to host cells.

Auteurs : Hajer Habouria^{1,2}, Hicham Bessaiah^{1,2}, Pravil Pokharel^{1,2}, Sabin Dhakal^{1,2}, Ségolène Maris^{1,2}, Julie Buron^{1,2}, Sébastien Houle^{1,2}, and Charles M. Dozois^{1,2}

1INRS-Centre Armand-Frappier Santé Biotechnologie, Laval, Québec, Canada

2CRIPA-Centre de recherche en infectiologie porcine et avicole, Saint-Hyacinthe, Québec, Canada

Nom du journal : Applied and Environmental Microbiology.

Volume 33 – Issue 13 | Reçu 19 Juillet 2021, Accepté 03 Juin 2022, Publié en ligne : 27 juin 2022.

DOI : <https://doi.org/10.1128/aem.01421-21>

Résumé en français

Les adhésines fimbriaires jouent un rôle essentiel dans l'adhérence bactérienne et la formation du biofilm. Le séquençage de la souche d'*Escherichia coli* pathogène aviaire (APEC) QT598 a révélé la présence d'un opéron fimbriaire appartenant au clade des π -fimbriae, que nous avons nommé fimbriae PL (P-like) puisqu'il présente des similarités avec les fimbriae P. Les analyses bio-informatiques ont montré que les gènes qui codent pour les fimbriae PL se trouvent sur des plasmides IncF. Ces plasmides sont présents chez les souches APEC, les souches responsables d'infections systémiques humaines et d'autres souches environnementales. Comparativement aux fimbriae P, il existe aussi une diversité génétique au niveau des gènes de l'adhésine PifG des fimbriae PL, ce qui nous a permis de les classer en cinq classes distinctes. Les gènes qui codent pour les fimbriae PL porteurs des adhésines PifG classe I et classe II ont été clonés dans une bactérie *E. coli* K-12. Les deux clones étaient capables d'exprimer des fimbriae à leur surface, d'agglutiner différents types d'érythrocytes, de former un biofilm et d'adhérer aux cellules rénales humaines et de la vessie. De plus, nous avons essayé de créer des clones hybrides composés d'un polymère de sous-unités majeures et mineures du fimbriae PL de QT598 (PifA, PifK, PifE et PifF) et au sommet, on trouve l'adhésine PapG classe I ou PapG classe II ou PapG classe III ou PifG_{UMEA-3703-1} classe I. Ces clones hybrides étaient également fonctionnels et étaient capables d'adhérer aux cellules épithéliales humaines (rénales et de la vessie). Cela montre que les fimbriae PL partagent une ressemblance et une compatibilité avec les fimbriae P. La délétion de l'opéron *pif* n'a pas réduit la colonisation des souches QT598 et UMEA-3703-1 au niveau des voies urinaires lors des expériences de mono-infection. Toutefois, la capacité compétitive du mutant *pif* de la souche QT598 était significativement moins importante au niveau des reins de souris CBA/J. Encore plus, l'expression de l'opéron *pif* de la souche QT598 était 40 fois plus dans les vessies infectées que dans le milieu de culture *in vitro*. Dans l'ensemble, les fimbriae PL représentent un nouveau groupe de fimbriae possédant des différences et des similitudes avec le fimbriae P. De plus, ils peuvent contribuer à l'augmentation de l'adhérence aux cellules de l'hôte et à la colonisation compétitive au niveau des reins de souris infectées.

Mots-clés : Fimbriae ; *E. coli* pathogènes ; Biofilm ; Adhérence ; infection de tractus urinaire ; Volaille.



A Newly Identified Group of P-like (PL) Fimbria Genes from Extraintestinal Pathogenic *Escherichia coli* (ExPEC) Encode Distinct Adhesin Subunits and Mediate Adherence to Host Cells

Hajer Habouria,^{a,b} Hicham Bessaiah,^{a,b} Pravil Pokharel,^{a,b} Sabin Dhakal,^{a,b} Ségolène Maris,^{a,b} Julie Buron,^{a,b} Sébastien Houle,^{a,b}
 Charles M. Dozois^{a,b}

^aINRS-Centre Armand-Frappier Santé Biotechnologie, Laval, Québec, Canada

^bCentre de Recherche en Infectiologie Porcine et Avicole (CRIPA), Saint-Hyacinthe, Québec, Canada

ABSTRACT Fimbrial adhesins promote bacterial adherence and biofilm formation. Sequencing of avian pathogenic *Escherichia coli* (APEC) strain QT598 identified new fimbriae belonging to the π group, which we named PL (P-like) fimbriae since the genetic organization and sequence are similar to those of P and related fimbriae. Genes encoding PL fimbriae located on IncF plasmids are present in diverse *E. coli* isolates from poultry, human systemic infections, and other sources. As with P fimbriae, PL fimbriae exhibit divergence in adhesin-encoding genes and could be divided into 5 classes based on sequence differences in the PlfG adhesin. *plf* genes from two predominant PlfG adhesin classes, PlfG class I (PlfGI) and PlfGI1, were cloned. PL fimbriae were visualized by electron microscopy, associated with increased biofilm, demonstrated distinct hemagglutination profiles, and promoted adherence to human bladder and kidney epithelial cells. The genes encoding hybrid fimbriae were comprised of genes from *plf*_{QT598}, wherein *plfG* was replaced by *papG*; the adhesin-encoding genes were also functional and mediated adherence to epithelial cells, demonstrating compatibility between the components of these two types of fimbriae. Deletion of *plf* genes did not reduce colonization of the mouse urinary tract in a single-strain infection model. In contrast, loss of *plf* genes significantly reduced competitive colonization in the mouse kidneys. Furthermore, *plf* gene expression was increased over 40-fold in the bladder compared to during *in vitro* culture. Overall, PL fimbriae represent a new group of fimbriae demonstrating both functional differences from and similarities to P fimbriae, which mediated adherence to host cells and improved competitive colonization of the mouse kidney.

IMPORTANCE Fimbriae are important colonization factors in many bacterial species. The identification of a new type of fimbriae encoded on some IncF plasmids in *E. coli* was investigated. Genomic sequences demonstrated these fimbrial gene clusters have genetic diversity, particularly in the adhesin-encoding *plfG* gene. Functional studies demonstrated differences in hemagglutination specificity, although both types of Plf adhesin under study mediated adherence to human urinary epithelial cells. A *plf* mutant also showed decreased colonization of the kidneys in a mouse competitive infection model. PL fimbriae may represent previously unrecognized adhesins that could contribute to host specificity and tissue tropism of some *E. coli* strains.

KEYWORDS fimbria, pili, pathogenic *E. coli*, hemagglutination, biofilm, adherence, urinary tract infection, poultry, adhesins, avian, extraintestinal diseases, fimbriae, pilus

Bacterial adherence to surfaces is an important survival mechanism. Attachment to host cells or extracellular matrix can provide access to specific niches and promote colonization of host tissues. Adhesins can also mediate biofilm formation through

Editor Christopher A. Elkins, Centers for Disease Control and Prevention

Copyright © 2022 American Society for Microbiology. All Rights Reserved.

Address correspondence to Charles M. Dozois, charles.dozois@inrs.ca.

The authors declare no conflict of interest.

Received 19 July 2021

Accepted 3 June 2022

bacterium-bacterium associations and improve survival in the environment. Bacterial adhesins include hair-like appendages (fimbriae or pili) as well as other molecules, including proteins or polysaccharides, which are displayed on the cell surface (1). Many types of fimbriae (pili) in Gram-negative bacteria are assembled by the chaperone/usher pathway (CUP) (2, 3). In *Escherichia coli*, numerous types of CUP fimbriae have been identified and characterized. Pathogenic *E. coli* often can produce multiple types of fimbrial adhesins, and genomic analyses indicate that some strains may contain 10 or more fimbrial gene clusters (3, 4). The ability to produce a variety of adhesins can provide a fitness advantage by expanding potential host receptor targets or promoting adherence to environmental substrates.

Two main types of fimbriae, type 1 and P fimbriae, from *E. coli* have been extensively investigated to determine aspects of their roles in disease, particularly urinary tract infections (UTIs) (5–7), as well as molecular aspects of biogenesis and assembly of these structures (8). Type 1 fimbriae mediate adherence to mannose-containing receptors and have been shown to be critical for virulence of extraintestinal pathogenic *E. coli* (ExPEC), including *E. coli* strains causing urinary tract infections (9) and neonatal meningitis (10). P fimbriae were first described in uropathogenic *E. coli* (UPEC) and were named based on receptor affinities for P blood group oligosaccharides, and they were also described as pyelonephritis-associated pili (Pap), since these fimbriae were more associated with *E. coli* strains from cases of pyelonephritis (11). P fimbriae have also been identified in other ExPEC isolates, including *E. coli* associated with systemic infections in swine (12) and some strains of avian pathogenic *E. coli* (APEC) (13–16). The P fimbrial gene clusters are commonly located on horizontally acquired chromosomal regions, which have been termed pathogenicity islands (14, 17, 18).

The P fimbrial gene cluster comprises 11 genes, including regulatory genes (*papI* and *papB*) and genes dedicated to fimbrial assembly and structure (*papAHCDKJFEG*). The *papA* gene encodes the major fimbrial subunit and has been used to class P fimbriae into serological variants (F₇₁ and F₇₂ through F16) (19). The adhesin specificity of P fimbriae is mediated by the *papG* gene product. The G adhesins of P fimbriae were grouped into 3 major classes based on sequence differences and receptor specificity to different Gal(α1-4)Gal-containing glycolipids (20, 21). PapG class I (PapGI) adhesins recognize globotriaosylceramide (Gb3), PapGII adhesins recognize globotetraosylceramide (Gb4), and PapGIII or PrsGIII adhesins recognize galactosylgloboside (Gb5). These glycolipid receptors are usually found on the surface of red blood cells and on human bladder and kidney cells. A distinct variant allele, *papG*_{F31}, which was termed class IV, was also reported, although receptor specificity for this fimbrial adhesin was not described (22).

P fimbriae are the archetype representatives of the π fimbrial family (2), which includes a number of other types of *E. coli* fimbriae, including Pix fimbriae present in some UPEC strains (23, 24) and Sfp fimbriae encoded on plasmids in some lineages of enterohemorrhagic *E. coli* (EHEC) (25, 26). This report describes a new type of *E. coli* fimbriae from the π group that we have named P-like (PL) fimbriae, since they share sequence similarity and genetic organization with P fimbriae. The PL fimbriae are distinct from other known members of the π fimbriae and are encoded on IncFIB plasmids containing numerous other virulence genes associated with ExPEC and APEC strains. As with P fimbriae (20), PL fimbriae have also diversified into a number of different G adhesin classes and major subunit variants, suggesting adaptive potential for host specificity and tissue tropism. Herein, we characterized two different types of PL fimbria genes encoding distinct G adhesins that were cloned from avian pathogenic *E. coli* strain QT598 and a urinary tract pathogenic *E. coli* (UPEC) strain, UMEA3703-1, and demonstrate these fimbriae can mediate adherence to host epithelial cells.

RESULTS

Genomic analysis identifies a new type of fimbria with a genetic organization similar to P fimbriae. Previously, we reported that ExPEC strain QT598, originally isolated from an infected turkey, contains a large ColV-type virulence plasmid, pEC598,

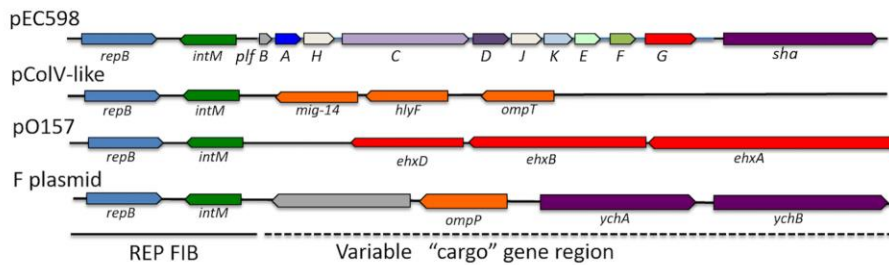


FIG 1 The P-like fimbrial (*plf*) gene cluster of plasmid pEC-598-1 is located adjacent to the RepFIB region. Comparison of the cargo genes adjacent to the RepFIB region on F and related conjugative plasmids. In other ColV-like plasmids, such as pAPEC-1 or pVM1, the cargo region contains *mig-14*, *hlyF*, and *ompT* virulence genes. In pO157 from *E. coli* O157:H7 strains, the genes encoding the Ehx hemolysin RTX toxin are within the cargo region. The F plasmid cargo gene region also contains virulence-associated genes, a gene coding for a protease, OmpP, that can degrade host defense peptides, and genes encoding self-associating AIDA-1-like autotransporters YchA and YchB. NCBI accession numbers: pEC598, [NZ_KP119165.1](#); "pColV-like" pAPEC-1, [CP000836](#); pVM01, [NC_010409.1](#); pO157, [AB011549](#); F plasmid, [NC_002483.1](#).

that encodes a novel autotransporter protein, the serine-protease hemagglutinin autotransporter, Sha (27). The region adjacent to the *sha* gene on pEC598 contains a fimbrial gene cluster (Fig. 1). Due to the close protein identity and genetic organization with P fimbriae (see below), we have named these genes *plf* (for P-like fimbriae) and called these adhesins PL fimbriae. The *plf* gene cluster on the pEC598 plasmid is inserted beside a site-specific integrase gene located adjacent to the RepFIB region, which is one of the known replicons of IncF plasmids (Fig. 1). The RepFIB region and *intM* integrase genes are also present in most IncFII plasmids and are also commonly flanked by other predicted virulence genes, such as *mig-14*, *hlyF*, and *ompT*, in other APEC virulence plasmids, suggesting this conserved region may have led to acquisition of different genes through integration/recombination. Interestingly, this region of F and related plasmids is considered a "hot spot" for insertion of diverse virulence and antibiotic resistance genes, which have been termed cargo genes (Fig. 1) (28).

Fimbriae have been classed into specific groups by a variety of criteria, including comparison of usher, chaperone, and major fimbrial subunits (2, 29). Fimbria classification using the usher-encoding protein sequences has placed P fimbriae within the π fimbrial clade (2). Phylogenetic analysis using the usher proteins indicated that the PL fimbriae also belong to the π fimbrial clade and cluster with Pix, Sfp, and Pap fimbriae (Fig. 2A). Phylogenetic comparison based on the chaperone proteins also indicated PL fimbriae were most closely related to P, Pix, and Sfp fimbriae (see Fig. S1 in the supplemental material).

Girardeau et al. (29) also classified fimbriae based on amino acid sequence motifs within the major subunit proteins. The subfamily Ic (PapA-like) group, which included PapA variants, SmfA (*Serratia marcescens*), PmpA (*Proteus mirabilis*), and MrpA (*Proteus mirabilis*), also includes PixA, SfpA, and PlfA major subunit proteins that share a conserved sequence signature motif in segment S1 of the fimbrial subunits: GxG[KT]V[TS]FxG[TS]V[V]DAP (Fig. 2B).

The PL fimbriae (*plf*) gene cluster contains 10 genes predicted to encode one regulatory protein and 9 structural/assembly proteins that share identity to equivalent proteins present in the *pap* gene cluster (Fig. 3). A predicted regulatory protein, PlfB, shares identity with members of the PapB regulatory family, which includes PapB (P fimbriae), PixB (Pix fimbriae), FocB/SfaB (F1C/S fimbriae), AfaA (Afa class III adhesin), Daa (F1845 fimbriae), and FanA/FanB (K99 pili) regulatory proteins (30). The highest identity was to PixB (57% identity, 76% similarity), followed by FanB (47% identity, 69% similarity) and PapB (45% identity, 67% similarity). No equivalent of the PapI regulatory gene was present. Some Plf proteins show higher identity to other *pap*-related fimbrial gene clusters, specifically from Pix fimbriae, identified in some *E. coli* urinary tract

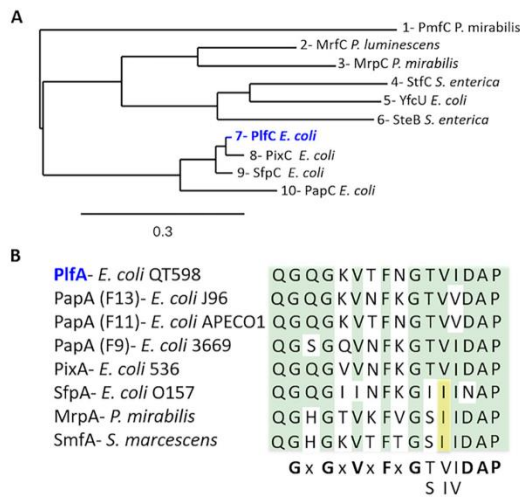
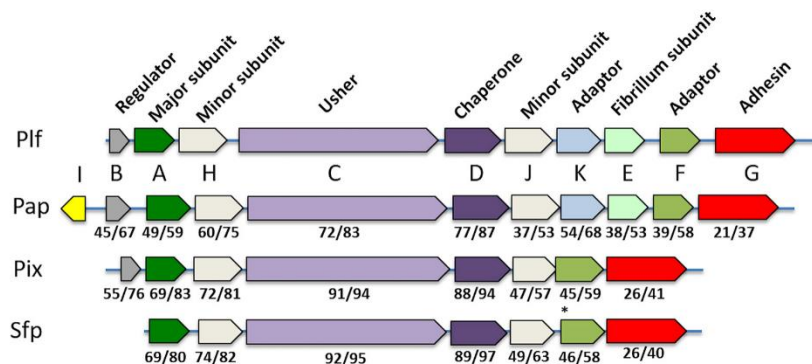


FIG 2 Phylogenetic relationship of PL fimbriae (Pif) with other types of fimbriae. (A) Phylogram using sequences of fimbrial usher proteins (FUPs) belonging to the π fimbriae clade based on the classification scheme of reference 2. The PifC protein, shown in blue, clusters with other FUPs, Sfp and Pix, more closely related to P fimbriae. (B) Alignment of segment S1 of the major subunit proteins also places PifA within the PapA-like subfamily (Ic) according to the scheme of Girardeau et al. (29). The alignment results in a consensus GxGxVxFxGTS[V]I[V]DAP motif. Accession numbers for the sequences shown in panel A were from UniProt (uniprot.org) or GenBank: 1, P53514; 2, Q93MT4; 3, Q51904; 4, H9L4A4; 5, P77196; 6, A0A3U8KZY6; 7, AKG46878.1; 8, A0A454A7L3; 9, B8RHG0; 10, P07110. Accession numbers for the sequences shown in panel B were from UniProt or GenBank: PifA, AKG46876.1; PapA (F13), X61239; PapA (F11), Q4FBG1; PapA (F9), M68059; PixA, A0A454A7E1; SfpA, W6JHT1; MrpA, Q03011; SmfA, P13421.

infection isolates (23, 24), and the plasmid-encoded Sfp fimbriae, present in sorbitol-fermenting diarrheagenic *E. coli* O157:H7 strains (25, 26) (Fig. 3). Despite demonstrating higher identity to certain proteins from these other fimbriae, only Pif demonstrates a complete set of structural/assembly protein genes equivalent to each in the *pap* gene



cluster, as Pix and Sfp fimbriae both lack PapE or PapK protein paralogs, which are known to code for a fibrillum subunit and adaptor protein, respectively (Fig. 3). The gene products showing the greatest degree of diversity among these fimbrial protein paralogs were the G adhesins, which exhibited less than 30% identity. Taken together, the PL fimbrial system is highly similar to Pix and Sfp fimbrial systems, but shares a genetic organization more akin to that of P fimbriae, as it includes the PapK and PapE paralogous proteins PlfK and PlfE, which are predicted to be part of a thin fibrillum structure.

PL fimbrial gene clusters contain different types of G adhesins. To best identify potential fimbrial genes that are very closely related to the Plf system of strain QT598, alignment searches of the NCBI database were done against the predicted adhesin-encoding gene product PlfG using a cutoff of >90% amino acid identity. The search revealed 105 samples (104 *E. coli* strains and one *Escherichia albertii* strain) containing a *plfG* allele with high identity to PlfG_{QT598}. Interestingly, among the sample sources, a majority were isolated from avian species as well as clinical isolates from urine or extraintestinal infections in humans (Table S1). However, some samples were also from a variety of livestock, healthy human fecal samples, exotic zoo animal fecal samples, and environmental sources. BLAST analyses against PlfG_{QT598} also identified a series of proteins demonstrating from 44% to 77% identity to PlfG that were all associated with fimbrial gene clusters belonging to the Plf family, since these fimbrial gene clusters shared the same genetic organization as the *plf* gene cluster and had highly conserved identity (>94%) to the PlfD_{QT598} gene products (data not shown).

The initial similarity search of the NCBI database was based on specifically identifying *plf* gene clusters with PlfG adhesin similar to that of strain QT598. In order to determine the general prevalence of PL fimbrial genes distributed among different isolates, a protein BLAST search specific to the predicted PlfK protein was used. We chose this predicted open reading frame for the search as it is not present in the other closely related π fimbria Pix or Sfp and since identity to the PapK protein ortholog is only 54% (Fig. 3). The search (conducted in November 2021) revealed 686 distinct isolates/strain entries containing *plf* sequences, most of which contained complete *plf* gene clusters. This included 171 entries from human infections (mainly urinary isolates or blood infections) and 66 human fecal source isolates. As well, many of the entries were associated with infections in dogs, samples from turkeys, chickens, and other avian species, and a variety of environmental sources. Details of the sources of entries containing *plf* sequences are presented in the supplemental material (Table S2). Taken together, sequence survey analyses indicate that the PL fimbria-encoding sequences are present in *E. coli* strains associated with extraintestinal infections as well as a variety of other animal and environmental sources.

Based on sequences in the database and identification of entries containing enough sequence data to span the length of the fimbrial gene clusters, a phylogenetic analysis of distinct protein entries for different PlfG adhesins was determined. In all, 21 protein entries sharing identity with PlfG were identified (Fig. 4). Analysis determined 5 distinct clades of the PlfG adhesins, including a group (class V) specific to some *Cronobacter* spp. The number of individual entries from the sequence database indicated that PlfG class I and PlfG class II families were predominant, whereas PlfG classes III to V were represented by only a few individual strains in the sequence database (Fig. 4). All of the *plf* gene clusters identified from *E. coli* strains regardless of G adhesin class were inserted adjacent to a site-specific integrase and RepFIB region (data not shown), suggesting that these fimbrial systems are likely to be plasmid encoded. Taken together, these results suggest that *plf* gene clusters are present in numerous *E. coli* strains and that the G adhesins of these fimbriae have diversified into distinct alleles.

Phylogenetic analysis based on the comparison of the PlfG adhesin sequences in the database demonstrated that two main classes of PlfG adhesins, class I and class II, are predominant in sampled genomes. Specific BLAST comparisons of the PlfG class II adhesin from *E. coli* strain QT598 with a representative gene encoding the class I

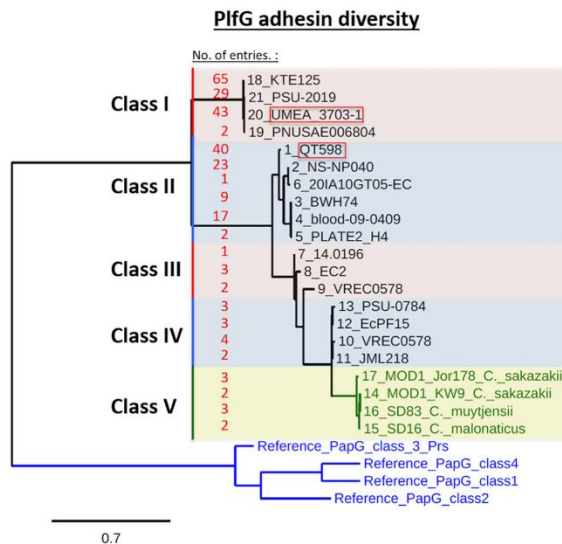


FIG 4 Phylogenetic analysis of different PfG adhesin proteins. Predicted PfG proteins from individual isolates were obtained from the sequence database at NCBI (<https://www.ncbi.nlm.nih.gov>) and one individual isolate sequence was selected based on sequence diversity and association with a complete *plf* fimbrial gene cluster. The total number of protein accessions for each group (no. of entries) at time of submission are listed on the left in red. Multiple-sequence alignment (MUSCLE) and phylogeny (PhyML) were generated using Phylogeny.fr (www.phylogeny.fr). Analysis determined 5 distinct clades of PfG adhesins. Included is a group (class V) specific to some *Citrobacter* spp. (indicated in green). The PapG reference adhesins from P fimbriae clustered as a distinct group from all of the PfG adhesin proteins. The *plf* gene clusters from two strains, UMEA-3703-1 (PfGI) and QT598 (PfGII), both boxed in red, were cloned for further analysis. The total numbers of protein accessions for each group, at time of submission, are listed on the right. Twenty-one different nonredundant entries were used: 1, WP_059331527.1; 2, WP_137488293.1; 3, WP_097732425.1; 4, WP_033555940.1; 5, WP_201475228.1; 6, EGW8442016.1; 7, MBB8123006.1; 8, WP_029305610.1; 9, WP_112039355.1; 10, WP_096965282.1; 11, WP_137504062.1; 12, WP_176323703.1; 13, EFO1491433.1; 14, WP_133116004.1; 15, WP_158696804.1; 16, WP_158685756.1; 17, WP_105536056.1; 18, WP_001523394.1; 19, EFB9349400.1; 20, WP_016233112.1; 21, WP_033549358.1. Alignment also included reference entries for the 4 established PapG alleles under the following accession numbers: PapGI (strain J96), CAA43570.1; PapGII (strain IA2), AAA24293.1; PapGIII (strain J96), P42188; PapGIV (AAK08949.1) strain BF31.

adhesin from strain UMEA-3703-1, showed 45% identity and 65% similarity. This sequence divergence is similar to the difference between P fimbriae class I and class II G adhesins (46% identity, 64% similarity). As such, and since these two PfG classes are the most common in the database, we cloned both of them for further investigation.

The *plf* class I and II gene clusters encode fimbriae with distinct hemagglutination activity. To demonstrate that the *plf*-containing clones produced fimbriae, the plasmids encoding *plf* genes were transformed into the afimbriated *E. coli* K-12 strain ORN172. Transmission electron microscopy (TEM) demonstrated that both *plf*_{QT598}⁻ and *plf*_{UMEA-3703-1}⁻-containing plasmids produced peritrichous fimbrial filaments at the surface of cells of strain ORN172 (Fig. 5; Fig. S2). In contrast, ORN172 containing the empty vector did not produce any fimbriae, as expected.

P fimbrial adhesins are known to be mannose-resistant hemagglutinins, and they demonstrate lectin activity specific to Gal(α 1-4)Gal-containing glycolipids present on the surface of erythrocytes and other host cells. To compare the hemagglutination activity of P fimbria reference clones with that of clones producing PL fimbriae, we tested hemagglutination activity of fimbriae expressing clones in the nonfimbriated *E. coli* strain ORN172 for a variety of erythrocytes from different species (Fig. 6). The reference clone encoding P fimbriae with the PapG class I adhesin from *E. coli* J96/pPap5

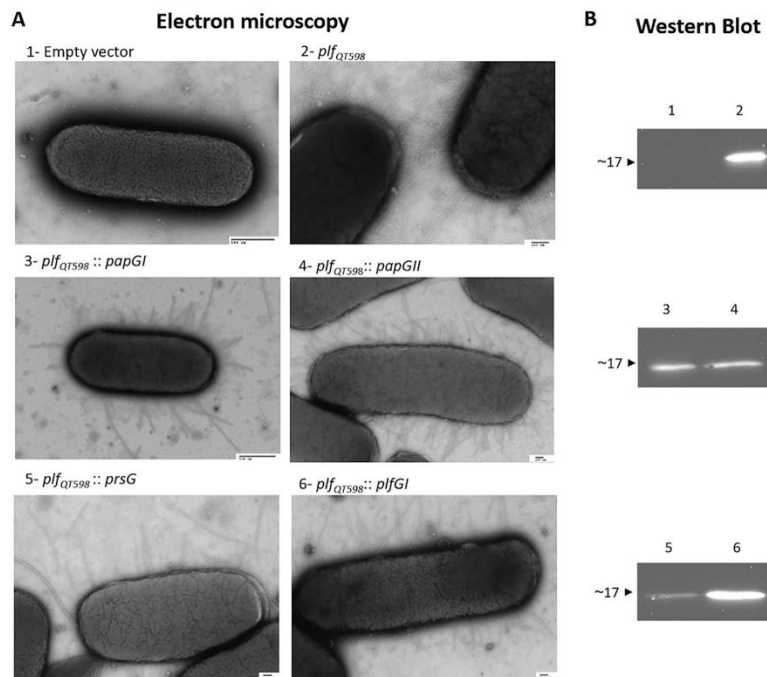


FIG 5 Cloning and expression of PL and Pap chimeric fimbriae. (A) Chimeric clones expressing different classes of PapG/PlfG were visualized by electron microscopy. The Plf class II clone was used as a template to generate chimeric clones. (B) Western blot analysis of heated surface protein extracts from *plf* and chimeric clones. The antibodies used were polyclonal rabbit antibodies raised against a peptide corresponding to the fimbrial major subunit protein PlfA.

demonstrated strong hemagglutinin activities with human, pig, dog, and rabbit erythrocytes. The reference clone encoding P fimbriae containing the Pap class II adhesin from *E. coli* IA2/pDC5 strongly agglutinated pig and human erythrocytes and, to a lesser extent, sheep and chicken erythrocytes. The reference clone encoding Prs fimbriae with the PapG class III (PrsG) adhesin from uropathogenic *E. coli* J96/pJFK102 agglutinated dog, pig, and sheep erythrocytes. The clone encoding PL fimbriae containing a Plf class I adhesin from *E. coli* UMEA-3703-1 agglutinated a broad range of erythrocytes from all species tested, except dog blood, although hemagglutinin titers were higher for human and sheep blood. Interestingly, the clone encoding the Plf class II adhesin from *E. coli* QT598 only strongly agglutinated human and turkey blood (Fig. 6). Taken together, these results indicate that, as with the P fimbrial classes of adhesins, PL fimbriae are hemagglutinins and that the PlfG class I and class II adhesins demonstrate distinct hemagglutination activities compared to the P fimbrial adhesins.

To determine whether lectin-based hemagglutination by PL adhesins was similar to those of Pap and related fimbriae, we tested whether globoseries glycolipids known to be recognized by P fimbriae (Gb3, Gb4, or Gb5) were able to inhibit PL fimbria-mediated hemagglutination. For the hemagglutination tests, we used a macroagglutination assay using human and turkey erythrocytes, since these were both recognized by the PlfG class I and class II fimbriae. In this assay, the PlfG class I and II adhesins strongly agglutinated human and turkey erythrocytes, and addition of either Gb3, Gb4, or Gb5 did not inhibit or reduce this agglutination. For the test controls, none of the PapG adhesins agglutinated turkey erythrocytes, but all strongly agglutinated human erythrocytes. For

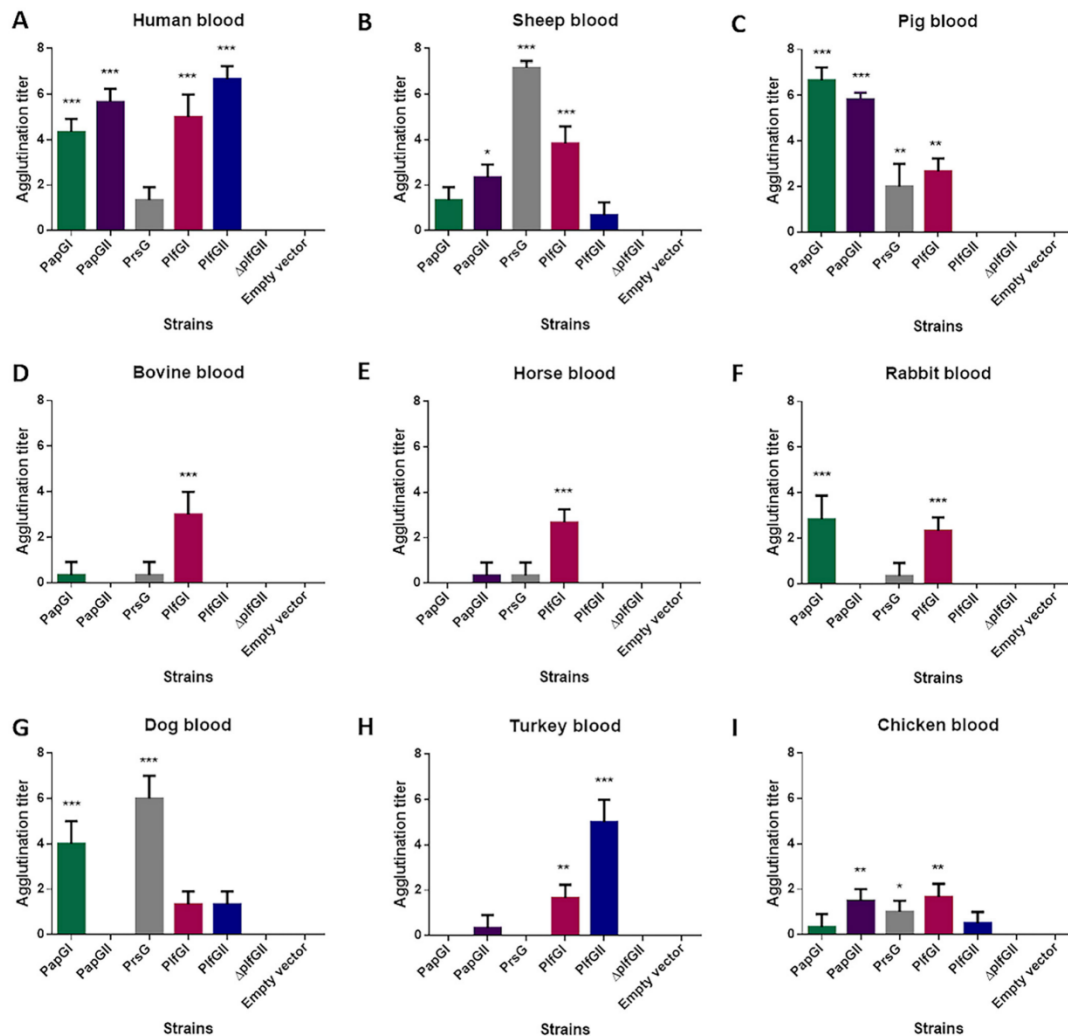


FIG 6 Hemagglutination activity of different clones expressing P or PL fimbriae. Clones were *E. coli* afimbriated *fim*-negative strain ORN172 expressing P or PL fimbriae. Cells were adjusted to an OD_{600} of 0.6 and then cells were centrifuged and concentrated 100-fold and then 2-fold serial dilutions were made in 96-well plates containing a final concentration of 3% erythrocytes from different species. Titers are the average maximal dilution showing agglutination. Erythrocytes tested were: (A) human, (B) sheep, (C) pig, (D) bovine, (E) horse, (F) rabbit, (G) dog, (H) turkey, and (I) chicken. Both human A and O blood types gave similar titers. Reference clones showed different hemagglutination activity. However, the Δ plfG clone as well as the empty vector showed no hemagglutination activity of any of the erythrocytes tested. *, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$, versus empty vector by one-way ANOVA. The plasmids used were pPap5 (*papGI*), pDC5 (*papGII*), pJFK102 (*prsG*), pJ523 (*plf*_{UMEA-3703-1};*plfI*), pJ507 (*plf*_{QT598};*plfII*), and pJ598 (*plf*_{QT598} Δ *plfG*).

PapG class I, Gb3 and Gb4 reduced this agglutination, but not Gb5. For PapG class II, only Gb4 was able to inhibit agglutination. For PapG class III (Prs), Gb4 and Gb5, but not Gb3, were able to inhibit agglutination (Table 1). Taken together, these results indicate that the PL fimbriae belonging to either class I or class II recognize receptors on both human and turkey erythrocytes, but these receptors are distinct from those recognized by P fimbrial adhesin classes.

Different Plf and Pap G adhesin alleles can be expressed by PL fimbriae. Since the PlfG class I and II adhesin sequences from *E. coli* strains QT598 and UMEA-3703-1,

TABLE 1 Effect of globosides on hemagglutination of turkey or human (O) erythrocytes by PL fimbriae or Pap fimbrial adhesins expressed on plasmids in *E. coli* fimbria-negative K-12 strain ORN172

Erythrocytes and strain	Hemagglutination result for ^a :					
	Globotriaosylceramide (Gb3)		Globotetraosylceramide (Gb4)		Galactosylgloboside (Gb5)	
	Present	Absent	Present	Absent	Present	Absent
Turkey erythrocytes						
QT5722 (PapGI P fimbriae)	–	–	–	–	–	–
QT5723 (PapGII P fimbriae)	–	–	–	–	–	–
QT5724 (PapGIII P fimbriae)	–	–	–	–	–	–
QT4741 (PlfGI PL fimbriae)	+++	+++	+++	+++	+++	+++
QT5725 (PlfGII PL fimbriae)	+++	+++	+++	+++	+++	+++
QT5727 (<i>plfB</i> - <i>F</i> Δ <i>plfGII</i>) ^b	–	–	–	–	–	–
QT5726 (empty vector)	–	–	–	–	–	–
PBS	–	–	–	–	–	–
Human (O) erythrocytes						
QT5722 (PapGI P fimbriae)	+	+++	+/-	+++	+++	+++
QT5723 (PapGII P fimbriae)	+++	+++	–	+++	+++	+++
QT5724 (PapGIII P fimbriae)	+++	+++	+/-	+++	–	+++
QT4741 (PlfGI PL fimbriae)	+++	+++	++	++	+++	+++
QT5725 (PlfGII PL fimbriae)	+++	+++	+++	+++	+++	+++
QT5727 (<i>plfB</i> - <i>F</i> Δ <i>plfGII</i>) ^b	–	–	–	–	–	–
QT5726 (empty vector)	–	–	–	–	–	–
PBS	–	–	–	–	–	–

^aMacroagglutination level: ++++, strong; ++, moderate; +, weak; +/-, very weak; –, absent. Tests were repeated twice. See Materials and Methods for details of the agglutination assay.

^bStrain QT5727 contains all *plf*_{QT598} fimbrial genes, except *plfG*, predicted to encode the adhesin.

respectively, are quite distinct from each other and as the *plf* gene clusters share close genetic organization with *pap* gene clusters, we generated chimeric fimbrial gene clusters encoding different G adhesin alleles. These chimeric clones were based on the *plf*_{QT598} gene cluster by generating a clone lacking the *plfG* gene (*pIJ598*) and then cloning the *plfG*_{UMEA-3703-1} or *papG* alleles from each of the three PapG adhesin classes (Fig. S3A and B). Electron microscopy demonstrated that each of the five chimeric clones introduced to nonfimbriated *E. coli* ORN172 produced fimbriae on the surface of the cells (Fig. 5A). The PlfA subunit protein was also detected from cell surface extracts, as shown in immunoblots, although the level of protein present was decreased when the *papG* recombinant alleles were expressed compared to the *plfG*_{QT598} and *plfG*_{UMEA-3703-1}-expressing clones (Fig. 5B). We also verified the hemagglutination capacity of hybrid fimbrial systems with some types of blood. The PL hybrid fimbriae with Pap class I or Pap class II adhesin could agglutinate human and pig blood, but were no longer able to agglutinate turkey blood. The PL fimbria hybrid with Pap class III Prs adhesin showed agglutination of sheep and human blood, but was also unable to agglutinate turkey erythrocytes (Table S3). This indicates that replacement of the PlfG allele with a Pap adhesin resulted in production of hybrid fimbriae with an altered binding specificity. It was also possible to replace the PapG adhesin in reference clone pDC1 (31) with the PlfG class II adhesin, resulting in production of fimbriae, as shown by electron microscopy (Fig. S3). However, this hybrid fimbrial system was unable to agglutinate either human or turkey erythrocytes. These results indicate that the PlfG adhesin can complement loss of the PapG adhesin for biogenesis of fimbriae, but suggest that an altered conformation of the adhesin in the hybrid system results in loss of hemagglutination capacity. As such, the Pap and PL fimbrial systems, although they share some characteristics, are not fully interchangeable with regard to adhesin function and specificity.

PL fimbriae mediate adherence to human epithelial cells. The adherence of bacteria to host epithelial cells such as bladder and kidney cells is an important step in colonization of the urinary tract. To investigate whether PL fimbriae can mediate adherence to host

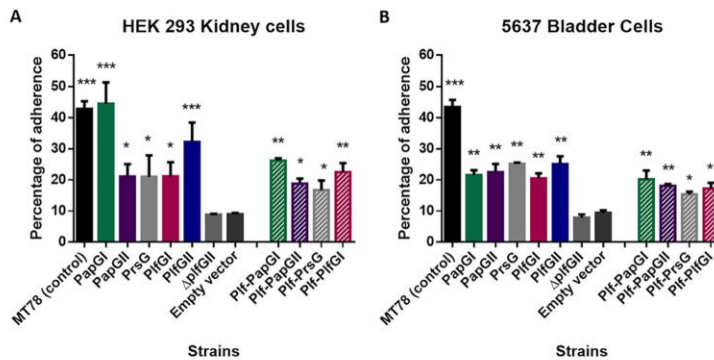


FIG 7 Reference and chimeric clones promote adherence to human kidney (strain HEK-293) and bladder (strain 5637) epithelial cells. Cell monolayers were infected with *E. coli* *fim*-negative ORN172 expressing P and PL fimbrial proteins at a multiplicity of infection (MOI) of 10 as described by Matter et al. (55) and incubated at 37°C at 5% CO₂ for 2 h. Adherent bacteria were enumerated by plating on LB agar. Empty vector was used as a negative control and APEC MT78 as a positive control for adherence to cell lines. All the clones encoding P or PL fimbrial adhesins (reference and chimeric clones) demonstrated increased adherence to cells of the human bladder 5637 and kidney HEK-293 epithelial cell lines compared to the strain containing the empty vector. The $\Delta plfG$ clone also did not adhere to human epithelial cells. Data are the averages from three independent experiments. Error bars represent standard errors of the means. *, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$, versus empty vector by one-way ANOVA.

cells, we used clones expressing PlfG class I, PlfG class II, or PapG class I, II, or III fimbrial adhesins.

All the clones encoding P or PL fimbrial adhesins (reference and chimeric clones) demonstrated increased adherence to the bladder 5637 and kidney HEK-293 epithelial cell lines compared to the strain containing the empty vector (Fig. 7). The chimeric clones containing the *plf* gene cluster with hybrid *pap* or *plfG* adhesin-encoding genes also promoted adherence to epithelial cells. However, the clone containing a *plf* gene cluster, lacking the *plfG* or *papG* gene showed no appreciable adherence compared to the empty vector-containing clone (Fig. 7). These results demonstrate that PL fimbriae producing distinct types of G adhesins can mediate adherence to urinary tract epithelial cells and suggest that these fimbriae may potentially play a role during host colonization.

PL fimbriae promote biofilm production. Since fimbriae can contribute to biofilm production, we tested for biofilm formation in polystyrene microtiter plates at different temperatures (25, 37, and 42°C) (Fig. 8). The clone expressing PL fimbriae with the PlfG class II adhesin showed a high level of biofilm production at all tested temperatures, even above that of a positive-control biofilm-forming *Serratia liquefaciens* reference strain. The clone expressing PL fimbriae with the PlfG class I adhesin also produced biofilm at 25 and 37°C at moderate levels compared to the clone producing the PlfG class II adhesin. However, biofilm production was very low at 42°C. The Pap class I- and III-producing reference clones were also able to form significantly more biofilm at 25, 37, and 42°C than the negative control. The Pap class II-expressing reference clone produced biofilm at higher temperatures (37 and 42°C), but biofilm levels were reduced at 25°C.

The chimeric clones that expressed different Pap or Plf adhesins fused to the *plf*_{QT598} gene cluster were all able to produce appreciable levels of biofilm at both 25 and 37°C, although biofilm was much reduced at 42°C. The clone expressing the *plf*_{QT598} gene cluster lacking a *papG* or *plfG* adhesin-encoding gene ($\Delta plfG$ clone) as well as clone containing the empty vector were not able to produce biofilm at all the tested temperatures (Fig. 8). Taken together, these results indicate that PL and P fimbriae expressing different types of G adhesins can mediate biofilm production in *E. coli* K-12 and that the PlfG class II adhesin in particular can contribute to strong biofilm formation.

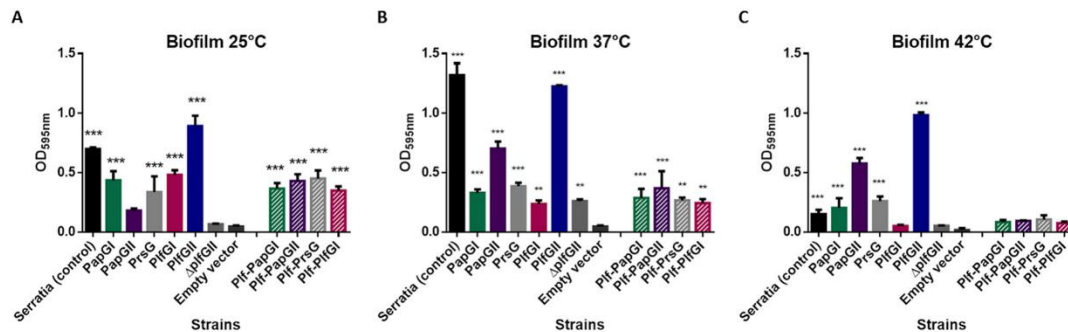


FIG 8 Biofilm production by clones expressing Pfl_{QT598} and $Pfl_{UMEA-3703-1}$ and reference and chimeric clones at different temperatures. Clones of *E. coli* strain ORN172 expressing P and P-like fimbriae proteins were grown at different temperatures (25, 37, and 42°C) in polystyrene plate wells for 48 h and then stained with crystal violet. The remaining crystal violet after washing with acetone was measured as absorbance at 595 nm. Data are the means from three independent experiments, and error bars represent standard errors of the means. *, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$, compared to empty vector using one-way ANOVA. Empty vector was used as a negative control and *S. liquefaciens* strain was used as a positive control for biofilm production.

PL fimbrial genes are upregulated in the bladders of infected mice and contribute to competitive colonization in a mouse urinary track model.

To investigate the potential role of the PL fimbriae in virulence in the UTI model, 6-week-old female CBA/J mice were infected with wild-type strains QT598 and UMEA-3703-1 or with mutant Δpfl strains lacking the genes encoding PL fimbriae. In the mouse infection model, loss of PL fimbriae did not have a significant effect on colonization of the bladder or kidneys by strain QT598 (Fig. S4). Strain UMEA-3703-1 and its Δpfl mutant also showed no significant differences in colonization. However, UMEA-3703-1 was only able to colonize at lower levels (10^2 to 10^3 CFU/g) compared to strain QT598 (10^5 to 10^6 CFU/g) (Fig. S4). Interestingly, the expression level of pfl_{QT598} was upregulated by more than 40-fold in the bladder of infected mice compared to expression following growth *in vitro* in LB medium (see Fig. 9 below). This suggests that the expression of these fimbriae is favored by environmental cues during infection in the urinary tract.

We also conducted a competitive infection model using the pfl mutant and a wild-type *lac*-negative derivative of strain QT598. In this infection model, there was no significant difference in colonization of the mouse bladders (Fig. 10A). In contrast, the bacterial numbers of the pfl mutant were significantly lower than those of the pfl -expressing strain ($P = 0.0084$) and showed a median decrease of more than 16-fold

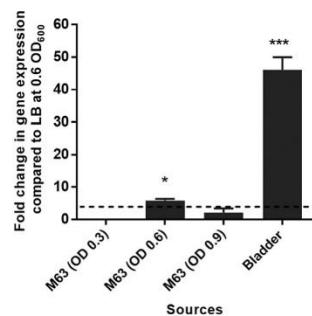


FIG 9 RT-PCR analysis of pfl expression by strain QT598. QT598 was grown in LB medium to OD_{600} of 0.6 and used as a standard to compare it with growth in M63 minimal medium (with glycerol as carbon) at different growth phases (OD_{600} values of 0.3, 0.6, and 0.9). RNAs were also extracted from infected bladders. Transcription of pfl was significantly upregulated in the mouse bladder. Error bars indicate standard deviations. *, $P < 0.05$, and ***, $P < 0.001$, by Student's *t* test. The dashed line corresponds to the cutoff for a significant difference in expression.

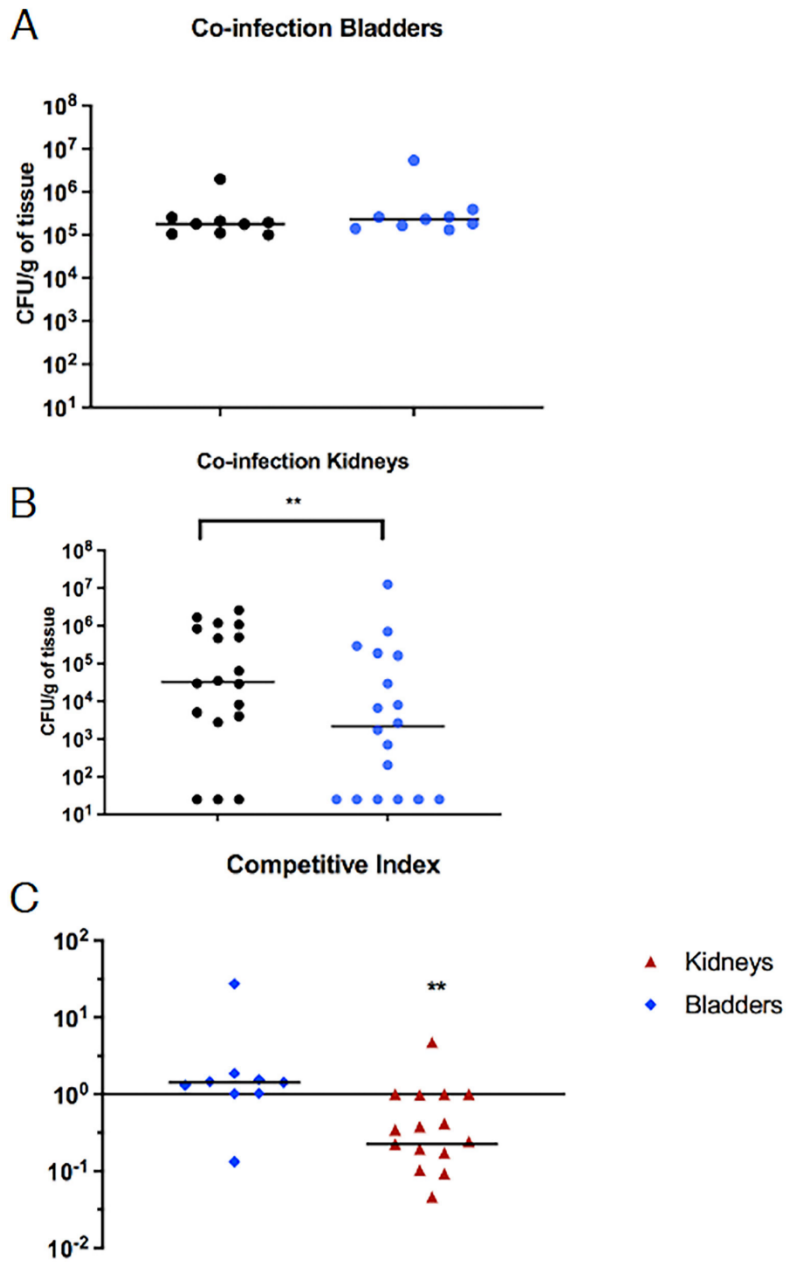


FIG 10 A *plf* mutant shows decreased colonization in mouse kidneys in a coinfection model. An equal number of cells each of the two strains was administered to CBA/J mice. Mice ($n = 9$) were euthanized after 48 h, and bladder and kidneys were harvested for colony counts. (A) Bacterial numbers in the bladders; (B) bacterial numbers in the kidneys. Mice were infected with the QT598 $\Delta lacZYA$ strain QT4567 (black circles) and the QT598 Δplf mutant QT4420 (blue circles). The horizontal lines indicate the median numbers of bacterial colonies in tissues. (C) Competitive index comparing the colonization in these organs by the QT598 (*lac*-negative) strain and the QT598 Δplf mutant. The Wilcoxon matched-pairs test was used for statistical analysis. *, $P < 0.05$; **, $P < 0.01$.

(Fig. 10B) and an overall competitive index of -0.65 (Fig. 10C) in the kidneys. Taken together, these results indicate that the *plf* fimbrial gene expression is increased in the mouse urinary tract and that in competitive infection, loss of the PL fimbriae decreases competitive fitness in the mouse kidneys.

DISCUSSION

A novel plasmid-carried fimbrial gene cluster was identified on the large colicin V plasmid of avian pathogenic *E. coli* strain QT598 (serotype O1:K1, sequence type ST1385). Strains from this and related sequence types such as ST91 are commonly associated with extraintestinal infections in poultry and urinary tract infections in humans (27) (<http://enterobase.warwick.ac.uk/>). The *plf* genes were shown to be adjacent to the RepFIB, and *intM* genes on plasmid pEC598. This is a common integration site for a diversity of genes on F and related plasmids, and collectively this region has been named the cargo gene region (28). The cargo gene region has been found to encode a diversity of accessory genes, insertion sequences, and integrons known to carry genes for resistance to antimicrobials and metals, microcins, and virulence genes (28, 32). It is therefore likely that the *plf* fimbrial gene cluster along with other genes was inherited by certain *E. coli* strains through a recombination/integration event and that it has since disseminated or been transferred into a diversity of *E. coli* isolates associated with different host or environmental sources (highlighted in Tables S1 and S2 in the supplemental material). As with the P fimbriae, PL fimbriae have also diversified considerably, and there has been notable divergence in the PlfG adhesin-encoding sequences into 5 distinct PlfG adhesin classes (Fig. 4). Such changes may have occurred to promote adherence and colonization to a variety of surfaces or host cell receptors in different niches or environments.

The PL fimbriae are new members of the π fimbrial family, which contains P-fimbria-like operons present in some *Betaproteobacteria* and *Gammaproteobacteria* (2). More specifically, based on comparison of the fimbrial usher proteins, the PL fimbriae are part of a subgroup which includes true P fimbriae, as well as closely related Sfp and Pix fimbriae (2) (Fig. 2), all of which have been shown to mediate mannose-resistant hemagglutination (MRHA) of erythrocytes from humans in addition to some distinct MRHA profiles for erythrocytes from other species. Pix fimbriae, which have been identified in some uropathogenic *E. coli* strains, were shown to agglutinate human erythrocytes, but not sheep or goat erythrocytes, and do not recognize the Gal-Gal sugars recognized by P fimbriae (23). Sfp fimbriae also mediate MRHA of human erythrocytes, which was dependent on the *sfpG* gene (26). However, to our knowledge, no tests for MRHA for erythrocytes from other species have been reported. Interestingly, the G adhesin proteins from Pix and Sfp fimbriae share amino acid homology between them (63% identity and 81% similarity), suggesting these G adhesin proteins are more closely related to each other than to PlfG or PapG adhesins, which share no more than 25% amino acid identity. Herein, we demonstrated that PL fimbriae producing the class I adhesin mediated MRHA for erythrocytes from different species, including equine, ovine, bovine, rabbit, and human erythrocytes, whereas PL fimbriae producing the class II adhesin mediated MRHA only to human and turkey erythrocytes (Fig. 6). Taken together, this subgroup of π fimbriae (true P fimbriae, Sfp, Pix, and PL fimbriae) have developed important differences in adhesin protein sequences that have expanded the capacity to adhere to a variety of receptors on erythrocytes and host cells from different species. It will be of interest to more specifically determine the lectin receptor specificity of this family of fimbriae.

The genetic organization of the *plf* gene cluster includes 9 predicted fimbrial subunit genes, which is the number of predicted structural genes encoding P fimbriae (Fig. 3). In contrast, both the Sfp and Pix fimbrial gene clusters comprise 7 structural genes and lack the genes corresponding to the *papK* and *papE* genes predicted to encode an adaptor and a minor fimbrial subunit (Fig. 3). From this standpoint, overall, PL fimbriae are most similar to true P fimbriae.

To further demonstrate potential complementarity between P and PL fimbriae, we also generated hybrid fimbrial gene clusters, wherein the *plfG*_{QT598} gene was replaced

by PapG adhesin-encoding genes belonging to class I, class II, or class III adhesins. Each of these clones was able to produce a functional fimbrial structure that also increased adherence to human urinary tract epithelial cells. This also further indicates that the PL fimbriae, despite having adhesins that are quite distinct in amino acid sequence from P fimbriae, also produce mannose-resistant hemagglutinins that can mediate adherence to human bladder and kidney cells and that the bioassembly of these fimbriae is compatible with that of P fimbrial G adhesins. It is interesting, however, that the production of the hybrid fimbriae from bacterial cells was substantially reduced compared to that of the PL fimbrial clones, suggesting that efficiency of biogenesis of the hybrid fimbriae is reduced.

As with the *plf* gene cluster, the location of the *sfp* genes is also on IncF plasmids, in close proximity to RepFIB of the pSFO157 plasmid (26). However, it is flanked on both sides by insertion sequences that are distinct from the region adjacent to *plf* genes on pEC598. The Sfp fimbriae were initially found not to be expressed by EHEC strains under normal laboratory conditions, and properties of these fimbriae were first determined using cloned fimbrial genes in *E. coli* K-12 (26). The *sfp* genes encoding a fimbrial system with mannose-resistant hemagglutinin activity have been identified on a subgroup of sorbitol-fermenting EHEC/Shiga-toxigenic *E. coli* (STEC) strains and some EHEC O165:H25/NM strains from humans and cattle, but are absent from most other types of *E. coli* (25, 26). This suggests that the *sfp* genes were likely acquired independently by horizontal transfer to both a nonmotile sorbitol O157 strain and independently to an O165:H25/NM strain and have since remained in these branches of EHEC (25). This is clearly in contrast to the *plf* gene cluster, which is present in a diversity of *E. coli* strains from multiple sources and has likely been transferred either through multiple conjugation and/or recombination events and has also diversified, since distinct G adhesin classes have emerged among strains.

DNA sequence comparisons of gene clusters that are highly similar to the *plf* fimbrial system of *E. coli* QT598 from nucleotide databases provided a means to identify subgroups of PL fimbria genes encoding 5 distinct classes of PlfG adhesins (Fig. 4). Since the PlfG class I- and class II-encoding alleles were predominant among isolates that notably included strains associated with human extraintestinal infections as well as infections from poultry, we focused our attention on functional characterization of one of each of the PL fimbriae belonging to these classes. It was also interesting to identify some variant alleles of the PlfG adhesin in other *E. coli* strains as well as a subgroup that was identified in some strains of *Cronobacter sakazakii* and other *Cronobacter* spp. (Fig. 4). Although *Cronobacter* strains containing the *plf* fimbrial clusters were sampled from spices, *C. sakazaki* and related *Cronobacter* spp. are important foodborne pathogens that can contaminate dehydrated milk and other products and cause serious extraintestinal infections, particularly in neonates (33, 34).

The capacity of PL fimbriae to form biofilms at different temperatures was also investigated, and both the class I and class II PL fimbriae promoted biofilm formation, with PlfG class I producing more biofilm at 25 and 37°C, but not 42°C. In contrast, the PlfG class II adhesin produced very high levels of biofilm at all temperatures tested. The presence of the PlfG adhesin was important for high-level biofilm production for PlfG class II, since the absence of the *plfG* adhesin gene greatly reduced biofilm formation. Notably, after growth at 37°C, the level of biofilm produced by the $\Delta plfG$ clone was significantly higher than that with the empty vector and comparable to levels of biofilm produced by Pap reference clones and the PlfG class I clone. This suggests that other factors in addition to the G adhesin may also contribute to increased biofilm formation associated with expression of *plf* or *pap* fimbrial genes.

Since both types of PL fimbriae conveyed increased adherence to human epithelial bladder and kidney cells (Fig. 7), we investigated the potential of these fimbriae to contribute to urinary tract colonization in a murine model. Deletion of the *plf* genes from either *E. coli* strain QT598 or strain UMEA-3703-1 did not have an appreciable effect on colonization of either the bladder or the kidneys. Furthermore, despite being isolated

from a human UTI, strain UMEA-3703-1 was not a strong colonizer in the mouse UTI model.

The mouse UTI model may not be as representative of a human infection when using certain bacterial strain backgrounds or when investigating specific mechanisms of virulence such as adherence and fimbrial adhesins. P fimbriae have been shown to play a role in urinary infection, particularly for pyelonephritis in cynomolgus monkeys (35), and these fimbriae alone can confer an asymptomatic *E. coli* urinary strain the capacity to elicit strong regulatory modulation in humans by acting as an IRF-7 agonist and reprogramming the immune response in the urinary tract (5). In the case of the murine model, it has been demonstrated that P fimbriae can reduce the immune response in the kidney by decreased production of polymeric Ig receptor and reduced secretion of IgA (36). However, the role of P fimbriae in bacterial colonization in the UTI mouse model has been less evident. Initially, *pap* genes cloned into an avirulent *E. coli* K-12 strain or an intestinal commensal *E. coli* strain were shown to increase colonization of the mouse kidney (37, 38). In contrast, deletion of *pap* from different UPEC strains did not alter colonization of the urinary tract in CBA/J mice (39). Reasons why PL fimbriae as much as P fimbriae may not play as critical role in the mouse UTI model may be due to differences in lectin receptor target specificity present on murine cells and/or the potential redundancy of adherence mechanisms due to production of multiple fimbrial adhesins in UPEC strains. Since a slight increase in colonization of the *plf* mutant of strain UMEA-3703-1 in the bladder was observed (Fig. S4), it is also possible that loss of the *plf* genes may alter expression of other types of fimbriae, since regulatory cross talk can occur and loss of one type of fimbriae can enhance expression of other types of fimbriae. The expression of *plf* was also upregulated more than 40-fold by strain QT598 in the bladder and was increased by 5-fold in minimal medium compared to rich medium (Fig. 9). This indicates that the expression of these fimbriae can be increased by cues during infection, which may include host factors or decreased nutrient availability. In the current study, we used an acute infection model with an endpoint of 48 h, and our detection of increased expression of *plf* in the urinary tract was only determined at this time point. As such, these fimbriae may potentially play a more important role in more prolonged infections, and it may be of interest to investigate colonization at later time points using other models, such as catheter-associated infections.

Since the competitive coinfection model demonstrated a potential fitness advantage in the mouse kidneys, this nevertheless supports a potential advantage for production of these fimbriae by strain QT598 during UTI in the mouse model (Fig. 10). The prevalence of *plf* genes was associated with strains from human infections, as well as canine infections and from samples in poultry, including turkeys and chickens. In future studies, it will be of interest to determine whether PL fimbriae or specific P1fG adhesin classes may contribute to infection in other animal models such as poultry and to further investigate PL fimbria receptor specificity, potential role in modulation of host immune response and the regulation of production of this newly identified group of fimbriae.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 2. ExPEC strain QT598 (a passaged derivative of strain MT156 [40]) is an O1:K1 sequence type ST1385 strain originally isolated from a turkey suffering from colibacillosis in France (27). UPEC strain CFT073 was isolated from the blood and urine of a woman suffering from urinary tract infection (41), and UMEA-3703-1 (NCBI Biosample no. [SAMN01885978](https://www.ncbi.nlm.nih.gov/biosample/SAMN01885978)) was isolated from the urine of a human with bacteremia. *E. coli* K-12 laboratory strains DH5 α and ORN172 (type 1 fimbria *fim*-negative strain) were used for cloning and protein expression. Reference clones expressing fimbria genes encoding different PapG adhesin classes were used as controls, including plasmids pPap5 (*papG*₅₆ class I) (42, 43), pDC5 (*papG*₁₆ class II) (31), and pJFK102 (*prsG*₉₆ class III) (44, 45).

Bacteria were routinely grown at 37°C on solid or liquid Luria-Bertani (LB) medium (Alpha Bioscience, Baltimore, MD). When required, antibiotics were added to a final concentration of 100 μ g/mL of ampicillin, 30 μ g/mL of chloramphenicol, or 50 μ g/mL of kanamycin.

Bioinformatics analysis. Identification and comparison of sequences in the databases were achieved by accessing data on completed genomes and BioProjects publicly available in the NCBI database (www.ncbi.nlm.nih.gov). Analyses included BLAST against both nucleotide and protein entries. Figures presenting the organization and comparison of genes and gene clusters were generated from the nucleotide

TABLE 2 Plasmids and strains used in this study

Plasmid, strain, or clone	Characteristic(s)	Source or reference(s)
Plasmids		
pKD4	Plasmid used for amplification of <i>kan</i> cassette	50
pKD13	Plasmid used for amplification of <i>kan</i> cassette	50
pKD46	λ Red plasmid; Amp ^r	50
pUCmT	Cloning vector; Amp ^r	Bio Basic, Inc., Markham, ON, Canada
pBC sk+	Cloning vector; Cm ^r	Stratagene, La Jolla, CA
pPap5	P fimbriae expressing PapGI from J96	42, 43
pDC5	P fimbriae expressing PapGII from IA2	31
pJFK102	P fimbriae expressing PrsG (PapGIII) from J96	44, 45
pIJ507	pUCmT:: <i>plf</i> _{QT598} , PL fimbriae PIfGII	This study
pIJ523	pBC sk+:: <i>plf</i> _{UMEA-3703-01} , PL fimbriae PIfGI	This study
pIJ594	pUCmT:: <i>plf</i> _{QT598} :: <i>papGI</i> , PL chimera <i>papGI</i>	This study
pIJ595	pUCmT:: <i>plf</i> _{QT598} :: <i>papGII</i> , PL chimera <i>papGII</i>	This study
pIJ596	pUCmT:: <i>plf</i> _{QT598} :: <i>prsG</i> , PL chimera <i>papGIII</i>	This study
pIJ597	pUCmT:: <i>plf</i> _{QT598} :: <i>plfGI</i> , PL _{QT598} hybrid PIfGI	This study
pIJ598	pUCmT:: <i>plf</i> _{QT598} Δ <i>plfGII</i> , clone lacks <i>plfG</i> gene	This study
Strains		
ORN172	<i>fim</i> -negative strain; <i>thr-1 leuB thi-1</i> Δ (<i>argF-lac</i>)U169 <i>xyl-7 ara-13 mlt-2 gal-6 rpsL tonA2 supE44</i> Δ (<i>fimBEACDFGH</i>):Km <i>pilG1</i>	60
MT78	APEC O2:H ⁻ :K1, ST95	61
QT2799	<i>Serratia liquefaciens</i>	ATCC 27592
QT598	APEC O1:K1, ST1385	27, 40
QT4420	QT598 Δ <i>plf</i> Km ^r	This study
UMEA-3703-1	UPEC strain from urine of patient with bacteremia	NCBI Biosample no. SAMN01885978
QT4598	UMEA-3703-1 Δ <i>plf</i> Km ^r	This study
Clones		
QT5722	ORN172/pPap5 (reference clone expressing P fimbriae with PapGI adhesin)	This study
QT5723	ORN172/pDC5 (reference clone expressing P fimbriae with PapGII adhesin)	This study
QT5724	ORN172/pJFK102 (reference clone expressing P fimbriae with PrsG (PapGIII) adhesin)	This study
QT4741	ORN172/pIJ523 (reference clone expressing PL fimbriae with PIfGI adhesin)	This study
QT5726	ORN172/pIJ507 (reference clone PL fimbriae with PIfGII adhesin)	This study
QT5727	ORN172/pIJ598 (reference clone expressing <i>plf</i> _{QT598} Δ <i>plfGII</i>)	This study
QT5732	ORN172/pUCmT empty vector	This study
QT5728	ORN172/pIJ594 (chimeric clone expressing <i>papGI</i>)	This study
QT5729	ORN172/pIJ595 (chimeric clone expressing <i>papGII</i>)	This study
QT5730	ORN172/pIJ596 (chimeric clone expressing <i>prsG</i>)	This study
QT5731	ORN172/pIJ597 (chimeric clone expressing <i>plfGI</i>)	This study

accession numbers and entries using SnapGene (version 5.2.1) (www.snapgene.com). For comparison of the protein sequences, entries were obtained from either the NCBI or the Universal Protein Resource (UniProt) (www.uniprot.org) website. Phylogenetic analyses of protein sequences were done with the platform at Phylogeny.fr (<http://www.phylogeny.fr>) using the default ("one click") parameters (46). Analyses consisted of multiple-sequence alignment with MUSCLE (47), alignment curation with GBlocks (48), maximum likelihood phylogeny analysis using PhyML 3.0 (49), and TreeDyn for generation and editing of trees (www.treedyn.org). Specific parameters are described at the Phylogeny.fr web site platform.

Construction of plasmids. Cloning of the *plf* gene clusters and *plfG* and *papG* genes encoding the different classes of adhesins was performed by PCR amplification using specific primers (Table 3) and Q5 high-fidelity DNA polymerase (New England Biolabs [NEB]). The A-Tailing kit (NEB) was then used to add additional deoxyadenosine (A) to the 3' end of the PCR products. The insert possessing the additional A at the 3' end was ligated to the linearized vector with additional deoxythymidine (T) residues using T4 DNA ligase (NEB). The *plf* gene cluster from strain QT598 (*plf*_{QT598}) was amplified using primers CMD1847_F and CMD1900_R and cloned into vector pUCm-T (Bio Basic, Markham, ON, Canada), generating plasmid pIJ507. This plasmid contains the full *plf* gene cluster with a PIfG class II adhesin. The *plf* gene cluster from strain UMEA-3703-1 (*plf*_{UMEA}) was amplified using primers CMD2119_F and CMD2120_R and cloned into vector pBC sk+, generating plasmid pIJ523. This plasmid contains the full *plf*_{UMEA} gene cluster with a PIfG class I adhesin. To generate chimeric gene clusters comprised of *plf*_{QT598} with different types of G adhesin-encoding genes, pIJ507 was used as a template. *plfG*_{QT598} was deleted using an inverse-PCR method with primers CMD2168_F and CMD2169_R, which introduced PmeI sites and amplified a linear fragment lacking the *plfG*_{QT598} gene. The linearized product was then treated with DpnI endonuclease (NEB) to cleave any methylated template sequence. The linear fragment was either ligated using T4 DNA ligase (NEB) to generate pIJ598, which contains *plfBAHCDJKEF* (*plf*_{QT598} Δ *plfG*), or

TABLE 3 Primers used in this study

Primers	Characteristic(s) ^a	Sequence
CMD1847	P1f _{QT598} _cloning_F	AGCTTAGCGGCCGCATCCGCACAAACGGTCTTAC
CMD1900	P1f _{QT598} _cloning/screening_R	ATGAACGGGCCACCCGCATGAACATTCTCC
CMD2119	P1f _{UMEA-3703-01} _cloning_F	TCCCCCGGGTCGAGGAATTCGAGGGAGGGCGTGAATTCTG
CMD2120	P1f _{UMEA-3703-1} _cloning_R	GGCGAATTGGTACCGGGCCCTCGCAGATGTCACCG
CMD2171	PapG1 _{J96} _cloning_F	CGATGATGAAGGTTTATGAAAAATGGTTCCTGCTT
CMD2172	PapG1 _{J96} _cloning_R	CATAATAAAAAATGTTTTCAGGGGAAACTCAGAACCA
CMD2174	PapG11 _{CFT073} _cloning_F	CGATGATGAAGGTTTATGAAAAATGGTTCCTGCTT
CMD2175	PapG11 _{CFT073} _cloning_R	CATAATAAAAAATGTTTTCAGGCAATATCATGAGCAGC
CMD2177	PrsG _{J96} _cloning_F	CGATGATGAAGGTTTATGAAAAATGGTTCCTGCTT
CMD2178	PrsG _{J96} _cloning_R	CATAATAAAAAATGTTTTCAGGCAATATCATGAGCAGCG
CMD2180	P1fG1 _{UMEA-3703-1} _cloning_F	CGATGATGAAGGTTTATGAAAAAGGTTATCCTTTTGTCTATTG
CMD2181	P1fG1 _{UMEA-3703-1} _cloning_R	CATAATAAAAAATGTTTTCATTAATATCAACCTTTTAAACAGCGC
CMD2168	Delete P1fG1 _{QT598} _KO_F	ACGCTAACTCAGTTTAAACATTTTTATTATGATGTAAATATTTGTGCGCTTTTG
CMD2169	Delete P1fG1 _{QT598} _KO_R	AAACGTGAGTTAGCGTTTAAACCTTACATATCGGATCATAAAAAACGCACGGGTGAC
CMD2112	P1f _{QT598/UMEA-3703-1} _KO_F	AGTAATAACTGACAGGATATTTTAACTATAATCAGGAGGTTATTTCCATGGTGAGGCTGGAGCTGCTTC
CMD2113	P1f _{QT598} _KO_R	CCCGACATGAACATTCCTCAGACTATATTACAGGAGAATTCACCAAGTTCATGGGAATTAGCCATGGTCC
CMD2114	P1f _{UMEA-3703-1} _KO_R	TGCTACGTGCCATCTCATGCTCTTTACTTTCTCTCTGGTTAATATCAAAAAATGGGAATTAGCCATGGTCC
CMD1849	P1f _{QT598} _screening_F	AGATGGGATCCCAACAACCAAGGTCGCTCAGGGC
CMD2115	P1f _{UMEA-3703-1} _screening_F	ATGTCCGTTGAGCACTTTCG
CMD2116	P1f _{UMEA-3703-1} _screening_R	TGAAATCACCCCTATGCACA
CMD2186	qPCR P1fA_F	CGGATCAGGGACAAGGTAAAG
CMD2187	qPCR P1fA_R	CAGCCAGATGAGCTTTGG

^aF, forward; R, reverse; KO, knockout.

used as a template to generate chimeric fimbrial gene clusters containing different G adhesin genes using the T4 DNA ligase (NEB). PCR fragments containing G adhesin genes were obtained using primer pairs CMD2171_F and CMD2172_R (for *papG* class I from strain J96), CMD2174_F and CMD2175_R (for *papG* class II from strain CFT073), CMD2177_F and CMD2178_R (for *prsG* [papG class III] from strain J96), and CMD2180_F and CMD2181_R (for *p1fG*_{UMEA-3703-1}; *p1fG1* from strain UMEA-3703-1). Cloning experiments to generate recombinant plasmids or subclones were first achieved using *E. coli* strain DH5 α . The plasmids were extracted using a Miniprep kit according to the manufacturer's recommendations (Bio Basic, Inc.) and then transformed into *E. coli* *fin*-negative strain ORN172 for phenotypic testing. Strains containing the following reference plasmids that contain full P fimbrial gene clusters were used as reference controls: pPap5 (encoding P fimbriae PapG class I from *E. coli* J96) (42, 43), pDC5 (encoding P fimbriae PapG class II from strain IA2) (31), and pJFK102 (encoding Prs fimbria PrsG [PapG class III] from *E. coli* J96) (44, 45).

Deletion of the *p1f* genes from strains QT598 and UMEA-3703-1. A *p1f* knockout mutant of APEC strain QT598 was obtained by the lambda red recombinase method (50). First, plasmid pKD46 expressing lambda red recombinase was transformed into QT598 by electroporation, and then the kanamycin resistance cassette was amplified from plasmid pKD4 by PCR with primers CMD2112_F and CMD2113_R and transformed into QT598 carrying plasmid pKD46 by electroporation. Mutants were selected at 37°C, and then the loss of genes was confirmed by PCR using screening primers CMD1849_F and CMD1900_R to obtain the QT598 $\Delta p1f$ strain (QT4420). The same method was used to create the $\Delta p1f$ deletion mutation in UMEA-3703-1 using primers CMD2112_F and CMD2114_R to generate strain QT4598 (UMEA-3703-1 $\Delta p1f$). The deletion was confirmed using primers CMD2115_F and CMD2116_R.

Extraction of fimbriae and Western blotting. Fimbriae were extracted using the heat extraction method as described previously by (51), with some modifications. Briefly, overnight cultures were incubated at 56°C for 1 h and harvested by centrifugation at 4,000 rpm for 15 min. Supernatants were incubated in 10% trichloroacetic acid (TCA) to precipitate proteins. Proteins were then concentrated by centrifugation at 12,000 rpm for 20 min, washed twice with Tris-EDTA (0.05 M) at pH 12 and 8.5, and resuspended in 0.1 mL of Tris-EDTA (0.05 M) at pH 8.5. Western blotting was performed as previously described (52), with some modifications. Proteins were separated using 15% polyacrylamide gel, transferred onto a nitrocellulose membrane (Bio-Rad Laboratories, CA, USA), and blocked with 15 mL of blocking buffer TBST (0.15 M NaCl, 0.025 M Tris, 0.05% Tween, 3% skim milk) for 1 h at 4°C. Fimbrial major subunit protein was detected using rabbit polyclonal antibodies provided by New England Peptide (1:1,000) against a peptide corresponding to part of the P1fA major fimbrial subunit (Ac-CAHLAADGISVKKD-amide) for 45 min at room temperature. The membrane was then washed 3 times with wash buffer (0.15 M NaCl, 0.025 M Tris, 0.05% Tween) and incubated with an anti-rabbit-conjugated secondary antibody (1:20,000) for 45 min at room temperature. After four washes with TBST, proteins were detected using SuperSignal West Pico chemiluminescent substrate (Pierce) according to the manufacturer's instructions.

Transmission electron microscopy. Bacteria for electron microscopy were grown overnight at 37°C. Cultures were then adsorbed onto a glow-discharged Formvar-coated copper grid for 1 min and stained with 1% phosphotungstic acid. The excess of liquid was removed with a filter paper. Samples were then dried and observed under a Hitachi H700 transmission electron microscope.

Hemagglutination assays. Hemagglutination was performed in 96-well round-bottom plates as described in reference 53. Briefly, different types of blood were tested for this assay, human (A and O), horse, bovine, sheep, pig, rabbit, chicken, turkey, and dog red blood cells (RBCs) were suspended in phosphate-buffered saline (PBS) at a final concentration of 3% and added to 96-well plates. Clones expressing different classes of PapG were grown in LB broth at 37°C and centrifuged at $3,000 \times g$ for 15 min, and pellets were suspended in PBS (pH 7.4) and adjusted to an optical density at 600 nm (OD_{600}) of 0.6, centrifuged, and then concentrated 100-fold in PBS (pH 7.4). The agglutinating titer was determined as the most diluted well with agglutination after 30 min of incubation on ice.

To investigate potential inhibition of hemagglutination by carbohydrates known to be recognized by P fimbriae, a macroagglutination assay was performed. Overnight bacterial cultures were centrifuged and concentrated $10 \times$ in PBS. Thirty microliters of 6% human O blood group or turkey erythrocyte suspensions were deposited into each well of a 10-well glass plate. Two microliters of the different inhibitors was added to the erythrocytes. Finally, 30 μ L of the different control and test strains was mixed using toothpicks. Samples were left on ice for 30 min followed by gentle agitation, then hemagglutination levels were noted. The carbohydrates used for inhibition tests were ceramide trihexosides (CTH) (globotriaosylceramides [Gb3]) and globotetrahexosylceramide-globosides (Gb4) obtained from Matreya (State College, PA, USA), and Forsman antigen pentaose (Gb5) from Biosynth Carbosynth (San Diego, CA, USA). Stock solutions of the carbohydrates were prepared in dimethyl sulfoxide (DMSO). Final concentrations used for hemagglutination assays were as follows: Gb3, 32 μ g/mL or 1.8 nmol per well; Gb4, 161 mg/mL or 7.5 nmol per well; and Gb5, 161 μ g/mL or 11 nmol per well. Samples were tested in duplicate and gave similar results.

Biofilm assays. Biofilm formation in 96-well microtiter plates was examined as previously described (54). Fimbrial clones were grown statically in LB at 25, 30, 37, and 42°C for 48 h. After 48 h of incubation, the liquid was discarded, and plates were washed and stained with 0.1% crystal violet (Sigma) for 15 min. Ethanol-acetone solution (80:20) was used to dissolve biofilm, and the optical density was measured at 595 nm to determine the production of biofilm.

Bacterial adherence to epithelial cell lines. Cells of the human bladder 5637 (ATCC HTB-9) and kidney HEK-293 (ATCC CRL-1573) epithelial cell lines were grown to confluence in 24-well plates in RPMI 1640 or Eagle's minimal essential medium (EMEM) (Wisent Bio Products, St-Bruno, Canada) supplemented with 10% fetal bovine serum (FBS) at 37°C in 5% CO_2 . Fimbrial clones expressing different classes of PapG or PIFG adhesins were grown in LB medium at 37°C, cultures were then centrifuged, resuspended in RPMI 1640 or EMEM with 10% FBS, and added to cells at a multiplicity of infection (MOI) of 10 for 2 h, as described previously (55). After 2 h, cells were washed three times with PBS, lysed with 1% Triton X-100, diluted, and plated onto LB agar plates supplemented with selective antibiotics.

Murine urinary tract infection models. To determine the potential role of PL fimbriae in virulence, wild-type strains QT598 and UMEA-3703-1 as well as the QT598 Δplf (QT4420) and UMEA-3703-1 Δplf (QT4598) mutants were tested in 6-week-old CBA/J female mice using an ascending UTI model adapted from reference 38. A total of 5 mice in each group were infected with 10^9 CFU/mL of bacteria. After 48 h, the infected mice were euthanized, and bladders and kidneys were harvested aseptically for the bacterial count on MacConkey's agar plates. To study the expression of *plf* *in vivo*, bladder samples after necropsy were homogenized with TRIzol LS reagent (Thermo Fisher Scientific) for RNA extractions.

A competitive infection model was also used as described in reference 27. Briefly, a murine ascending UTI model with 9 mice in each group was used for coinfection, in which a virulent $\Delta lacZ$ derivative of QT598 (QT4567) was coinfecting with the Δplf strain (QT4420). Twenty-five microliters (10^9 CFU) of a mixed culture containing equal amounts of each strain were inoculated through a catheter in 6-week-old CBA/J female mice. Mice were euthanized after 48 h, and bladders and kidneys were harvested aseptically, homogenized, diluted, and plated on MacConkey agar plates.

qRT-PCR to measure PL fimbrial gene expression levels. We compared the expression of PL fimbriae by comparing RNA levels of the *plfA* gene under different conditions: LB medium, minimal M63 medium, and during infection in bladders of mice. For *in vitro* analysis, total RNAs from bacterial samples were extracted according to the manufacturer's protocol EZ-10 spin column total RNA Miniprep kit (BioBasic). For *in vivo* analysis, bladder samples were homogenized with TRIzol LS reagent (Thermo Fisher Scientific), incubated with chloroform followed by centrifugation and incubation in ethanol (95 to 100%) to separate the aqueous phase that contains RNA. Then, RNA samples were extracted using a Direct-zol RNA Miniprep kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's recommendations. All RNA samples were treated with Turbo DNase (Ambion), to eliminate any DNA contamination. The Iscript reverse transcription supermix (Bio-Rad Life Science, Mississauga, ON, Canada) was used to synthesize cDNAs from samples according to the manufacturer's protocol. Primers were specific to the *plfA* gene and the RNA polymerase sigma factor *rpoD* (housekeeping control). Quantitative reverse transcription-PCR (qRT-PCR) was performed in the Corbett Rotorgene (Thermo Fisher) instrument using 50 ng of cDNA, 100 nM each primer, and 10 μ L of SsoFast Evagreen supermix (Bio-Rad). Data were analyzed using the threshold cycle ($2^{-\Delta\Delta CT}$) method (56).

Statistical analyses. All data were analyzed with the Graph Pad Prism 6 software (GraphPad Software, San Diego, CA, USA). A Mann-Whitney test was used for mouse infection experiments to determine statistical significance. Analysis of variance (ANOVA) was used to compare the means of samples. Differences between groups were considered significant for P values of <0.05 .

Ethics. Protocols for mouse urinary tract infection were approved by the animal ethics evaluation committee (Comité Institutionnel de Protection des Animaux; CIPA no. 1608-02) of the INRS-Centre Armand-Frappier Santé Biotechnologie.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.3 MB.

SUPPLEMENTAL FILE 3, PDF file, 1.3 MB.

ACKNOWLEDGMENTS

We thank James R. Johnson, University of Minnesota, for providing reference clones carrying different classes of Pap and related fimbrial adhesins and Niels Frimodt-Møller, University of Copenhagen, for providing UPEC strain UMEA-3703-1. We thank Micheline Letarte and Arnaldo Nakamura for assistance with electron microscopy.

Funding for this work was supported by Natural Sciences and Engineering Research Council (NSERC) Canada Discovery Grant 2019-06642.

NSERC funds were for a Discovery grant to C.M. Dozois. FRQNT funds for the multi institutional Swine and poultry infectious diseases research centre (CRIPA) grant.

REFERENCES

- Soto GE, Hultgren SJ. 1999. Bacterial adhesins: common themes and variations in architecture and assembly. *J Bacteriol* 181:1059–1071. <https://doi.org/10.1128/JB.181.4.1059-1071.1999>.
- Nuccio SP, Baumberg AJ. 2007. Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol Mol Biol Rev* 71:551–575. <https://doi.org/10.1128/MMBR.00014-07>.
- Werneburg GT, Thanassi DG. 13 March 2018, posting date. Pili assembled by the chaperone/usher pathway in *Escherichia coli* and *Salmonella*. *EcoSal Plus* 2018 <https://doi.org/10.1128/ecosalplus.ESP-0007-2017>.
- Wurpel DJ, Beatson SA, Totsika M, Petty NK, Schembri MA. 2013. Chaperone-usher fimbriae of *Escherichia coli*. *PLoS One* 8:e52835. <https://doi.org/10.1371/journal.pone.0052835>.
- Ambite I, Butler DSC, Stork C, Gronberg-Hernandez J, Kovcs B, Zdzinski J, Pinkner J, Hultgren SJ, Dobrindt U, Wullt B, Svanborg C. 2019. Fimbriae reprogram host gene expression—divergent effects of P and type 1 fimbriae. *PLoS Pathog* 15:e1007671. <https://doi.org/10.1371/journal.ppat.1007671>.
- Connell I, Agace W, Klemm P, Schembri M, Märlid S, Svanborg C. 1996. Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proc Natl Acad Sci U S A* 93:9827–9832. <https://doi.org/10.1073/pnas.93.18.9827>.
- Sokurenko EV, Courtney HS, Abraham SN, Klemm P, Hasty DL. 1992. Functional heterogeneity of type 1 fimbriae of *Escherichia coli*. *Infect Immun* 60:4709–4719. <https://doi.org/10.1128/iai.60.11.4709-4719.1992>.
- Lillington J, Geibel S, Waksman G. 2014. Biogenesis and adhesion of type 1 and P pili. *Biochim Biophys Acta* 1840:2783–2793. <https://doi.org/10.1016/j.bbagen.2014.04.021>.
- Gunther NW, IV, Snyder JA, Lockatell V, Blomfield I, Johnson DE, Mobley HL. 2002. Assessment of virulence of uropathogenic *Escherichia coli* type 1 fimbrial mutants in which the invertible element is phase-locked on or off. *Infect Immun* 70:3344–3354. <https://doi.org/10.1128/IAI.70.7.3344-3354.2002>.
- Khan NA, Kim Y, Shin S, Kim KS. 2007. FimH-mediated *Escherichia coli* K1 invasion of human brain microvascular endothelial cells. *Cell Microbiol* 9:169–178. <https://doi.org/10.1111/j.1462-5822.2006.00779.x>.
- Plos K, Carter T, Hull S, Hull R, Svanborg EC. 1990. Frequency and organization of pap homologous DNA in relation to clinical origin of uropathogenic *Escherichia coli*. *J Infect Dis* 161:518–524. <https://doi.org/10.1093/infdis/161.3.518>.
- Dezfulian H, Batisson I, Fairbrother JM, Lau PC, Nassar A, Szatmari G, Harel J. 2003. Presence and characterization of extraintestinal pathogenic *Escherichia coli* virulence genes in F165-positive *E. coli* strains isolated from diseased calves and pigs. *J Clin Microbiol* 41:1375–1385. <https://doi.org/10.1128/JCM.41.4.1375-1385.2003>.
- Dozois CM, Pourbakhsh SA, Fairbrother JM. 1995. Expression of P and type 1 (F1) fimbriae in pathogenic *Escherichia coli* from poultry. *Vet Microbiol* 45:297–309. [https://doi.org/10.1016/0378-1135\(94\)00127-i](https://doi.org/10.1016/0378-1135(94)00127-i).
- Kariyawasam S, Johnson TJ, Nolan LK. 2006. The pap operon of avian pathogenic *Escherichia coli* strain O1:K1 is located on a novel pathogenicity island. *Infect Immun* 74:744–749. <https://doi.org/10.1128/IAI.74.1.744-749.2006>.
- Mellata M, Dho-Moulin M, Dozois CM, Curtiss R, III, Lehoux B, Fairbrother JM. 2003. Role of avian pathogenic *Escherichia coli* virulence factors in bacterial interaction with chicken heterophils and macrophages. *Infect Immun* 71:494–503. <https://doi.org/10.1128/IAI.71.1.494-503.2003>.
- van den Bosch JF, Hendriks JH, Gladigau I, Willems HM, Storm PK, de Graaf FK. 1993. Identification of F11 fimbriae on chicken *Escherichia coli* strains. *Infect Immun* 61:800–806. <https://doi.org/10.1128/iai.61.3.800-806.1993>.
- Blum G, Falbo V, Caprioli A, Hacker J. 1995. Gene clusters encoding the cytotoxic necrotizing factor type 1, Prs-fimbriae and alpha-hemolysin form the pathogenicity island II of the uropathogenic *Escherichia coli* strain J96. *FEMS Microbiol Lett* 126:189–195. <https://doi.org/10.1111/j.1574-6968.1995.tb07415.x>.
- Guyer DM, Kao JS, Mobley HL. 1998. Genomic analysis of a pathogenicity island in uropathogenic *Escherichia coli* CFT073: distribution of homologous sequences among isolates from patients with pyelonephritis, cystitis, and catheter-associated bacteriuria and from fecal samples. *Infect Immun* 66:4411–4417. <https://doi.org/10.1128/IAI.66.9.4411-4417.1998>.
- Johnson JR, Stell AL, Scheutz F, O'Bryan TT, Russo TA, Carlino UB, Fasching C, Kavle J, Van Dijk L, Gaastra W. 2000. Analysis of the F antigen-specific papA alleles of extraintestinal pathogenic *Escherichia coli* using a novel multiplex PCR-based assay. *Infect Immun* 68:1587–1599. <https://doi.org/10.1128/IAI.68.3.1587-1599.2000>.
- Marklund BI, Tennent JM, Garcia E, Hamers A, Baga M, Lindberg F, Gaastra W, Normark S. 1992. Horizontal gene transfer of the *Escherichia coli* pap and prs pili operons as a mechanism for the development of tissue-specific adhesive properties. *Mol Microbiol* 6:2225–2242. <https://doi.org/10.1111/j.1365-2958.1992.tb01399.x>.
- Stromberg N, Marklund BI, Lund B, Ilver D, Hamers A, Gaastra W, Karlsson KA, Normark S. 1990. Host-specificity of uropathogenic *Escherichia coli* depends on differences in binding specificity to Gal alpha 1-4Gal-containing isoreceptors. *EMBO J* 9:2001–2010. <https://doi.org/10.1002/j.1460-2075.1990.tb08328.x>.
- Manning SD, Zhang L, Foxman B, Spindler A, Tallman P, Marrs CF. 2001. Prevalence of known P-fimbrial G alleles in *Escherichia coli* and identification of a new adhesin class. *Clin Diagn Lab Immunol* 8:637–640. <https://doi.org/10.1128/CDLI.8.3.637-640.2001>.
- Lugering A, Benz I, Knochenhauer S, Ruffing M, Schmidt MA. 2003. The Pix pilus adhesin of the uropathogenic *Escherichia coli* strain X2194 (O2:K(-):H6) is related to Pap pili but exhibits a truncated regulatory region. *Microbiology (Reading)* 149:1387–1397. <https://doi.org/10.1099/mic.0.26266-0>.
- Schneider G, Dobrindt U, Bruggemann H, Nagy G, Janke B, Blum-Oehler G, Buchrieser C, Gottschalk G, Emody L, Hacker J. 2004. The pathogenicity island-associated K15 capsule determinant exhibits a novel genetic structure and correlates with virulence in uropathogenic *Escherichia coli* strain 536. *Infect Immun* 72:5993–6001. <https://doi.org/10.1128/IAI.72.10.5993-6001.2004>.
- Bielaszewska M, Prager R, Vandivinit L, Musken A, Mellmann A, Holt NJ, Tarr PI, Karch H, Zhang W. 2009. Detection and characterization of the fimbrial sfp cluster in enterohemorrhagic *Escherichia coli* O165:H25/NM isolates from humans and cattle. *Appl Environ Microbiol* 75:64–71. <https://doi.org/10.1128/AEM.01815-08>.
- Brunder W, Khan AS, Hacker J, Karch H. 2001. Novel type of fimbriae encoded by the large plasmid of sorbitol-fermenting enterohemorrhagic

- Escherichia coli O157:H(-). Infect Immun 69:4447–4457. <https://doi.org/10.1128/IAI.69.7.4447-4457.2001>.
27. Habouria H, Pokharel P, Maris S, Garenaux A, Bessaiah H, Houle S, Veyrier FJ, Guyomard-Rabenirina S, Talamain A, Dozois CM. 2019. Three new serine-protease autotransporters of Enterobacteriaceae (SPATEs) from extra-intestinal pathogenic Escherichia coli and combined role of SPATEs for cytotoxicity and colonization of the mouse kidney. Virulence 10:568–587. <https://doi.org/10.1080/21505594.2019.1624102>.
 28. Koraimann G. 17 July 2018, posting date. Spread and persistence of virulence and antibiotic resistance genes: a ride on the F plasmid conjugation module. EcoSal Plus 2018 <https://doi.org/10.1128/ecosalplus.ESP-0003-2018>.
 29. Girardeau JP, Bertin Y, Callebaut I. 2000. Conserved structural features in class I major fimbrial subunits (pili) in Gram-negative bacteria. Molecular basis of classification in seven subfamilies and identification of intrasubfamily sequence signature motifs which might be implicated in quaternary structure. J Mol Evol 50:424–442. <https://doi.org/10.1007/s002390010045>.
 30. Holden NJ, Uhlin BE, Gally DL. 2001. PapB paralogues and their effect on the phase variation of type 1 fimbriae in Escherichia coli. Mol Microbiol 42:319–330. <https://doi.org/10.1046/j.1365-2958.2001.02656.x>.
 31. Clegg S. 1982. Cloning of genes determining the production of mannose-resistant fimbriae in a uropathogenic strain of Escherichia coli belonging to serogroup O6. Infect Immun 38:739–744. <https://doi.org/10.1128/iai.38.2.739-744.1982>.
 32. Lanza VF, de Toro M, Garcillan-Barcia MP, Mora A, Blanco J, Coque TM, de la Cruz F. 2014. Plasmid flux in Escherichia coli ST131 sublineages, analyzed by plasmid constellation network (PLACNET), a new method for plasmid reconstruction from whole genome sequences. PLoS Genet 10:e1004766. <https://doi.org/10.1371/journal.pgen.1004766>.
 33. Lee IPA, Andam CP. 2019. Pan-genome diversification and recombination in Cronobacter sakazakii, an opportunistic pathogen in neonates, and insights to its xerotolerant lifestyle. BMC Microbiol 19:306. <https://doi.org/10.1186/s12866-019-1664-7>.
 34. Healy B, Cooney S, O'Brien S, Iversen C, Whyte P, Nally J, Callanan JJ, Fanning S. 2010. Cronobacter (Enterobacter sakazakii): an opportunistic foodborne pathogen. Foodborne Pathog Dis 7:339–350. <https://doi.org/10.1089/fpd.2009.0379>.
 35. Roberts JA, Marklund BI, Ilver D, Haslam D, Kaack MB, Baskin G, Louis M, Mollby R, Winberg J, Normark S. 1994. The Gal(alpha 1-4)Gal-specific tip adhesin of Escherichia coli P-fimbriae is needed for pyelonephritis to occur in the normal urinary tract. Proc Natl Acad Sci U S A 91:11889–11893. <https://doi.org/10.1073/pnas.91.25.11889>.
 36. Rice JC, Peng T, Spence JS, Wang HQ, Goldblum RM, Corthesy B, Nowicki BJ. 2005. Pyelonephritic Escherichia coli expressing P fimbriae decrease immune response of the mouse kidney. J Am Soc Nephrol 16:3583–3591. <https://doi.org/10.1681/ASN.2005.030243>.
 37. O'Hanley P, Lark D, Falkow S, Schoolnik G. 1985. Molecular basis of Escherichia coli colonization of the upper urinary tract in BALB/c mice. Gal-Gal pili immunization prevents Escherichia coli pyelonephritis in the BALB/c mouse model of human pyelonephritis. J Clin Invest 75:347–360. <https://doi.org/10.1172/JCI111707>.
 38. Hagberg L, Hull R, Hull S, Falkow S, Freter R, Svanborg EC. 1983. Contribution of adhesion to bacterial persistence in the mouse urinary tract. Infect Immun 40:265–272. <https://doi.org/10.1128/iai.40.1.265-272.1983>.
 39. Mobley HL, Jarvis KG, Elwood JP, Whittle DI, Lockett CV, Russell RG, Johnson DE, Donnenberg MS, Warren JW. 1993. Isogenic P-fimbrial deletion mutants of pyelonephritogenic Escherichia coli: the role of alpha Gal(1-4) beta Gal binding in virulence of a wild-type strain. Mol Microbiol 10:143–155. <https://doi.org/10.1111/j.1365-2958.1993.tb00911.x>.
 40. Marc D, Dho-Moulin M. 1996. Analysis of the fim cluster of an avian O2 strain of Escherichia coli: serogroup-specific sites within fimA and nucleotide sequence of fimL. J Med Microbiol 44:444–452. <https://doi.org/10.1099/00222615-44-6-444>.
 41. Mobley HL, Green DM, Trifillis AL, Johnson DE, Chippendale GR, Lockett CV, Jones BD, Warren JW. 1990. Pyelonephritogenic Escherichia coli and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. Infect Immun 58:1281–1289. <https://doi.org/10.1128/iai.58.5.1281-1289.1990>.
 42. Hull RA, Gill RE, Hsu P, Minshew BH, Falkow S. 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection Escherichia coli isolate. Infect Immun 33:933–938. <https://doi.org/10.1128/iai.33.3.933-938.1981>.
 43. Lindberg FP, Lund B, Normark S. 1984. Genes of pyelonephritogenic E. coli required for digalactoside-specific agglutination of human cells. EMBO J 3:1167–1173. <https://doi.org/10.1002/j.1460-2075.1984.tb01946.x>.
 44. Karr JF, Nowicki B, Truong LD, Hull RA, Hull SI. 1989. Purified P fimbriae from two cloned gene clusters of a single pyelonephritogenic strain adhere to unique structures in the human kidney. Infect Immun 57:3594–3600. <https://doi.org/10.1128/iai.57.11.3594-3600.1989>.
 45. Lindstedt R, Baker N, Falk P, Hull R, Hull S, Karr J, Leffler H, Svanborg Edén C, Larson G. 1989. Binding specificities of wild-type and cloned Escherichia coli strains that recognize globo-A. Infect Immun 57:3389–3394. <https://doi.org/10.1128/iai.57.11.3389-3394.1989>.
 46. Dereeper A, Guignon V, Blanc G, Audic S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36:W465–W469. <https://doi.org/10.1093/nar/gkn180>.
 47. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
 48. Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17:540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>.
 49. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. <https://doi.org/10.1093/sysbio/syq010>.
 50. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 97:6640–6645. <https://doi.org/10.1073/pnas.120163297>.
 51. Lymberopoulos MH, Houle S, Daigle F, Leveille S, Bree A, Moulin-Schouleur M, Johnson JR, Dozois CM. 2006. Characterization of Stg fimbriae from an avian pathogenic Escherichia coli O78:K80 strain and assessment of their contribution to colonization of the chicken respiratory tract. J Bacteriol 188:6449–6459. <https://doi.org/10.1128/JB.00453-06>.
 52. Crépin S, Lamarche MG, Garneau P, Séguin J, Proulx J, Dozois CM, Harel J. 2008. Genome-wide transcriptional response of an avian pathogenic Escherichia coli (APEC) pst mutant. BMC Genomics 9:568–571. <https://doi.org/10.1186/1471-2164-9-568>.
 53. Provence DL, Curtis R. 1994. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic Escherichia coli strain. Infect Immun 62:1369–1380. <https://doi.org/10.1128/iai.62.4.1369-1380.1994>.
 54. Genevoux P, Muller S, Bauda P. 1996. A rapid screening procedure to identify mini-Tn10 insertion mutants of Escherichia coli K-12 with altered adhesion properties. FEMS Microbiol Lett 142:27–30. <https://doi.org/10.1111/j.1574-6968.1996.tb08402.x>.
 55. Matter LB, Barbieri NL, Nordhoff M, Ewers C, Horn F. 2011. Avian pathogenic Escherichia coli MT78 invades chicken fibroblasts. Vet Microbiol 148:51–59. <https://doi.org/10.1016/j.vetmic.2010.08.006>.
 56. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25:402–408. <https://doi.org/10.1006/meth.2001.1262>.
 57. Hochhut B, Wilde C, Balling G, Middendorf B, Dobrindt U, Brzuszkiewicz E, Gottschalk G, Carniel E, Hacker J. 2006. Role of pathogenicity island-associated integrases in the genome plasticity of uropathogenic Escherichia coli strain 536. Mol Microbiol 61:584–595. <https://doi.org/10.1111/j.1365-2958.2006.05255.x>.
 58. Jacob-Dubuisson F, Heuser J, Dodson K, Normark S, Hultgren S. 1993. Initiation of assembly and association of the structural elements of a bacterial pilus depend on two specialized tip proteins. EMBO J 12:837–847. <https://doi.org/10.1002/j.1460-2075.1993.tb05724.x>.
 59. Kuehn MJ, Heuser J, Normark S, Hultgren SJ. 1992. P pili in uropathogenic E. coli are composite fibres with distinct fibrillar adhesive tips. Nature 356:252–255. <https://doi.org/10.1038/356252a0>.
 60. Woodall LD, Russell PW, Harris SL, Orndorff PE. 1993. Rapid, synchronous, and stable induction of type 1 piliation in Escherichia coli by using a chromosomal lacUV5 promoter. J Bacteriol 175:2770–2778. <https://doi.org/10.1128/jb.175.9.2770-2778.1993>.
 61. Dho M, Lafont JP. 1982. Escherichia coli colonization of the trachea in poultry: comparison of virulent and avirulent strains in gnotoxenic chickens. Avian Dis 26:787–797. <https://doi.org/10.2307/1589865>.

DISCUSSION GÉNÉRALE

7 DISCUSSION GÉNÉRALE

Dans cette étude, nous avons caractérisé des nouveaux facteurs de virulence qui ont été identifiés pour la première fois chez la souche APEC QT598 (O1:K1, ST1385). Cette souche, isolée en France en 1982 à partir du foie d'un dindonneau atteint de colibacillose, est capable de produire le sidérophore yersiniabactine en très grande quantité. Suite à ces résultats, nous avons procédé au séquençage de son génome et avons révélé la présence d'un grand plasmide de virulence d'environ 200 kb, que nous avons nommé pEC598. Ce plasmide possède toute la machinerie nécessaire pour son transfert conjugatif, ainsi que les gènes typiques de virulence associés aux souches APEC comme les gènes codant pour la colicine V, l'hémolysine *hlyF*, l'AT *tsh*, et les sidérophores : salmochélines et l'aérobactine (non fonctionnel parce qu'il est interrompu par un codon stop au niveau du gène de synthèse *iucD*) (**figure 17**).

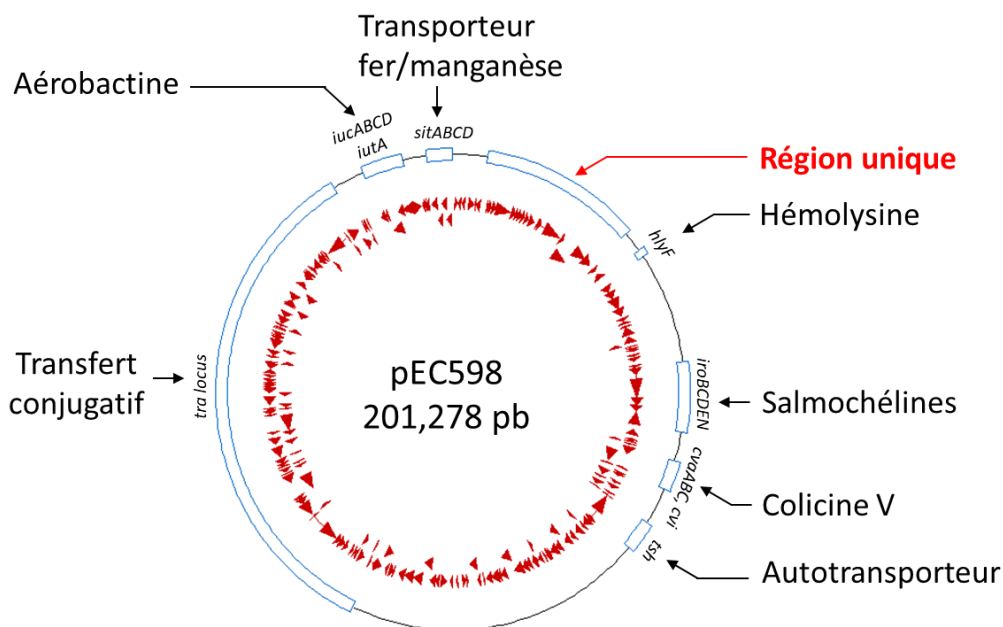


Figure 17: Représentation schématique du plasmide de virulence pEC598 retrouvé chez la souche QT598.

Le plasmide code pour plusieurs facteurs de virulence, tels que l'hémolysine (HlyF), la colicine V, le transporteur Fer/Manganèse (Sit), l'AT (Tsh), les sidérophores (salmochélines et l'aérobactine) et la région unique (résultats non publiés, figure créée par notre équipe).

En plus de ces facteurs, nous avons pu identifier une région d'environ 25 kb dite « région unique » dans laquelle se trouve les nouveaux gènes qui codent pour un P-like fimbriae (PL) et un serine-protease hemagglutinin autotransporteur (Sha) et une nouvelle toxine Prt (**figure 18**).

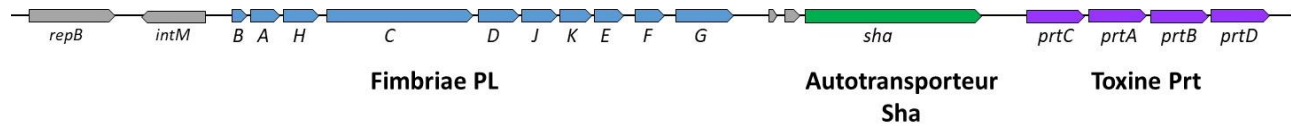


Figure 18: Organisation génétique de la région unique de 25kb retrouvée dans le plasmide de virulence pEC598.

La région contient des nouveaux gènes qui codent pour un nouveau fimbriae PL, un nouvel autotransporteur Sha, et une nouvelle toxine Prt (résultats non publiés, figure créée par notre équipe).

L'identification de ces nouveaux facteurs soulève la question de leurs rôles dans la virulence et leurs prévalences chez d'autres souches ExPEC. Pour répondre à ces questions, nous avons fixé deux objectifs principaux qui ont été traités au cours des travaux réalisés pendant ce projet de thèse. Le premier objectif de ce travail était donc de caractériser l'AT Sha.

7.1 Autotransporteur Sha

7.1.1 Caractérisation des nouveaux membres de la famille des SPATEs

Les AT de la famille SPATE constituent un groupe de protéines extracellulaires produites par la plupart des *E. coli* pathogènes. Ils possèdent un domaine passager qui porte le site actif sérine protéase (Pokharel *et al.*, 2019, Ruiz-Perez *et al.*, 2014). Chez la souche QT598, nous avons pu identifier cinq gènes qui codent pour des ATs de la famille des SPATEs.

Deux nouveaux gènes codant pour *tagB* et *tagC* sont retrouvés sur un îlot de pathogénicité adjacent des gènes *yjdI* et *yjdK*. Cette région génomique est aussi bordée par des séquences répétées directes (DR) qui correspondent à la duplication du gène *yjdJ* bordant chaque côté de cet îlot. Ils sont constitués par :

- Un peptide signal (entre les résidus 1 et 58 pour TagB et entre les résidus 1 et 53 pour TagC),
- Un domaine passager (entre les résidus 59 et 1006 pour TagB et entre les résidus 54 et 1032 pour TagC) contenant le site actif sérine protéase GDSGS.

- Un site de clivage entre le domaine passager et domaine β qui est situé entre deux asparagines (EIN¹⁰⁰⁶NLNDRM pour TagB et EVN¹⁰³²NLNKRM pour TagC, respectivement).
- Un domaine β (entre les résidus 1007 et 1283 pour TagB et entre les résidus 1033 et 1309 pour TagC)

Les protéines TagB et TagC appartiennent aux classes II et classe I SPATE avec une homologie de séquence avec EaaA (46% d'identité) et EspC (60% d'identité) respectivement.

On trouve aussi le gène *tsh* qui est présent sur le plasmide de virulence, et le gène *vat* qui se trouve sur un îlot de pathogénicité dans le génome bactérien.

Dans cette étude, on va s'intéresser à caractériser l'AT Sha. Pour ce faire, nous avons d'abord identifié le gène qui se trouve dans la région unique au niveau du plasmide de virulence pEC598. Le gène *sha* code pour une protéine de 1302 acides aminés qui possède une structure générale typique d'un AT de la famille SPATE avec trois domaines : un peptide signal assez long (1 à 51 acides aminés), un domaine passager portant le site actif sérine protéase (²⁵⁶GDSGS) et un domaine β . Généralement, après translocation de l'AT dans le milieu externe, le domaine passager est clivé par un mécanisme d'autoprotéolyse qui implique la présence d'un site spécifique de reconnaissance (EVNNLNK) situé entre le domaine passager et le domaine β . Le clivage se fait alors entre les deux acides aminés polaires asparagines (NN) et le domaine passager qui porte le site actif sera libéré dans le milieu externe pour exercer sa fonction (Pokharel *et al.*, 2019). Dans le cas de l'AT Sha, nous avons montré que le site de clivage typique est absent. Une telle modification n'a pas eu d'impact sur la sécrétion de Sha dans le milieu externe, ceci a été vérifié par la détection de la protéine dans le surnageant des cultures bactériennes (par gel SDS PAGE). L'absence de ce site n'était pas spécifique à Sha. En effet, l'AT RpeA identifié chez la souche entéropathogène du lapin (REPEC) ne possède pas ce site et contrairement à Sha, son domaine passager reste lié à la bactérie (Leyton *et al.*, 2007). Par la suite, nous avons procédé aux analyses par spectrométrie de masse afin de déterminer le nouveau site de clivage tout en respectant les caractéristiques générales d'un AT de la famille SPATE (domaine β conservé et/ou clivage protéolytique entre deux acides aminés polaires). Ceci nous a permis de prédire le site de clivage potentiel qui est situé entre les deux acides aminés polaires Ser¹⁰¹⁵ et Asp¹⁰¹⁶. Par conséquent, nous avons pu prédire la taille du domaine passager (52 à 1015 acides aminés) et le domaine β (1016 à 1302).

Après détermination préalable des trois domaines de l'AT Sha, nous avons effectué des analyses de séquences afin de classer la protéine parmi la famille des SPATEs. Ceci nous a permis de montrer que Sha partage une forte homologie avec Tsh (66% d'identité), TleA (60%

d'identité) et Vat (54% d'identité). Ces trois ATs appartiennent aux classes II lectine-like immunomodulateurs (Ruiz-Perez *et al.*, 2014). Nous avons ainsi émis l'hypothèse que Sha pourrait appartenir aux classes II SPATEs.

Pour vérifier cette hypothèse nous avons procédé, dans un premier temps, à la prédiction de la structure 3D de l'AT Sha en la comparant avec le modèle déjà établi de Hbp/Tsh (PDB 1WXR). La structure 3D générée pour Sha a montré une grande similarité avec celle de Hbp/Tsh. De plus, nous avons pu identifier un « domaine-2 » qui est spécifique pour les ATs de la classe II SPATE. Ce domaine d'environ 75 acides aminés, retrouvé dans le domaine passager, possède une région qui ressemble au domaine de fixation de la chitine chez les chitinase (enzyme qui dégrade la chitine), ce qui pourrait jouer un rôle dans la reconnaissance du substrat (Ruiz-Perez *et al.*, 2014). Toutefois, le rôle attribué à ce domaine reste à vérifier puisqu'il n'est présent que chez les classes II SPATE.

Dans un second temps, nous avons construit un arbre phylogénétique en comparant les séquences en acide aminé du domaine passager de différents ATs déjà caractérisés (16 membres de classe I et classe II) avec celle de Sha. En se basant sur le pourcentage d'identité et les résultats générés par l'arbre phylogénétique, on peut confirmer notre hypothèse que Sha est relativement proche de Tsh, TleA puis Vat et donc il appartient à la classe II SPATE.

Après avoir identifié et classé l'AT Sha, il a été intéressant par la suite de déterminer sa prévalence parmi une sélection de 697 souches UPEC et 299 souches APEC (Dozois *et al.*, 2000, Guyomard-Rabenirina *et al.*, 2016) en utilisant des amorces spécifiques pour le gène en question. Ceci nous a permis d'identifier sa présence dans environ 20% des souches APEC (61 souches) dont 42 souches appartenant au phylogroupe A, 11 souches appartenant au phylogroupe B1, 5 souches appartenant au phylogroupe B2 et 3 souches appartenant au phylogroupe D. La majorité de ces souches était de sérotype O78 et d'autres sérotypes, moins fréquents, ont été aussi identifiés parmi lesquels O1, O11 O54, O21 et O8. Toutefois, la prévalence de l'AT Sha chez les souches UPEC était faible, seulement 6 souches UPEC de groupe B2 possédaient le gène dans leur génome. De même pour Tsh, sa prévalence chez les UPEC était très faible (5.9%) par rapport à celle de Vat (47.8%) qui était plus fréquente chez les souches UPEC. Ceci peut s'expliquer par le fait que les ATs Sha et Tsh sont portés par des plasmides de virulence de type ColV qui sont généralement moins fréquents chez les souches UPEC que les souches APEC (à noter que les plasmides ColV sont présents chez la majorité des souches APEC) (Timothy J Johnson *et al.*, 2006).

La prévalence des SPATEs chez les souches pathogènes explique bien leur importance dans la virulence. En effet, le domaine passager des ATs peut jouer un rôle important dans l'adhésion, l'invasion, l'auto-agrégation et la formation de biofilm. De plus, il peut avoir un effet protéolytique ou cytotoxique grâce à la présence d'une activité serine protéase portée par le site actif (Dautin, 2010).

Dans une seconde partie du projet de thèse, nous nous sommes intéressés à caractériser le rôle de Sha dans la virulence. Nous avons émis l'hypothèse qu'il pourrait avoir un effet adhérent et cytotoxique. Cette hypothèse est d'autant plus soutenue par le fait que l'AT Sha appartient à la classe II immunomodulateurs et possède une homologie avec Tsh et Vat, ainsi il pourrait probablement avoir le même rôle que Tsh et Vat.

Afin de confirmer cette hypothèse, nous avons procédé à une série de tests phénotypiques.

7.1.2 Rôle de l'AT Sha dans la virulence

La présence de plusieurs facteurs de virulence chez la souche QT598 (ATs, fimbriae, toxines, etc.) rend la détermination de l'effet de Sha *in vitro* difficile (à cause du phénomène de redondance). Par conséquent dans le but d'évaluer son rôle, nous avons décidé de cloner le gène dans une bactérie *E. coli* K-12 qui n'est pas pathogène et qui ne possède pas les facteurs de virulence décrits précédemment.

7.1.2.1 Effet adhérent de l'AT Sha

L'adhérence bactérienne est une propriété que possède *E. coli* afin de se lier à la surface des cellules de l'hôte et initier l'infection. Ce phénomène est généralement médié par les fimbriae. Cependant, certains ATs peuvent aussi jouer un rôle dans l'adhérence comme par exemple TleA qui a montré une capacité à adhérer aux cellules Caco-2 (Gutiérrez *et al.*, 2015). Dans cette étude, nous avons pu montrer que Vat et Tsh sont capables d'adhérer aux cellules rénales et de la vessie, respectivement. De même pour Sha, nous avons montré qu'il est capable d'adhérer aux cellules épithéliales humaines (vessie et reins) et aviaires ce qui peut suggérer son rôle dans l'adhérence de la souche QT598 au cours des premiers stades de l'infection.

Ensuite nous avons tenté de voir si Sha est impliqué dans l'hémagglutination qui pourrait être un phénotype secondaire découlant de l'adhésion comme l'AT Tsh qui est capable d'agglutiner les érythrocytes et les leucocytes de différentes espèces (Ayala-Lujan *et al.*, 2014,

Kostakioti *et al.*, 2004). D'après les essais d'hémagglutination que nous avons effectués, nous avons montré que Sha possède aussi une grande capacité à agglutiner les érythrocytes de neuf espèces différentes y compris les érythrocytes humains, d'où le nom « serine-protease hemagglutinin autotransporter, Sha ». De plus, Tsh et Vat ont montré une activité similaire, ils étaient capables d'agglutiner les érythrocytes de toutes les espèces testées. Cette forte affinité des ATs vis-à-vis des érythrocytes peut être un moyen utilisé par la bactérie pour séquestrer le « fer » présent dans l'hémoglobine. En effet, grâce à leur activité protéase, ces ATs peuvent d'abord se lier aux érythrocytes avec une forte affinité puis les dégrader par un mécanisme dépendant de leur activité sérine protéase afin de libérer le fer dans le milieu et faciliter leur utilisation par la bactérie. Ce phénomène a été suggéré pour Tsh/Hbp (Henderson *et al.*, 2004, Provence *et al.*, 1994), et d'après nos résultats, on peut suggérer que Sha et Vat sont aussi impliqués dans l'hémagglutination et pourraient avoir un rôle dans l'acquisition du fer.

En plus de son activité hémagglutinante, nous avons tenté d'évaluer si Sha est impliqué ou non dans l'auto-agrégation. C'est un mécanisme de défense, utilisé par les bactéries pour s'échapper de l'action lytique du complément, qui implique la formation d'agrégats observables dans une culture statique. Ainsi, lors des essais *in vitro*, seules les bactéries exprimant Sha forment des agrégations au fond du tube ce qui suggère qu'il interagit directement avec lui-même pour former ces agrégats. Cet effet a été observé notamment chez l'AT adhesin involved in diffuse adherence (AIDA) qui est un AT monomérique associé à la pathogenèse des souches d'*E.coli* par des mécanismes qui impliquent l'adhésion et l'invasion de cellules épithéliales, l'auto-agrégation des bactéries et la formation de biofilm (Charbonneau *et al.*, 2007). Par association, nous avons vérifié si les propriétés d'auto-agrégation de Sha permettent la formation de biofilm, ce qui constitue aussi une forme de résistance adoptée par les bactéries pour survivre dans des conditions environnementales hostiles. Ce mécanisme implique quatre étapes principales : (1) attachement des bactéries à une surface abiotique ; (2) interaction bactéries-bactéries pour former des microcolonies et produire des exo-polysaccharides ; (3) multiplication bactérienne et maturation du biofilm ; (4) détachement et dispersion des bactéries pour démarrer un nouveau cycle (Van Houdt *et al.*, 2005). Ainsi, nous avons pu observer que Sha, Tsh et Vat formait des biofilms à des températures qui ne dépasse pas 30°C suggérant qu'ils sont impliqués dans l'interaction entre les bactéries pour former un biofilm.

Après avoir montré l'effet adhérent de Sha, il a été intéressant de vérifier un possible effet protéolytique et/ou cytotoxique, deux rôles supplémentaires qui ont été attribués aux ATs de la famille SPATE.

7.1.2.2 Effet protéolytique et cytotoxique

Les ATs de la famille SPATE possèdent une activité protéase comme fonction principale. Cette activité est médiée par la triade catalytique des trois résidus d'acides aminés Aspartate (**D**), Histidine (**H**) et Sérine (G**D**S**G**S) qui interviennent dans le site actif pour catalyser l'hydrolyse de certaines liaisons peptidiques. Au niveau de la séquence peptidique de l'AT, ces trois acides aminés sont éloignés les uns des autres. Cependant, ils se retrouvent à proximité au sein du site actif en raison du repliement de la protéine. Cette configuration permet aux résidus acide (D) et base (H) d'activer et le résidu nucléophile (S), lequel va attaquer le groupe carboxyle du substrat (Carter *et al.*, 1988, Van Houdt *et al.*, 2005). Les sérines protéases les plus étudiées sont :

- La trypsine (hydrolyse les protéines après les acides aminés basique) (Van Houdt *et al.*, 2005).
- La chymotrypsine (hydrolyse les protéines après les acides aminés aromatiques) (Van Houdt *et al.*, 2005).
- L'élastase (dégrade l'élastine) (Van Houdt *et al.*, 2005).

Dans cette étude, nous avons montré que les ATs : Sha, Tsh et Vat possèdent une activité de type élastase. Cette activité a été aussi observée chez la souche sauvage QT598, en plus, d'une autre activité de type trypsine. Ceci peut s'expliquer par la présence des ATs TagB et TagC qui ont montré une activité similaire. Par conséquent, la présence des 5 ATs chez la souche QT598 produit une souche à large spectre d'activité qui peut cibler plusieurs substrats. Toutefois, la mutation du site actif des ATs ou la présence des inhibiteurs « sérine protéase » (PMSF) dans le surnageant des cultures a inhibé l'hydrolyse des oligopeptides testés. Ceci a montré que l'activité protéolytique des ATs dépend essentiellement du site actif.

En plus de l'activité protéase, Sha a montré un effet cytopathique sur les cellules épithéliales de la vessie. Cet effet se traduit par une altération morphologique au niveau des cellules infectées après 12h d'incubation (cellules rondes). Ceci peut suggérer que l'AT Sha cible les microfilaments d'actine ce qui entraîne une destruction de l'intégrité des cellules. De plus, l'effet cytopathique de Sha est probablement dû à une internalisation cellulaire par pinocytose. Cette internalisation dépend aussi de l'activité protéase de Sha pour dégrader les protéines de la matrice extracellulaire comme la mucine, ce qui facilite leur internalisation (et donc leur effet cytopathique) (Pokharel *et al.*, 2020). Il sera donc intéressant de déterminer davantage le mécanisme d'internalisation des autres SPATEs.

Après avoir étudié le nouveau AT Sha, il sera intéressant de caractériser le deuxième facteur de virulence, le nouveau fimbriae PL, qui se trouve dans la région unique. Comme décrit précédemment, les fimbriae jouent un rôle important dans l'adhérence, qui est une étape primordiale pour la colonisation des tissus de l'hôte (Soto *et al.*, 1999). La présence de ce fimbriae chez la souche QT598 pourrait augmenter l'adhérence bactérienne. Dans ce contexte, nous avons fixé un deuxième objectif qui consiste à caractériser les fimbriae PL et de déterminer leurs rôles dans la virulence.

7.2 Fimbriae PL

7.2.1 Caractérisation d'un nouveau membre de la famille des π -fimbriae

Les gènes impliqués dans la production des fimbriae PL sont regroupés en un opéron *plf* de 10 gènes. Nous avons pu identifier un gène de régulation (*plfB*) et neuf gènes qui codent pour des protéines de structure et d'assemblage du fimbriae (*plfA*, *plfH*, *plfC*, *plfD*, *plfJ*, *plfK*, *plfE*, *plfF* et *plfG*). Les gènes *plfH*, *plfJ*, *plfK*, *plfE*, *plfF* codent pour les sous-unités mineures, le gène *plfA* code pour la sous-unité majeure, le gène *plfG* code pour l'adhésine, le gène *plfC* code pour la protéine d'ancrage ou « usher » et le gène *plfD* code pour la protéine chaperon.

L'organisation génétique du fimbriae PL est similaire à celle du fimbriae P (codé par un opéron de 11 gènes : 2 gènes de régulation et 9 gènes qui codent pour les sous-unités de structure et d'assemblage) (Marklund *et al.*, 1992), à l'exception du deuxième gène de régulation (*plfI*) qui est absent. Les autres fimbriae P-like comme les fimbriae Pix (codés par un opéron de 8 gènes *pixBAHCDJFG* : 1 gène de régulation et 7 gènes qui codent pour les sous-unités de structure et d'assemblage) (Lügering *et al.*, 2003) et les fimbriae Sfp (codés par un opéron de 7 gènes de structure et d'assemblage, *sfpAHCDJFG*) (Brunner *et al.*, 2001) ne possèdent pas les gènes qui codent soit pour les sous-unités mineures, soit pour les gènes de régulation. Suite aux résultats de cette partie, nous avons décidé de nommer le nouveau fimbriae « fimbriae P-like » ou « fimbriae PL ».

D'après les analyses des séquences de différentes sous-unités fimbriaires et les données phylogénétiques de la protéine usher, nous avons pu montrer que le fimbriae PL partage une forte homologie avec les fimbriae Sfp et Pix (91% et 92% d'identité avec PlfC respectivement) alors qu'il partageait environ 72% d'identité avec PapC du fimbriae P (72% d'identité avec PlfC). Comme décrit précédemment, les fimbriae P ainsi que les fimbriae P-like (Sfp et Pix) appartiennent tous au clade des π -fimbriae (Nuccio *et al.*, 2007), ceci nous a permis d'adopter la

classification par voie CU et de classer le fimbriae PL comme un nouveau membre de la famille des π -fimbriae.

Si l'on s'intéresse maintenant à la région en amont de l'opéron *plf*, on note la présence des gènes *intM* et *repB* qui représente le réplicon RepFIB. Cette région est généralement présente chez les plasmides d'incompatibilité de type IncFII (les plasmides du groupe d'incompatibilité F possède le facteur de fertilité F qui permet le transfert à d'autres bactéries par conjugaison) (Gibbs *et al.*, 1993, Partridge *et al.*, 2018). L'incompatibilité des plasmides est généralement due à une ressemblance génétique au niveau des régions impliquées dans la réplication. Ainsi, les plasmides qui appartiennent au même groupe ne peuvent pas coexister de façon stable dans la même bactérie ce qui mène à la perte d'un des plasmides au cours des divisions cellulaires. Ces plasmides portent généralement des gènes de résistance aux antibiotiques, ce qui peut aboutir à leur acquisition par une bactérie donnée. Ces plasmides jouent ainsi un rôle dans l'émergence et la diffusion de la résistance aux antibiotiques en les transférant d'une bactérie à une autre (Gibbs *et al.*, 1993). Le réplicon RepFIB est généralement conservé parmi les plasmides de type IncFII et il est localisé à proximité d'une région appelée « cargo gene region » qui est un site de recombinaison et d'intégration de différents gènes de virulence, de microcines, de résistances aux antibiotiques, etc. (Koraimann, 2018, Lanza *et al.*, 2014). Dans le cas de l'opéron *plf*, on suggère qu'il s'est intégré dans cette région suite à un transfert horizontal à partir d'autres souches pathogènes.

La localisation de l'opéron *plf* au niveau du plasmide de virulence n'est pas spécifique aux fimbriae PL, L'opéron *sfp* se trouve aussi à proximité de la région RepFIB sur le plasmide de virulence de type « sorbitol-fermenting » pSFO157 chez la souche EHEC O157:NM (Brunner *et al.*, 2001). Toutefois, les séquences d'insertion flanquant leurs extrémités sont distinctes de celles retrouvées sur le plasmide de virulence pEC598 de la souche QT598. Contrairement à l'opéron *plf* qui a été identifié chez une diversité de souches d'*E. coli* pathogène (souches pathogènes qui infectent les humains, les poulets, les dindes, les chiens, les bovins, les ovins, etc.), l'opéron *sfp* a été identifié seulement chez les souches EHEC/STEC fermentant le sorbitol et certaines souches EHEC O165:H25/NM humaines et bovines (Bielaszewska *et al.*, 2009, Brunner *et al.*, 2001). Ceci peut suggérer que l'opéron *sfp* a été acquis suite à un transfert horizontal, mais son émergence était limitée au groupe des souches EHEC alors que l'opéron *plf* a été diffusé à une diversité des souches.

Contrairement aux fimbriae PL qui ont été identifiés sur le plasmide de virulence, la majorité des fimbriae P et même les fimbriae Pix ont été identifiés sur des PAI chromosomiques et parfois

ils se retrouvent présents à plusieurs copies dans le génome bactérien (Blum *et al.*, 1995, Guyer *et al.*, 1998, Hochhut *et al.*, 2006, Lügering *et al.*, 2003). L'exemple le plus étudié est celui de la souche UPEC J96 qui possède deux copies de l'opéron *pap/prs*. Le premier PAI-I_{J96} de 170kb est porteur de l'opéron *pap* et de l'hémolysine α (*hly*). Le deuxième PAI-II_{J96} de 110 kb est porteur de l'opéron *prs*, l'opéron *hly* et la toxine (*cnf1*). De plus, la souche UPEC CFT073 possède aussi deux copies de l'opéron *pap*, le premier se trouve sur le PAI-I_{CFT073} porteur de l'opéron *pap*, l'hémolysine α (*hly*) et les gènes de transport du fer et des glucides. Le deuxième opéron se trouve sur le PAI-II_{CFT073}, porteur de l'opéron *pap* et d'autres gènes à fonction inconnue (Kao *et al.*, 1997, Schmidt *et al.*, 2004). Dans le cas des fimbriae Pix, il a été identifié sur le PAI I4787 en association avec d'autres gènes de virulence (Lügering *et al.*, 2003). Toutes ces données suggèrent que les gènes fimbriaires sont transférés de façon horizontale soit par des PAI ou par des plasmides de virulence et sont intégrés ensuite dans des sites spécifiques d'intégration/de recombinaison. Lors de ce processus, des variations aléatoires peuvent apparaître dans les allèles d'une population (mutations). Ceci a été observé avec une grande diversité au niveau des gènes qui codent pour l'adhésine PapG et PlfG ce qui donne aux bactéries un tropisme différent (une grande capacité à adhérer aux différents récepteurs présents sur les cellules hôtes de différentes espèces) (Marklund *et al.*, 1992).

Il a été démontré qu'il existe au moins quatre allèles différents de l'adhésine *papG* du fimbriae P (responsable de la reconnaissance des récepteurs glycoprotéiques membranaires de type Gal-(α 1-4)-Gal) :

- PapG classe I (codé par l'opéron *pap*) reconnaît le récepteur globotriosylcéramides (GbO3) (Lund *et al.*, 1987, Strömberg *et al.*, 1990).
- PapG classe II (codé par l'opéron *pap*) reconnaît le récepteur globotétraosylcéramides (GbO4) (Strömberg *et al.*, 1990).
- PapG/PrsG classe III (codé par l'opéron *prs*) reconnaît le récepteur globopentaosylcéramides (GbO5) (Strömberg *et al.*, 1990).
- Le récepteur pour l'adhésine PapG classe IV reste inconnu.

Cette spécificité reflète souvent la capacité des souches à coloniser différents organes lors de l'infection puisque ces récepteurs sont présents surtout sur les érythrocytes des humains et des animaux (porc, pigeon, volaille, mouton, chien), ainsi que sur les cellules épithéliales du tractus urinaire chez l'homme (vessie pour GbO3, rein pour GbO4) et les chiens (GbO5) (Strömberg *et al.*, 1990). Cela permettrait aux bactéries d'exprimer, au moment opportun, le type de fimbriae approprié afin d'adhérer aux tissus spécifiques de l'hôte. Dans ce projet, nous avons

pu montrer qu'il existe aussi une diversité génétique au niveau des gènes de l'adhésine PflG des fimbriae PL, ce qui nous a permis de les classer en cinq classes distinctes :

- PflG classe I représentée par la souche UMEA-3703-1, souche UPEC isolée à partir des urines d'un patient souffrant d'une bactériémie. celle-ci possède des gènes qui codent pour le fimbriae PL et qui sont très semblables aux gènes présents chez la souche QT598 (environ 98% d'identité avec PflC_{QT598}) à l'exception de l'adhésine PflG_{UMEA-3703-1} qui est différente. De plus, des études ont montré que les APEC et les UPEC peuvent partager les mêmes facteurs de virulence puisqu'ils sont proches génétiquement et par conséquent la possibilité d'avoir un risque zoonotique.

Il était donc intéressant de caractériser le fimbriae PL de la souche UPEC UMEA-3703-1 et de comparer son phénotype avec celui de la souche APEC QT598 pour voir si les fimbriae PL possèdent les mêmes affinités vis-à-vis des récepteurs de l'hôte ou s'ils sont différents.

- PflG classe II représentée par la souche APEC QT598. Cette adhésine est très spécifique, elle partage une homologie de séquence d'environ 24% d'identité avec PapG classe I, 27% d'identité avec PapG classe II, 31% d'identité avec PapG/PrsG classe III, 22% d'identité avec PapG classe IV, 26% d'identité avec SfpG et 25% d'identité avec PixG,
- PflG classe III, PflG classe IV et PflG classe V sont représentées par quelques souches dans la base de données NCBI. Toutefois, l'adhésine PflG classe V semble être spécifique à certaines souches de *Cronobacter* spp qui sont des pathogènes d'origine alimentaires pouvant contaminer le lait et d'autres produits alimentaires ce qui peut induire des infections extra-intestinales chez les humains et en particulier chez les nouveau-nés (Healy *et al.*, 2010, Lee *et al.*, 2019).

Les allèles qui codent pour les adhésines PflG classe I et classe II sont prédominants parmi les souches testées (souches associées à des infections extra-intestinales chez les humains et les volailles) et ceux qui codent pour les adhésines PflG classe III, classe IV et classe V sont représentés seulement par quelques souches. Toute cette diversité génétique pourrait augmenter l'adhérence à la bactérie vis-à-vis des cellules cibles. À ce jour, les récepteurs des adhésines PflG restent inconnus pour la grande majorité d'entre eux, malgré toutes nos tentatives à déterminer leurs récepteurs spécifiques. Par ailleurs, il sera intéressant de déterminer quels sont les différents récepteurs glycoprotéiques ou glycolipidiques reconnus par les fimbriae PL.

Dans la partie suivante, nous nous sommes intéressés à caractériser le rôle des fimbriae PL dans la virulence. Pour cela, nous avons émis l'hypothèse qu'ils pourraient augmenter l'adhérence vis-à-vis des cellules hôtes. Afin de confirmer cette hypothèse, nous avons procédé à une série de tests phénotypiques.

7.2.2 Rôle du Fimbriae PL dans la virulence

Nous avons tenté de caractériser les fimbriae PL de la souche APEC QT598 et de la souche UPEC UMEA-3703-1. Pour ce faire, nous avons cloné les gènes *plf* dans une bactérie *E. coli* K-12 qui est non-pathogène et ne possède aucun facteur de virulence. De plus, nous avons essayé de créer des clones chimères ou hybrides composés d'un polymère de sous-unités majeures et mineures (PlfA, PlfK, PlfE et PlfF) et au sommet, on trouve l'adhésine PapG classe I ou PapG classe II ou PapG classe III ou Plf_{UMEA-3703-1} classe I. D'après nos résultats, la machinerie d'assemblage du fimbriae PL est capable de reconnaître et de former un fimbriae avec un orthologue d'adhésine PapG des fimbriae P, malgré la faible homologie entre ces adhésines (entre 20% - 30% d'identité). Cela montre que les fimbriae PL partagent une ressemblance et une compatibilité avec les fimbriae P pour former un fimbriae hybride à la surface des cellules bactériennes.

Dans cette partie, la capacité de ces adhésines à agglutiner les érythrocytes de différentes espèces, de former un biofilm et d'adhérer aux cellules rénales et vésicales a été testée.

- **Les fimbriae PL possèdent une activité d'hémagglutination**

Il a été démontré que les fimbriae P et les fimbriae P-like (Sfp et Pix) sont capables d'agglutiner les érythrocytes de différentes espèces humaine et animale (Bielaszewska *et al.*, 2009, Lindstedt *et al.*, 1989, Lügering *et al.*, 2003, Marklund *et al.*, 1992). Par conséquent, nous avons évalué l'activité hémagglutinante des fimbriae PL. Nous avons montré que le fimbriae PL porteur de l'adhésine PlfGII (QT598) était capable d'agglutiner seulement les érythrocytes de dinde et les érythrocytes humains. Par contre, le fimbriae PL porteur de l'adhésine PlfGI (UMEA-3703-1) était capable d'agglutiner les érythrocytes de toutes les espèces testées à l'exception les érythrocytes de chien. Cet effet était absent chez le mutant $\Delta plfG$ et le témoin négatif (vecteur vide). Aussi nous avons pu confirmer que les clones de référence (fimbriae P porteur de PapGI, PapGII et PapGIII) sont capables d'agglutiner les érythrocytes humains, du porc, du lapin et du mouton (Lindstedt *et al.*, 1989, Marklund *et al.*, 1992). Dans l'ensemble, il apparaît que les fimbriae PL et les fimbriae P ont développé des changements au niveau de leurs séquences en

acides aminés (PlfG/PapG) ce qui leur permettent d'adhérer à une variété de récepteurs qui se trouvent sur les érythrocytes et les cellules hôtes de différentes espèces.

Par la suite, nous avons tenté de déterminer les récepteurs spécifiques pour les fimbriae PL. Pour ce faire, nous avons donc testé GbO3, GbO4 et GbO5 qui sont des récepteurs pour les fimbriae P (Strömberg *et al.*, 1990). Nos résultats ont confirmé la spécificité de ces récepteurs vis-à-vis des clones de références portant soit PapGI, PapGII ou PapGIII. Toutefois, ces récepteurs n'étaient pas spécifiques ni pour PlfGI ou PlfGII (résultats non publiés). Cela suggère que les fimbriae PL reconnaissent des récepteurs présents sur les érythrocytes humains et de dinde qui sont différents de ceux reconnus par les fimbriae P. Il sera donc intéressant de déterminer quels sont les récepteurs spécifiques pour les fimbriae PL.

- **Les fimbriae PL peuvent former un biofilm**

Après avoir déterminé l'activité hémagglutinante des fimbriae P et PL, nous avons vérifié si les clones sont capables de former un biofilm à différentes températures. Comme décrit précédemment, la formation de biofilm (qui est généralement médiée par les fimbriae) est une forme de résistance adoptée par certaines bactéries pour survivre dans des conditions environnementales hostiles (Van Houdt *et al.*, 2005). Nous avons montré que le fimbriae PL porteur de l'adhésine PlfGII (QT598) était capable de former un biofilm à des températures qui dépassent 40°C alors que le fimbriae PL porteur de l'adhésine PlfGI (UMEA-3703-1) était capable de former un biofilm à des températures qui ne dépassent pas 37°C. De même, les clones de références et les clones hybrides étaient tous capables de former un biofilm à des températures qui ne dépassent pas 37°C. Cela peut suggérer que les fimbriae P et PL sont impliqués dans l'interaction entre les bactéries pour former un biofilm. Cet effet était absent chez le témoin négatif et le mutant $\Delta plfG$ (sauf à 37°C ce qui peut suggérer la présence d'autres facteurs qui peuvent également contribuer à l'augmentation de la formation de biofilm).

- **Les fimbriae PL peuvent adhérer aux cellules épithéliales humaines**

Dans cette étude, nous avons pu montrer que les fimbriae PL ainsi que les fimbriae P sont capables d'adhérer aux cellules épithéliales humaines (rénales et de la vessie) ce qui peut suggérer leurs rôles dans l'adhérence des souches APEC et UPEC au cours des premiers stades de l'infection (Cet effet était atténué chez le mutant $\Delta plfG$ et le témoin négatif).

Dans l'ensemble les fimbriae PL étaient capables d'agglutiner les érythrocytes (humains, dinde, etc.) de former un biofilm à différentes températures et en plus d'adhérer aux cellules rénales et vésicales. Il en va de même pour les clones de référence (fimbriae P) et les clones

hybrides. Cet effet était atténué chez le mutant $\Delta pIfG$ et le témoin négatif. Ceci peut suggérer que les fimbriae PL pourraient augmenter l'adhérence des bactéries vis-à-vis des cellules de l'hôte lors des ITU. De plus, nous avons pu confirmer que la fonction adhésine est portée par la sous-unité PIfG puisque la perte de cette adhésine a atténué l'effet de fimbriae PL dans la majorité des analyses *in vitro*.

7.3 Rôle de l'autotransporteur Sha et le Fimbriae PL dans la virulence des souches d'*E. coli* pathogènes extra-intestinales

Le but de cette partie était d'étudier la virulence potentielle des nouveaux gènes d'AT et du fimbriae qui sont portés par une région dite unique retrouvée sur le plasmide pEC598 de la souche APEC QT598. Pour ce faire, nous avons utilisé deux modèles d'infection :

- Le modèle de mono-infection qui a servi à déterminer la capacité des souches à coloniser les organes et à déterminer l'expression des gènes *in vivo*
- Le modèle de co-infection (inoculation en quantité égale de la souche sauvage et la souche mutante) qui a servi à déterminer si l'une des 2 souches aura un avantage compétitif sur l'autre.

L'expression de ces gènes a été d'abord étudiée en modèle d'infection respiratoire de poulets de 3 semaines. Toutefois, la souche QT598 n'était pas capable de coloniser les organes et causer une infection. Par conséquent, aucune différence significative n'a été observée entre la souche sauvage et les mutants (AT et fimbriae PL). Ceci peut s'expliquer par le fait que la souche QT598 a été isolée d'un cas de colibacillose chez le dindonneau (âgés de 1 jour). Il semble donc que les poulets utilisés dans notre modèle (âgés de 3 semaines) possèdent un système immunitaire assez développé, leur conférant une résistance accrue à cette souche. Par conséquent, nous avons utilisé un autre modèle d'infection respiratoire avec des poulets de 5 jours. De même, nous n'avons pas observé une différence significative au niveau de la colonisation entre la souche sauvage et les mutants (résultats non publiés). Il est donc probable que le modèle d'infection utilisé dans cette étude ne soit pas idéal pour la souche QT598. Elle pourrait plutôt infecter les dindes de façon préférentielle que les poulets.

Par la suite, nous avons essayé de transférer par conjugaison le plasmide pEC598 qui possède la région unique à la souche APEC CH138. Cette souche possède le même sérotype que la souche QT598, soit O1 : K1, en plus elle est assez virulente dans le modèle d'infection respiratoire de poulets. D'après nos résultats, la présence du plasmide de virulence pEC598 confère à la souche une grande capacité à coloniser les différents organes tels que le foie, la

rate, les sacs aériens, les poumons, le sang et même causer la mort des poussins comparativement à la souche sans la région unique qui était moins virulente. Ces résultats suggèrent que les nouveaux gènes qui se trouvent au niveau de la région unique augmentent considérablement la virulence de la souche APEC CH138 (résultats non publiés, Annexe III). La région unique a été identifiée chez une variété de souches ExPEC entre autres, la souche UPEC UMEA-3703-1. Ceci peut suggérer que les gènes qui se trouvent au niveau de la région unique (le fimbriae PL, l'AT Sha, la toxine Prt) peuvent être impliqués dans plusieurs types d'infection. De plus, nous avons montré que les fimbriae PL ainsi que l'AT Sha sont capables d'adhérer aux cellules rénales et vésicales ce qui pourrait améliorer l'adhérence de la souche QT598 au cours de l'infection du tractus urinaire. Pour cela, nous avons testé un autre modèle d'infection qui est le modèle murin d'infection. La souche UPEC UMEA-3703-1 testée n'était pas capable de coloniser la vessie et les reins de façon significative. Bien qu'elle fût isolée d'un patient souffrant d'ITU, il semble que le modèle de souris utilisé ne soit pas spécifique à cette souche et par conséquent, elle ne peut pas coloniser les organes de façon significative. Par contre, la souche QT598 était capable de coloniser les organes du tractus urinaire après 48h d'infection et le niveau d'infection était comparable à celle d'une souche UPEC CFT073 de référence (souche utilisée comme témoin positif) (Kao *et al.*, 1997). Ces résultats illustrent bien que les souches APEC et UPEC sont proches génétiquement de façon à ce que les APEC puissent coloniser le même hôte que les UPEC et donc puissent présenter un potentiel risque zoonotique de transmission de l'infection des animaux aux humains (Kathayat *et al.*, 2021). Aussi, le « Sequence Type » de la souche QT598 est fréquemment associé à des infections des voies urinaires chez l'homme et des infections extra-intestinales chez la volaille, ceci peut expliquer la virulence de la souche APEC dans le modèle d'infection de souris (Kathayat *et al.*, 2021).

- Rôle des ATs de la famille des SPATEs

L'absence de l'AT Sha ne semble pas affecter le niveau de colonisation de la souche QT598 dans les vessies et les reins des souris et par conséquent, aucune différence significative n'a été observée entre les organes infectés par la souche sauvage et ceux infectés par le mutant QT598 Δsha . Il est donc possible que la perte du gène *sha* soit compensée par d'autres gènes de virulence comme les fimbriae type 1, les toxines, les ATs (Tsh, Vat, TagB ou TagC), etc. ce qui permet de rétablir l'infection. D'après ces résultats, nous souhaitons déterminer le rôle cumulatif des 5 ATs de la famille des SPATEs qui se trouvent sur le génome de la souche QT598. Pour ce faire, nous avons testé le mutant $\Delta 5AT$ dans un modèle d'infection de souris. Nous avons pu

montrer une diminution significative de la capacité du mutant $\Delta 5AT$ à coloniser les reins si on le compare avec la souche sauvage. Cela peut suggérer que l'effet cumulatif des SPATEs pourrait jouer un rôle dans la colonisation de la souche QT598 dans le modèle ITU de souris. Fait intéressant, seulement le gène *sha* a été régulé positivement (six fois plus que dans le milieu LB) dans les vessies des souris infectées. L'expression des autres ATs n'a pas été détectée. Il est donc possible que l'AT Sha puisse jouer un effet sur l'induction du système immunitaire de l'hôte. Il sera donc intéressant de vérifier la réponse immunitaire de l'hôte lors des ITU de souris.

Bien que l'expression des gènes *vat*, *tsh*, *tagB* et *tagC* était faible dans les vessies des souris infectées, nous avons pu montrer que le gène *vat* a été exprimé dix fois plus dans le milieu minimum M63-glycérol. De même pour Tsh, nous avons montré qu'il est préférentiellement exprimé en milieu LB. Cela peut suggérer qu'il existe une régulation croisée entre les différents ATs de la famille des SPATEs. Il sera donc intéressant d'étudier la régulation de ces ATs et les conditions qui favorisent leur expression.

- Rôle des fimbriae PL

L'absence de l'opéron *plf* ne semble pas affecter le niveau de colonisation de la souche QT598 dans les vessies et les reins des souris et par conséquent, aucune différence significative n'a été observée entre les organes infectés par la souche sauvage et ceux infectés par le mutant QT598 Δplf . Il en va de même pour les infections avec la souche UPEC UMEA-3703-1 Δplf . Des résultats semblables ont été obtenus avec le fimbriae P : l'absence de l'opéron *pap* n'a pas affecté le niveau de colonisation de la souche UPEC76 dans les voies urinaires chez les souris CBA/J (Moblely *et al.*, 1993). Plusieurs hypothèses pourraient expliquer de tels résultats. En effet, il est possible que la similitude des résultats entre la souche sauvage et les mutants soit due à l'expression d'autres gènes de virulence retrouvés sur le génome bactérien (phénomène de redondance), ce qui permet le rétablissement de l'infection. Ou encore, le modèle d'ITU de souris utilisé ne soit pas idéal pour étudier l'adhérence médiée par certains fimbriae comme dans le cas des fimbriae P ou les fimbriae PL (la spécificité des récepteurs cibles est différente d'une espèce à une autre).

Toutefois, il a été démontré que les fimbriae P peuvent jouer un rôle dans l'induction d'une réponse immunitaire de l'hôte. Ainsi, une étude a montré que ces fimbriae sont importants dans l'établissement d'une pyélonéphrite chez le modèle d'infection urinaire de primates, ce qui permet d'induire une réponse inflammatoire accompagnée d'une augmentation de la production d'interleukine 6 (Roberts *et al.*, 1994). Une autre étude a montré que la réintroduction des fimbriae P chez une souche urinaire asymptomatique d'*E. coli* 83972, a changé l'expression de certains

gènes de l'hôte en agissant comme un agoniste du facteur 7 régulateur d'interféron (IRF-7) dont l'activation déclenche une cascade de signaux qui sont associés à la pyélonéphrite aiguë (Ambite *et al.*, 2019). De plus, il a été démontré que les fimbriae P peuvent réduire la réponse immunitaire de l'hôte dans le modèle d'infection des souris. Ils jouent, ainsi, un rôle dans la diminution de l'expression des gènes qui codent pour le « polymeric immunoglobulin receptor (pIgR) » ce qui provoque une diminution de taux sérique d'IgA dans les urines des souris infectées (Rice *et al.*, 2005). Bien que les fimbriae PL n'aient pas un effet direct sur la colonisation de la souche QT598, son expression dans les vessies infectées était 40 fois plus élevée que dans le milieu LB. D'après ces données, il est donc possible que le fimbriae PL puisse jouer un rôle dans l'induction du système immunitaire de l'hôte. Il sera donc intéressant d'étudier la réponse immunitaire de l'hôte lors des ITU de souris, ou encore, de tester d'autres modèles d'infection puisque l'opéron *plf* a été identifié chez une diversité de souches ExPEC.

Par la suite, afin de déterminer si l'une des deux souches (sauvage et Δplf) aura un avantage compétitif sur l'autre, nous avons procédé à la co-infection des souris. D'après les résultats obtenus, nous avons pu montrer une diminution significative de la capacité du mutant Δplf à coloniser les reins comparativement à la souche sauvage (résultats non publiés, Annexe IV). Cela peut suggérer le rôle potentiel des fimbriae PL dans les infections systémiques comme les infections des voies urinaires (ils peuvent jouer un rôle dans le tropisme tissulaire qui permet à la bactérie de coloniser un organe de façon préférentielle grâce à la présence des récepteurs reconnu par ces fimbriae (Roberts, 1996)).

CONCLUSION ET PERSPECTIVES

8 CONCLUSION ET PERSPECTIVES

Le but de cette thèse était de caractériser les nouveaux facteurs de virulence à savoir l'AT Sha et le fimbriae PL afin de définir leurs rôles dans la virulence des souches d'*E. coli* pathogènes extra-intestinales.

Dans un premier temps, nous avons caractérisé l'AT Sha qui se trouve dans la région unique au niveau du plasmide de virulence pEC598. C'est un nouveau membre de la famille des SPATEs avec un site de clivage atypique. Il appartient à la classe II immunomodulateurs et possède une homologie de séquence avec Tsh, TleA et Vat. Nous avons également montré qu'il possède un effet adhérent (capable d'agglutiner le sang de différentes espèces, de se lier aux cellules épithéliales rénales et vésicales, de former des agrégations et un biofilm), un effet protéolytique de type estérase qui est médié par la triade catalytique (Aspartate (**D**), Histidine (**H**) et Sérine (GDSGS)) et un effet cytotoxique sur les cellules épithéliales vésicales après 12h d'incubation.

Le gène qui code pour l'AT Sha était présent chez une variété des souches APEC et UPEC, ce qui nous a permis d'étudier son rôle dans différents modèles animaux. Nous avons montré que l'absence du gène *sha* n'a pas affecté le niveau de colonisation de la souche QT598 dans les vessies et les reins des souris. Toutefois, le gène *sha* a été six fois plus exprimé *in vivo* que dans le milieu LB. Il pourrait donc avoir un effet sur la réponse immunitaire de l'hôte. Nous avons également montré que la délétion des 5 SPATEs a considérablement réduit la colonisation bactérienne au niveau des reins, ce qui peut suggérer un rôle cumulatif des SPATE chez la souche QT598 dans le modèle ITU de souris.

Dans un second temps, nous avons caractérisé le fimbriae PL qui se trouve dans la région unique au niveau du plasmide de virulence pEC598 chez la souche APEC QT598. Nous avons montré qu'il est également présent dans le génome de la souche UPEC UMEA-3703-1. C'est un nouveau membre de la famille π -fimbriae. Il partage une homologie avec les fimbriae P, les fimbriae Pix et les fimbriae Sfp. Nous avons pu identifier un gène de régulation (*plfB*) et neuf gènes qui codent pour des protéines de structure et d'assemblage du fimbriae (*plfA*, *plfH*, *plfC*, *plfD*, *plfJ*, *plfK*, *plfE*, *plfF* et *plfG*). Les gènes *plfH*, *plfJ*, *plfK*, *plfE*, *plfF* codent pour les sous-unités mineures, le gène *plfA* code pour la sous-unité majeure, le gène *plfC* code pour la protéine d'ancrage ou usher et le gène *plfD* code pour la protéine chaperon et le gène *plfG* code pour l'adhésine. Nous avons aussi montré qu'il existe cinq allèles différents de l'adhésine *plfG* qui sont PlfG classe I représentée par la souche UMEA-3703-1, PlfG classe II représentée par la souche

APEC QT598. PflG classe III, PflG classe IV et PflG classe V sont représentées par quelques souches dans la base de donnée NCBI. Les allèles qui codent pour les adhésines PflG classe I et classe II sont prédominants et se retrouvent chez une variété des souches ExPEC qui infectent les humains et les animaux. Nos travaux ont permis de mettre en évidence le rôle des fimbriae PL dans la virulence. En effet, nous avons montré que le fimbriae PL porteur de l'adhésine PflG II était capable d'agglutiner seulement les érythrocytes de dinde et les érythrocytes humains, de former un biofilm à différentes températures et d'adhérer aux cellules rénales et vésicales. Par contre, le fimbriae PL porteur de l'adhésine PflG I était capable d'agglutiner les érythrocytes de 8 espèces différentes, de former un biofilm à différentes températures et d'adhérer aux cellules rénales et vésicales.

Lors des essais *in vivo*, nous avons montré que l'absence de l'opéron *plf* n'a pas affecté le niveau de colonisation de la souche QT598 dans les vessies et les reins des souris. Toutefois, nous avons montré une diminution significative de la capacité du mutant Δplf à coloniser les reins comparativement à la souche sauvage dans le modèle de co-infection des souris. Cela peut suggérer le rôle potentiel des fimbriae PL dans les infections systémiques comme les infections des voies urinaires.

- **Perspectives**

Il reste toutefois, de nombreuses parties à éclaircir, notamment approfondir les données au niveau de la régulation de *sha* et *plf* et étudier les conditions qui affectent leurs expressions chez la souche APEC QT598. Pour ce faire, nous proposons de construire une fusion transcriptionnelle par l'opéron *lux* de *Photobacterium luminescens*, qui sera fusionnée à la région promotrice du gène *sha* et de l'opéron *plf* ce qui va permettre de mesurer l'activité des promoteurs. Ensuite, nous allons construire une banque de mutants chez la souche sauvage à l'aide du transposon aléatoire (transposon Tn10). Ces transposons vont s'intégrer de manière aléatoire dans le génome bactérien ce qui va entraîner soit une augmentation, soit une diminution de l'expression du fimbriae et/ou de l'AT. Les clones qui présentent une augmentation ou une diminution importante de l'émission de signaux de luminescence seront testés pour leur capacité à agglutiner les érythrocytes humains. Une fois que l'on aura sélectionné les clones, ils seront séquencés pour déterminer le site d'insertion du transposon afin de déterminer les gènes qui affectent l'expression des fimbriae PL et de l'AT Sha, cela permettra de déterminer les régulateurs de ces deux facteurs.

Il sera aussi intéressant d'approfondir les données sur l'expression des gènes *sha* et *plf* *in vivo* dans le modèle d'infection urinaire chez les souris (puisque les deux facteurs sont régulés positivement) ou chez les mammifères (comme les primates) et d'étudier l'impact de ces facteurs

sur la réponse immunitaire de l'hôte comme l'expression des cytokines et chimiokines. Dans le même contexte, on pourrait étudier l'expression des gènes *sha* et *plf* en utilisant le modèle d'infection aviaire avec des dindonneaux afin de déterminer si la souche QT598 est capable de causer une colibacillose aviaire sans oublier d'étudier la réponse immunitaire de l'hôte. Pour ce faire, nous pouvons utiliser les deux modèles :

- Le modèle de mono-infection qui servira à déterminer la capacité des souches (souche sauvage et mutante) à coloniser les organes *in vivo*.
- Le modèle de co-infection qui servira à déterminer si l'une des 2 souches (souche sauvage et mutante) aura un avantage compétitif sur l'autre.

Afin de confirmer les résultats de l'infection, la microscopie multiphotonique (MPM) sera utilisée sur des animaux vivants afin de visualiser en temps réel le processus d'infection bactérienne et aussi de vérifier le niveau d'expression des fimbriae PL ou de l'AT Sha.

Par ailleurs, il est aussi primordial de déterminer quels seront les différents récepteurs glycoprotéiques ou glycolipidiques (glucosylceramide, glucose, galactose, fucose, galactosylceramide, lactosylceramide, et N-acetylgalactosamine) reconnus par l'AT Sha et/ou les fimbriae PL. Pour ce faire, nous proposons de réaliser l'hémagglutination avec les érythrocytes humains et de dinde, à laquelle nous allons ajouter les différents substrats. Toute inhibition de l'hémagglutination signifie que le récepteur testé est une cible potentielle pour l'AT Sha et/ou les fimbriae PL.

9 REFERENCES

- Abadia G (2008) Zoonoses et santé au travail. *Archives des Maladies Professionnelles et de l'Environnement* 69(2):337-340.
- Alber A, Morris KM, Bryson KJ, Sutton KM, Monson MS, Chintoan-Uta C, Borowska D, Lamont SJ, Schouler C & Kaiser P (2020) Avian pathogenic Escherichia coli (APEC) strain-dependent immunomodulation of respiratory granulocytes and mononuclear phagocytes in CSF1R-reporter transgenic chickens. *Frontiers in immunology* 10:3055.
- Ambite I, Butler DSC, Stork C, Gronberg-Hernandez J, Koves B, Zdziarski J, Pinkner J, Hultgren SJ, Dobrindt U, Wullt B & Svanborg C (2019) Fimbriae reprogram host gene expression - Divergent effects of P and type 1 fimbriae. *PLoS Pathog* 15(6):e1007671.
- Andersen PS, Stegger M, Aziz M, Contente-Cuomo T, Gibbons HS, Keim P, Sokurenko EV, Johnson JR & Price LB (2013) Complete genome sequence of the epidemic and highly virulent CTX-M-15-producing H30-Rx subclone of Escherichia coli ST131. *Genome announcements* 1(6):e00988-00913.
- Anderson GG, Dodson KW, Hooton TM & Hultgren SJ (2004) Intracellular bacterial communities of uropathogenic Escherichia coli in urinary tract pathogenesis. *Trends in microbiology* 12(9):424-430.
- Ayala-Lujan JL, Vijayakumar V, Gong M, Smith R, Santiago AE & Ruiz-Perez F (2014) Broad spectrum activity of a lectin-like bacterial serine protease family on human leukocytes. *PLoS One* 9(9):e107920.
- Bacher A, Eberhardt S, Fischer M, Kis K & Richter G (2000) Biosynthesis of vitamin B2 (riboflavin). *Annual review of nutrition* 20(1):153-167.
- Barrington K, pédiatrie Scd & nouveau-né Cdédfed (2007) La prise en charge du nourrisson plus vulnérable à la septicémie. *Paediatrics & Child Health* 12(10):899-905.
- Bentley R & Meganathan R (1982) Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiological reviews* 46(3):241.
- Benz I & Schmidt MA (2011) Structures and functions of autotransporter proteins in microbial pathogens. *International Journal of Medical Microbiology* 301(6):461-468.
- Bielaszewska M, Prager R, Vandivinit L, Müsken A, Mellmann A, Holt NJ, Tarr PI, Karch H & Zhang W (2009) Detection and characterization of the fimbrial sfp cluster in enterohemorrhagic Escherichia coli O165: H25/NM isolates from humans and cattle. *Applied and environmental microbiology* 75(1):64-71.

- Bingen E (1999) Physiopathologie et diagnostic bactériologique des méningites néonatales à *Escherichia coli*. *Médecine thérapeutique/Pédiatrie* 2(1):19-25.
- Bingen E (2001) *Méningites bactériennes communautaires*. Elsevier.
<https://books.google.ca/books?id=lyPOt5T1IzoC>
- Blount ZD (2015) The natural history of model organisms: The unexhausted potential of *E. coli*. *Elife* 4:e05826.
- Blum G, Falbo V, Caprioli A & Hacker J (1995) Gene clusters encoding the cytotoxic necrotizing factor type 1, Prs-fimbriae and α -hemolysin form the pathogenicity island II of the uropathogenic *Escherichia coli* strain J96. *FEMS Microbiology Letters* 126(2):189-195.
- Bonacorsi S, Bidet P, Geslain G, Cointe A, Doit C, Biran V & Mariani-Kurkdjian P (2018) Spécificités des examens bactériologiques du nouveau-né suspect d'infection. *Revue Francophone des Laboratoires* 2018(500):55-62.
- Bonacorsi S & Bingen E (2005a) Molecular epidemiology of *Escherichia coli* causing neonatal meningitis. *Int J Med Microbiol* 295(6-7):373-381.
- Bonacorsi S, Houdoin V & Bingen E (2001) Facteurs de virulence associés à *E. coli* responsable de méningite néonatale. *Archives de pédiatrie* 8:726-731.
- Bonacorsi S, Lefèvre S, Clermont O, Houdouin V, Bourrillon A, Loirat C, Aujard Y & Bingen E (2005b) *Escherichia coli* strains causing urinary tract infection in uncircumcised infants resemble urosepsis-like adult strains. *The Journal of urology* 173(1):195-197.
- Boone DR, Castenholz RW, Garrity GM, Brenner DJ, Krieg NR & Staley JT (2005) *Bergey's Manual® of Systematic Bacteriology*. Springer Science & Business Media,
- Boudeau J, Glasser AL, Masseret E, Joly B & Darfeuille-Michaud A (1999) Invasive ability of an *Escherichia coli* strain isolated from the ileal mucosa of a patient with Crohn's disease. *Infection and Immunity* 67(9):4499-4509.
- Boyd EF & Hartl DL (1998) Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. *J Bacteriol* 180(5):1159-1165.
- Brinton Jr C (1959) Non-flagellar appendages of bacteria. *Nature* 183:782-786.
- Brugère-Picoux J (2017) *Mes poules en bonne santé: comment reconnaître, prévenir et traiter leurs maladies*. Aurélie Mercier,
- Brunder W, Khan AS, Hacker Jr & Karch H (2001) Novel type of fimbriae encoded by the large plasmid of sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157: H-. *Infection and immunity* 69(7):4447-4457.

- Busch A & Waksman G (2012) Chaperone–usher pathways: diversity and pilus assembly mechanism. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1592):1112-1122.
- Carter P & Wells JA (1988) Dissecting the catalytic triad of a serine protease. *Nature* 332(6164):564-568.
- Castellani A & Chambers AJ (1919) *Manual of tropical medicine*. William Wood,
- Caza M, Lépine F, Milot S & Dozois CM (2008) Specific roles of the iroBCDEN genes in virulence of an avian pathogenic Escherichia coli O78 strain and in production of salmochelins. *Infection and Immunity* 76(8):3539-3549.
- Chahales P & Thanassi DG (2015) Structure, Function, and Assembly of Adhesive Organelles by Uropathogenic Bacteria. *Microbiology spectrum* 3(5).
- Charbonneau M-È & Mourez M (2007) Functional organization of the autotransporter adhesin involved in diffuse adherence. *Journal of bacteriology* 189(24):9020-9029.
- Clermont O, Christenson JK, Denamur E & Gordon DM (2013) The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylogroups. *Environmental microbiology reports* 5(1):58-65.
- Clermont O, Dixit OV, Vangchhia B, Condamine B, Dion S, Bridier-Nahmias A, Denamur E & Gordon D (2019) Characterization and rapid identification of phylogroup G in Escherichia coli, a lineage with high virulence and antibiotic resistance potential. *Environmental microbiology* 21(8):3107-3117.
- Croxen MA & Finlay BB (2010) Molecular mechanisms of Escherichia coli pathogenicity. *Nature Reviews Microbiology* 8(1):26-38.
- Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M & Finlay BB (2013) Recent advances in understanding enteric pathogenic Escherichia coli. *Clinical microbiology reviews* 26(4):822-880.
- Dautin N (2010) Serine Protease Autotransporters of Enterobacteriaceae (SPATEs): Biogenesis and Function. *Toxins* 2(6):1179-1206.
- De Kraker ME, Davey PG, Grundmann H & Group BS (2011) Mortality and hospital stay associated with resistant Staphylococcus aureus and Escherichia coli bacteremia: estimating the burden of antibiotic resistance in Europe. *PLoS medicine* 8(10):e1001104.
- Delannoy S, Beutin L, Mariani-Kurkdjian P, Fleiss A, Bonacorsi S & Fach P (2017) The Escherichia coli serogroup O1 and O2 lipopolysaccharides are encoded by multiple O-antigen gene clusters. *Frontiers in cellular and infection microbiology* 7:30.

- Denis F, Cattoir V, Martin C, Ploy M-C & Poyart C (2016) *Bactériologie médicale: techniques usuelles*. Elsevier Masson,
- Dho-Moulin M & Fairbrother JM (1999) Avian pathogenic *Escherichia coli* (APEC). *Vet Res* 30(2-3):299-316.
- Donnenberg M (2013) *Escherichia coli: pathotypes and principles of pathogenesis*. Academic Press,
- Donnenberg MS & Welch RA (1996) *Virulence determinants of uropathogenic Escherichia coli*. Urinary tract infections: molecular pathogenesis and clinical management. ASM Press, Washington, DC,
- Dozois CM, Daigle F & Curtiss R (2003) Identification of pathogen-specific and conserved genes expressed in vivo by an avian pathogenic *Escherichia coli* strain. *Proceedings of the National Academy of Sciences* 100(1):247-252.
- Dozois CM, Dho-Moulin M, Brée A, Fairbrother JM, Desautels C & Curtiss III R (2000) Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the Tsh genetic region. *Infection and immunity* 68(7):4145-4154.
- Dozois CM, Pourbakhsh SA & Fairbrother JM (1995) Expression of P and type 1 (F1) fimbriae in pathogenic *Escherichia coli* from poultry. *Veterinary microbiology* 45(4):297-309.
- Dubrovsky AS, Foster BJ, Jednak R, Mok E & McGillivray D (2012) Visibility of the urethral meatus and risk of urinary tract infections in uncircumcised boys. *CMAJ* 184(15):E796-E803.
- Dufies O, Doye A, Courjon J, Torre C, Michel G, Loubatier C, Jacquel A, Chaintreuil P, Majoor A & Guinamard RR (2021) *Escherichia coli* Rho GTPase-activating toxin CNF1 mediates NLRP3 inflammasome activation via p21-activated kinases-1/2 during bacteraemia in mice. *Nature Microbiology* 6(3):401-412.
- Duguid J (1959) Fimbriae and adhesive properties in *Klebsiella* strains. *Microbiology* 21(1):271-286.
- Duguid JP, Smith IW, Dempster G & Edmunds P (1955) Non-Flagellar Filamentous Appendages (" Fimbriae") and Haemagglutinating Activity in *Bacterium coli*. *Journal of Pathology and Bacteriology* 70(2):335-348.
- Durrmeyer X, Cohen R, Bingen E & Aujard Y (2012) Stratégies thérapeutiques des méningites néonatales à *Escherichia coli*. *Archives de pédiatrie* 19:S140-S144.
- El-Naggar W, Afifi J, McMillan D, Toye J, Ting J, Yoon EW & Shah PS (2019) Epidemiology of meningitis in Canadian neonatal intensive care Units. *The Pediatric infectious disease journal* 38(5):476-480.

- Elmund GK, Allen MJ & Rice EW (1999) Comparison of Escherichia coli, total coliform, and fecal coliform populations as indicators of wastewater treatment efficiency. *Water Environment Research* 71(3):332-339.
- Flatau G, Lemichez E, Gauthier M, Chardin P, Paris S, Fiorentini C & Boquet P (1997) Toxin-induced activation of the G protein p21 Rho by deamidation of glutamine. *Nature* 387(6634):729-733.
- Flores-Mireles AL, Walker JN, Caparon M & Hultgren SJ (2015) Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nature reviews microbiology* 13(5):269-284.
- Forde BM, Zakour NLB, Stanton-Cook M, Phan M-D, Totsika M, Peters KM, Chan KG, Schembri MA, Upton M & Beatson SA (2014) The complete genome sequence of Escherichia coli EC958: a high quality reference sequence for the globally disseminated multidrug resistant E. coli O25b: H4-ST131 clone. *PLoS One* 9(8):e104400.
- Forsythe SJ (2020) *The microbiology of safe food*. John Wiley & Sons,
- Foxman B (2014) Urinary tract infection syndromes: occurrence, recurrence, bacteriology, risk factors, and disease burden. *Infectious Disease Clinics* 28(1):1-13.
- Frank C, Faber MS, Askar M, Bernard H, Fruth A, Gilsdorf A, Hohle M, Karch H, Krause G, Prager R, Spode A, Stark K, Werber D & team Hi (2011) Large and ongoing outbreak of haemolytic uraemic syndrome, Germany, May 2011. *Euro Surveill* 16(21).
- Fratamico PM, Liu Y & Sommers H (2018) *Pathogenic Escherichia coli: evolution, omics, detection and control*. Caister Academic Press,
- Gandon S (2016) Why be temperate: lessons from bacteriophage λ . *Trends in microbiology* 24(5):356-365.
- Garenaux A, Caza M & Dozois CM (2011) The Ins and Outs of siderophore mediated iron uptake by extra-intestinal pathogenic Escherichia coli. *Veterinary microbiology* 153(1-2):89-98.
- Gibbs M, Spiers A & Bergquist P (1993) RepFIB: a basic replicon of large plasmids. *Plasmid* 29(3):165-179.
- Gouali M & Weill F-X (2013) Les Escherichia coli entérohémorragiques: des entérobactéries d'actualité. *La Presse Médicale* 42(1):68-75.
- Graveline R, Mourez M, Hancock MA, Martin C, Boisclair S & Harel J (2011) Lrp–DNA complex stability determines the level of ON cells in type P fimbriae phase variation. *Molecular microbiology* 81(5):1286-1299.
- Guabiraba R & Schouler C (2015) Avian colibacillosis: still many black holes. *FEMS microbiology letters* 362(15):fnn118.

- Gutiérrez D, Pardo M, Montero D, Oñate A, Farfán MJ, Ruiz-Pérez F, Del Canto F & Vidal R (2015) TleA, a Tsh-like autotransporter identified in a human enterotoxigenic *Escherichia coli* strain. *Infection and immunity* 83(5):1893-1903.
- Guyer DM, Kao J-S & Mobley HL (1998) Genomic analysis of a pathogenicity island in uropathogenic *Escherichia coli* CFT073: distribution of homologous sequences among isolates from patients with pyelonephritis, cystitis, and catheter-associated bacteriuria and from fecal samples. *Infection and immunity* 66(9):4411-4417.
- Guyomard-Rabenirina S, Malespine J, Ducat C, Sadikalay S, Falord M, Harrois D, Richard V, Dozois C, Breurec S & Talarmin A (2016) Temporal trends and risk factors for antimicrobial resistant Enterobacteriaceae urinary isolates from outpatients in Guadeloupe. *BMC microbiology* 16(1):1-8.
- Hacker J & Carniel E (2001) Ecological fitness, genomic islands and bacterial pathogenicity. *EMBO reports* 2(5):376-381.
- Hacker J & Kaper JB (2000) Pathogenicity islands and the evolution of microbes. *Annual Reviews in Microbiology* 54(1):641-679.
- Harry L. T. Mobley MSDaECH (2009) Uropathogenic *Escherichia coli*. in *EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology* (ASM Press, Washington, DC).
- Healy B, Cooney S, O'Brien S, Iversen C, Whyte P, Nally J, Callanan JJ & Fanning S (2010) Cronobacter (*Enterobacter sakazakii*): an opportunistic foodborne pathogen. *Foodborne Pathogens and Disease* 7(4):339-350.
- Henderson IR & Nataro JP (2001) Virulence functions of autotransporter proteins. *Infection and immunity* 69(3):1231-1243.
- Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC & Ala'Aldeen D (2004) Type V protein secretion pathway: the autotransporter story. *Microbiology and molecular biology reviews* 68(4):692-744.
- Hernday AD, Braaten BA, Broitman-Maduro G, Engelberts P & Low DA (2004) Regulation of the pap epigenetic switch by CpxAR: phosphorylated CpxR inhibits transition to the phase ON state by competition with Lrp. *Molecular cell* 16(4):537-547.
- Hicks S, Candy DC & Phillips AD (1996) Adhesion of enteroaggregative *Escherichia coli* to pediatric intestinal mucosa in vitro. *Infection and Immunity* 64(11):4751-4760.
- Hochhut B, Wilde C, Balling G, Middendorf B, Dobrindt U, Brzuszkiewicz E, Gottschalk G, Carniel E & Hacker J (2006) Role of pathogenicity island-associated integrases in the genome plasticity of uropathogenic *Escherichia coli* strain 536. *Molecular microbiology* 61(3):584-595.

- Holden N, Totsika M, Dixon L, Catherwood K & Gally DL (2007) Regulation of P-fimbrial phase variation frequencies in *Escherichia coli* CFT073. *Infection and immunity* 75(7):3325-3334.
- Hospenthal MK, Costa TR & Waksman G (2017) A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nature reviews microbiology* 15(6):365-379.
- Huja S, Oren Y, Trost E, Brzuszkiewicz E, Biran D, Blom J, Goesmann A, Gottschalk G, Hacker J & Ron EZ (2015) Genomic avenue to avian colisepticemia. *MBio* 6(1):e01681-01614.
- Janda JM, Abbott SL & Janda M (2006) *The enterobacteria*. American Society for Microbiology (ASM),
- Jensen BH, Olsen KE, Struve C, Krogfelt KA & Petersen AM (2014) Epidemiology and Clinical Manifestations of Enteroaggregative *Escherichia coli*. *Clinical microbiology reviews* 27(3):614-630.
- Johanson I-M, Plos K, Marklund B-I & Svanborg C (1993) Pap, papG and prsG DNA sequences in *Escherichia coli* from the fecal flora and the urinary tract. *Microbial pathogenesis* 15(2):121-129.
- Johnson JR & Russo TA (2002) Extraintestinal pathogenic *Escherichia coli*: "the other bad *E. coli*". *J Lab Clin Med* 139(3):155-162.
- Johnson TJ, Aziz M, Liu CM, Sokurenko E, Kisiela DI, Paul S, Andersen P, Johnson JR & Price LB (2016) Complete genome sequence of a CTX-M-15-producing *Escherichia coli* strain from the H30Rx subclone of sequence type 131 from a patient with recurrent urinary tract infections, closely related to a lethal urosepsis isolate from the patient's sister. *Genome announcements* 4(3):e00334-00316.
- Johnson TJ, Hargreaves M, Shaw K, Snippes P, Lynfield R, Aziz M & Price LB (2015) Complete genome sequence of a carbapenem-resistant extraintestinal pathogenic *Escherichia coli* strain belonging to the sequence type 131 H30R subclade. *Genome announcements* 3(2):e00272-00215.
- Johnson TJ, Siek KE, Johnson SJ & Nolan LK (2006) DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. *Journal of bacteriology* 188(2):745-758.
- Jørgensen SL, Stegger M, Kudirkiene E, Lilje B, Poulsen LL, Ronco T, Pires Dos Santos T, Kiil K, Bisgaard M & Pedersen K (2019) Diversity and population overlap between avian and human *Escherichia coli* belonging to sequence type 95. *MSphere* 4(1):e00333-00318.

- Justice SS, Hung C, Theriot JA, Fletcher DA, Anderson GG, Footer MJ & Hultgren SJ (2004) Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Proceedings of the National Academy of Sciences* 101(5):1333-1338.
- Kabir S (2010) Avian colibacillosis and salmonellosis: a closer look at epidemiology, pathogenesis, diagnosis, control and public health concerns. *International journal of environmental research and public health* 7(1):89-114.
- Källenius G, Svenson S, Hultberg H, Möllby R, Helin I, Cedergren B & Winberg J (1981) Occurrence of P-fimbriated *Escherichia coli* in urinary tract infections. *The Lancet* 318(8260):1369-1372.
- Kao JS, Stucker DM, Warren JW & Mobley HL (1997) Pathogenicity island sequences of pyelonephritogenic *Escherichia coli* CFT073 are associated with virulent uropathogenic strains. *Infect Immun* 65(7):2812-2820.
- Kaper JB, Nataro JP & Mobley HL (2004) Pathogenic *Escherichia coli*. *Nature Reviews Microbiology* 2(2):123-140.
- Kariyawasam S & Han J (2019) Avian Pathogenic *Escherichia coli*: Link to Foodborne Urinary Tract Infections in Humans. *Food Safety in Poultry Meat Production*, Springer. p 261-292.
- Karmali MA, Steele BT, Petric M & Lim C (1983) Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* 1(8325):619-620.
- Kathayat D, Lokesh D, Ranjit S & Rajashekara G (2021) Avian Pathogenic *Escherichia coli* (APEC): An Overview of Virulence and Pathogenesis Factors, Zoonotic Potential, and Control Strategies. *Pathogens* 10(4):467.
- Kauffmann F (1947) The serology of the coli group. *J Immunol* 57(1):71-100.
- Khandige S, Kronborg T, Uhlin BE & Møller-Jensen J (2015) sRNA-mediated regulation of P-fimbriae phase variation in uropathogenic *Escherichia coli*. *PLoS pathogens* 11(8):e1005109.
- Kim KS (2003) Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury. *Nature Reviews Neuroscience* 4(5):376-385.
- Kim KS (2016) Human meningitis-associated *Escherichia coli*. *EcoSal Plus* 7(1).
- Klemm P & Schembri M (2004) Type 1 Fimbriae, Curli, and Antigen 43: adhesion, colonization, and biofilm formation. *EcoSal Plus* 1(1).
- Koraimann G (2018) Spread and persistence of virulence and antibiotic resistance genes: a ride on the F plasmid conjugation module. *EcoSal Plus* 8(1).

- Korhonen TK, Virkola R & Holthöfer H (1986) Localization of binding sites for purified Escherichia coli P fimbriae in the human kidney. *Infection and Immunity* 54(2):328-332.
- Kostakioti M & Stathopoulos C (2004) Functional analysis of the Tsh autotransporter from an avian pathogenic Escherichia coli strain. *Infection and immunity* 72(10):5548-5554.
- Kunert Filho H, Carvalho D, Grassotti T, Soares B, Rossato J, Cunha A, Brito K, Cavalli L & Brito B (2015) Avian pathogenic Escherichia coli-methods for improved diagnosis. *World's Poultry Science Journal* 71(2):249-258.
- La Ragione R, Cooley W & Woodward MJ (2000) The role of fimbriae and flagella in the adherence of avian strains of Escherichia coli O78: K80 to tissue culture cells and tracheal and gut explants. *Journal of Medical Microbiology* 49(4):327-338.
- Lanza VF, de Toro M, Garcillán-Barcia MP, Mora A, Blanco J, Coque TM & de la Cruz F (2014) Plasmid flux in Escherichia coli ST131 sublineages, analyzed by plasmid constellation network (PLACNET), a new method for plasmid reconstruction from whole genome sequences. *PLoS genetics* 10(12):e1004766.
- Lee IPA & Andam CP (2019) Pan-genome diversification and recombination in Cronobacter sakazakii, an opportunistic pathogen in neonates, and insights to its xerotolerant lifestyle. *BMC microbiology* 19(1):1-14.
- Leimbach A, Hacker J & Dobrindt U (2013) E. coli as an all-rounder: the thin line between commensalism and pathogenicity. *Between pathogenicity and commensalism*, Springer. p 3-32.
- Leperchois-Loritte M (2019) Infections néonatales bactériennes et conseils à l'officine.
- Leyton DL, Adams LM, Kelly M, Sloan J, Tauschek M, Robins-Browne RM & Hartland EL (2007) Contribution of a novel gene, rpeA, encoding a putative autotransporter adhesin to intestinal colonization by rabbit-specific enteropathogenic Escherichia coli. *Infection and immunity* 75(9):4664-4669.
- Lillington J, Geibel S & Waksman G (2014) Biogenesis and adhesion of type 1 and P pili. *Biochimica et Biophysica Acta (BBA)-General Subjects* 1840(9):2783-2793.
- Lindstedt R, Baker N, Falk P, Hull R, Hull S, Karr J, Leffler H, Svanborg Edén C & Larson G (1989) Binding specificities of wild-type and cloned Escherichia coli strains that recognize globo-A. *Infection and immunity* 57(11):3389-3394.
- Liu C, Zheng H, Yang M, Xu Z, Wang X, Wei L, Tang B, Liu F, Zhang Y & Ding Y (2015) Genome analysis and in vivo virulence of porcine extraintestinal pathogenic Escherichia coli strain PCN033. *BMC genomics* 16(1):717.
- Lobel B & Soussy C (2007) *Les infections urinaires*. Springer science & Business media,

- Lügering A, Benz I, Knochenhauer S, Ruffing M & Schmidt MA (2003) The Pix pilus adhesin of the uropathogenic *Escherichia coli* strain X2194 (O2: K-: H6) is related to Pap pili but exhibits a truncated regulatory region. *Microbiology* 149(6):1387-1397.
- Lund B, Lindberg F, Marklund B-I & Normark S (1987) The PapG protein is the alpha-D-galactopyranosyl-(1----4)-beta-D-galactopyranose-binding adhesin of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences* 84(16):5898-5902.
- Macnab RM (1992) Genetics and biogenesis of bacterial flagella. *Annual review of genetics* 26(1):131-158.
- Mahillon J & Chandler M (1998) Insertion sequences. *Microbiology and molecular biology reviews* 62(3):725-774.
- Mainil J (2003) Facteurs de virulence et propriétés spécifiques des souches invasives d'*Escherichia coli*: I) les adhésines et facteurs de colonisation. *Ann Med Vet* 147:105-126.
- Manning SD, Zhang L, Foxman B, Spindler A, Tallman P & Marrs CF (2001) Prevalence of known P-fimbrial G alleles in *Escherichia coli* and identification of a new adhesin class. *Clinical Diagnostic Laboratory Immunology* 8(3):637-640.
- Marklund BI, Tennent JM, Garcia E, Hamers A, Baga M, Lindberg F, Gaastra W & Normark Sa (1992) Horizontal gene transfer of the *Escherichia coli* pap and prs pili operons as a mechanism for the development of tissue-specific adhesive properties. *Molecular microbiology* 6(16):2225-2242.
- Mehershahi KS, Abraham SN & Chen SL (2015) Complete genome sequence of uropathogenic *Escherichia coli* strain C15. *Genome announcements* 3(3):e00558-00515.
- Mellata M (2013) Human and avian extraintestinal pathogenic *Escherichia coli*: infections, zoonotic risks, and antibiotic resistance trends. *Foodborne pathogens and disease* 10(11):916-932.
- Mellata M, Dho-Moulin M, Dozois CM, Curtiss III R, Lehoux B & Fairbrother JM (2003) Role of avian pathogenic *Escherichia coli* virulence factors in bacterial interaction with chicken heterophils and macrophages. *Infection and immunity* 71(1):494-503.
- Mellata M, Johnson J & Curtiss III R (2018) *Escherichia coli* isolates from commercial chicken meat and eggs cause sepsis, meningitis and urinary tract infection in rodent models of human infections. *Zoonoses and public health* 65(1):103-113.
- Miquel S (2010) Facteurs de virulence de *Escherichia coli* adhérents et invasifs associés à la maladie de Crohn caractérisation et régulation de leur expression. Édité Laurent C & Arlette D-M ([s. n.], [S. I.]).

- Mobley HL, Donnenberg MS & Hagan EC (2009) Uropathogenic *Escherichia coli*. *EcoSal Plus* 3(2).
- Mobley HL, Jarvis KG, Elwood JP, Whittle DI, Lockatell C, Russell RG, Johnson DE, Donnenberg MS & Warren JW (1993) Isogenic P-fimbrial deletion mutants of pyelonephritogenic *Escherichia coli*: the role of α Gal (1–4) β Gal binding in virulence of a wild-type strain. *Molecular microbiology* 10(1):143-155.
- Moxley RA & Francis DH (1986) Natural and experimental infection with an attaching and effacing strain of *Escherichia coli* in calves. *Infection and immunity* 53(2):339-346.
- Mulvey MA, Lopez-Boado YS, Wilson CL, Roth R, Parks WC, Heuser J & Hultgren SJ (1998) Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. *Science* 282(5393):1494-1497.
- Nataro JP & Kaper JB (1998) Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 11(1):142-201.
- Nielubowicz GR & Mobley HL (2010) Host–pathogen interactions in urinary tract infection. *Nature Reviews Urology* 7(8):430.
- Nolan LK, Horne S, Giddings C, Foley S, Johnson T, Lynne A & Skyberg J (2003) Resistance to serum complement, iss, and virulence of avian *Escherichia coli*. *Veterinary research communications* 27(2):101-110.
- Nowicki B, Rhen M, Väisänen-Rhen V, Pere A & Korhonen T (1984) Immunofluorescence study of fimbrial phase variation in *Escherichia coli* KS71. *Journal of bacteriology* 160(2):691-695.
- Nuccio S-P & Baumler AJ (2007) Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiology and Molecular Biology Reviews* 71(4):551-575.
- Ochman H, Lawrence JG & Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *nature* 405(6784):299-304.
- Ochoa TJ & Contreras CA (2011) Enteropathogenic *E. coli* (EPEC) infection in children. *Current opinion in infectious diseases* 24(5):478.
- Organization WH (1992) Les zoonoses.).
- Ørskov F & Ørskov I (1992) *Escherichia coli* serotyping and disease in man and animals. *Canadian journal of microbiology* 38(7):699-704.
- Ørskov I & Ørskov F (1983) Serology of *Escherichia coli* fimbriae. *Prog Allergy* 33:80-105.
- Ørskov I & Ørskov F (1985) *Escherichia coli* in extra-intestinal infections. *Epidemiology & Infection* 95(3):551-575.
- Ottow J (1975) Ecology, physiology, and genetics of fimbriae and pili. *Annual review of microbiology* 29(1):79-108.

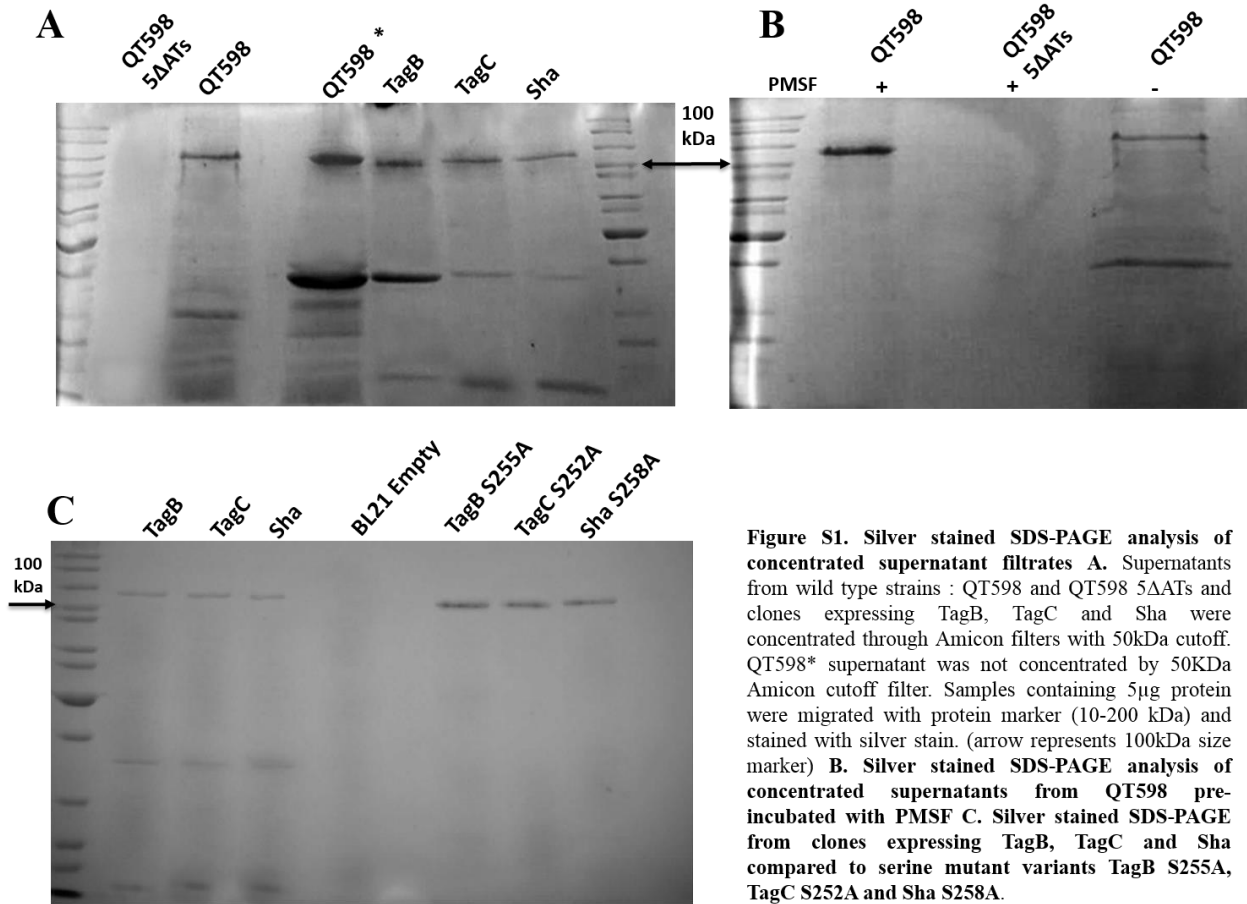
- Paletta AC, Castro VS & Conte-Junior CA (2020) Shiga toxin-producing and enteroaggregative *Escherichia coli* in animal, foods, and humans: Pathogenicity mechanisms, detection methods, and epidemiology. *Current microbiology* 77(4):612-620.
- Paolozzi L, Roose-Amsaleg C, Liébart JC, Bauda P, Bodilis J, Dussurget O, Forterre P, Gutierrez C, LATIFI A & Ona-Nguema G (2021) *Microbiologie - 2e éd.: Biologie des procaryotes et de leurs virus*. Dunod. <https://books.google.ca/books?id=-HwnEAAAQBAJ>
- Partridge SR, Kwong SM, Firth N & Jensen SO (2018) Mobile genetic elements associated with antimicrobial resistance. *Clinical microbiology reviews* 31(4):e00088-00017.
- Perna NT, Plunkett G, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J & Kirkpatrick HA (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157: H7. *Nature* 409(6819):529-533.
- Pokharel P, Díaz JM, Bessaiah H, Houle S, Guerrero-Barrera AL & Dozois CM (2020) The serine protease autotransporters tagb, tagc, and sha from extraintestinal pathogenic *Escherichia coli* are internalized by human bladder epithelial cells and cause actin cytoskeletal disruption. *International journal of molecular sciences* 21(9):3047.
- Pokharel P, Houria H, Bessaiah H & Dozois CM (2019) Serine protease autotransporters of the Enterobacteriaceae (SPATEs): out and about and chopping it up. *Microorganisms* 7(12):594.
- Poolman JT & Wacker M (2016) Extraintestinal pathogenic *Escherichia coli*, a common human pathogen: challenges for vaccine development and progress in the field. *The Journal of infectious diseases* 213(1):6-13.
- Pourbakhsh SA, Dho-Moulin M, Brée A, Desautels C, Martineau-Doize B & Fairbrother JM (1997) Localization of the vivoexpression of P and F1 fimbriae in chickens experimentally inoculated with pathogenic *Escherichia coli*. *Microbial pathogenesis* 22(6):331-341.
- Provence DL & Curtiss 3rd R (1994) Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic *Escherichia coli* strain. *Infection and immunity* 62(4):1369-1380.
- Rankin DJ, Rocha EP & Brown SP (2011) What traits are carried on mobile genetic elements, and why? *Heredity* 106(1):1-10.
- Rasko DA, Rosovitz M, Myers GS, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebahia M, Thomson NR & Chaudhuri R (2008) The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *Journal of bacteriology* 190(20):6881-6893.

- Rice JC, Peng T, Spence JS, Wang HQ, Goldblum RM, Corthesy B & Nowicki BJ (2005) Pyelonephritic Escherichia coli expressing P fimbriae decrease immune response of the mouse kidney. *J Am Soc Nephrol* 16(12):3583-3591.
- Roberts JA (1996) Tropism in bacterial infections: urinary tract infections. *The Journal of urology* 156(5):1552-1559.
- Roberts JA, Marklund B-I, Ilver D, Haslam D, Kaack MB, Baskin G, Louis M, Möllby R, Winberg J & Normark S (1994) The Gal (alpha 1-4) Gal-specific tip adhesin of Escherichia coli P-fimbriae is needed for pyelonephritis to occur in the normal urinary tract. *Proceedings of the National Academy of Sciences* 91(25):11889-11893.
- Robinson JL, Finlay JC, Lang ME, Bortolussi R & pédiatrie Scd (2014) Le diagnostic et la prise en charge des infections urinaires chez les nourrissons et les enfants. *Paediatrics & Child Health* 19(6):320-325.
- Rodriguez-Siek KE, Giddings CW, Doetkott C, Johnson TJ & Nolan LK (2005) Characterizing the APEC pathotype. *Vet Res* 36(2):241-256.
- Ronald A (2003) The etiology of urinary tract infection: traditional and emerging pathogens. *Disease-a-month* 49(2):71-82.
- Rosen DA, Hooton TM, Stamm WE, Humphrey PA & Hultgren SJ (2007) Detection of intracellular bacterial communities in human urinary tract infection. *PLoS medicine* 4(12):e329.
- Ruiz-Perez F & Nataro JP (2014) Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cellular and molecular life sciences* 71(5):745-770.
- Russo TA & Johnson JR (2000) Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J Infect Dis* 181(5):1753-1754.
- Russo TA & Johnson JR (2003) Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and infection* 5(5):449-456.
- Saif YM, Fadly AM, Glisson JR, McDougald LR, Nolan LK & Swayne DE (2011) *Diseases of Poultry*. Wiley. <https://books.google.ca/books?id=XWleGIZcZakC>
- Sarowska J, Futoma-Koloch B, Jama-Kmiecik A, Frej-Madrzak M, Ksiazczyk M, Bugla-Ploskonska G & Choroszy-Krol I (2019) Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: recent reports. *Gut pathogens* 11(1):1-16.
- Schembri MA & Klemm P (2001) Biofilm formation in a hydrodynamic environment by novel FimH variants and ramifications for virulence. *Infection and immunity* 69(3):1322-1328.

- Scheutz F, Nielsen EM, Frimodt-Møller J, Boisen N, Morabito S, Tozzoli R, Nataro JP & Caprioli A (2011) Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. *Euro Surveill* 16(24).
- Schmidt H & Hensel M (2004) Pathogenicity islands in bacterial pathogenesis. *Clinical microbiology reviews* 17(1):14-56.
- Servin AL (2005) Pathogenesis of Afa/Dr diffusely adhering *Escherichia coli*. *Clinical microbiology reviews* 18(2):264-292.
- Soto GE & Hultgren SJ (1999) Bacterial adhesins: common themes and variations in architecture and assembly. *Journal of Bacteriology* 181(4):1059-1071.
- Stanley T & Wilson IG (2003) Multilocus enzyme electrophoresis. *Molecular biotechnology* 24(2):203-220.
- Stoesser N, Sheppard AE, Pankhurst L, De Maio N, Moore CE, Sebra R, Turner P, Anson LW, Kasarskis A & Batty EM (2016) Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *MBio* 7(2):e02162-02115.
- Stordeur P & Mainil J (2002) La colibacillose aviaire. *Annales de médecine vétérinaire*. Annales Medecine Veterinaire, p 11-18.
- Strömberg N, Marklund B, Lund B, Ilver D, Hamers A, Gaastra W, Karlsson K & Normark S (1990) Host-specificity of uropathogenic *Escherichia coli* depends on differences in binding specificity to Gal alpha 1-4Gal-containing isoreceptors. *The EMBO journal* 9(6):2001-2010.
- Thanassi DG, Nuccio S-P, Shu Kin So S & Bäumler AJ (2007) Fimbriae: classification and biochemistry. *EcoSal Plus* 2(2).
- Toh H, Oshima K, Toyoda A, Ogura Y, Ooka T, Sasamoto H, Park S-H, Iyoda S, Kurokawa K & Morita H (2010) Complete genome sequence of the wild-type commensal *Escherichia coli* strain SE15, belonging to phylogenetic group B2. *Journal of bacteriology* 192(4):1165-1166.
- Tony Mazzulli M (2012) Diagnosis and management of simple and complicated urinary tract infections (UTIs). *Can J Urol* 19(1):42-48.
- Van den Beld M & Reubsæet F (2012) Differentiation between *Shigella*, enteroinvasive *Escherichia coli* (EIEC) and noninvasive *Escherichia coli*. *European journal of clinical microbiology & infectious diseases* 31(6):899-904.
- Van Der Woude MW & Bäumler AJ (2004) Phase and antigenic variation in bacteria. *Clinical microbiology reviews* 17(3):581-611.

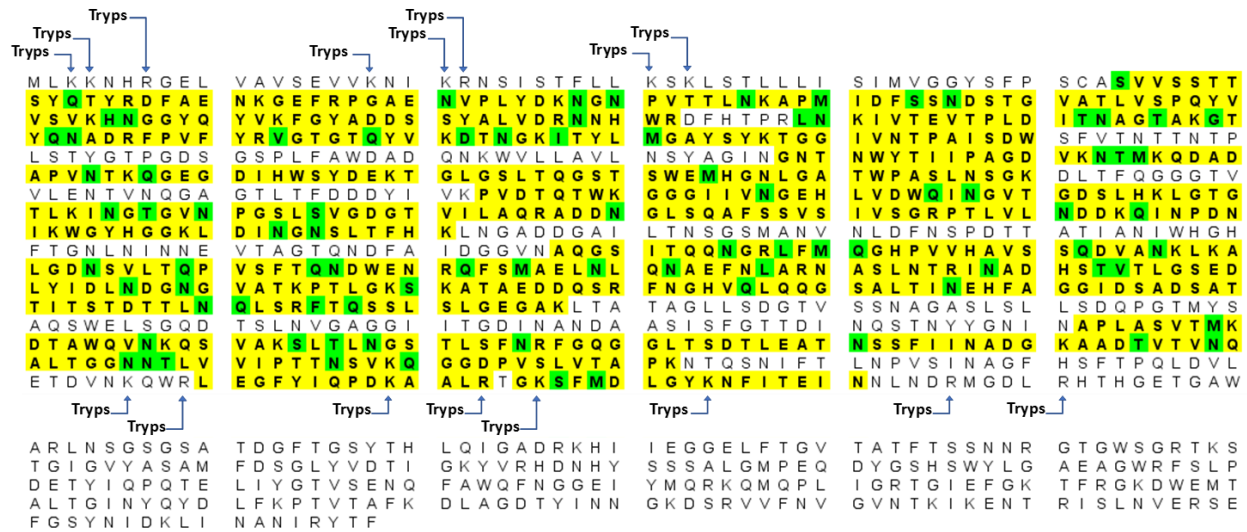
- Van Elsas JD, Semenov AV, Costa R & Trevors JT (2011) Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *The ISME journal* 5(2):173-183.
- Van Houdt R & Michiels CW (2005) Role of bacterial cell surface structures in *Escherichia coli* biofilm formation. *Research in microbiology* 156(5-6):626-633.
- Wagner PL & Waldor MK (2002) Bacteriophage control of bacterial virulence. *Infection and immunity* 70(8):3985-3993.
- Welch RA, Burland V, Plunkett G, Redford P, Roesch P, Rasko D, Buckles E, Liou S-R, Boutin A & Hackett J (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences* 99(26):17020-17024.
- White-Ziegler CA, Villapakkam A, Ronaszeki K & Young S (2000) H-NS controls pap and daa fimbrial transcription in *Escherichia coli* in response to multiple environmental cues. *Journal of Bacteriology* 182(22):6391-6400.
- Whitman WB (2015) *Bergey's manual of systematics of Archaea and Bacteria*. Wiley Online Library,
- Whittam TS, Wolfe ML, Wachsmuth IK, Orskov F, Orskov I & Wilson RA (1993) Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infection and Immunity* 61(5):1619-1629.
- Wijetunge D, Gongati S, DebRoy C, Kim K, Couraud P, Romero I, Weksler B & Kariyawasam S (2015) Characterizing the pathotype of neonatal meningitis causing *Escherichia coli* (NMEC). *BMC microbiology* 15(1):1-15.
- Yan X, Fratamico PM, Bono JL, Baranzoni GM & Chen C-Y (2015) Genome sequencing and comparative genomics provides insights on the evolutionary dynamics and pathogenic potential of different H-serotypes of Shiga toxin-producing *Escherichia coli* O104. *BMC microbiology* 15(1):83.
- Zamora M, Ziegler CA, Freddolino PL & Wolfe AJ (2020) A thermosensitive, phase-variable epigenetic switch: pap revisited. *Microbiology and Molecular Biology Reviews* 84(3):e00030-00017.
- Zheng B, Dong H, Xu H, Lv J, Zhang J, Jiang X, Du Y, Xiao Y & Li L (2016) Coexistence of MCR-1 and NDM-1 in clinical *Escherichia coli* isolates. *Clinical Infectious Diseases* 63(10):1393-1395.

10 ANNEXE I : FIGURES SUPPLEMENTAIRES DE L'ARTICLE 1



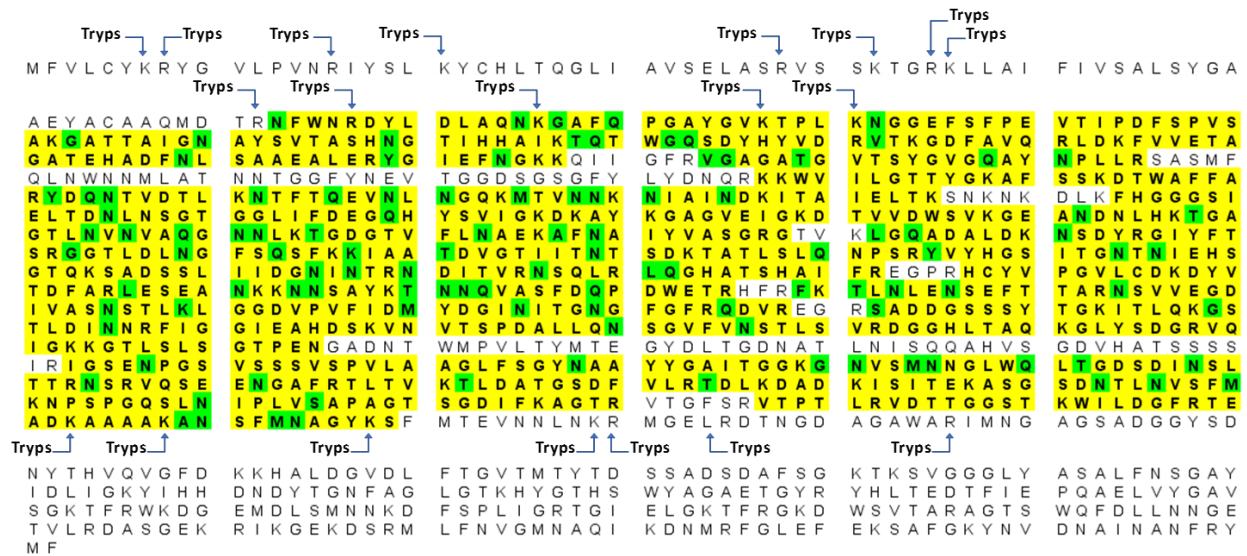
S2A

TagB



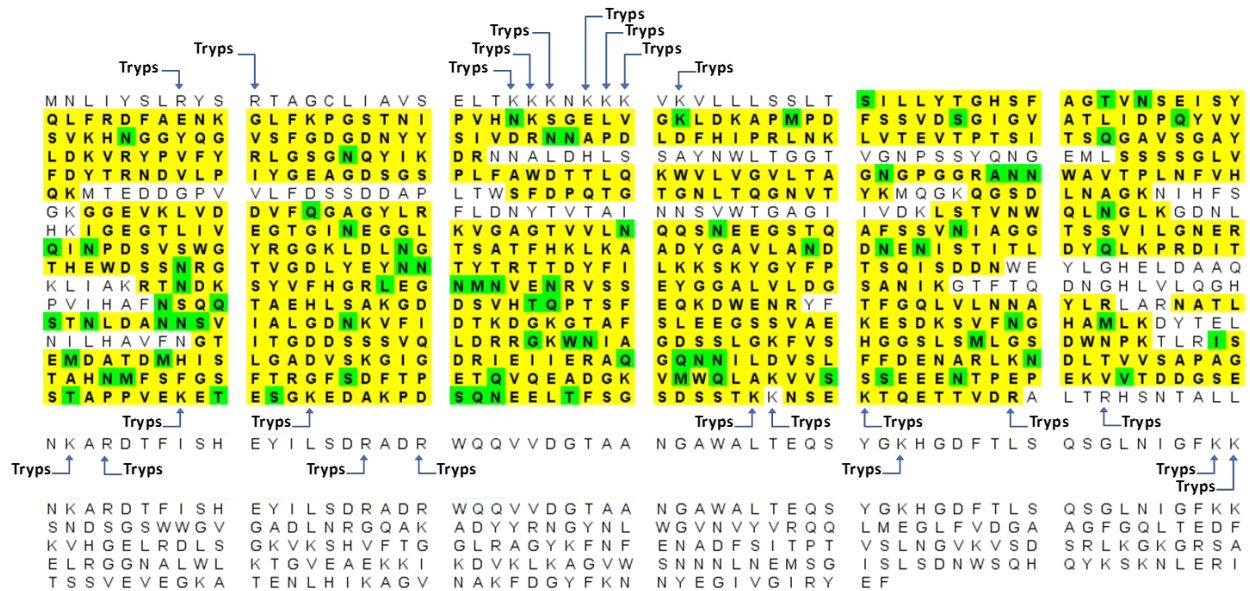
S2B

TagC



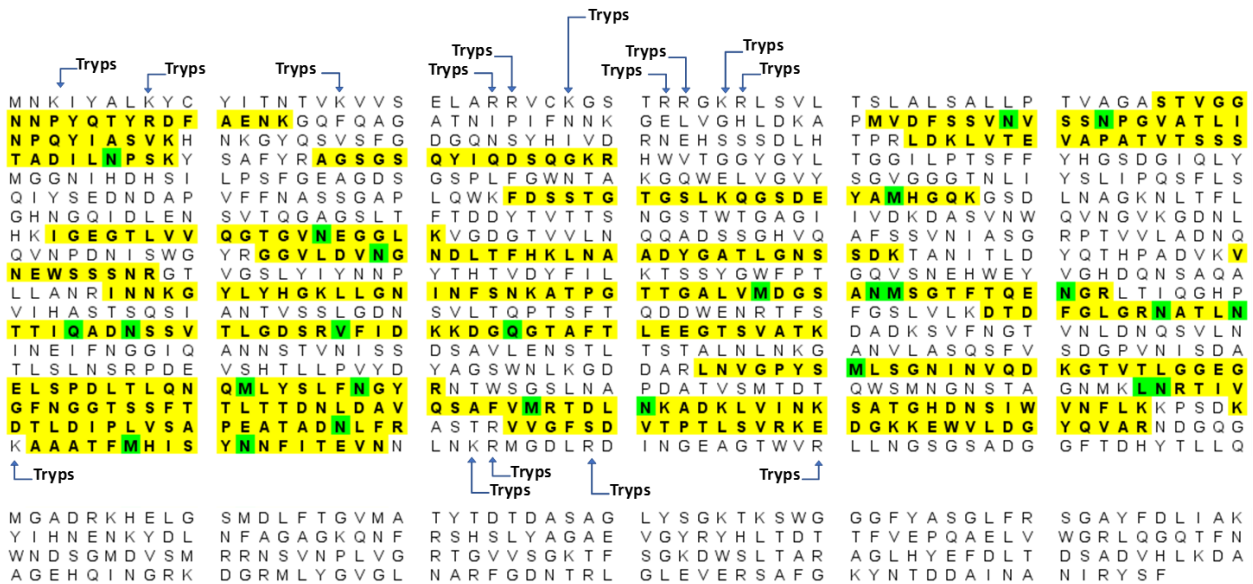
S2C

Sha



S2D

Vat



S2E

Tsh

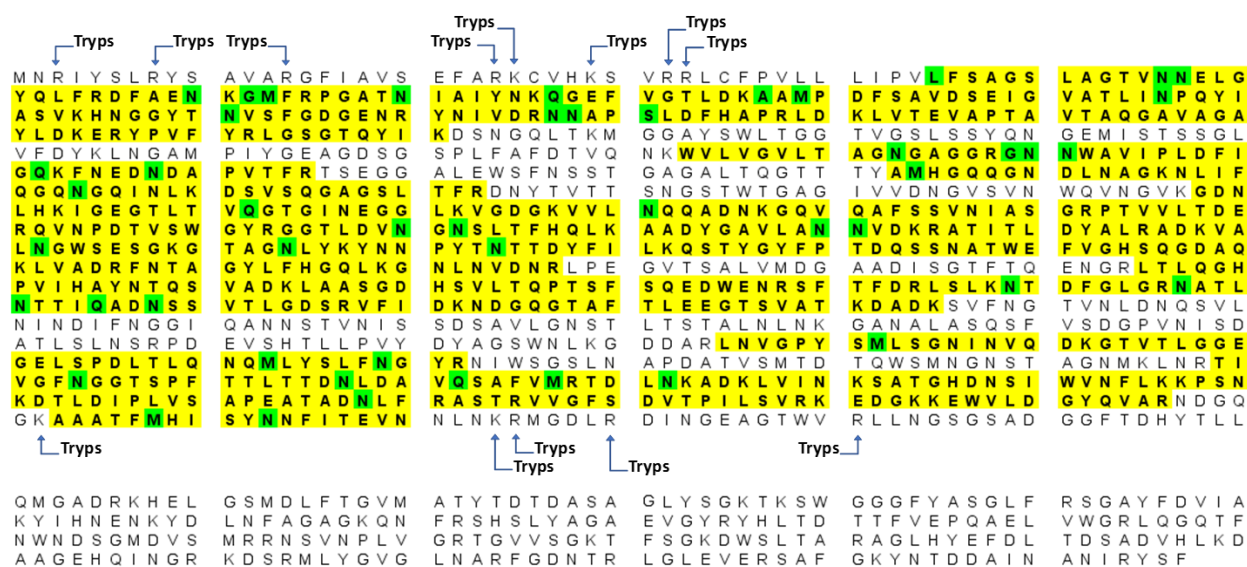


Figure S2. Sequence coverage of autotransporters. LC-MS analysis of each sample. The peptides sequence that have been identified are highlighted in yellow and modifications are highlighted in green **A. TagB**. A total of 676 of 1278 amino acids was detected (53% of coverage) **B. TagC**. A total of 864 of 1322 amino acids was detected (64% of coverage)., **C. Sha**. A total of 816 of 1302 amino acids was detected (63% of coverage). **D. Vat**. A total of 450 of 1376 amino acids was detected (33% of coverage) from QT598 WT. **E. Tsh**. A total of 746 of 1377 amino acids was detected (54% of coverage) from QT598 WT. Blue arrows are the predicted Trypsin sites in both N-terminal and C-terminal of each proteins.

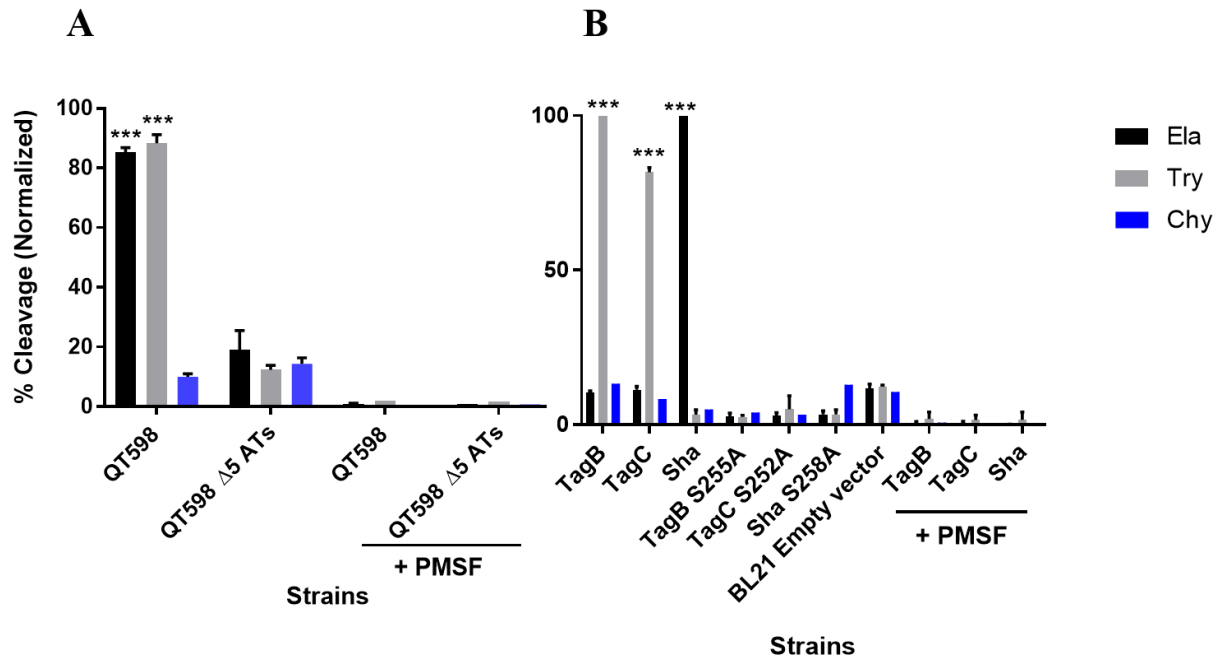


Figure S3. Oligopeptide cleavage profiles of SPATEs from wild-type (A) and clonal backgrounds (B). 5 μ g of each SPATE-containing supernatant filtrate from clones in BL21 / supernatant from QT598 and its SPATE-free $\Delta 5AT$ mutant was incubated at 37°C for 3 h with 1mM of synthetic oligopeptide specifically recognized by the following enzymatic activities: **Elastase (Ela)**-(N-Suc-Ala-Ala-Ala-pNA) ; **Trypsin (Try)**-(N-Ben-L-arginine-pNA); or **Chymotrypsin (Chy)**-(N-Suc-Ala-Ala-Pro-Phe-pNA). Absorbance at 410 nm was normalized to the maximum absorbance value. Data are the means of three independent experiments and error bars represent the standard errors of the means. (* p<0.05, **p<0.01, ***p<0.001 one - way ANOVA with multiple comparisons vs QT598/ $\Delta 5AT$ (A) or BL21 empty vector (B))

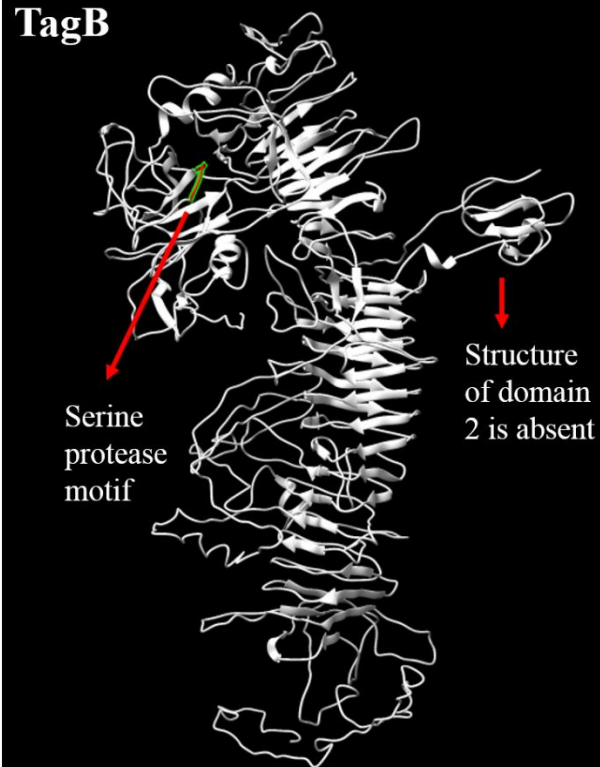
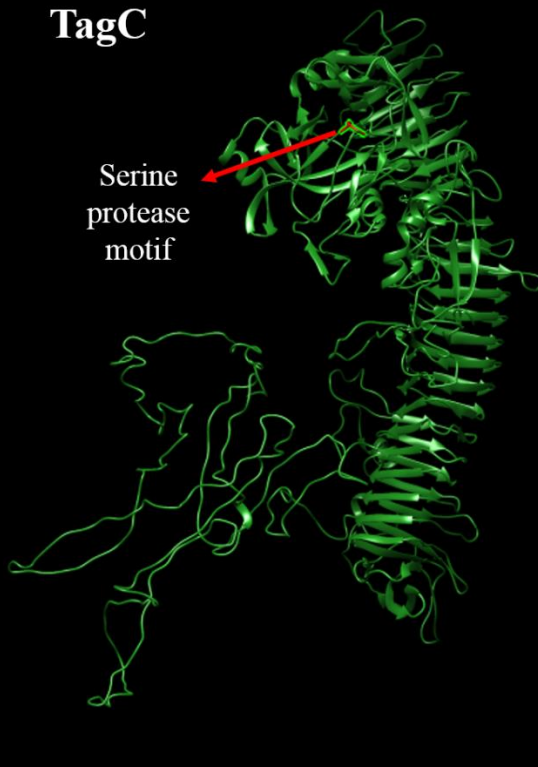
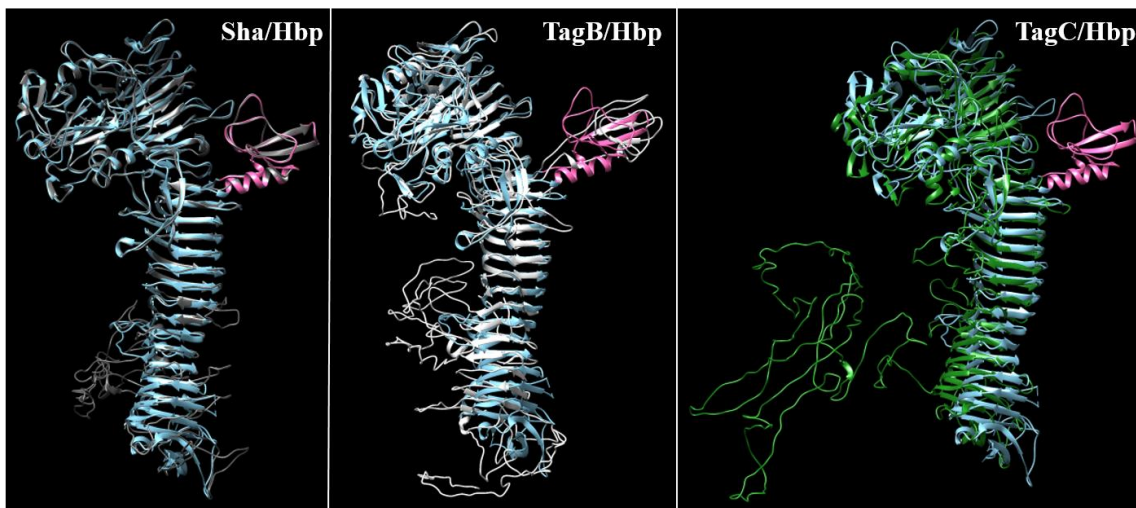
S4A**TagB****S4B****TagC****S4C****S4D****S4 E**

Figure S4. Stereo ribbon diagram showing the predicted 3-dimensional structures of TagB and TagC passenger domains derived from the Hbp/Tsh and EspP crystal structures. A. Predicted structure of TagB. The model was generated using the I-TASSER server with Hbp (PDB 1WXR) template. The structure of domain 2 is absent in TagB. **B.** Predicted structure of TagC. The model was generated using the the I-TASSER server with EspP (extracellular serine protease plasmid-encoded, PDB 3SZE) template. **C.** Overlap of Hbp (blue) with Sha (grey) . The structure of the domain 2 is present in Sha.. **D.** Overlap of Hbp (blue) with TagB (white). The domain 2 structure is not conserved in TagB **E.** Overlap of Hbp (blue) with TagC (green) The domain 2 (pink) is absent in TagC.

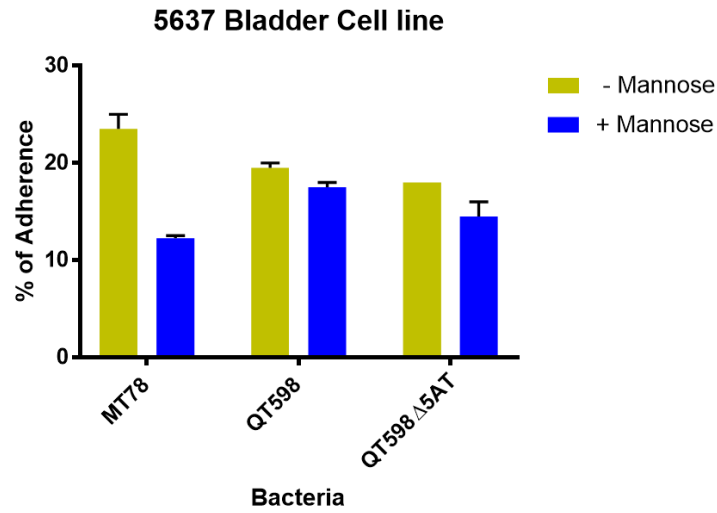


Figure S5. No difference in percentage of adherence to human bladder (5637) epithelial cell line by QT598 and its SPATEs free mutant. Cell monolayers were infected with QT598 and QT598 Δ 5ATs at a multiplicity of infection (MOI) of 10 and incubated at 37°C at 5% CO₂ for 2 hours. Adherent bacteria were enumerated by plating on LB agar. Empty vector (pBCsk+) was used as a negative-control and APEC MT78 as a positive control for adherence to cell lines. Data are the averages of three independent experiments. Error bars represent standard errors of the means.

Supplemental Table 1.

List of *E. coli* strains and accession numbers for Vat proteins identical to Vat_{QT598}

Strain/isolate	Accession number	Information or source
KTE28	ELC53090.1	UTI
UMEA 4076-1	ERA51660.1	UTI
004PP2015	OCS73020.1	Turkey colibacillosis
FEX675	PPE49807.1	Retail chicken meat potential ExPEC
VRES0524	SQW13100.1	Turkey feces
VRES0524	SQV09766.1	Turkey feces
VRES0540	SQX23694.1	Turkey feces
VRES0561	SQW81848.1	Turkey feces
VRES0558	SQW95192.1	Turkey feces
VRES0563	SQX99582.1	Turkey feces
VRES0557	SQZ03425.1	Turkey feces
VRES0564	SRZ55961.1	Turkey feces
VRES0569	SQX13605.1	Turkey feces
VRES0574	SQX78012.1	Turkey feces
VRES0567	SQX85253.1	Turkey feces
VRES0560	SRB39602.1	Turkey feces
VRES0571	SQX68131.1	Turkey feces

VRES0528	<u>SRB74388.1</u>	Turkey feces
VRES0550	<u>SRA14460.1</u>	Turkey feces
VRES0527	<u>SQV24988.1</u>	Turkey feces
VRES0541	<u>SQV90344.1</u>	Turkey feces
VRES0531	<u>SRZ02671.1</u>	Turkey feces
VRES0529	<u>SQW37114.1</u>	Turkey feces
VRES0584	<u>SQX38949.1</u>	Turkey feces
VRES0585	<u>SQY07345.1</u>	Turkey feces
VRES0533	<u>SRY84987.1</u>	Turkey feces
VRES0536	<u>SQW01088.1</u>	Turkey feces
VRES0580	<u>SQW89426.1</u>	Turkey feces
VRES0535	<u>SQW64783.1</u>	Turkey feces
VRES0668	<u>SQZ94674.1</u>	Turkey feces
VRES0568	<u>SQX57971.1</u>	Turkey feces
VRES0552	<u>SRA59934.1</u>	Turkey feces
VRES0570	<u>SQX27250.1</u>	Turkey feces
VRES0573	<u>SQY17530.1</u>	Turkey feces
VRES0581	<u>SQV73047.1</u>	Turkey feces
VRES0559	<u>SQY77711.1</u>	Turkey feces
VRES0572	<u>SQY89347.1</u>	Turkey feces

VRES0532	<u>SQW23350.1</u>	Turkey feces
SC371	<u>RDQ09216.1</u>	Surface water lake Superior
VRES0562	<u>SVF57037.1</u>	Turkey feces
VRES0586	<u>SVF61232.1</u>	Turkey feces

Supplemental Table 2.

List of *E. coli* strains that contain Tag genomic islands - (*tagB* and *tagC* genes)

Strain/isolate	Accession number	Information or source
DA33135	CP029576.1	Clinical source, Sweden
Ecol_448	CP015076.1	Clinical source, Argentina
Ecol_743	CP015069.1	Human isolate, Dubai
SE15	AP009378.1	Human commensal, B2 group (Toh <i>et al.</i> , 2010)
Ecol_745	CP015074.2	Human isolate, Morocco
Eco 889	CP015159.1	NIH, human urine
MVAST0167	CP014492.1	ST131, Minnesota
55989	CP028304.1	UTI, Pakistan
AR_451	CP030337.1	Antimicrobial resistant strain
E41-1	CP028483.1	Human sputum, Shanghai
AR_0081	CP027534.1	Antimicrobial resistant
NQ3	CP024720.1	Domestic yak
LS4	CP024717.1	Domestic yak
AR_0055	CP021935.1	Antimicrobial resistant
H105	CP021454.1	ST131, vaginal swab Germany
AR_0058	CP021689.1	Antimicrobial resistant

81009	CP021179.1	O25b:H4 urine isolate United Arab Emirates
AR_0104	CP020116.1	Antimicrobial resistant
Ecol_AZ162	CP019015.1	Antimicrobial resistant, human, Boston
Ecol_867	CP019000.1	Antimicrobial resistant, Toronto, Canada
Ecol_656	CP018979.1	Human Beijing, China
Ecol_542	CP018970.1	Vietnam
Ecol_276	CP018953.1	Human Antimicrobial resistant, U.S.A.
NCTC 13441	LT632320.1	O25b:H4 ST131 uropathogenic strain, CTX-M-15
O25b:H4	CP015085.1	Urinary tract pathogenic, Saudi Arabia
Ecol 732	CP015138.1	Human AR strain, Bangkok
JJ1887	CP014316.1	UTI, ST131 strain (Timothy J Johnson <i>et al.</i> , 2016)
ZH193	CP014497.1	Human ST131, New York
JJ2434	CP013835.1	Human Minneapolis
CD306	CP013831.1	ST131 from a Cat, New York
P46212	CP013658.1	Oxford, U.K. human urine (Stoesser <i>et al.</i> , 2016)
EC958	HG941718.1	UTI, blood, England (Forde <i>et al.</i> , 2014)
JJ1886	CP006784.1	USA, fatal sepsis (Andersen <i>et al.</i> , 2013)

Z247	CP021207.1	Human blood, China (Zheng <i>et al.</i> , 2016)
Ecol 244	CP019020.1	Argentina, human clinical
G749	CP014488.1	Human clinical, Seattle Washington
MNCRE44	CP010876.1	Human ST131 sputum (Timothy J Johnson <i>et al.</i> , 2015)
ZH063	CP014522.1	Human ST131, Winnipeg, Canada
Ecol AZ159	CP019008.1	Human Bogota, Columbia Human ST131, Minneapolis
SaT040	CP014495.1	Human ST131, Burlingtonm, VT
JJ1897	CP013837.1	Human ST131, Minneapolis
PCN033	CP014488.1	O11 group D, PorcineExPEC (Liu <i>et al.</i> , 2015)
FHI40	LM996283.1	Non-O157 STEC human Norway
RM9387	CP009104.1	O104 STEC, cattle feces (Yan <i>et al.</i> , 2015)
M18	CP010219.1	Mouse feces, China
CI5	CP011018.1	Uropathogenic strain, pyelonephritis (Mehershahi <i>et al.</i> , 2015)
M3	CP010183.1	Mouse feces, China
M1	CP010180.1	Mouse feces, China
M8	CP010191.1	Mouse feces, China

References

1. Toh, H., et al., *Complete genome sequence of the wild-type commensal Escherichia coli strain SE15, belonging to phylogenetic group B2*. Journal of bacteriology, 2010. **192**(4): p. 1165-1166.
2. Johnson, T.J., et al., *Complete genome sequence of a CTX-M-15-producing Escherichia coli strain from the H30Rx subclone of sequence type 131 from a patient with recurrent urinary tract infections, closely related to a lethal urosepsis isolate from the patient's sister*. Genome announcements, 2016. **4**(3): p. e00334-16.
3. Stoesser, N., et al., *Evolutionary history of the global emergence of the Escherichia coli epidemic clone ST131*. MBio, 2016. **7**(2): p. e02162-15.
4. Forde, B.M., et al., *The complete genome sequence of Escherichia coli EC958: a high quality reference sequence for the globally disseminated multidrug resistant E. coli O25b: H4-ST131 clone*. PLoS One, 2014. **9**(8): p. e104400.
5. Andersen, P.S., et al., *Complete genome sequence of the epidemic and highly virulent CTX-M-15-producing H30-Rx subclone of Escherichia coli ST131*. Genome announcements, 2013. **1**(6): p. e00988-13.
6. Zheng, B., et al., *Coexistence of MCR-1 and NDM-1 in clinical Escherichia coli isolates*. Clinical Infectious Diseases, 2016. **63**(10): p. 1393-1395.
7. Johnson, T.J., et al., *Complete genome sequence of a carbapenem-resistant extraintestinal pathogenic Escherichia coli strain belonging to the sequence type 131 H30R subclade*. Genome announcements, 2015. **3**(2): p. e00272-15.
8. Liu, C., et al., *Genome analysis and in vivo virulence of porcine extraintestinal pathogenic Escherichia coli strain PCN033*. BMC genomics, 2015. **16**(1): p. 717.
9. Yan, X., et al., *Genome sequencing and comparative genomics provides insights on the evolutionary dynamics and pathogenic potential of different H-serotypes of Shiga toxin-producing Escherichia coli O104*. BMC microbiology, 2015. **15**(1): p. 83.
10. Mehershahi, K.S., S.N. Abraham, and S.L. Chen, *Complete genome sequence of uropathogenic Escherichia coli strain CI5*. Genome announcements, 2015. **3**(3): p. e00558-15.

11 ANNEXE II : FIGURES SUPPLEMENTAIRES DE L'ARTICLE 2

Habouria et al. Supplementary Fig. S1 New group of P-like (PL) fimbriae in *E. coli*

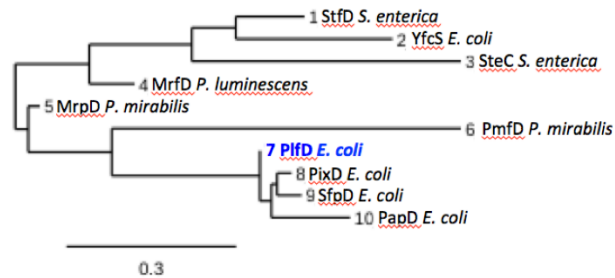


Fig. S1. Pap-Like fimbriae (Plf) Phylogram using the chaperone proteins with other members of the π -fimbriae group. Phylogram using sequences of fimbrial chaperone proteins belonging to the π -fimbriae clade based on (Nuccio *et al.*, 2007). The PlfD protein, shown in blue, clusters with other fimbriae showing closer identity to P fimbriae including Pix, Sfp, and Pap. Sequences used were UniProt (uniprot.org): **1-** O87659, **2-** P77599, **3-** Q8Z454, **4-** Q93MT3, **5-** Q51905, **6-** P53520, **7-** Genbank-NCBI: AKG46876.1, **8-** A0A454A7I7, **9-** Q933Y4, **10-** A0A454AB62. Tree was generated using Phylogeny.fr Platform (phylogeny.fr) (see methods).

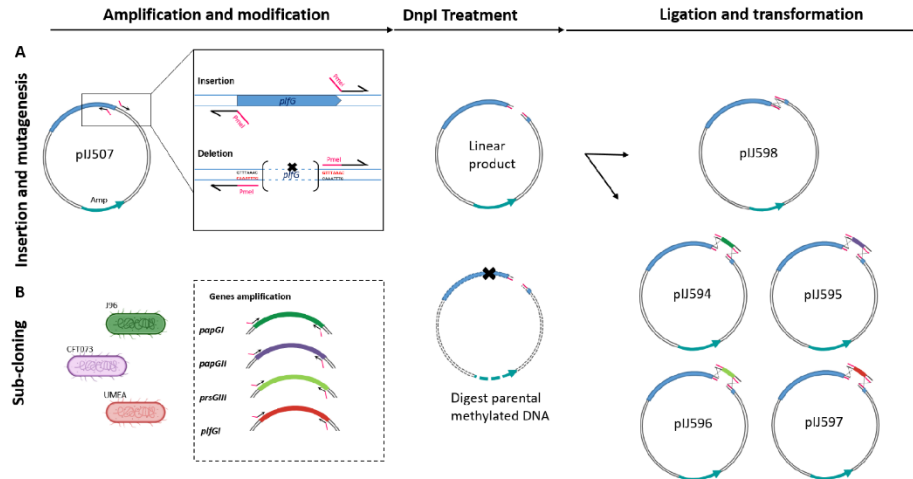
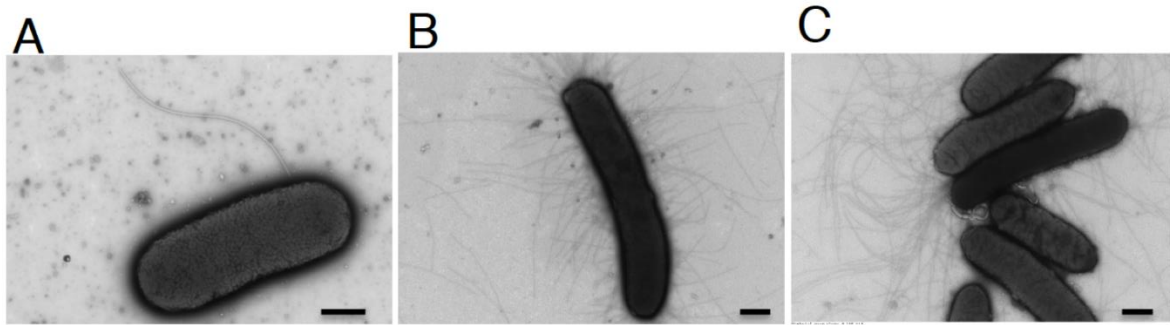


Fig. S2. Cloning of different chimeric Pif fimbriae. **A.** Schematic representation of the mutation of the adhesin *plfG* protocol that consisting of an inverse PCR reaction, using specific primers with PmeI ends (pink), followed by DpnI digestion, ligation and transformation. **B.** Subcloning involves the amplification of different class of *papG* along with *plfGI* and incorporation into the $\Delta plfG$ background.

Supplementary Fig. S3



Supplementary Figure S3. Transmission Electron Microscopy (TEM) of bacterial cells containing plasmids encoding different *plf* and *pap* fimbrial genes. The *fim*-K-12 strain ORN 172 was used as the recipient strain. A) Deletion of the *plfG* gene from plasmid pIJ507 (*plfBAHCDJKEFG*)=pIJ598 (*plfBAHCDJKEF*) resulted in lack of production of fimbriae at the surface. B) Replacement of the *papG* class II adhesin in pDC1 with the *plfG* class II gene resulted in a chimeric fimbriae (*papAHCDJKEF::plfG*) produced by bacteria (pIJ602) C) Higher magnification of ORN172 containing pIJ602. Although the hybrid fimbriae could be produced at the cell surface, these fimbriae did not demonstrate any functional hemagglutination capacity. Bar= 500 nm

Supplementary Fig. S4

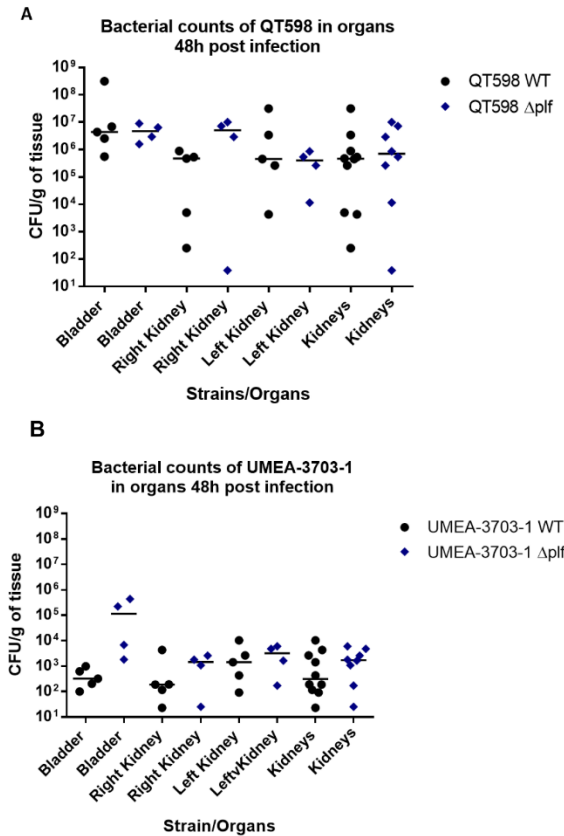


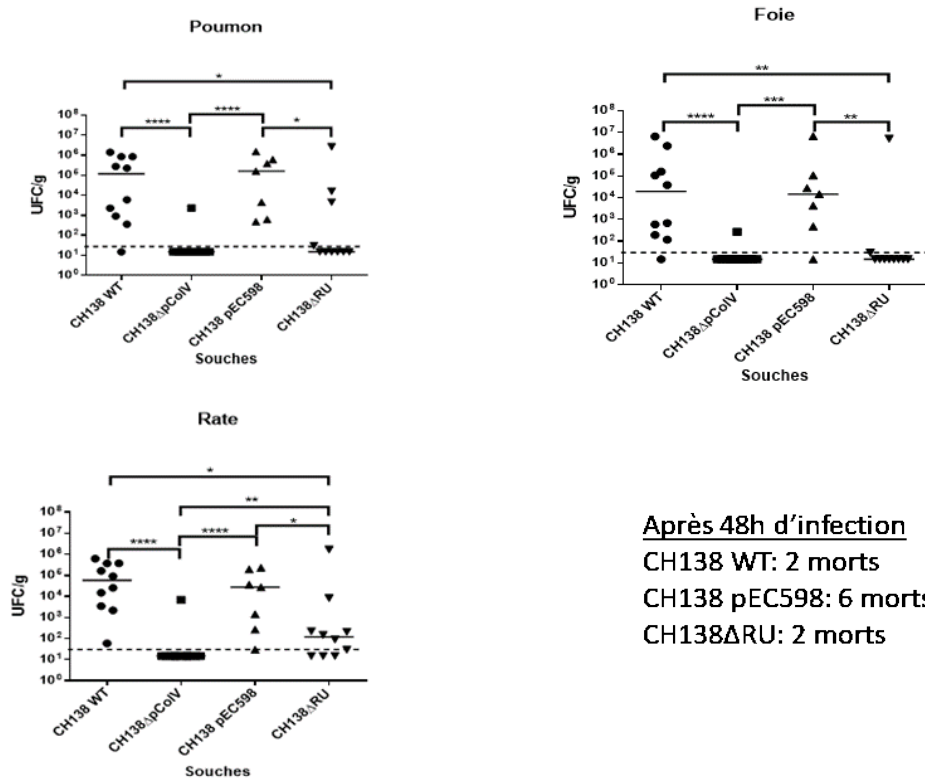
Fig. S4. Comparison of UTI colonization in single strain-infection model was performed to compare wild-type strain QT598 (A) and UMEA-3703-1 (B) to the Δ plf mutant. 48h post-infection, CBA/J female mice were euthanized and bladder and kidneys were harvested for colony counts. There was no significant difference in colonization of the bladder or kidneys (Data are means \pm standard errors of the means of 5 mice. Mann-Whitney Test was used for statistical analysis).

Supplemental Table 3:**Hemagglutination titer of different clones expressing P and P-like (PL) fimbriae***

Erythrocyte:	Sheep	Bovine	Pig	Horse	Dog	Human (A and O)	Rabbit	Chicken	Turkey
Clones									
QT5722- PapGI	-	-	7	-	4.5	4.5	6	-	-
QT5723- PapGII	-	-	5.5	-	-	6	-	2	-
QT5724- PrsGIII	6	-	3	3	6	2.5	1	-	-
QT5721- PlfG-II	-	-	-	-	1	6.5		1	6
QT4741- PlfG- I	5.5	5	2	5	1	3	5	2	2.5
QT5727- PlfΔplfG	-	-	-	-	-	-	-	-	-
QT5732- Vector only	-	-	-	-	-	-	-	-	-
Chemirec clones									
QT5728- Plf::PapGI	-	-	5.5	-	-	4	-	-	-
QT5729- Plf::PapGII	-	-	4	-	-	3	-	-	-
QT5730- Plf::PrsGIII	4	1.5	1.5	-	-	2	-	-	-

*Numbers are the highest average titers showing hemagglutination (HA) of a in a microtiter assay after duplicate assays were performed as described in the methods. Titres of 2 or less may be background aggregation. (-) indicates no HA was observed.

12 ANNEXE III



Colonisation d'organes internes de poulets de 3 semaines par *Escherichia coli* CH138, pathogène aviaire et ses mutants.

Après 48h d'infection, les animaux ont été euthanasiés et autopsiés. Les organes (poumon, foie et rate) ont été récupérés, homogénéisés puisensemencés sur géloses MacConkey pour le dénombrement des bactéries. Les résultats sont exprimés en unités formant des colonies (CFU) par gramme de tissu. Les résultats montrent que les nouveaux gènes qui se trouvent au niveau de la région unique augmentent considérablement la virulence de la souche APEC CH138. La barre horizontale représente la médiane des résultats (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Mann-Whitney Test).

Légende :

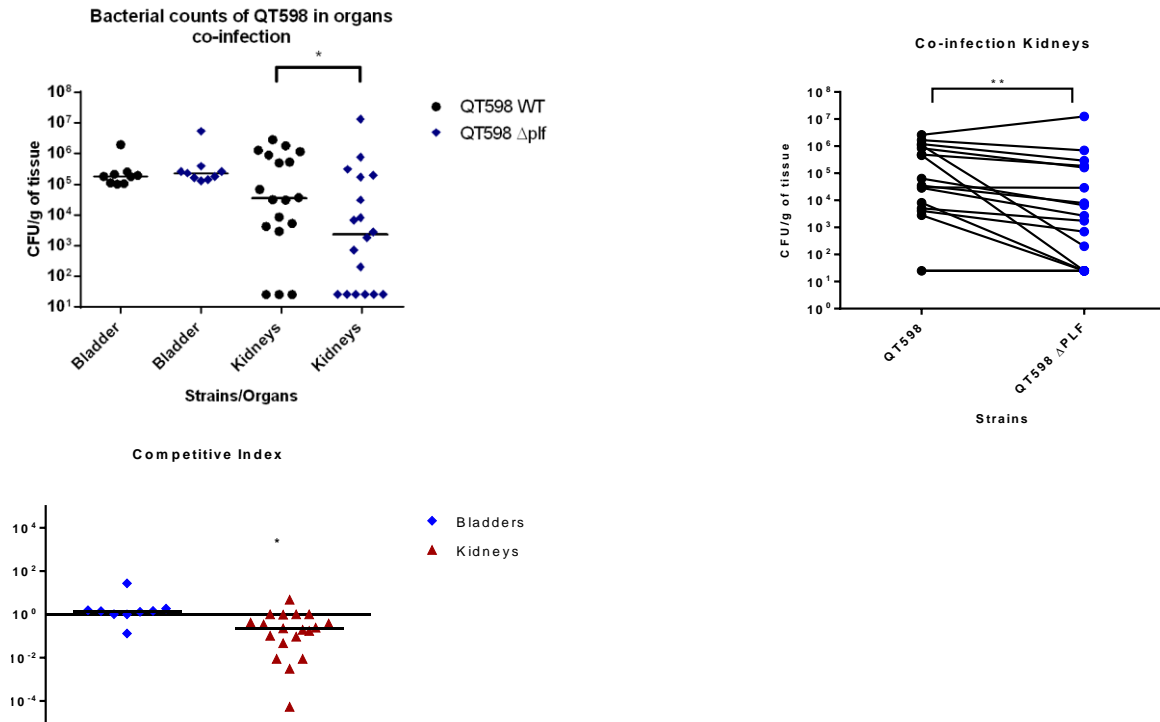
CH138 WT= Souche APEC de référence (sauvage).

CH138 Δ pColV= CH138 sans son plasmide de virulence.

CH138 pEC598= CH138 avec le plasmide de virulence pEC598 de la souche QT598.

CH138 Δ RU= CH138 avec le plasmide de virulence pEC598 où les gènes de la région unique ont été mutés.

13 ANNEXE IV



Role of PL fimbriae in the murine model.

Co-infections experiments were performed between the QT598 Δ lac and the Δ plf mutant. The Δ plf mutant colonized the bladder at similar levels to the wild-type (Δ lac) strain; however, the Δ plf strain was outcompeted in the kidneys by over 10-fold (Data are means \pm standard errors of the means of 10 mice (* $p < 0.05$, ** $p < 0.01$, Mann–Whitney Test).

LISTE DE PUBLICATIONS

Articles publiés

Three new serine-protease autotransporters (SPATEs) from extra-intestinal pathogenic *Escherichia coli* and combined role of SPATEs for cytotoxicity and colonization of the mouse kidney.

Hajer Habouria, Pravil Pokharel, Segolène Maris, Amélie Garénaux, Hicham Bessaiah, Sébastien Houle, Frédéric J. Veyrier, Stéphanie Guyomard-Rabenirina, Antoine Talarmin, and Charles M. Dozois

Serine protease autotransporters of the *Enterobacteriaceae* (SPATEs): Out and about and chopping it up

Hajer Habouria, Pravil Pokharel, Hicham Bessaiah, Charles M. Dozois

A newly identified group of P-like (PL) fimbriae from extra-intestinal pathogenic *Escherichia coli* (ExPEC) encode distinct adhesin subunits and mediate adherence to host cells.

Hajer Habouria, Amélie Garénaux, Sébastien Houle et Charles M. Dozois

yqhG Contributes to Oxidative Stress Resistance and Virulence of Uropathogenic *Escherichia coli* and Identification of Other Genes Altering Expression of Type 1 Fimbriae

Hicham Bessaiah, Pravil Pokharel, **Hajer Habouria**, Sébastien Houle and Charles Martin Dozois

The RyfA small RNA serves as a master regulator of stress responses and virulence in uropathogenic *Escherichia coli*

Hicham Bessaiah; Pravil Pokharel; Hamza Loucif; Merve Kulbay; Charles Sasseville; **Hajer Habouria**; Sébastien Houle; Jacques Bernier; Éric Massé; Julien Van Grevenynghe; Charles Dozois